different names (e.g., *intensity fluctuation spectroscopy*, *quasi-elastic light scattering*, etc.) are used in the literature depending on what is measured and how it is measured (the *diffusing wave spectroscopy* [DWS] mentioned below is an example of such a variation). However, we focus on the most essential (and standard) in this chapter and leave the rest to advanced books cited below in this chapter.

The techniques discussed in this chapter generally assume that multiple scattering is negligible (i.e., each photon is scattered only once as it passes through the sample). Multiple scattering is a problem in this sense. However, one recent advance in light scattering turns this problem into an opportunity and models the random path of a multiply scattered photon as a random walk (see Chapter 2). This technique, known as *diffusing wave spectroscopy*, is highlighted in Vignette V as an example of recent advances that take the basic idea of light scattering one step (or many steps!) further from what we discuss in this chapter.

Multiple scattering is also avoided to a large extent when x-rays or neutrons are used in place of light. In fact, many of the concepts we discuss in this chapter on the use of light scattering carry over to the scattering of x-rays and neutrons. Moreover, x-rays and neutrons can probe shorter length scales because of their smaller wavelengths. As a result, x-ray and neutron scattering have become valuable adjuncts to light scattering in colloid science. Although a full discussion of x-ray and neutron scattering is beyond our scope, where appropriate we have pointed out their similarities and complementary nature.

The theory of light scattering has changed rapidly in recent years. A general relationship for the intensity of light scattered by a spherical particle was derived by Lorenz in the latter part of the 19th century and applied to colloids by Mie in 1908. Some quotations from two of the references at the end of this chapter give an indication of the historical development of light scattering since Mie's complicated theory was presented. In a book published in 1956, Stacey remarks that scattering patterns "have been tabulated in only a few cases because the computation is so laborious" (p. 56). In his 1969 book, Kerker writes of the same patterns that so many have been published "that these can hardly be coped with in the usual tabular form, much less published in the normal way" (p. 77).

During the period of slightly more than a decade between these books, the computer arrived on the scene. Things have not been the same since! A corresponding level of rapid advances has occurred in instrumental capabilities and design with the advent of lasers, fiber-optic probes, and digital electronics (see Zare et al.'s 1995 work for an introduction to the use of lasers in applications of interest here). The availability of lasers makes measurement of very small frequency shifts possible, and fiber-optic probes provide ways to limit sample volumes and to minimize multiple scattering in concentrated systems so that the theories discussed in this chapter can be used without modification. Similarly, advances in digital electronics and software have led to fast and powerful correlators for dynamic analysis (see Sections 5.4 and 5.8).

Our concern in this chapter is not primarily with complicated systems that require elaborate calculations, although we briefly discuss several specific examples of such systems. Instead, we focus attention on systems for which simplifying approximations to the general theory can be applied.

VIGNETTE V STRUCTURE AND STRUCTURAL TRANSITIONS IN DENSE SYSTEMS: Single and Multiple Scattering— The Art and Science of Seeing the Invisible

Standard static and dynamic light scattering methods assume that there is very little multiple scattering by the particles, that is, the dispersion has to be sufficiently dilute so that the photons are scattered only once as they pass through the sample. Is there a way to "look" inside a dispersion that is cloudy or milky, such as a foam, and to extract information on the local structure and its kinetics and relaxation? Or, is it possible to "tailor" a dispersion so that

some particles may be selectively made "visible," while others are invisible (and therefore do not scatter light)?

New innovations in theoretical analysis and surface chemistry provide affirmative answers to the above questions and have opened new ways to study interactions in colloidal systems.

For example, an important variant of the techniques discussed in this chapter, known as *diffusing wave spectroscopy* (DWS), takes a conceptually different view of multiple scattering in concentrated systems such as foams and turns it into an asset. The traditional methods we discuss in this chapter require single scattering so that shifts in frequency and phase in the scattered light can be related to information about the particles. In contrast, in DWS the tortuous path of a multiply scattered photon (see Figure 5.1) is approximated by a random walk, very much the same way we have modeled the random walk of a diffusing particle in Chapter 2! For instance, in the case of foams, the static transmission of light, modeled through the solution of the diffusion equation (with the diffusion coefficient written in terms of the mean free path using classical results), can be used to obtain the average bubble size and its variation as the foam coarsens (Weitz and Pine 1993). The fluctuations in the intensity of the diffusing wave provides, with a corresponding diffusion-based analysis, information on the dynamics.

In contrast to traditional dynamic light scattering (which, as we discuss in Section 5.8, probes the continuous random motion of the particles), DWS here probes spatially localized, temporally intermittent events that are characterized by relatively larger length scales. Despite this, the basic quantity measured ($g_1(t_d)$ in the notation of Section 5.8; there is no angular dependence in DWS because of multiple scattering) is essentially the same, only the method of analysis is different.

Even the traditional methods discussed in this chapter can be used for concentrated dispersions through contrast matching. For example, silica particles coated with silane coupling agents in a refractive index-matched mixture of ethanol and toluene can be used in combination with visible "probe" particles to study the dynamics of particles in dense systems. In the case of microemulsions (Chapter 8), selective deuteration of a component (oil, water, or surfactant) can be used in neutron scattering experiments even to measure the curvature of the oil-water interface.

Scattering techniques have thus become a powerful and indispensable tool for providing both routine information such as molecular weight and size and detailed structural information previously considered inaccessible. This chapter introduces the basic concepts (although restricted largely to light scattering) necessary for understanding the standard, as well as the emerging, uses of scattering in colloid science.

5.1c Focus of This Chapter

The primary focus of this chapter is on classical, static light scattering since a good understanding of this is essential for many of routine uses, such as determination of molecular weight and second virial coefficient. Static light scattering also sets the stage for the introduction of dynamic light scattering, as well as other forms of radiation scattering.

1. First we focus on static scattering. We begin with a review of some basic concepts from the theory of electromagnetic radiation (Section 5.2) since a background in this area and in the interaction of radiation with matter forms the first step in understanding scattering techniques. This review, however, is restricted to interaction of light with matter, although the chapter does discuss, briefly, the use of x-ray and neutron scattering in colloid science.

2. Next, we introduce the theory of Rayleigh scattering (Section 5.3), the first of many models covered in the chapter. The Rayleigh theory for dilute systems and solutions is developed here, with illustrative examples of the determination of molecular weight and the second virial coefficient. This is followed by a brief description of some of the basic experimental considerations and an introduction to absorbance and turbidity (Section 5.4).

3. Section 5.5 moves on to an extension of the Rayleigh theory essential for colloid science, namely, the Debye theory for particles of the order of the wavelength of the radiation source. The important concept of interference effects, the form factor, the Zimm plot, and

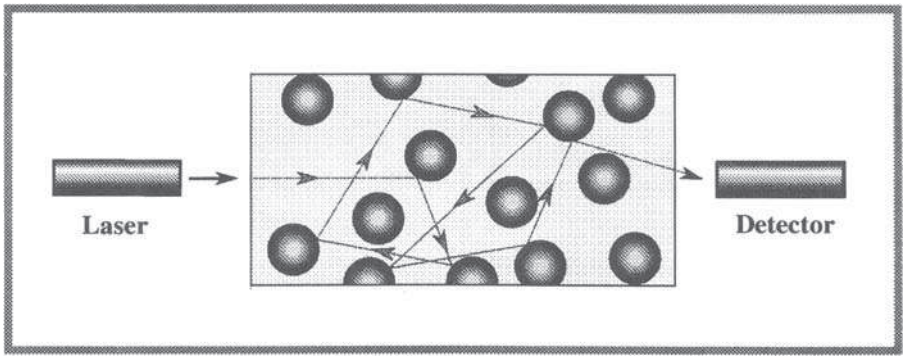


FIG. 5.1 Multiple scattering is viewed as a random walk of the photon in diffusing wave spectroscopy (DWS).

the dissymmetry ratio are introduced. Examples of the use of the form factor (often known as the *intra*particle structure factor) and its limiting behavior and the extension of the interference concept to the *inter*particle structure factor and its uses are introduced in Section 5.6. Since a combination of light scattering and x-ray or neutron scattering is needed for the latter, Section 5.6 also serves as a natural place to compare these complementary scattering techniques.

4. A brief treatment of scattering by large, absorbing particles and the concept of absorption and scattering cross sections are presented in Section 5.7 along with two examples of applications of the Mie theory (to absorbing, but small, particles) and a discussion of Tyndall spectra.

5. The final section (Section 5.8) introduces dynamic light scattering with a particular focus on determination of diffusion coefficients (self-diffusion as well as mutual diffusion), particle size (using the Stokes-Einstein equation for the diffusion coefficient), and size distribution.

This chapter is designed to provide the basic know-how for using light scattering and to set up the foundation needed for learning more advanced concepts and recent developments. For more advanced material, a number of advanced monographs (some containing state-of-the-art reviews), textbooks, and research publications containing details and new applications that are beyond the undergraduate level have been cited throughout the chapter.

5.2 INTERACTION OF RADIATION WITH MATTER

The phenomena with which we are concerned in this chapter are displayed by the entire spectrum of electromagnetic radiation. For those applications done in the visible part of the spectrum, the common designation *light* scattering is used. Visible light shares a variety of parameters and descriptive relationships with other regions of the electromagnetic spectrum. The purpose of this section is to examine briefly some of the characteristics of electromagnetic radiation, particularly those that are needed for an understanding of light scattering.

5.2a Elements of the Theory of Electromagnetic Radiation

Electromagnetic radiation consists of oscillating electrical (**E**) and magnetic (**H**) fields that are perpendicular to each other and perpendicular to the direction of propagation of the wave, as shown in Figure 5.2. Under vacuum, the velocity of propagation of an electromagnetic wave c is about $3 \cdot 10^8$ m s^{-1} and is independent of the wavelength of the radiation. The frequency ν, wavelength λ_0, and velocity of the radiation are related through the familiar equation

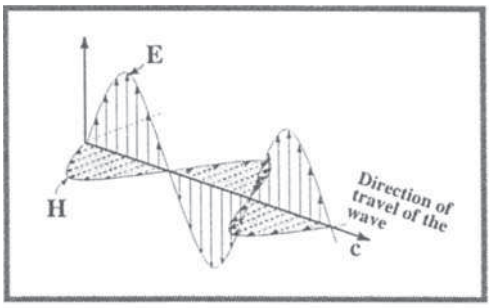


FIG. 5.2 The relationship between the electric and magnetic fields and the direction of propagation of electromagnetic radiation.

$$c = \lambda_0 \nu \tag{1}$$

If the radiation is passing through a medium other than a vacuum, its velocity and wavelength are both diminished by a factor n, the refractive index of the medium. Then Equation (1) becomes

$$\frac{c}{n} = \frac{\lambda_0}{n} \nu \tag{2}$$

or

$$v = \lambda \nu \tag{3}$$

We use the symbols c and λ_0 to refer, respectively, to the velocity and wavelength under vacuum only. In this chapter the symbol λ, without a subscript, always refers to the wavelength of the radiation in the medium.

5.2a.1 Coulomb's Law

Our discussion of light scattering centers on the oscillations of the electric field. Before turning to the oscillating aspect of the field, let us first review a few points about electric fields per se. We begin by retreating to Coulomb's law, which states that the force **F** between two charges q_1 and q_2 that are separated by a distance r is proportional to $q_1 q_2 / r^2$. In SI units, the q's are measured in coulombs, r in meters, and **F** in newtons; the proportionality factor in Coulomb's law must be dimensionally consistent. For charges under vacuum, it is traditional to write the proportionality factor as $1/(4\pi\varepsilon_0)$, where ε_0 is called the *permittivity of vacuum*. Thus Coulomb's law for **F** along the r-direction ($\mathbf{i}_r$ = unit vector) is written as

$$\mathbf{F} = \frac{1}{4\pi\varepsilon_0} \frac{q_1 q_2}{r^2} \mathbf{i}_r \tag{4}$$

where the permittivity of vacuum ε_0 is equal to $8.854 \cdot 10^{-12}$ C^2 N^{-1} m^{-2}. (The SI units C^2 N^{-1} m^{-2} are equivalent to C^2 J^{-1} m^{-1} or kg^{-1} m^{-3} s^2.) The proportionality constant $1/(4\pi\varepsilon_0)$ is equal to $8.988 \cdot 10^9$ N m^2 C^{-2}. The older literature uses cgs units, for which the proportionality constant between **F** and $q_1 q_2 / r^2$ equals 1.00 dyne cm^2 $(esu)^{-2}$, where the esu is the electrostatic unit of charge defined to make this proportionality factor equal to unity.

If the charges are embedded in a medium, the electrical properties of the intervening molecules decrease the force from the value calculated by Equation (4). The relative dielectric constant of the medium ε_r measures this effect quantitatively. In surroundings other than a vacuum, the force between two charges is given by

$$\mathbf{F} = \frac{1}{4\pi\varepsilon_0\varepsilon_r} \frac{q_1 q_2}{r^2} \mathbf{i}_r \tag{5}$$

Note that the force is a vector. Its direction is along the line connecting the two charges, but Equation (5) specifies only the magnitude of the force. It is important to remember several aspects of Equation (5):

1. The relative dielectric constant ε_r is dimensionless. This quantity is also known as the *relative permittivity*.
2. The product $\varepsilon_0\varepsilon_r$ is sometimes written ε (without subscripts), and ε has the units of ε_0.
3. The relative permittivity ε_r appears in the denominator of Coulomb's law even when cgs units are used.

5.2a.2 Electric Field

Next, let us apply these ideas to the electric field. By definition, an electric field **E** describes the force experienced by a unit test charge $q_t = 1$. Thus, if we let one of the charges in Equation (5) equal q_t the following expression is obtained for the field associated with the remaining charge q:

$$\mathrm{E} = \frac{1}{4\pi\varepsilon}\frac{q}{r^2} \tag{6}$$

The electric field, being a vector, has a direction as well as a magnitude; the direction of the field is perpendicular to the direction of propagation of the field.

The dominant characteristic of the electrical and magnetic fields that comprise electromagnetic radiation is their periodically oscillating nature, a fact that enables us to describe them by the mathematics of waves. For light scattering, it is the electric field that is of interest. The oscillating nature of an electric field propagating in the positive x direction is described by the equation

$$\mathbf{E} = \mathbf{E}_0 \cos\left[2\pi\left(\nu t - \frac{x}{\lambda}\right)\right] \tag{7}$$

in which $\mathbf{E}_0$ is a vector with a magnitude equal to the maximum amplitude of the field. Since we have postulated the x direction as the direction of propagation, the electric field lies in the yz plane and may, in general, be resolved into y and z components since **E** is a vector. Both the y and z components of the field are described by Equation (7) when the equation is modified by the inclusion of phase angles. This is because the two components need not be in phase with each other. Accordingly, we write

$$E_y = E_{0y} \cos\left[2\pi\left(\nu t - \frac{x}{\lambda}\right) + \delta_y\right] \tag{8}$$

and

$$E_z = E_{0z} \cos\left[2\pi\left(\nu t - \frac{x}{\lambda}\right) + \delta_z\right] \tag{9}$$

in which the δ terms are the phase angles of the two components.

5.2a.3 Polarization of Light

In the most general case, the above two equations mean that the electric field vector traces an ellipse in the yz plane. There are two special cases of note in this general situation. If the phase difference between the two components of the field ($\delta_y - \delta_z$) is zero or some integral multiple of π, the ellipse flattens to a line. If the phase difference is $\pi/2$ or any odd integral multiple of $\pi/2$ and the amplitudes of the two components are equal, the ellipse is rounded to a circle. In the former case we speak of the radiation as being *plane polarized*, and in the latter case as being *circularly polarized*.

Ordinary light is said to be unpolarized. This last term is somewhat unfortunate because all light displays some form of polarization. In ordinary light, however, all forms of polarization are present, so the individual effects cancel out. The use of various filters makes it possible to conduct experiments with radiation that show a unique state of polarization. Polaroid filters, for example, transmit plane-polarized light only. In discussing light scattering, we are concerned primarily with unpolarized light and occasionally with plane-polarized light.

An interesting example of polarization arises in the study of light reflected from a surface. Suppose we consider a beam of light incident on the planar surface of some material having a higher refractive index than the medium from which the beam approaches. At the surface some of the light will be refracted, and some will be reflected. Figure 5.3 illustrates this for the case in which the reflected and the refracted beams are separated by an angle of 90°.

In this situation, two very different results are obtained, depending on whether the incident light is linearly polarized in the plane of the figure or perpendicular to it. If the light is polarized in the plane of the figure, no light will be reflected at all. On the other hand, if it is initially polarized perpendicular to the plane of the figure, it will be reflected with the same polarization. If ordinary light—a mixture of the two types of light just mentioned—is used for the incident radiation, only one of the components contributes to the reflection. Furthermore, the reflected light is polarized perpendicular to the plane of the figure. Polaroid filters are used in photography and in sunglasses to reduce the glare of reflected light since this light is polarized.

This behavior is not observed uniquely when the angle between the two beams is 90°; rather, the intensity of the reflected beam varies continuously with the angle. At 90°, however, the polarization effect is most pronounced. We shall see that some scattering phenomena also show an angular dependence, as well as the fact that scattered light displays maximum polarization at 90°.

5.2b Interaction Between an Electric Field and a Charge

In this section we discuss the interaction between an electric field and a charge that is free to move with the field. Such a charge experiences a force that accelerates it with the field. If the field is oscillating, the acceleration of the charge will also oscillate. One of the basic results of classical electromagnetics is that the acceleration of a charge leads to the emission of radiation.

It was the apparent violation of this requirement that led to the postulate of quantization in the Bohr theory of the hydrogen atom. However, we are concerned here with the classical result in which the charge does radiate. Our objective is to describe the emitted radiation some distance r from the emitter.

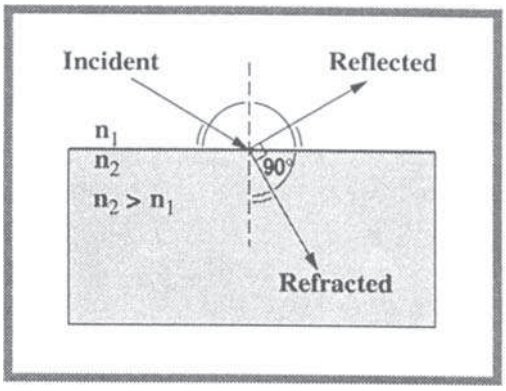


FIG. 5.3 The relationship between the incident, reflected, and refracted beams of radiation at a plane surface.

5.2b.1 Emission of Radiation by a Charge

The radiation emitted by an oscillating charge may be described by its electric field vector, which is given by

$$\mathbf{E} = \frac{q\,\mathbf{a}_p \sin \phi_z}{4\pi\varepsilon_0 c^2 r} \tag{10}$$

according to electromagnetic theory. In this equation, q equals the magnitude of the charge, $\mathbf{a}_p$ is its periodic acceleration, and c is the velocity of light. The coordinate system defined by Figure 5.4a will help describe this field. The origin of the coordinates is located at the emitting charge. The angles between the line of sight—along which r is measured—and the x, y, and z axes are designated ϕ_x, ϕ_y, and ϕ_z, respectively.

The oscillating charge behaves like an antenna, and Equation (10) describes the field of such an antenna as long as r is large compared to the wavelength of the radiation that induces the oscillation. It should also be noted that the antenna to which Equation (10) applies is aligned vertically (z axis) and is therefore "driven" by vertically polarized radiation.

We are using Equation (10), presented without proof, as the point of departure for our discussion of light scattering. Therefore, it is important that we find its predictions reasonable.

1. First, let us consider the plausibility of the $\sin \phi_z$ factor. This factor ranges between 0 and 1 as ϕ_z varies from 0 to $\pi/2$. This means that the maximum field will be observed at right angles to the oscillating charge, and no field will be observed along the axis of the oscillation. It is the projection of the acceleration in the plane perpendicular to the line of sight that induces the field. The strength of the field is proportional to this projection at any location, as shown in Figure 5.4b. This factor describes the entire angular dependence of the induced field produced by a vertical driving field. Since it depends on the angle ϕ_z alone, the induced field is seen to be symmetrical with respect to the z axis.

2. Next we note that the induced field varies inversely with r. It makes sense that the field should decrease as we get farther from the antenna, but the inverse first-power dependence may be unexpected since we are more familiar with inverse-square laws. However, it is the energy or intensity of the light that varies according to an inverse-square law. In the next section we convert this expression for $\mathbf{E}$ to an expression for energy; the more familiar r^{-2} functionality appears then.

3. Finally, it is sufficient for our purposes to think of the remaining factors in Equation (10), $(4\pi\varepsilon_0 c^2)^{-1}$, as providing dimensional consistency to the expression. Taking a look at the SI units of the right-hand side of Equation (10), we obtain

$$\frac{(C)(\mathrm{m\,s^{-2}})}{(C^2 N^{-1} \mathrm{m^{-2}})(\mathrm{m\,s^{-1}})^2 (\mathrm{m})} = \frac{N}{C}$$

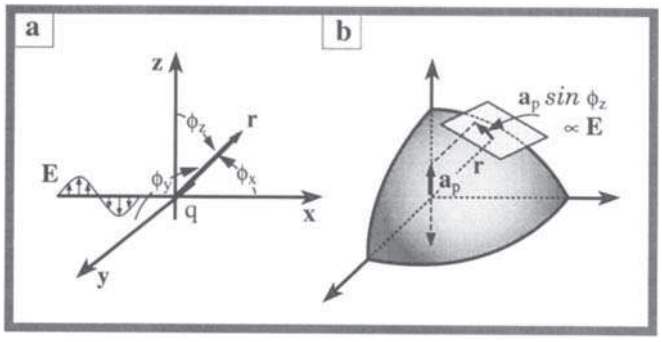


FIG. 5.4 Coordinates and acceleration relevant to the interaction of an electric field with a charge: (a) the coordinates of an electric field $\mathbf{E}$ relative to an oscillating charge located at the origin; (b) projection of the acceleration in the plane perpendicular to the line of sight.

Newtons coulomb $^{-1}$ are units of force per charge as required. Note that multiplication of both the numerator and denominator of this dimensional expression converts N C $^{-1}$ to V m $^{-1}$ since Nm = J and J/C = V. This last set of units for **E** is particularly useful for describing the field between electrodes and, as such, will be encountered in Chapter 12, in which we discuss electrokinetic phenomena.

Equation (10) describes the field emitted by an antenna that, in turn, is driven by another field. The oscillation of one field promotes the oscillation of a charge in the antenna, and this induces another electric field. The frequency is the same for all three.

5.2b.2 Induced Dipole Due to the Field

This description of antennas may seem more appropriate to a discussion of radio or television waves. We must realize, however, that at the molecular level dipoles behave exactly like antennas. Since molecules are made up of charged parts, a *dipole moment* μ is induced by the electric field of the radiation in any material through which radiation passes. In this discussion, the dipole moment equals the product of the effective charge displaced by the field and its distance of separation from the opposite charge. In SI, μ has units C m. We consider isotropic materials characterized by a *polarizability* α. As the name implies, this property measures the ease with which charge separation—polarity—is induced in a molecule by an electric field. For isotropic substances, the dipole moment and the field are related by the expression

$$\boldsymbol{\mu} = \alpha \mathbf{E}_0 \tag{11}$$

Using Equation (11) as the basis of a dimensional analysis of α shows α has SI units given by

$$\frac{C\,m}{N\,C^{-1}} = \frac{C^2\,m}{N}$$

The quantity $\alpha/4\pi\varepsilon_0$—which is the polarizability value used in cgs units—is informative. It is examined in Example 5.1.

* * *

EXAMPLE 5.1 *Polarizability of Particles.* Criticize or defend the following proposition on the basis of the units of $\alpha/4\pi\varepsilon_0$: The larger the volume of a particle is, the easier it is to induce polarity in that particle.

Solution: The ratio $\alpha/4\pi\varepsilon_0$ has the units $(C^2\ m\ N^{-1})/(C^2\ m^{-2}\ N^{-1}) = m^3$, which are units of volume. In fact, polarizabilities of actual molecules are on the order of 10^{-29} m^3 molecule^{-1} = 0.01 nm^3 molecule^{-1}, which is the magnitude of molecular volumes. For individual atoms, we expect the polarizability to increase with atomic volume since the outermost electrons are less tightly restrained by the nucleus in such cases. Extension of this principle to covalently bonded species must be done cautiously, however, since the bonding affects the overall picture. The accuracy of the proposition, then, depends on the nature of the "particle" under consideration. Even when the principle stated in the proposition is not literally true, it offers a convenient mnemonic for the definition of polarizability. ■

* * *

5.2b.3 Electric Field Emitted by a Charge

If we imagine the molecule to lie at the origin of a coordinate system so that $x = 0$, we may substitute Equation (7) into (11) to obtain

$$\boldsymbol{\mu} = \alpha \mathbf{E}_0 \cos(2\pi\nu t) \tag{12}$$

A dipole moment may be regarded as the product of the distance ξ between two charges and the magnitude of the charge q. A useful way of looking at Equation (12) is to identify the charge as

$$q = \alpha E_0 \tag{13}$$

where E_0 is the magnitude of the vector $\mathbf{E}_0$, and the separation of the charges as

$$\xi = \cos(2\pi\nu t) \tag{14}$$

Then the magnitude of the periodic acceleration of the charge is given by

$$a_p = \frac{d^2\xi}{dt^2} = -4\pi^2\nu^2\cos(2\pi\nu t) \tag{15}$$

Equations (13) and (15) may now be used in Equation (10) to give

$$E = -\frac{\alpha E_0 4\pi^2 \nu^2 \cos(2\pi\nu t)\sin\phi_z}{4\pi\varepsilon_0 c^2 r} \tag{16}$$

For maximum generality, we must remember that the field is periodic in space as well as in time; hence the cosine factor in Equation (16) is corrected by analogy with Equation (7):

$$E = -\frac{\pi\nu^2 \alpha E_0 \cos[2\pi(\nu t - r/\lambda)]\sin\phi_z}{\varepsilon_0 c^2 r} \tag{17}$$

Equation (17) describes the induced field a distance r from the dipole.

5.3 SCATTERING BY SMALL PARTICLES: THEORY OF RAYLEIGH SCATTERING

In this section we discuss the first of several light scattering theories to be considered in this chapter, Rayleigh scattering. We shall see presently that Rayleigh scattering applies only when the scattering centers are small in dimension (i.e., when the "characteristic dimension" of the scatterers is small) compared to the wavelength of the radiation, the "yardstick" used in the measurements. As such, it is severely limited in its applicability to colloidal particles, at least when visible light is the radiation involved. Rayleigh scattering is the easiest of the scattering theories to understand, however, so it is a logical place to begin. Furthermore, we shall extend its applicability to larger particles in other sections by introducing suitable correction factors.

5.3a Scattering by Single Molecules and Gases

When a beam of radiation is incident on a molecule, a certain fraction of that radiation will undergo the process described in the preceding section and be emitted by the dipole. Any light that does not interact this way will continue past the molecule along the original path. This undeviated or transmitted light will be attenuated compared to the incident light since some of the original beam of light is scattered from its initial path. Note that this attenuation has nothing to do with absorption; the effect we are considering is a classical result and does not involve transitions between quantum states.

Since Rayleigh scattering does not apply to particles in the colloidal size range, we do not present the derivation in detail; instead, Table 5.1 summarizes some key steps in the development of the *Rayleigh equation*:

$$\frac{i_s}{I_{0,u}} = \frac{2\pi^2 M}{r^2 \lambda_0^4 N_A \rho}(n-1)^2(1+\cos^2\phi_x) \tag{18}$$

in which i_s is the intensity, as measured at r and ϕ_x (see Figure 5.4), of the light scattered per unit volume by a gas of molecular weight, density, and refractive index given by M, ρ, and n, respectively. The incident light is unpolarized (subscript "u") and has an intensity $I_{0,u}$. Equation (18) was derived by Lord Rayleigh in 1871.

An interesting application of the *Rayleigh equation* is the explanation it offers for why the sky appears blue. This arises from the inverse fourth-power dependence on λ_0 for i_s. Suppose, for example, that two radiations are compared that have wavelengths that differ by a factor of 2. Then the scattered intensity of the shorter wavelength will be 16 times as great as that of the longer wavelength. Although red and blue light do not differ by quite this much in wavelength, the blue component of white light, say, sunlight, is scattered very much more

LE 5.1 Steps Involved in the Derivation of the Rayleigh Equation

	Justification	Cumulative effect
Evaluate the intensity of light scattered at ϕ_z	Intensity of light is proportional to the square of the electrical field	$i \propto \dfrac{\pi^2 \nu^4 \alpha^2 \lvert \mathbf{E}_0 \rvert^2 \cos^2[2\pi(\nu t - r/\lambda)] \sin^2\phi_z}{\varepsilon_0^2 c^4 r^2}$
Evaluate the intensity of incident light	No scattering factors needed in previous result	$I_0 \propto \lvert \mathbf{E}_0 \rvert^2 \cos^2 \left\lvert 2\pi\left(\nu t - \dfrac{r}{\lambda}\right)\right\rvert$
Evaluate the intensity ratio for light polarized in vertical plane	Ratio of preceding factors	$\dfrac{i_v}{I_{0,v}} = \dfrac{\pi^2 \nu^4 \alpha^2 \sin^2 \phi_z}{\varepsilon_0^2 c^4 r^2}$
Evaluate the intensity ratio for light polarized in horizontal plane	Replace $\sin \phi_z$ by $\sin \phi_y$ in the previous equation	$\dfrac{i_h}{I_{0,h}} = \dfrac{\pi^2 \nu^4 \alpha^2 \sin^2 \phi_y}{\varepsilon_0^2 c^4 r^2}$
Evaluate the intensity ratio for unpolarized light	Equal contributions from vertical and horizontal components	$\dfrac{i}{I_{0,u}} = \dfrac{\frac{1}{2}(i_v + i_h)}{I_{0,u}} = \dfrac{1}{2}\dfrac{\pi^2 \nu^4 \alpha^2}{\varepsilon_0^2 c^4 r^2}(\sin^2 \phi_z + \text{si}$
Replace $\sin^2 \phi_y + \sin^2 \phi_z$ with $(1 + \cos^2 \phi_x)$ and introduce λ_0 through Equation (1)	$r \cos \phi_i$ is the projection of r on the i axis; therefore, $r^2 (\cos^2 \phi_x + \cos^2 \phi_y + \cos^2 \phi_z)$ is equal to r^2	$\dfrac{i}{I_{0,u}} = \dfrac{\pi^2 \alpha^2}{2 \varepsilon_0^2 \lambda_0^4 r^2}(1 + \cos^2 \phi_x)$
Scale up for independent scatterers: Multiply by number of sites per unit volume (subscript "s" on i)	For ideal gases, this factor equals $(N_A \rho / M)$	$\dfrac{i_s}{I_{0,u}} = \dfrac{\pi^2 N_A \rho \alpha^2}{2 \varepsilon_0^2 r^2 \lambda_0^4 M}(1 + \cos^2 \phi_x)$
Use the Clausius-Mosotti equation from physical chemistry (Atkins 1994, Chapter 22) to introduce refractive index n	For n close to unity (gases), the Clausius-Mosotti equation becomes $\alpha = \dfrac{3 M \varepsilon_0}{N_A \rho}\dfrac{n^2 - 1}{n^2 + 2} \approx \dfrac{2 M \varepsilon_0}{\rho N_A}(n - 1)$	$\dfrac{i_s}{I_{0,u}} = \dfrac{2 \pi^2 M}{r^2 \lambda_0^4 N_A \rho}(n - 1)^2(1 + \cos^2 \phi_x)$

than the red. Accordingly, the sky overhead appears blue. At sunset, we see mostly transmitted light. Since the blue has been most extensively removed from this by scattering, the sky appears red at sunset. One of the early uses of the Rayleigh equation was in the determination of the value of the Avogadro's number, as illustrated through Problem 2 at the end of this chapter.

The Rayleigh theory does not apply when the scattering molecules are absorbing or when the atmosphere contains dust particles, water drops, or other particles with dimensions that are larger than ordinary gas molecules.

5.3b Rayleigh Scattering Applied to Solutions: Fluctuations

A crucial aspect of the transition from Step 6 to Step 7 in Table 5.1 is the requirement that the individual molecules are far enough apart to be treated as independent sources. This assumption is justified for gases, but in liquids the molecules are close enough together so that interference occurs between the waves emitted from different centers. As a matter of fact, there would be complete destructive interference of all scattered light in liquids if the molecules were randomly arranged and stationary. Nevertheless, pure liquids do scatter light. It is not the individual molecules that are the scattering centers in this case, but rather the small domains of compression or rarefaction that arise from fluctuations.

We saw in Chapter 2 that molecular motion results in small fluctuations in density at the molecular level. Although the average density of a liquid is a constant equaling the experimental density, there will be small transient domains within it that have densities larger or smaller than the mean value.

Liquid solutions also scatter light by a similar mechanism. In the case of a solution, the scattering may be traced to two sources: fluctuations in solvent density and fluctuations in solute concentration. The former are most easily handled empirically by subtracting a solvent "blank" correction from measurements of the intensity of light scattered from solutions. What we are concerned with in this section, then, is the remaining scattering, which is due to fluctuations in the solute concentration in the solution.

A fluctuation in the concentration of a small volume element δV of solution will result in a change in the properties of that volume element. We begin the analysis of this situation by defining δc and $\delta\alpha$ as the fluctuations in solution concentration and polarizability, respectively, in this element. The first thing that we recognize about these quantities is that their average values are zero since both positive and negative fluctuations occur.

Although their averages may be zero, the averages of their squares are not zero. Remember that similar situations were encountered in Chapter 2, Sections 2.6 and 2.7, in which we discussed particle displacements due to diffusion and segment displacements in a random coil. In both of these cases it was by considering the average values of the square of the displacements that meaningful quantities could be obtained. Similarly, the average values of δc and $\delta\alpha$ are zero, but their mean square values will be different from zero.

The equation in Step 6 in Table 5.1 shows that the intensity of scattered light depends on the square of polarizability. We conclude, therefore, that the way to adapt this equation to the scattering by solutions is to replace α^2 in the above equation by $\overline{(\delta\alpha)^2}$

$$\frac{i}{I_0} = \frac{\pi^2 \overline{(\delta\alpha)^2}}{2\varepsilon_0^2 r^2 \lambda^4} (1 + \cos^2 \phi_x) \tag{19}$$

Note that the wavelength λ of light in the medium rather than the value under vacuum is used in this expression since this is the light that reaches the scattering center.

Next we must consider how to extend this result to a unit volume of solution and how to relate $\delta\alpha$ to δc. The steps involved in these extensions are not difficult, but they are lengthy. Accordingly, we shall not develop the entire argument in detail. Instead, some of the key steps in the development along with a brief justification for each are summarized in Table 5.2. In this table, each major substitution is presented along with variations of Equation (19) that reflect the cumulative effects of all the substitutions. The first entry in the table, for example,

TABLE 5.2 Some Key Substitutions and Their Justification for the Transformation of Equation (19) to Equation (20)

Substitution	Justification	Cumulative effect on Equation (19)
$\dfrac{\overline{\delta\alpha^2}}{\text{volume}} = \dfrac{1}{\delta V}\dfrac{\overline{\delta\alpha^2}}{\text{fluctuation}}$	1. $\dfrac{1}{\delta V}$ domains of volume δV can fit into 1 cm^3 of solution	$\dfrac{i_s}{I_0} = \dfrac{\pi^2\delta\alpha^2}{r^2 2\varepsilon_0^2\lambda^4\,\delta V}(1 + \cos^2\phi_x)$
$\overline{\delta\alpha^2} = \varepsilon_0^2\delta V^2\left(2n\dfrac{dn}{dc}\right)^2\overline{\delta c^2}$	1. δV replaces $\dfrac{M}{\rho N_A}$ in Clausius-Mosotti equation (Table 5.1) 2. For $n \simeq 1$, $\alpha = \varepsilon_0\,\delta V(n^2 - 1)$ 3. $\delta\alpha \simeq \varepsilon_0\,\delta V\,2n\,dn$ and $dn = \dfrac{dn}{dc}\delta c$	$\dfrac{i_s}{I_0} = \dfrac{2\pi^2\delta V[n(dn/dc)]^2\overline{\delta c^2}}{r^2\lambda^4}(1 + \cos^2\phi_x)$
$\overline{\delta c^2} = \dfrac{k_B T}{(\partial^2 G/\partial c^2)_0}$ Subscript 0: evaluated at equilibrium	1. Taylor series (see Appendix A) expansion of G around equilibrium value: $G = G_0 + \left(\dfrac{\partial G}{\partial c}\right)_0\delta c + \dfrac{1}{2}\left(\dfrac{\partial^2 G}{\partial c^2}\right)_0\delta c^2$ 2. $\left(\dfrac{\partial G}{\partial c}\right)_0 = 0$ at equilibrium 3. $G - G_0 \simeq \dfrac{1}{2}k_B T$, since fluctuation is due to thermal energy $\left(\dfrac{1}{2}k_B T \text{ per degree of freedom}\right)$	$\dfrac{i_s}{I_0} = \dfrac{2\pi^2\delta V[n(dn/dc)]^2 k_B T}{r^2\lambda^4(\partial^2 G/\partial c^2)_0}(1 + \cos^2\phi_x)$
$\dfrac{\partial^2 G}{\partial c^2} = \dfrac{\delta V}{\overline{V}_1 c}\left(-\dfrac{\partial\mu_1}{\partial c}\right)_0$	1. $dG = \mu_1 dn_1 + \mu_2 dn_2$ 2. Since $\overline{V}_1 dn_1 = -\overline{V}_2 dn_2$, $dG = \left(\mu_2 - \dfrac{\overline{V}_2}{\overline{V}_1}\mu_1\right)dn_2$ 3. Since $M\,dn_2 = dc\,\delta V$, $\dfrac{dG}{dc} = \left(\mu_2 - \dfrac{\overline{V}_2}{\overline{V}_1}\mu_1\right)\dfrac{\partial V}{M}$ 4. By the Gibbs-Duhem equation, $\dfrac{\partial^2 G}{\partial c^2} = -\dfrac{\delta V}{M}\left(\dfrac{\overline{V}_2}{\overline{V}_1} + \dfrac{n_1}{n_2}\right)\dfrac{d\mu_1}{dc}$ 5. $c = \dfrac{n_2 M}{n_1\overline{V}_1} + n_2\overline{V}_2$	$\dfrac{i_s}{I_0} = \dfrac{2\pi^2[n(dn/dc)]^2 k_B T\overline{V}_1 c}{r^2\lambda^4(-\partial\mu_1/\partial c)_0}(1 + \cos^2\phi_x)$
$\left(\dfrac{\partial\mu_1}{\partial c}\right)_0 = -\overline{V}_1\left(\dfrac{\partial\pi_{osm}}{\partial c}\right)_0$	1. By Equation (3.21)	$\dfrac{i_s}{I_0} = \dfrac{2\pi^2[n(dn/dc)]^2 k_B T c}{r^2\lambda^4(\partial\pi_{osm}/\partial c)_0}(1 + \cos^2\phi_x)$

converts Equation (19) to an expression for the relative light scattered by a unit of volume i_s/I_0 by multiplying Equation (19) by the number of volume elements in 1 cm^3: $1/\delta V$. Although the justifications in Table 5.2 are sketchy, they provide hints that will show the interested reader how to proceed in order to develop the required relationship in detail. Only the third entry in the table, the connection between δc^2 and $\partial^2 G/\partial c^2$ requires a more elaborate proof than the qualitative justification supplied.

The cumulative result of these substitutions is to replace $\overline{(\delta\alpha)^2}$ in Equation (19) by a number of other factors, all of which are experimentally measurable:

1. The refractive index gradient (dn/dc). This is simply the local slope of a plot of the refractive index of a solution versus its concentration.
2. The concentration c of the solution. This is expressed in units as grams per volume.
3. The quantity $(\partial\pi_{osm}/\partial c)_{0,T}$ is evaluated for an equilibrium solution. This is the significance of the subscript 0. The subscript T denotes that the derivative is taken at isothermal conditions.

In view of these substitutions, Equation (19) becomes

$$\frac{i_s}{I_0} = \frac{2\pi^2 [n\,(dn/dc)]^2 k_B T c}{r^2 \lambda^4 (\partial\pi_{osm}/\partial c)_{0,T}} (1 + \cos^2\phi_x) \tag{20}$$

In Chapter 3 we developed expressions for the equilibrium osmotic pressure of a solution as a function of its concentration. Equation (3.34) may be written

$$\pi_{osm} = RT\left(\frac{c}{M} + Bc^2\right) \tag{21}$$

Since Equation (21) applies at equilibrium, we may evaluate $(\partial\pi_{osm}/\partial c)_{0,T}$ from Equation (21):

$$\left(\frac{\partial\pi_{osm}}{\partial c}\right)_{0,T} = RT\left(\frac{1}{M} + 2Bc\right) \tag{22}$$

Combining Equations (20) and (22) yields

$$\frac{i_s}{I_0} = \frac{2\pi^2 [n(dn/dc)]^2 c}{N_A r^2 \lambda^4 (1/M + 2Bc)} (1 + \cos^2\phi_x) \tag{23}$$

Before looking at the experimental aspects of light scattering, it is convenient to define several more quantities. First, a quantity known as the Rayleigh ratio R_θ is defined as

$$R_\theta = \frac{i_s r^2}{I_0 (1 + \cos^2\theta)} \tag{24}$$

where θ is the value of ϕ_x measured in the horizontal plane. According to Equation (23), the Rayleigh ratio should be independent of both θ and r. An experimental verification of this is one way of testing the applicability of the Rayleigh theory to the experimental data. Next it is convenient to identify the numerical and optical constants in Equation (23) as follows:

$$K = \frac{2\pi^2 n^2 (dn/dc)^2}{N_A \lambda^4} \tag{25}$$

With these changes in notation, Equation (23) becomes

$$R_\theta = \frac{Kc}{1/M + 2Bc} \tag{26}$$

or

$$\frac{Kc}{R_\theta} = \frac{1}{M} + 2Bc \tag{27}$$

This suggests that a plot of (Kc/R_θ) versus c should be a straight line for which the intercept and slope have the following significance:

$$\text{Intercept} = 1/M \tag{28}$$

$$\text{Slope} = 2B \tag{29}$$

Comparing Equations (28) and (29) with Equations (3.35) and (3.36) reveals that plots of (π_{osm}/RTc) versus c and (Kc/R_θ) versus c have identical intercepts, at least for monodisperse colloids (see Section 5.4d for a discussion of the average obtained for polydisperse systems), and that the slopes differ by a factor of 2, with the light-scattering results having the larger slope.

The Rayleigh ratio as defined by Equation (24) has a precise meaning, yet it is a quantity somewhat difficult to visualize physically. After we have discussed the experimental aspects of light scattering, we shall see that R_θ is directly proportional to the turbidity of the solution when turbidity is the same as the absorbance determined spectrophotometrically.

5.4 EXPERIMENTAL ASPECTS OF LIGHT SCATTERING

With the advances in highly monochromatic laser light sources and powerful desktop computers, light scattering instruments have attained a corresponding level of sophistication (see Zare et al. 1995). In general terms, all light scattering instruments contain the same components: a light source (usually a laser), a spectrometer (containing the optical components for defining the scattering angle volume), a detector (usually a photomultiplier), a signal analyzer such as a spectrum analyzer or a correlator (for dynamic measurements; see Section 5.8), and a computer with software for analysis are standard for dynamic measurements. A schematic representation of a typical arrangement is shown in Figure 5.5.

Commercial instruments vary in features and capabilities; some, for example, are restricted to a single angle (very low θ or $\theta = 45°$ or 90°), and some have capabilities for simultaneous measurements at a fixed number of different angles. Fiber-optic cables for source and detection paths can also be used for minimizing alignment problems and multiple scattering effects in concentrated dispersions, but these are usually found in specially built research instruments rather than commercial ones. Excellent introductions to most of the hardware and software features are available in specialized books (Chu 1991; Pecora 1985), and we focus only on some general background here. Zare et al. (1995), cited above, is an excellent source of relatively simple experiments meant for the beginner.

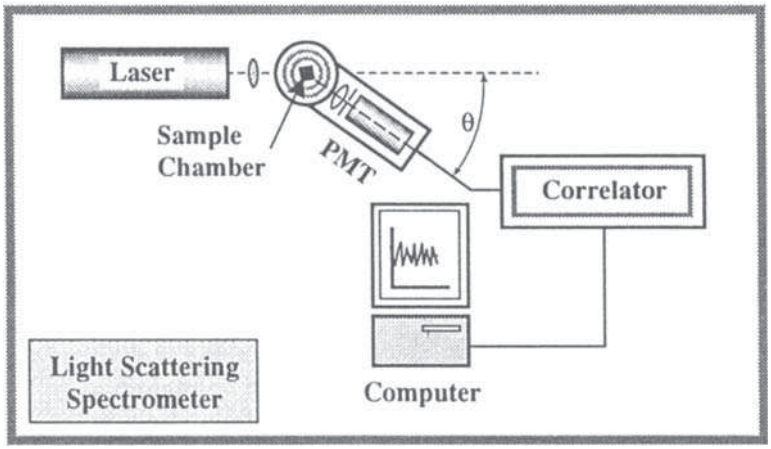


FIG. 5.5 Schematic top view of a typical light scattering instrument showing the different components and the definition of θ.

5.4a Some Preliminary Considerations

In order to determine M and B by means of Equation (27), it is clear that all the other quantities in the equation must be measured. It is convenient to group these factors into two categories, concentration and optical terms, for the purposes of our discussion.

5.4a.1 Concentration Terms

To begin, concentration enters the light scattering expressions primarily through the equations for osmotic pressure. Hence the same conditions apply in this application as in osmometry (see Chapter 3). Specifically, light scattering should be measured under isothermal conditions. Although concentration units other than weight per unit volume (the units of c) may be used, few of the alternatives are as useful as these. In consulting the literature, however, one should be attentive to the possibility that various workers may use slightly different units for c.

5.4a.2 Optical Terms

All the remaining variables in Equation (27) are optical in origin. The factors to be provided with numerical values are the refractive index of the solution and the refractive index gradient, dn/dc in K from Equation (25), and the Rayleigh ratio from Equation (24). All these optical parameters are wavelength dependent; therefore, each should be measured at the same wavelength. It is the value of this working wavelength that is used as the numerical substitution for λ in K.

The actual measurement of the refractive index of the solution poses no difficulty, but the evaluation of the refractive index gradient is more troublesome. The assumptions of the derivation of Equation (23) restrict its applicability to dilute solutions. The refractive index of a dilute solution changes very gradually with concentration; hence a plot of n versus c, the slope of which equals dn/dc, will be nearly horizontal. Since the intensity ratio depends on the square of dn/dc, it is clear that successful interpretation of Equation (23) depends on the accuracy with which this small quantity is evaluated. Measuring the absolute refractive indices of various solutions and determining dn/dc by difference or graphically would introduce an unacceptable error. A more precise method must be used to measure this quantity.

A differential refractometer is a device that specifically measures differences in refractive indices. By means of a differential refractometer, the difference between the refractive index of a solution and that of a solvent may be measured directly with the necessary precision. A variety of instrument designs is available for doing this. Most involve directing a light beam toward a two-compartment chamber, one portion of which contains the solvent, while the other contains the solution. The light is deviated differently by each, and position-sensitive photodetectors measure the small differences in deflection and translate them into a refractive index difference. Differences as small as 10^{-7} refractive index units can be measured in this way. The search for detection methods suitable for liquid chromatography has contributed to the development of accurate techniques for measuring small changes in refractive index. Note that we have also cited refractive index measurements in connection with size-exclusion chromatography in Chapter 1, Section 1.6b, and with sedimentation/diffusion studies in Chapter 2. A differential refractometer is an indispensable part of any laboratory in which light-scattering experiments are conducted.

5.4b Intensity Measurements

Now let us consider the actual measurement of light intensity. A light scattering photometer differs from an ordinary spectrophotometer primarily in the fact that the photoelectric cell that measures the scattered light is mounted on an arm that permits it to be located at various angular positions relative to the sample. In commercial light scattering devices, the detector rests on a turntable, the center of which coincides with the center of the sample. Thus the angle ϕ_x is measured in the horizontal plane. From here, we use the symbol θ to signify the angle of observation in the horizontal plane. The angle θ is measured from the direction of the transmitted beam where $\theta = 0°$. The incident beam therefore strikes the sample at $\theta = 180°$.

The cells in which the scattering solutions are measured should have flat windows at the

angle at which the scattering is measured. Cells with octagonal cross sections (actually, only half an octagon is used) are especially convenient since they present flat faces at 0, 45, 90, 135, and 180°. Cylindrical cells have been used, but they must be corrected for reflections from the walls. Regardless of the cell geometry, it is imperative that the cells be clean; otherwise the scattering from a fingerprint may exceed that from the solute! The solvent must also be purified of all extraneous matter. Filtration through sintered glass or centrifugation is usually employed to remove any dust particles, which would also invalidate the measurement.

The easiest way to calibrate a light scattering photometer is to use a suitable standard as a reference. Although polymer solutions and dispersions of colloidal silica have been used for this purpose, commercial photometers are equipped with opal glass reference standards.

5.4c Relating Intensities to Absorbance and Turbidity

Except for the movable photomultiplier tube, a light scattering photometer is very nearly identical to an ordinary spectrophotometer, which measures the ratio of the intensity of transmitted light to the intensity of incident light I_t/I_0. The *absorbance* per unit optical path ε_{abs} is defined in terms of this quantity as

$$\varepsilon_{abs} = -\ln (I_t/I_0) \tag{30}$$

(Note that ε here does *not* stand for permittivity.) Now let us examine the relationship between absorbance and the intensity of scattered light. In a light scattering experiment with nonabsorbing materials, the intensity of the transmitted light equals the initial intensity minus the intensity of the light scattered in all directions I_s:

$$I_t = I_0 - I_s \tag{31}$$

Combining Equations (30) and (31) leads to the result

$$\varepsilon_{abs} = -\ln\left(\frac{I_0 - I_s}{I_0}\right) = -\ln\left(1 - \frac{I_s}{I_0}\right) \approx \frac{I_s}{I_0} \tag{32}$$

where the approximation arises from retaining the first term of the series expansion of the logarithm (see Appendix A). The entire development of Section 5.3 is limited to dilute solutions and small n values. Therefore the approximation in Equation (32) is applicable to the systems we have been discussing. When the light attenuation is due to scattering, the ratio (I_s/I_0) is called the *turbidity*, τ, instead of the absorbance.

The quantity I_s in Equation (32) is not the same as the light scattered to a particular point (r,ϕ_x), but equals the summation of these contributions, totaled over all angles:

$$\frac{I_s}{I_0} = \sum_{\substack{\text{all} \\ \text{angles}}} \frac{i_s}{I_0} \tag{33}$$

This summation may be replaced by an integral as follows. An element of area dA on the surface of a sphere of radius r and making an angle ϕ_x with the horizontal is

$$dA = 2\pi r \sin \phi_x (r d\phi_x) \tag{34}$$

as shown by Figure 5.6. Therefore, the total scattered intensity ratio is given by

$$\frac{I_s}{I_0} = \int_0^{\pi} \left(\frac{i_s}{I_0}\right) 2\pi r^2 \sin \phi_x \, d\phi_x \tag{35}$$

Substituting Equations (23) and (25) into this expression yields

$$\frac{I_s}{I_0} = \int_0^{\pi} \frac{Kc(1 + \cos^2 \phi_x) 2\pi r^2 \sin \phi_x d\phi_x}{r^2(1/M + 2Bc)} \tag{36}$$

The factor r^2 cancels out of Equation (36) and the value of the integral over ϕ_x is (8/3). Therefore

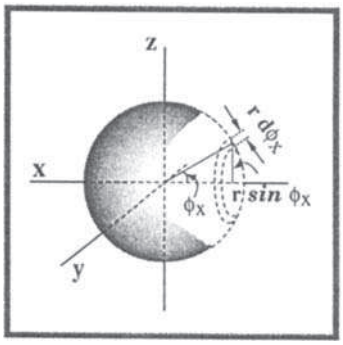


FIG. 5.6 Definition of an element of area required for the summation over all angles of the intensity of scattered light.

$$\frac{I_s}{I_0} = \tau = \frac{16\,\pi K c}{3(1/M + 2Bc)} \tag{37}$$

The parameter H is defined to equal the cluster of constants:

$$H = \frac{16\,\pi K}{3} = \frac{32\,\pi^3 n^2 (dn/dc)^2}{3 N_A \lambda^4} \tag{38}$$

and in terms of this quantity, Equation (37) becomes

$$\frac{Hc}{\tau} = \frac{1}{M} + 2Bc \tag{39}$$

The formal similarity between Equations (27) and (39) helps us understand somewhat better the physical significance of the Rayleigh ratio R_θ. It is directly proportional to the attenuation of the light per unit optical path, measured as absorbance, when the attenuation is due to scattering alone. In this case absorbance is more properly called *turbidity*.

Now let us consider some actual results from light scattering experiments on systems that satisfy the assumptions of the theory.

5.4d Results from Light Scattering Experiments: Weight-Average Molecular Weights

In the preceding section we saw how turbidity measurements are made and how they may be analyzed to yield numerical values for some of the parameters of interest in colloid chemistry. Figure 5.7 shows a plot of Hc/τ versus c for three different fractions of polystyrene in methylethyl ketone. The measurements shown in the figure were made at 25°C and at a wavelength of 436 nm. The molecular weights for the three fractions may be determined from the intercepts of the lines as illustrated in Example 5.2.

* * *

EXAMPLE 5.2 *Determination of Molecular Weights Using Rayleigh Ratios*: Assuming that the same system of units is used throughout, what are the units for R_θ and τ and also for K and H? Verify that these lead to reciprocal molecular weight units for the combinations Hc/τ and Kc/R_θ. To what molecular weights do the three intercepts in Figure 5.7 correspond? Do the three fractions behave as expected with respect to their second virial coefficients?

Solution: Since refractive index is dimensionless, dn/dc has reciprocal concentration units. In Equation (20) the factors $\{[n(dn/dc)]^2 c/(d\pi/dc)\}$ have the units pressure^{-1} since the units of c's cancel. The product $r^2\lambda^4$ in Equation (20) has the units length^6 or volume^2; therefore, the denominator of Equation (20) has the units pressure volume^2. The product $k_B T$ in the numerator

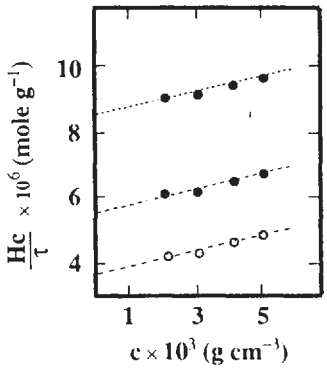


FIG. 5.7 Plots of Hc/τ versus c for three different fractions of polystyrene in methylethyl ketone. (Redrawn from B. A. Brice, M. Halwer, and R. Speiser, *J. Opt. Soc. Am.*, **40**, 768 (1950).)

can also be expressed in pressure volume units. Therefore i_s/I_0 has the units volume^{-1} and describes the intensity ratio per unit volume as described.

Equation (24) shows that the units of R_θ are those of $r^2(i_s/I_0)$. The units of the first factor are length^2 and those of the second length^{-3}. Therefore R_θ has the units length^{-1}. By Equations (35) and (37), the units of τ are the same as R_θ. Each of these parameters measures the light attenuation per unit path length.

By Equation (25), the units of K are concentration^{-2} length^{-4} = length^6 mass^{-2} length^{-4} = length^2 mass^{-2}. The fact that this quantity is divided by N_A places (mole^{-1})$^{-1}$ in the units. Equation (38) shows that H has the same units.

In terms of units, $Hc/\tau = Kc/R_\theta$ = (length^2 mass^{-2} mole)(mass length^{-3})/(length^{-1}) = mole mass^{-1}, reciprocal molecular weight units if mass is expressed in grams, as is the case in practical concentration units.

The three intercepts in Figure 5.7 are $3.70 \cdot 10^{-6}$, $5.56 \cdot 10^{-6}$, and $8.62 \cdot 10^{-6}$ mole g^{-1}. The reciprocals of these numbers give the molecular weights directly: 116,000, 180,000, and 270,000 g mole^{-1}.

The lines in Figure 5.7 appear to be parallel and hence characterized by a single B value. This would be expected for a single polymer-solvent-temperature system. ■

* * *

At first glance it appears that light scattering experiments as described so far provide no information that is not already available from osmometry. Indeed, for monodisperse colloids this is true, at least for the experiments we discussed above. The apparent redundancy between osmotic pressure and light scattering results should not be interpreted to mean that the two procedures duplicate one another entirely. For one thing, light scattering is free from the limitations imposed on osmometry by the availability of a suitable semipermeable membrane. Furthermore, turbidity measurements do not require time for equilibration, and hence they may be used for systems that change with time in a manner that is not possible with osmometry.

In addition to these practical considerations, there are other ways in which light scattering and osmometry differ. Some of these will become apparent only in subsequent sections in which additional characteristics of light scattering are developed. Another important difference, however, arises in the type of average that is measured for polydisperse systems.

In Chapter 3 we saw that osmometry enables us to measure the number-average molecular weight for a polydisperse colloid. In view of the way the osmotic pressure enters the development of Equations (27) and (39), it appears that the same type of average is obtained from turbidity experiments also. This is not the case. The following argument shows that light scattering measures the weight-average molecular weight.

For a polydisperse system, Equation (39) relates the experimental concentration, the experimental turbidity, and the average molecular weight:

$$\frac{H c_{exp}}{\tau_{exp}} = \frac{1}{\overline{M}} \tag{40}$$

It is sufficient to consider only the leading term of Equation (39) in writing Equation (40) since the molecular weight is evaluated from the intercept at infinite dilution. Likewise, we expect that Equation (39) will also apply to each molecular weight fraction in the polydisperse system

$$\frac{H c_i}{\tau_i} = \frac{1}{M_i} \tag{41}$$

It is the relationship between the value of $\overline{M}$ and the distribution of M_i values that we wish to determine. To accomplish this we note that

$$c_{exp} = \sum_i c_i \tag{42}$$

and

$$\tau_{exp} = \sum_i \tau_i \tag{43}$$

Combining the last four equations gives

$$\overline{M} = \frac{\tau_{exp}}{H c_{exp}} = \frac{\Sigma_i \tau_i}{H \Sigma_i c_i} = \frac{H \Sigma_i c_i M_i}{H \Sigma_i c_i} = \frac{\Sigma_i c_i M_i}{\Sigma_i c_i} \tag{44}$$

Now, recalling that $c_i = n_i M_i / V$ enables us to write

$$\overline{M} = \frac{\Sigma_i n_i M_i^2}{\Sigma_i n_i M_i} = \overline{M}_w \tag{45}$$

Equation (45) corresponds to the weight-average molecular weight $\overline{M}_w$ as defined by Equation (1.16).

It will be recalled from Chapter 1 that the number of particles in a molecular weight class provides the weighting factor used to compute the number-average molecular weight. The weight of particles in a class gives the weighting factor for the weight-average molecular weight. For this reason the weight average is especially influenced by the larger particles in a distribution. Therefore the weight-average molecular weight is always larger than the number average for a polydisperse system. As we saw in Chapter 1, the ratio of the two different molecular weights is a useful measure of the polydispersity of a sample. The authors of the research shown in Figure 5.7 also measured $\overline{M}_n$ for the same samples; the average value of $\overline{M}_w/\overline{M}_n$ for the three samples was 1.16. From polymerization theory this ratio is expected to be closer to 2 for polystyrene as synthesized, suggesting that these samples had been fractionated prior to molecular weight determination.

Thus we see that the redundancy between osmometry and light scattering is only an apparent effect for polydisperse systems. In fact, the combination of the two analyses provides additional information about the characteristics of the system.

Equation (39) shows that the slope of a light scattering plot is twice the value of the slope of a comparable plot from osmometry. In addition to the factor of 2, there is a more subtle difference between the slopes arising from a difference in the two values of the second virial coefficient. Examination of Equation (3.36) reveals that the second virial coefficient is inversely proportional to the square of the molecular weight of the solute. For polydisperse systems it is the average molecular weight that appears in this expression. Since the weight-average molecular weight is larger than the number average, the second virial coefficient B will be somewhat smaller as determined by light scattering than by osmometry after the factor of 2 has been taken into account.

Aside from the difference just noted, the interpretation of the second virial coefficient in light scattering is exactly the same as that developed in Section 3.4. It should be noted, however, that Equation (39) does not apply to charged systems. The reason for this lies in the

fact that the charge of macroions is also a fluctuating quantity, and this must also be considered in developing a scattering theory for charged particles. The resulting analysis shows that it is a plot of Hc/τ versus $c^{1/2}$ which is linear in this case, with the limiting slope proportional to z^2, the average value of the square of the charge. The slope is also predicted to be negative in this situation.

In concluding this section, it should be emphasized that the turbidity values plotted to interpret light scattering experiments are the solution turbidities, corrected for scattering by the solvent. Also, the entire theoretical development leading to Equation (39) is based on the assumption that the scatterers are isotropic. In this case, unpolarized incident light will produce a scattered beam that is totally polarized at $\theta = 90°$. When anisotropic particles are present, there is a depolarization of the light scattered at 90°. The ratio of the horizontally to vertically polarized scattered light may be determined by inserting a Polaroid filter between the sample and the photomultiplier. From the measured value of this depolarization ratio, a correction factor (called the *Cabannes factor*) may be introduced to allow for anisotropy. In the sample with $M = 116{,}000$ in Figure 5.7, for example, the ratio of the two different polarizations at 90° has a value of 0.013, for which the Cabannes factor equals 0.98. The turbidity should be multiplied by this factor to correct for the fact that the anisotropy enhances the amount of scattered light.

5.5 EXTENSION TO LARGER PARTICLES AND TO INTRAPARTICLE INTERFERENCE EFFECTS

In the remainder of this chapter we see that a good deal more information about scattering particles may be deduced, at least under some circumstances, from the study of the light scattered by a sample. In developing the Rayleigh theory and applying it to solutions, a definite model was postulated, and certain variables emerged as factors that affect the intensity of the scattered light. Before we extend light scattering theory to more complex systems, it is convenient to review the assumptions of the Rayleigh model:

1. The scattering centers are isotropic, dielectric, and nonabsorbing.
2. The scatterers have a refractive index that is not too large (see Step 8 in Table 5.1).
3. The particles are small in dimension compared to the wavelength of the light.

This last assumption is fundamental to the theory and originates as early as Equation (17), in which it is assumed that the field that drives the oscillating dipole is the same throughout the scatterer. It is generally held that particles must have no dimension larger than about $\lambda/20$ for this assumption to apply.

In general, in all radiation scattering techniques, the magnitude of the wavelength of the radiation relative to a suitable characteristic dimension, say, L_{ch}, appropriate for the particles plays an important role (as do other relevant properties of the system such as the absorptivity, refractive index, etc.). The wavelength of the radiation λ (or a combination of λ, the refractive index n, and the angle θ at which the scattered radiation is measured) may be thought of as the "yardstick" (say, L_{yd}) used in the scattering measurements. What properties of the dispersions or of the particles one can access through the scattering measurements are determined by the ratio (L_{ch}/L_{yd}).

The different theories (or "models") we discuss in this chapter and the different types of properties (e.g., osmotic pressure, particle size and shape, fractal structure of aggregates, etc.) we seek to measure using scattering correspond to appropriately different definitions of L_{ch} and L_{yd} and different ranges of magnitudes of the ratio (L_{ch}/L_{yd}). It is good to keep this in mind as we develop more complicated theories of scattering. As an aid, Table 5.3 presents a preliminary overview of some of the different definitions of L_{ch} and L_{yd} we encounter in subsequent sections of this chapter and their relation to the types of information that can be obtained using scattering techniques.

For example, Equation (39) suggests that light scattering is a technique ideally suited to the study of particles in the colloidal size range since the turbidity increases with the molecular

TABLE 5.3 Examples of the "Yardsticks" L_{yd} and the "Characteristic Lengths" L_{ch} Used in Different Theories of Scattering and the Different Properties Accessed through the Theories

Theory	L_{yd}	L_{ch}	Remarks
Rayleigh	λ	R_s, particle "size"	Applicable for $(R_s/\lambda) < 1/20$; extension of the Rayleigh equation to solutions allows the measurement of osmotic pressure, molecular weight, and turbidity of colloidal or polymer solutions; see Section 5.3
Debye	$\lambda/[4\pi \sin(\theta/2)] = s^{-1}$	Particle size or aggregate size; see Equation (57)	The information obtained depends on the magnitude of the ratio (L_{ch}/L_{yd}). Depending on the value of this ratio, details on particle shape or "fractal" dimension can be obtained; see Section 5.6
Mie	λ	R_s, particle "size"	Efficiencies of scattering and absorption are obtained as functions of $\beta = 2\pi R_s/\lambda$; see Section 5.7

Note: Consult the appropriate sections of the chapter for more comprehensive information on the assumptions, restrictions and uses of the theories.

weight of the particle. However, the assumptions underlying the derivation of Equation (39) impose a limitation. The Rayleigh theory shows that turbidity increases with molecular weight, at least until the particle is large enough to have some dimension exceeding about $\lambda/20$. In terms of the theory presented so far, we have no way of interpreting the light scattered by larger particles. The *Debye theory*, which we examine in the following section, will show us how to overcome this limitation.

The Rayleigh approximation shows that the intensity of scattered light depends on the wavelength of the light, the refractive index of the system (subject to the limitation already cited), the angle of observation, and the concentration of the solution (which is also restricted to dilute solutions). In the Rayleigh theory, the size and shape of the scatterers (M and B) enter the picture through *thermodynamic* rather than optical considerations.

A fully developed theory of light scattering that allows all the variables, including particle size and shape, to take on a full range of values is extremely complex. Because of this complexity, many treatments, such as the Rayleigh theory, are approximations that apply only to a narrow range of values of the parameters. In the Debye approximation, most of the preceding restrictions will continue to apply, except that the limitation on particle size will be relaxed considerably. At the same time, however, the stipulation of low values of the refractive index becomes even more stringent. As we see below, the Debye approximation introduces some additional complexity to the theory of light scattering and trades off some range in refractive index for extra range in particle size. From this, a positive dividend emerges: We shall be able to determine a characteristic linear dimension of the scattering particles without any assumptions about the shape of the particles. In the cases to which it applies, this information is definitely worth the price of a little additional complexity.

We omit most of the mathematical details in developing the additional theory; instead, we

emphasize the major concepts of the theory, its range of applicability, and its plausibility in limiting cases.

5.5a The Debye Scattering Theory

5.5a.1 Intraparticle Interference and the Form Factor

As emphasized above, the Rayleigh approximation is restricted to particles with dimensions that are small compared to the wavelength of light. Suppose we now relax this restriction in our model of a scattering particle, allowing the particles to take on dimensions comparable to λ. Under these circumstances different regions of the same particle will behave as scattering centers. Because the distances between these various scattering centers are of the same magnitude as the wavelength of light, there will be interference between the waves of light scattered from different parts of the same particle.

Figure 5.8 shows how this comes about. The electric field of the incident light beam is out of phase when it strikes two different portions of the scatterer, designated i and j. Although the figure shows i and j as polymer segments, they could be a pair of volume elements in any material. In fact, applications of the Debye theory to polymers are the most widely encountered. The light scattered from i and j is characterized by the same field that induces the oscillation at these locations. Therefore the light scattered from the two sites will be out of phase and display interference when observed at a large distance compared to Δx, say, along BB' in Figure 5.8. As shown in the figure, this interference effect may be constructive or destructive, depending on the value of θ. We propose, therefore, that the Rayleigh ratio defined by Equation (24) must be multiplied by a correction factor $P(\theta)$, known as the *form factor* (or the *intraparticle structure factor*), to correct for the interference effects not considered previously. Several things may be anticipated about this factor:

1. As the dimensions of the particle become negligible compared to the wavelength of light ($\Delta x \to 0$), $P(\theta) \to 1$—under these conditions, no correction is necessary.
2. The correction factor is a function of the angle of observation —as implied by the notation $P(\theta)$—since it is essentially an interference effect.
3. Multiplying the denominator of both sides of Equation (27) by $P(\theta)$ corrects i_s (a theoretical quantity calculated without interference) for interference:

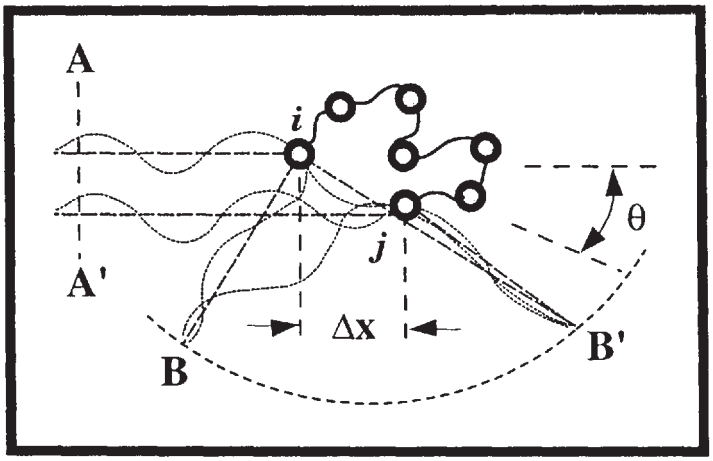


FIG. 5.8 Interference of light rays scattered by segments i and j in a polymer chain. (Redrawn with permission of P. C. Hiemenz, *Polymer Chemistry: The Basic Concepts*, Marcel Dekker, New York, 1984.)

$$\frac{Kc}{[i_s P(\theta) r^2]/[I_0(1 + \cos^2\theta)]} = \frac{1}{P(\theta)}\left(\frac{1}{M} + 2Bc\right) \tag{46}$$

4. The actual intensity of scattered light measured at a point C, denoted by i_C, equals the product i_s times $P(\theta)$. Therefore the experimental Rayleigh ratio obeys Equation (27) modified as follows:

$$\frac{Kc}{R_\theta} = \frac{1}{P(\theta)}\left(\frac{1}{M} + 2Bc\right) \tag{47}$$

In view of these considerations, it is apparent that the way to extend the previous theories to larger particles is to evaluate $P(\theta)$. To do this in a general way, two complications must be introduced into Figure 5.8:

1. Two scattering centers in a large particle are not simply displaced from one another in the x direction by an amount Δx—rather, the coordinates of one relative to the other must be described by a radial distance and two angles, for example, θ and ϕ.
2. A large particle does not consist of merely two scattering centers, but may be subdivided into several centers, the number of which increases with the size of the particle.

These two considerations must be incorporated into any general expression for $P(\theta)$.

5.5a.2 Expressions for the Form Factor

We begin our summary of the derivation of $P(\theta)$, however, by considering only the pair of scattering centers shown in Figure 5.8. To do this quantitatively, imagine that region i is at the origin of a coordinate system so the light that reaches j has to travel an additional distance Δx. The field at i and j is now represented by the following formulations of Equation (7):

$$\mathbf{E}_i = \mathbf{E}_0 \cos(2\pi\nu t) \tag{48}$$

and

$$\mathbf{E}_j = \mathbf{E}_0 \cos\left[2\pi\left(\nu t - \frac{\Delta x}{\lambda}\right)\right] \tag{49}$$

The light scattered from each of these sites will be characterized by the same field as the one that induces the oscillation. Therefore the light scattered from i and j will be out of phase by an amount $2\pi\,\Delta x/\lambda$.

Let us now consider the net scattered light that reaches a point C—located somewhere along BB'—a (large) distance r from the scatterer. The field at C is the sum of the fields emerging from i and j:

$$\mathbf{E}_C = \mathbf{E}_i + \mathbf{E}_j = \mathbf{E}_0\left\{\cos(2\pi\nu t) + \cos\left[2\pi\left(\nu t - \frac{\Delta x}{\lambda}\right)\right]\right\} \tag{50}$$

If we use the appropriate trigonometric formula for the sum of two cosines, this may be written as

$$\mathbf{E}_C = \left[2\cos\left(\frac{\pi\,\Delta x}{\lambda}\right)\right]\mathbf{E}_0 \cos\left[2\pi\left(\nu t - \frac{\Delta x}{2\lambda}\right)\right] \tag{51}$$

This equation shows that the electric field of light scattered to C is not altered in frequency or wavelength, but that the amplitude is modified by the factor $2\cos(\pi\Delta x/\lambda)$. The intensity of light depends on the square of the field amplitude; therefore, with interference,

$$i_C \propto \left[2\cos\left(\frac{\pi\,\Delta x}{\lambda}\right)\right]^2 |\mathbf{E}_0|^2 \tag{52}$$

and without interference ($\Delta x \to 0$),

$$i_s \propto 2^2 \, |\mathbf{E}_0|^2 \tag{53}$$

Combining Equations (52) and (53) leads to

$$\frac{i_C}{i_s} = \cos^2\left(\frac{\pi \Delta x}{\lambda}\right) \tag{54}$$

Since $i_c = i_s P(\theta)$, the factor $\cos^2(\pi \Delta x/\lambda)$ must equal $P(\theta)$ for this simple case.

A fair amount of straightforward but tedious trigonometry is necessary to establish the relationship between Δx and r, ϕ_x, and θ for the general case of any orientation between i and j. The result of this analysis is the expression (given without proof)

$$\Delta x = 2r \cos \phi_x \sin(\theta/2) \tag{55}$$

where ϕ_x is defined the same as in Figure 5.6 and θ continues to be the scattering angle measured in the horizontal plane.

It is not any specific value of ϕ_x in which we are interested, but in all possible values. This means that Equation (55) is to be integrated over all values of ϕ_x, a procedure that leads to the result (given without proof)

$$\frac{i_C}{i_s} = 1 + \frac{\sin[(4\pi r/\lambda) \sin(\theta/2)]}{(4\pi r/\lambda) \sin(\theta/2)} \tag{56}$$

Equation (56) benefits considerably from some simplification in notation. Accordingly, we define

$$s = \frac{4\pi}{\lambda} \sin(\theta/2) \tag{57}$$

This quantity has units of "1/length" and is the magnitude of a vector known as the *scattering vector*. For our purpose here, it is sufficient to note that s^{-1} is the "yardstick" used for measuring the distance between the scattering centers when the intensity of the scattered radiation of wavelength λ is recorded at an angle θ from the direction of the incident radiation.

Using the above notation permits us to rewrite Equation (56) as

$$\frac{i_C}{i_s} = 1 + \frac{\sin(sr)}{sr} \tag{58}$$

The right-hand side of Equation (58) correctly defines $P(\theta)$ for interference between two regions of a large particle.

The next question we must consider is how this result may be applied to a particle that consists of N, not just two, scattering centers. If the composition of the particle is such that the scattering from one part of the particle has no effect (other than interference) on the scattering from another part, then the particle may be subdivided into N scattering elements, and Equation (58) can be applied to all possible pairs. If we define r_{ij} as the distance between the ith and jth members of the set of N scattering elements, then this argument leads to the result (given without proof)

$$\frac{i_C}{i_s} \propto \sum_i \sum_j \left(\frac{\sin(sr_{ij})}{sr_{ij}}\right) \tag{59}$$

Expression (59) is written as a proportionality rather than an equation because the summation requires that a normalization factor be introduced. The ratio i_C/i_s must equal unity when r_{ij} is small since it explicitly corrects for interference effects that vanish under these circumstances. For small values of sr_{ij}, $\sin(sr_{ij})/sr_{ij}$ equals unity and $\Sigma_i \Sigma_j \sin(sr_{ij})/sr_{ij}$ equals N^2. Therefore normalization requires that we write

$$\frac{i_C}{i_s} = \frac{1}{N^2} \sum_i \sum_j \frac{\sin(sr_{ij})}{sr_{ij}} = P(\theta) \tag{60}$$

Note that the first term of Equation (58) is implicitly present in Equation (60) when r_{ij} equals zero. Equation (60) provides the general expression for $P(\theta)$ that we have sought. It also shows that $P(\theta)$ is more generally a function of s since one can vary s without changing θ by using a radiation source of a different wavelength λ. We return to this point in Section 5.6.

It is possible that the reader will recognize Equation (60) from another context. It is exactly the same expression that describes the diffraction of x-rays by polyatomic molecules, the situation for which it was derived by Debye (Nobel Prize, 1936). An important insight emerges from this realization. All interference phenomena between electromagnetic radiation and matter follow the same mathematical laws. Interference becomes important when there is some characteristic distance L_{ch} in the material under consideration that is of the same magnitude as the wavelength of the available radiation. The interference phenomena will be identical for identical values of L_{ch}/λ regardless of the separate values of L_{ch} and λ. Thus x-ray diffraction is an experimental technique that uses x-rays (for which $\lambda \approx 0.1$ nm) to measure interatomic distances that are on the order of nanometers. Likewise, we may use visible light (for which $\lambda \approx 500$ nm) to measure particles with dimensions in the colloidal size range. The scattering of microwaves by atmospheric rain and snow particles is an example of still another situation in which interference phenomena may be observed by scaling the wavelength of the radiation to suit the particle sizes under investigation. A brief comparison of x-ray, neutron, and light scattering is made in Section 5.6c.

Before turning to the applications of the Debye approximation, we should elaborate more fully on a point that was glossed over. This is the assumption—made at the outset, but explicated in going from Equation (58) to Equation (59)—that the scattering behavior of each scattering element is independent of what happens elsewhere in the particle. The approximation that the phase difference between scattered waves depends only on their location in the particle and is independent of any material property of the particle is valid as long as

$$\frac{2\pi L_{ch}}{\lambda}(n - 1) \ll 1 \tag{61}$$

It is the condition expressed by the inequality (61) that requires $(n - 1)$ to become smaller and smaller as the theory is applied to progressively larger particles. For example, when $\theta = 10°$, the Debye approximation is good to within 10% for spheres with a radius R_s that is about 62, 37, and 25 times λ at $n = 1.1$, 1.2, and 1.3, respectively. As we shall see presently, the approximation applies to a somewhat wider range of R_s and n values at smaller angles of observation and to a narrower range at larger angles.

5.5b Zimm Plots

Although Equation (60) may correct Equation (47) for the turbidity of systems in which the scattering centers are not negligible in size compared to the wavelength of light, it is not a very promising looking result. A little more mathematical manipulation will correct this impression. Specifically, suppose we examine Equation (57) for the case in which $\theta/2$ is small. This will make s small regardless of the value of r_{ij} so that sin (sr_{ij}) in Equation (60) may be expressed as a power series (see Appendix A):

$$P(\theta) = \frac{1}{N^2}\sum_i\sum_j \frac{\sin(sr_{ij})}{sr_{ij}} = \frac{1}{N^2}\sum_i\sum_j \frac{sr_{ij} - (sr_{ij})^3/3! + \cdots}{sr_{ij}} \tag{62}$$

For small values of s (more accurately, for small values of sr_{ij}; we come back to this later), the expansion may be limited to the first two terms:

$$P(\theta) = \frac{1}{N^2}\sum_i\sum_j 1 - \frac{(sr_{ij})^2}{6} = 1 - \frac{s^2}{6N^2}\sum_i\sum_j r_{ij}^2 \tag{63}$$

It is actually $1/P(\theta)$ in which we are interested, so we may again take advantage of the fact that s is small to write Equation (63) as

$$\frac{1}{P(\theta)} \approx 1 + \frac{s^2}{6N^2} \sum_i \sum_j r_{ij}^2 \tag{64}$$

when the second term on the right-hand side is much smaller than unity. In spite of various simplifying approximations, Equation (64) still does not appear particularly useful; the problem is the double summation of r_{ij} terms. We have encountered summations like this before in discussing the radius of gyration in Chapter 2, Section 2.7. A little manipulation of the results of that section will show us how to replace the summations in Equation (64):

1. Equation (2.76) gives an expression for $(R_g^2)_j$, the square of the radius of gyration of a swarm of masses around the jth, assumed to be the center of mass. (Note, also, the difference in notations; in particular, we have used n instead of N in Equation (2.76) for the number of segments.)
2. The two summations in Equation (2.76) can be consolidated into one—spanning all n of the mass elements—for the purposes of this discussion. Also, the product $P(r)r^2$ in Equation (2.76) is equivalent to a specific value of r_{ij}^2.
3. In going from Equation (2.76) to Equation (2.77), we added together (i.e., integrated) n terms like those in Equation (2.76) to allow for the fact that any of the n mass elements may play the role of the jth term in the derivation. Doing this in the case of the double sum in Equation (64) counts all mass elements twice, so a factor of 1/2 must be inserted.

We summarize these observations by writing

$$\overline{R_g^2} = \frac{1}{2N^2} \sum_i \sum_j r_{ij}^2 \tag{65}$$

which may be substituted into Equation (64) to give

$$\frac{1}{P(\theta)} = 1 + \frac{1}{3} \overline{R_g^2} s^2 \tag{66}$$

Equation (66) is valid in the limit of small values of $\theta/2$. Of course, if the scattering particle is not too large, the expansions of Equations (63) and (64) will be valid at larger θ, assisted by the fact that the values of r_{ij} will be small. This explains why the range of parameters to which the Debye equation applies depends on the angle of observation, becoming narrower for larger angles and broader at small angles.

Substitution of Equation (66) into Equation (47) yields

$$\frac{Kc}{R_\theta} = \left(\frac{1}{M} + 2Bc\right)\left|1 + \frac{16\pi^2 \overline{R_g^2}}{3\lambda^2} \sin^2\left(\frac{\theta}{2}\right)\right| \tag{67}$$

Let us consider Equation (67) in three important limiting cases with the objective of developing a graphical technique for using Equation (67):

1. In the limit of $\theta = 0$, Equation (67) reduces to Equation (27); that is, there is no interference effect in the scattered light.
2. In the limit of $c \to 0$, Kc/R_θ is proportional to $\sin^2(\theta/2)$.
3. If both c and $\theta \to 0$, Kc/R_θ equals $1/M$.

These limits suggest how experimental data might be collected, plotted, extrapolated, and interpreted. The resulting graph is known as the *Zimm plot* after its originator.

Equation (67) shows clearly that i should be measured as a function of both concentration *and* angle of observation in order to take full advantage of the Debye theory. The light scattering photometer described in Section 5.4 is designed with this capability, so this requirement introduces no new experimental difficulties. The data collected then consist of an array of i/I_0 values (i needs no subscript since it now applies to small and large particles) measured

over a range of c and θ values. By means of Equation (24), the i/I_0 ratios are converted to R_θ values. The results are plotted with Kc/R_θ as the ordinate and $\sin^2(\theta/2) + c$ as the abscissa.

Figure 5.9 shows this sort of plot for light scattering data collected from solutions of cellulose nitrate in acetone at 25°C. The measurements were made using 436-nm light of mercury. In the figure, $[\sin^2(\theta/2) + kc]$ has been used for the abscissa. The numerical constant k – equal to 2000 in this case – spreads out the points and results in a more intelligible display. As with any scale factor in graphing, it is found by trial and error. Each of the points in Figure 5.9 corresponds to a particular pair of c,θ values. When the points measured at the same values of c and those measured at the same values of θ are connected, a grid of lines is obtained like that sketched in the figure. If all the experimental c and θ values are small enough for the theories to hold exactly, then the grid consists of two sets of parallel straight lines with different slopes. In general, however, the range of experimental c and θ values exceeds the range of validity of the theory, and the lines show some curvature.

The next step in the treatment of the data is the extrapolation of the curves drawn at constant θ to $c = 0$ and the extrapolation of those drawn at constant c to $\theta = 0$. This is done by placing a mark (the triangles in Figure 5.9) on the smooth lines drawn through the experimental points at the value of the abscissa that corresponds to the value of that coordinate at the desired limit. For example, when the limit for $c = 0$ of the $\theta = \theta_1$ line is located, it will lie on the θ_1 line, and the value of the abscissa will be $\sin^2(\theta_1/2)$. Likewise, when the limit for $\theta = 0$ of the $c = c_1$ line is located, it will lie on the c_1 line, and the value of the abscissa will be kc_1. The triangles in Figure 5.9 were positioned on this basis. Note that the triangles describe two straight lines that have a common intercept.

Let us now consider the interpretation of this graph. First, it should be remembered that interference effects vanish at $\theta = 0$. The line so labeled in Figure 5.9 corresponds to values of Kc/R_θ at different values of c, all expressed at $\theta = 0$, where Equation (27) is valid. According to that equation, the slope of this line equals $2B$, and the intercept equals $1/M$. This is the same result we obtained previously for small particles. By extrapolating to $\theta = 0$, the procedure now applies to larger particles as well. In formulas, then, we write for the $\theta = 0$ line

$$(\text{Slope})_{\theta=0} = 2B \tag{68}$$

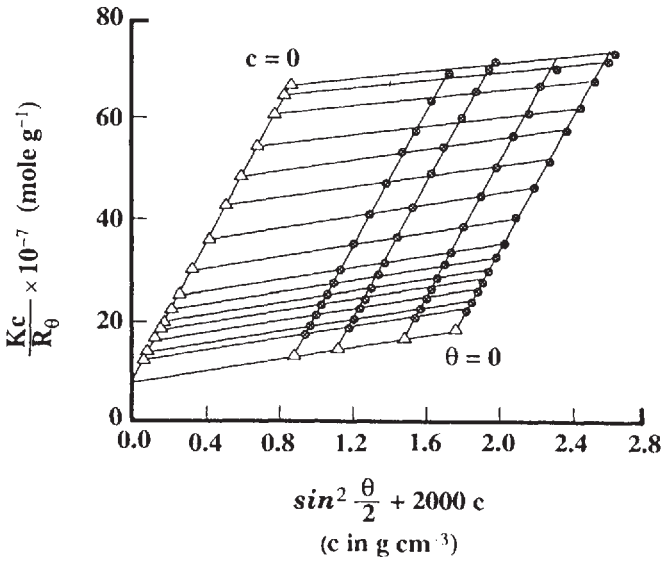


FIG. 5.9 Experimental Zimm plot for cellulose nitrate in acetone. (Redrawn from H. Benoit, A. M. Holtzer, and P. Doty, *J. Phys. Chem.*, **58**, 635 (1954).)

and

$$(\text{Intercept})_{\theta=0} = 1/M \tag{69}$$

The new feature of the Zimm plot is the second extrapolated line, corresponding to $c = 0$. This line connects values of Kc/R_θ measured at different values of θ and extrapolated to $c = 0$. Accordingly, it is described by Equation (67) with $c = 0$; that is, theory predicts the $c = 0$ line to have an intercept of $1/M$ and a slope equal to $(16\pi^2\overline{R_g^2}/3\lambda^2 M)$. It is expected, then, that the two lines will extrapolate to a common intercept $1/M$, and that the slope of the $c = 0$ line will be proportional to $\overline{R_g^2}$. Summarizing in formulas, we write for the $c = 0$ line

$$(\text{Slope})_{c=0} = \frac{16\pi^2 \overline{R_g^2}}{3\lambda^2 M} \tag{70}$$

and

$$(\text{Intercept})_{c=0} = 1/M \tag{71}$$

It follows, therefore, that the square of the radius of gyration equals

$$\overline{R_g^2} = \frac{3\lambda^2}{16\pi^2}\left(\frac{\text{Slope}}{\text{Intercept}}\right)_{c=0} \tag{72}$$

The Zimm plot shown in Figure 5.9 is analyzed according to these relationships in Example 5.3.

* * *

EXAMPLE 5.3 *The Use of Zimm Plots: Determination of Molecular Weight, Second Virial Coefficient, and Radius of Gyration of Cellulose Nitrate in Acetone.* The following are the values of the intercept and slope from Figure 5.9:

Intercept $= 7.87 \cdot 10^{-7}$ mole g^{-1}

Slope of the $\theta = 0$ line $= 5.70 \cdot 10^{-7}$ mole g^{-1}

Slope of the $c = 0$ line $= 6.78 \cdot 10^{-6}$ mole g^{-1}

Evaluate M, B, and R_g for cellulose nitrate in acetone at 25°C from these data. Use 1.359 for the refractive index of acetone at this wavelength.

Solution: The molecular weight is the reciprocal of the common intercept and equals $1.27 \cdot 10^6$ g mole^{-1}.

The slope of the $\theta = 0$ line includes the factor k, which must have the units concentration^{-1} for dimensional consistency. Hence the slope given must be multiplied by 2000 cm^3 g^{-1} to give the true slope to which the analysis applies. Since the true slope equals $2B$ according to Equation (68), we obtain

$$B = (1/2)(2000 \text{ cm}^3 \text{ g}^{-1})(5.70 \cdot 10^{-7} \text{ mole g}^{-1}) = 5.70 \cdot 10^{-4} \text{ cm}^3 \text{ mole g}^{-2}$$

The slope to intercept ratio for the $c = 0$ line equals $(6.78 \cdot 10^{-6})/(7.87 \cdot 10^{-7}) = 8.61$; this must be multiplied by $3\lambda^2/16\pi^2$ to obtain $\overline{R_g^2}$. Remember to convert $\lambda_0 = 436$ nm to the wavelength in acetone, which is the value used in this calculation: $436/1.359 = 321$ nm. Therefore $\overline{R_g^2} = 3(321)^2(8.61)/16\pi^2 = 1.69 \cdot 10^4$ nm^2 and the root-mean-square value is $R_g = 130$ nm. ■

* * *

It is important to remember that the radius of gyration provides an unambiguous measure of a particle's extension in space. This quantity is evaluated from experimental data—as illustrated in Example 5.3—with no assumptions as to particle geometry. If we happen to know the shape of a particle as well, R_g can be translated into a geometrical dimension of the particle. We examine this further in the next section.

5.5c The Dissymmetry Ratio

If the intensity of light scattered by a colloidal dispersion is measured as a function of c and θ, the Zimm method enables us to convert this information into several parameters that characterize the colloid: M, B, and R_g. In some situations this is more information than is actually needed. If spatial extension is the only information sought, a simpler method for evaluating it employs the so-called dissymmetry ratio.

In this technique, one measures i/I_0 at $\theta = 45°$ and $\theta = 135°$ for dispersions at several different concentrations. It should be noted that the factor $(1 + \cos^2\theta)$ in Equation (24) has the same value for these two angles of observation. Therefore any deviation of the ratio of the intensities z from unity must measure the ratio of the $P(\theta)$ values at these two angles [See Equation (47)]:

$$z = \frac{i_{45°}}{i_{135°}} = \frac{P(45°)}{P(135°)} \tag{73}$$

For example, we may examine the approximate form for $P(\theta)$ provided by Equation (66) to obtain

$$z = \frac{1 + (16\pi^2/3)(R_g/\lambda)^2 \sin 67.5°}{1 + (16\pi^2/3)(R_g/\lambda)^2 \sin 22.5°} \tag{74}$$

With this result—or its equivalent made with less restrictive approximations—values of z can be evaluated as a function of R_g/λ. From tables or plots of such calculated results, experimental dissymmetry ratios—extrapolated to $c = 0$ to eliminate the effects of solution nonideality—can be directly interpreted in terms of R_g.

We have noted previously that R_g is related to the geometrical dimensions of a body through expressions that are specific for the particle shape. Table 5.4 lists some of these relationships for shapes pertinent to colloidal systems. For a selected shape, one of the tabulated relationships can be used to replace R_g in Equation (74). What results is an expression that interprets z in terms of actual particle dimensions for the geometry chosen.

Figure 5.10 shows a plot of curves such as this drawn for spheres, coils, and rods. The characteristic dimension L_{ch} used in the abscissa of the figure is R_s for a sphere, r_{rms} for a coil, and L for a rod; the pertinent wavelength is λ_0/n. It will be noted from the ordinate values in Figure 5.10 that the scattering is always larger in the forward direction for particles showing interference effects. This is one reason why the presence of dust particles raises such havoc in a scattering experiment on particles that lie in the Rayleigh region.

TABLE 5.4 Relationships Between the Radius of Gyration and the Geometrical Dimensions of Some Bodies Having Shapes Pertinent to Colloid Chemistry

Geometry	Definition of parameters used as L_{ch}	Radius of gyration through the center of gravity
Random coil	$\overline{r^2}$: mean square end-to-end distance ($\propto n$)	$\overline{R_g^2} = \dfrac{\overline{r^2}}{6}$
Sphere	R_s: radius of sphere	$\overline{R_g^2} = \dfrac{3R_s^2}{5}$
Thin rod	L: length of rod (approximation for prolate ellipsoids for which $a/b \gg 1$)	$\overline{R_g^2} = \dfrac{L^2}{12}$
Cylindrical disk	R_d: radius of disk (approximation for oblate ellipsoids for which $a/b \ll 1$)	$\overline{R_g^2} = \dfrac{R_d^2}{2}$

Source: P. C. Hiemenz, *Polymer Chemistry: The Basic Concepts*, Marcel Dekker, New York, 1984.

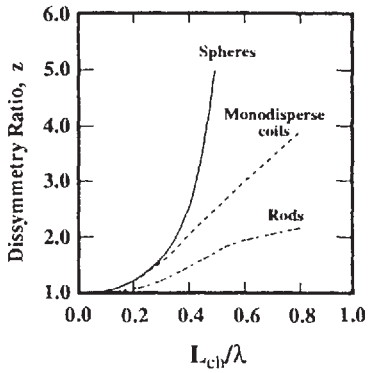


FIG. 5.10 Values of the dissymmetry ratio z versus the size parameter L_{ch}/λ for spheres, random coils, and rods. (Data from Stacey 1956.)

The development leading to Equation (66) in this section related the form factor $P(\theta)$ to the radius of gyration R_g, which is *one* measure of the "structure" of a particle. We can actually get much more information from the form factor. In the following section, we discuss this and illustrate the use of $P(\theta)$ for measuring the fractal dimension (defined in Chapter 1, Section 1.5b.1) of an aggregate.

5.6 INTERFERENCE EFFECTS AND STRUCTURE OF PARTICLES

The facts that we have explicitly included the intraparticle interference function $P(\theta)$ in the analysis of scattering intensities and that it is accessible experimentally allow us to characterize colloidal dispersions structurally in more detail than we have been able to so far. In order to understand this, we need to understand clearly what we mean by "small" or "large" values of θ or s and how they affect the behavior of $P(\theta)$. This will also help us to understand how (and why) it is possible to combine light scattering with x-ray or neutron scattering to study structures of particles and their aggregates.

5.6a The Form Factor and Particle Structure

The function $P(\theta)$ contains information on the *internal* structural details of the scatterers; this is the reason for its name "*form* factor" or "*intra*particle" structure factor, as defined in Section 5.5a.1. The length scales of the structural details contained in $P(\theta)$ depends roughly on the range of s^{-1} over which $P(\theta)$ is measured since, as we emphasized above, s^{-1} is the "yardstick" we use in the measurements. Notice that s^{-1} has units of length. We have so far alluded to the range of experimental measurements in terms of the magnitude of θ or s (i.e., as θ is "small" or "large" or as s is "small" or "large"). Before we discuss how $P(\theta)$ is used to obtain information about the geometrical structure of particles and in order to appreciate the physical significance of what is meant by "small" or "large," we need to make these notions more quantitative. For this, it is convenient to work with s rather than θ. The use of s is more general, as we see at the end of Section 5.6a.1.

5.6a.1 Physical Significance of the Magnitude of the Scattering Vector s

Since s^{-1} is a *dimensional* quantity, when we characterize it as "small" or "large" we implicitly do so relative to a suitable reference dimension L_{ch} (which is the characteristic dimension that represents the size of the scatterers). Otherwise, we can choose the units of s^{-1} to be angstrom or kilometer (or something even larger) and change the numerical value of s to something arbitrarily large or small! Let us define a new, dimensionless quantity Q by

$$Q = L_{ch}/s^{-1} = sL_{ch} \tag{75}$$

where L_{ch} is a characteristic length suitable for the particles in the dispersion. For example, we may choose L_{ch} to be equal to the radius R_s in the case of spherical particles or equal to the radius of gyration R_g or the root mean square radius of gyration $(\overline{R_g^2})^{1/2}$ in the case of polymer coils. With this definition, "small" or "large" s implies, respectively, that the yardstick s^{-1} is "large" or "small" compared to the characteristic dimension L_{ch}, i.e., $sL_{ch} \ll 1$ or $sL_{ch} \gg 1$. (Equivalently, we can say that the characteristic dimension of the particles—*or*, the length scale of the structural details of the particles—is "small" or "large" compared to the yardstick s^{-1}.) In addition to "small" or "large" values of s, we can also consider "intermediate" values of s, which are represented by $Q = sL_{ch} \sim 1$.

Let us now examine the physical significance of these limits or restrictions. First, the magnitude of s^{-1} is a measure of the distance of separation between the scattering centers that contribute to the intensity observed at that value of s.

1. When s^{-1} is large compared to L_{ch}, i.e., $s^{-1} \gg L_{ch}$ or $sL_{ch} = Q \ll 1$, it implies that the scattering centers are separated by a distance larger than L_{ch} and therefore lie on different particles. As a result, there is no intraparticle interference in the intensity at that value of s, i.e., $P(\theta)$, or more accurately, $P(Q)$, tends to unity for $Q \ll 1$. We already noted this in Section 5.5a.1, in which we compared the wavelength of light λ (which is proportional to s^{-1}; see Equation (57)) directly to particle dimension ("L_{ch}") and noted that $P(\theta) \to 1$ for $\lambda \gg L_{ch}$.
2. In the opposite limit, when s^{-1} is "small" or, equivalently, Q is large (typically, $Q \sim 1$ or larger), the measured $P(Q)$ consists of interference between closely lying scattering centers *within the same particle* (since the "yardstick" is small) and, therefore, contains information on the shape and intraparticle structure. We have more to say about this in Section 5.6b.

In general, the information contained in $P(Q)$ reflects the existence (or the lack) of interference effects *in the range of Q considered*. When the intensity measurements are obtained over a range of s^{-1}, we are, in effect, using a range of yardsticks to obtain the structural details of the particles over a number of length scales. However, comments (1) and (2) above highlight the fact that it is Q, the ratio of the reference dimension L_{ch} to the yardstick s^{-1}, that determines how we should interpret the information contained in $P(Q)$ (or, equivalently, in the intensity measured). This is, in fact, the rationale for defining the ratio L_{ch}/s^{-1} in Equation (75), which combines the two independent linear dimensions L_{ch} and s^{-1} into one variable.

The above discussion also draws our attention to another important way of looking at the measurements. Sometimes, we have more than one option for the characteristic dimension L_{ch}. One such example is a dispersion of colloidal aggregates (of diameter, say, d_{agg}) of smaller, primary particles (of diameter d_p). We can pick either d_{agg} or d_p as the characteristic dimension L_{ch}. Then, whether we call the $P(Q)$ measurements *inter*particle structure or *intra*particle structure depends on the choice of L_{ch}, i.e., our definition of "particle." For example, if we choose L_{ch} to be the diameter of the aggregate d_{agg}, then, for $Q\ (= d_{agg}/s^{-1})$ of the order of 1 or larger, $P(Q)$ describes the structure of the *aggregates*; i.e., the *intra*particle structure. We illustrate this in Example 5.4. If, on the other hand, we choose the diameter of the *primary* particle d_p as L_{ch}, the same range of s^{-1} corresponds to $Q \ll 1$, and $P(Q)$ is more appropriately referred to as *inter*particle structure (since our definition of "particle" now is the *primary* particle). In both cases, the information we have is the same, but our reference or perspective has changed. We return to this point in Section 5.6b.

As mentioned above, s (or s^{-1} or its dimensionless analog Q) is a more appropriate variable than θ to use in discussing the intraparticle (and interparticle) interference effects since θ is not the only variable that determines the range over which $P(Q)$ is measured. One can change the range by varying λ also (in fact, over a much larger range than possible by changing θ alone). This is the reason why sometimes other forms of radiation such as x-ray or neutron are used. We come back to this point in Section 5.6c.

5.6a.2 Different Regions of Interference Effects

Analytical expressions for $P(Q)$ can be derived for simple geometries of interest in colloid science (see, for example, Schmitz 1990), but it is sufficient for us to consider only the qualitative behavior of $P(Q)$ here. Such a picture of variation of $P(Q)$ with Q for aggregates of (primary) spherical particles of radius R_s is illustrated in Figure 5.11. One can identify four

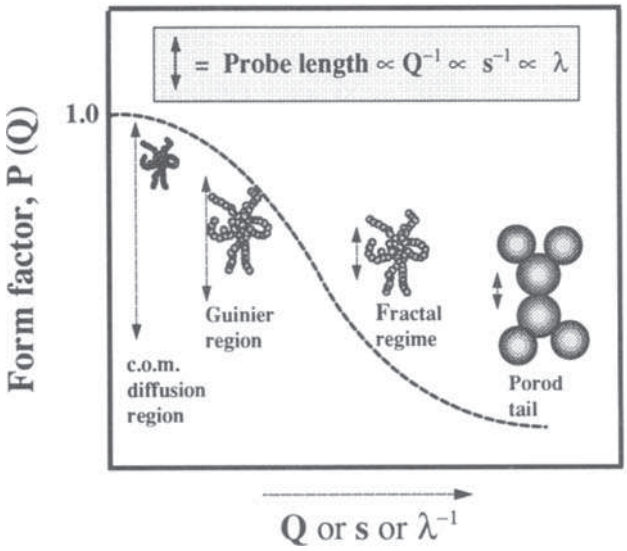


FIG. 5.11 Schematic representation of $P(Q)$ versus Q for fractal objects. The different parts of the curve corresponding to (a) the center-of-mass region, (b) the Guinier region, (c) the fractal region, and (d) the Porod region are indicated. The "probe length" is the "yardstick" corresponding to the measurement. (Adapted from Schmitz 1990.)

regions in this figure depending on the magnitude of Q, i.e., depending on the relative magnitudes of s^{-1} (suggested by the lengths of the arrows in the figure) and L_{ch}, which is taken to be the dimension of the aggregate (sketched in the figure):

1. As discussed above, in the limit of very small Q ($Q \ll 1$), $P(Q)$ approaches unity (see also Equation (63)), and there is no intraparticle interference since the "probe length" s^{-1} is much larger than R_s and the dimensions of the aggregate. This region is often called the "center-of-mass" region since the scattering centers act roughly as "centers of masses" that are uncorrelated with each other (i.e., $P(Q) \to 1$).

2. The next region, for Q's larger than the above limit but still small, signifies the onset of intraparticle interference. As we have already seen from Equations (63) and (65), the function $P(Q)$ here is of the form

$$P(Q) \approx 1 - \frac{1}{3}Q^2 \tag{76}$$

This region is known as the *Guinier region*.

3. Let us skip the next region temporarily and consider the range $Q \gg 1$. As mentioned above, intraparticle interference within the *primary* particles determines the function $P(Q)$ in this case, and the functional form of $P(Q)$ is determined by the shape of the primary particle (assumed here to be spherical). For spheres, one can show that

$$P(Q) \sim Q^{-4} \tag{77}$$

in this last region, which is known as the *Porod region*.

4. The third region is particularly important for dispersions of aggregates. As illustrated in the figure, Q in this case corresponds to the probe length s^{-1} larger than R_s, the radius of the primary particle, but comparable to the dimensions of the aggregates. The functional form of $P(Q)$ in this range depends on the detailed structure of the *aggregates* (i.e., distribution of the primary particles within the aggregates). If we think of the aggregate as the basic unit, $P(Q)$ describes the *intraaggregate* structure. For *fractal* aggregates (see Chapter 1, Section 1.5b), $P(Q)$ varies as

$$P(Q) \sim Q^{-d_f} \tag{78}$$

where d_f is known as the *fractal dimension* of the aggregate and is a measure of how compact the aggregate is, as discussed in Section 1.5b.

The above regions of $P(Q)$ stand out more clearly when $P(Q)$ versus Q is plotted on a *log-log* scale instead of on a linear scale as shown in Figure 5.11. We illustrate this and the above concepts in Example 5.4 using an experimental study of silica aggregates.

* * *

EXAMPLE 5.4 *Structure of Silica Aggregates Probed by Small-Angle Scattering.* The form factor $P(s)$ for a silica sol consisting of spherical primary particles of radius $R_s = 2.7$ nm, obtained from light and x-ray scattering, is plotted in Figure 5.12 as a function of s (Schaefer et al. 1984). What can you infer from the data about the structure of the aggregates?

Solution: Instead of choosing a characteristic dimension L_{ch} to define Q, let us look at the data in terms of s directly. An examination of the data in light of the discussion above leads to the following conclusions:

1. In the range of roughly $1 \cdot 10^{-3}$ nm$^{-1} \leq s \leq 2 \cdot 10^{-1}$ nm^{-1}, one observes from the figure that $P(s)$ is of the form

$$P(s) \sim s^{-2.1}$$

(remember the slope of a straight line in a log-log plot gives the exponent). This indicates that the dispersion consists of aggregates with a fractal dimension equal to 2.1 (a fairly loose aggregate). The lower limit of s in the fractal region is of the order of d_{agg}^{-1} (where, again, d_{agg} is the diameter of the aggregates), and the upper limit is related to the size of the primary particle, as we discuss below using the schematic representation of $i(s)$ versus s presented in Figure 5.13.

2. For s^{-1} smaller than the radius of the silica particles (i.e., for sR_s of the order of unity or larger), $P(s)$ appears to decrease as s^{-4} with increasing s, thereby confirming that the primary particles are spherical as given.

3. The data shown provide an example of how two different forms of radiation can be used effectively to cover a broad range of s^{-1}.

Note that the term *particle* in the intraparticle structure discussed here really refers to the aggregates. It is the internal structure of the aggregates that we have examined through the

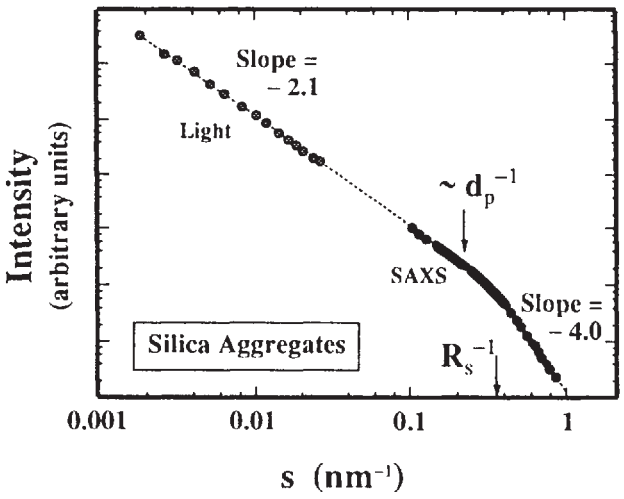


FIG. 5.12 Light scattering and small-angle x-ray scattering (SAXS) data for a dispersion of aggregates. The primary particles in the aggregates are monosize, spherical silica particles. The upper limit of s in the fractal region is roughly 0.2 nm^{-1}. Note log·log axes. (Data from Schaefer et al., 1984.)

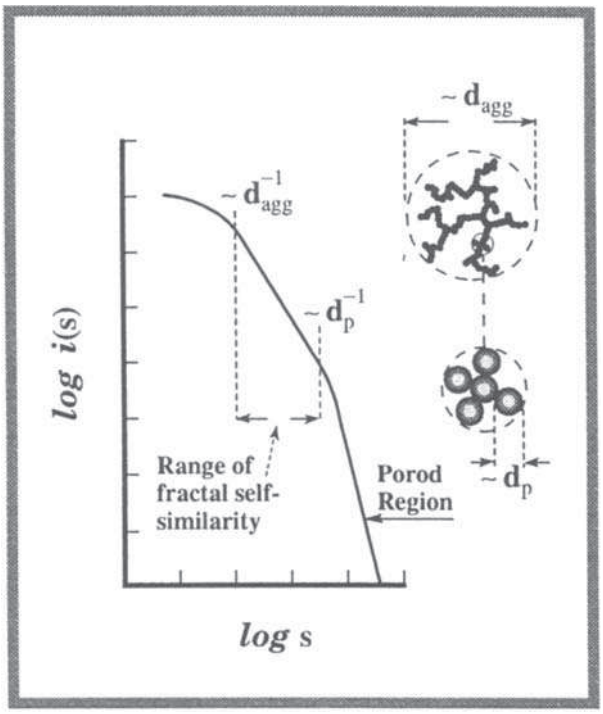


FIG. 5.13 A schematic illustration of the physical significance of the end points of the fractal region (see the text for the details). (Adapted from J. D. F. Ramsay and M. Scanlon, *Colloids and Surfaces*, **18**, 207 (1986).)

fractal region in the scattering curve. The "particles" with structures described by the Porod region are the primary particles, which are spherical silica particles in this case. ■

* * *

Comment (2) in the solution to Example 5.4 may appear to be superfluous since the example already states that the primary particles are spherical. However, that point is emphasized here in order to illustrate that scattering data can be used to follow the changes in the shapes of the primary particles if the physical situation may cause such changes. For example, if the primary particles are such that they coalesce and change their sizes or shapes, such changes may be followed using scattering experiments.

Moreover, as already pointed out in item (1) of the solution to Example 5.4, additional information about the aggregates and the primary particles may be obtained from plots such as Figure 5.12. This is illustrated schematically in Figure 5.13. This figure notes that the lower limit of the range of s in the fractal region is related to size of the aggregate; i.e., this limit is roughly d_{agg}^{-1}. The upper limit, on the other hand, is of the order of d_p^{-1}, where d_p is the diameter of the primary particle. As we noted in Chapter 1 (see Section 1.5b.1), fractal objects are *self-similar* objects consisting of similar structural features over a range of length scales. In the sketch shown in Figure 5.13, the aggregates look "similar" when we view them on length scales anywhere from roughly d_p to d_{agg}.

The above observations lead to the conclusion that the size of the aggregates in the dispersion represented by the data in Figure 5.12 is about 1000 nm, as already pointed out in item (1) in the solution to Example 5.4. The upper limit of s in the fractal region in Figure 5.12, i.e., $s = d_p^{-1} \approx 0.2$ nm $^{-1}$, is consistent with the fact that the primary particles in the example are spheres of diameter 5.5 nm.

5.6b Interparticle Structure

We have so far focused our attention on *dilute* systems so that we could avoid dealing with interference of scattering from different particles. The interference effects considered until now are restricted to interference due to scattering centers from within the same particle. When we have a fairly concentrated dispersion or even a dilute dispersion of charged particles that influence the position of each other through their interactions, the scattering data may have to be corrected for interparticle interference effects. Extending the previous discussion to *inter*particle interference is not difficult, but the subsequent analysis of the information obtained is not trivial. We shall not go into the details of these here, but just make some brief remarks to establish the connection between interparticle effects and what we have described so far for dilute systems.

Following the line of arguments we used in arriving at Equation (20) and at its modification, Equation (60), to account for interference effects, we can in general write the total intensity $i(Q)$ of the scattered light as

$$i(Q) \propto P(Q)S(Q) \tag{79}$$

The function $S(Q)$ in the above equation is known as the (*interparticle*) *structure factor* or *static structure factor* (the label *static* is used to differentiate $S(Q)$ from its time-dependent version, known as the *dynamic* structure factor) and contains information on how the particles are spatially distributed in the dispersion. When there are no interparticle effects, $S(Q)$ becomes unity and we recover Equation (60) from Equation (79).

A detailed discussion of $S(Q)$ requires statistical thermodynamics, which is beyond our scope here. We conclude this discussion by just noting three reasons for the importance of interparticle structure in colloid science:

1. In cases such as micellar solutions (discussed in Chapter 8) and in certain types of biological dispersions, interparticle interactions cannot be avoided. Surfactants in solutions form self-assembled structures (see Vignette I.3 on liposomes in Chapter 1) such as micelles, vesicles, and others at quite low surfactant concentrations. Vignette I.6 on polymer composites discusses another example of self-assembled colloidal structures. One can use scattering methods (particularly x-ray or neutron) to study the shape, size, and structure of such colloidal systems. In the case of micelles (especially those formed with ionic surfactants), the "particles" show significant interparticle interactions even at low concentrations. One cannot circumvent (as in the Zimm method) the interparticle interactions by diluting these dispersions for scattering experiments since these particles are thermodynamic entities and will dissociate (i.e., disintegrate) when diluted. Therefore, one is forced to account for the contribution of the interparticle interference to the measured intensities. If $S(Q)$'s can be established independently using suitable statistical mechanical models, then the form factor can be extracted for further analysis of shape and structure of the individual units. More commonly, one assumes models for both $P(Q)$ and $S(Q)$ and fits the observed intensities with the models in order to obtain structural information on the dispersion.
2. Measurement of the static structure factor using scattering techniques also provides us with a *nonintrusive* method to probe the structure of dispersions and the nature of interaction forces in colloids. Structural changes in colloids are particularly of interest in colloid-based techniques for fabrication of structural and special-purpose ceramics.
3. We have focused on only the *static* (i.e., the *time-averaged*) measurements of intensities so far. However, one can also obtain the *dynamic* structure factor (i.e., as a function of time) from scattering experiments. The dynamic structure factor can then be used as a probe of the rheological behavior of the dispersions discussed in Chapter 4.

In discussing the measurement and uses of form factors and static structure factors, we have, without stating explicitly, combined a discussion of light scattering with the use of other forms of radiation. In the following section we comment briefly on some of the similarities and differences between light scattering and x-ray and neutron scattering.

TABLE 5.5 Comparison of the Range Covered by Various Radiation Scattering Methods

Method	Typical wavelength (nm)	Range of s (nm^{-1})
Laser light scattering	500	$1 \cdot 10^{-3}$–$4 \cdot 10^{-2}$
Small-angle x-ray scattering	0.15	$2 \cdot 10^{-2}$–$4 \cdot 10^{-1}$
Small-angle neutron scattering	0.4	$7 \cdot 10^{-3}$–$9 \cdot 10^{-1}$
Wide-angle neutron scattering	0.4	$1 \cdot 10^{1}$–$5 \cdot 10^{1}$

5.6c X-Ray and Neutron Scattering as Complements to Light Scattering

This is a good place to draw attention to x-ray and neutron scattering since we have already introduced a combination of x-ray and light scattering to examine the fractal structure of silica aggregates in Example 5.4.

Much of the general formalism treated above remains valid for scattering of x-rays and neutrons by particles, although the mechanisms by which x-rays, neutrons, and light are scattered differ from each other. We have already seen that light scattering is a consequence of the interaction of photons with the electronic structure of atoms. X-rays are scattered by electron clouds, whereas neutrons are scattered by the nuclei of the atoms and by the magnetic moments of the atoms. A comparison of the wavelengths and the magnitudes of Q's accessible through x-ray, neutron, and light scattering is given in Table 5.5. From the point of view of colloid science, x-ray and neutron scattering techniques have two major advantages:

1. The particles and the fluid are effectively transparent to x-rays and neutrons; i.e., their effective refractive indices are nearly the same. Therefore, the criterion we specified in Equation (61) is easily satisfied, and we can avoid the need for the more complicated Mie theory (see Section 5.7b) and use the Rayleigh-Debye theories.
2. The very small values of the wavelengths of x-rays and neutrons allow one to reach large values of the parameter s. Since s^{-1} is the yardstick for the range of details accessible through scattering techniques, x-rays and neutrons allow us to probe structure at much shorter length scales than possible with light. Thus, x-ray scattering and neutron scattering complement light scattering measurements. Light scattering is indispensable for studying structure of particles and clusters in the micrometer size range. X-rays and neutrons help us to probe smaller particles.

Finally, it is also important to note that the term *small-angle scattering* does not necessarily specify the same range of Q; small-angle light scattering probes a region of Q different from the one probed by, for example, small-angle x-ray scattering because of the differences in the wavelengths of light and x-rays (see Example 5.4 and Table 5.5).

5.7 SCATTERING BY LARGE, ABSORBING PARTICLES

All the applications of *light* scattering that have been discussed so far have been restricted to very small particles and fairly small indices of refraction or to fairly small particles and very small indices of refraction. We have finally reached the point at which it seems appropriate to relax all these restrictions and consider the scattering by a particle of arbitrary size and index of refraction.

5.7a Scattering and Absorption

The general problem here is to solve Maxwell's electromagnetic equations both inside and outside a particle, for which the indices of refraction are different in the two regions. The problem has been solved for both spherical and cylindrical particles; we limit our considerations to the former.

5.7a.1 Accounting for Absorption Through Complex Refractive Index

In the most general case some provision must be made for the possibility that the particle absorbs as well as scatters light. This contingency is introduced by defining the refractive index of an absorbing material as a complex number $(n - ik)$ for which $i = \sqrt{-1}$. For nonabsorbing materials, $k = 0$. Both n and k are wavelength-dependent characteristics of the material; k obviously increases as the wavelength of an absorption peak is approached.

The idea of representing the refractive index of an absorbing material by a complex number may seem strange, so the following analysis will be helpful. Suppose we consider the passage of a beam of light through a layer (in the yz plane) of unspecified material of thickness Δx. If the layer contains a vacuum ($n = 1$), then the electric field transmitted through the layer will be given by Equation (7):

$$\mathbf{E}_{n=1} = \mathbf{E}_0 \cos [2\pi(\nu t - x/\lambda)] \tag{80}$$

On the other hand, suppose the layer consists of a material of refractive index n. The light will now take an increment of time Δt longer to pass through the layer owing to the delaying effect of the medium. In this case the emerging field would be given by

$$\mathbf{E}_n = \mathbf{E}_0 \cos [2\pi(\nu(t + \Delta t) - x/\lambda)] \tag{81}$$

The delay may be related to the thickness of the layer and the refractive index as follows:

$$\Delta t = t_n - t_{n=1} = \frac{\Delta x}{c/n} - \frac{\Delta x}{c} = (n - 1)\frac{\Delta x}{c} \tag{82}$$

where c is the speed of light in vacuum. This means Equation (81) may be written as

$$\mathbf{E}_n = \mathbf{E}_0 \cos \left| 2\pi\left(\nu t + (n - 1)\frac{\Delta x}{\lambda} - \frac{x}{\lambda}\right)\right| \tag{83}$$

which shows that, compared to Equation (80), the field experiences a phase shift in the medium.

A very important relationship involving complex numbers is

$$e^{i\theta} = \cos\theta + i \sin\theta \tag{84}$$

Cosine trigonometric functions, in other words, are given by the real part of the function $e^{i\theta}$. This means that Equations (80) and (83) may be written

$$\mathbf{E}_{n=1} = \mathbf{E}_0 \exp [2\pi i(\nu t - x/\lambda)] \tag{85}$$

and

$$\mathbf{E}_n = \mathbf{E}_0 \exp \left| 2\pi i\left(\nu t + (n - 1)\frac{\Delta x}{\lambda} - \frac{x}{\lambda}\right)\right| \tag{86}$$

It is only the real part of the complex number that is of interest to us. Equation (86) may be written as

$$\mathbf{E}_n = \mathbf{E}_0 \exp[2\pi i(\nu t - x/\lambda)] \exp\left(2\pi i(n - 1)\frac{\Delta x}{\lambda}\right) \tag{87}$$

$$= \mathbf{E}_{n=1} \exp\left(2\pi i(n - 1)\frac{\Delta x}{\lambda}\right)$$

As we have seen repeatedly in this chapter, the intensity of light is proportional to the field amplitude squared; therefore, we write

$$\frac{I_n}{I_{n=1}} = \exp\left(4\pi i(n - 1)\frac{\Delta x}{\lambda}\right) \tag{88}$$

Now let us consider the case in which the refractive index of the layer is a complex number. In that case, Equation (88) becomes

$$\frac{I_n}{I_{n=1}} = \exp\left(4\pi i(n + ik - 1)\frac{\Delta x}{\lambda}\right) = \exp\left(4\pi i(n - 1)\frac{\Delta x}{\lambda}\right)\exp\left(-4\pi k\frac{\Delta x}{\lambda}\right) \tag{89}$$

The first factor of Equation (89) is exactly what the ratio $(I_n/I_{n=1})$ would be if the material were nonabsorbing. It is modified, however, by a second term, $\exp(-4\pi k\Delta x/\lambda)$, which is real but contains the imaginary part of the index of refraction.

5.7a.2 The Beer-Lambert Formalism

The *Beer-Lambert equation* is another formalism that might describe the arrangement we have been discussing. In the Beer-Lambert equation, the intensity of the transmitted light I_t relative to the intensity of the incident light I_0 is given by Equation (30), which may be written

$$\left(\frac{I_n}{I_0}\right)_{abs} = \exp(-\varepsilon_{abs}\Delta x) \tag{90}$$

where ε_{abs} is the absorbance of the material. Because of its usefulness in analytical chemistry, chemistry students are more likely to be familiar with this formalism than with the one that leads to Equation (89). Comparison of Equations (89) and (90) reveals that both describe the attenuation of light due to absorption in terms of a negative exponential that is proportional to the path length through the absorbing material. Since the two approaches describe the same situation in the same functional form, the two proportionality factors must also be equal. Therefore the imaginary part of the complex refractive index and the absorbance must be related as follows:

$$\varepsilon_{abs} = 4\pi k/\lambda \tag{91}$$

We observed in Section 5.4c that, for nonabsorbing systems, the turbidity is a concept analogous to the absorbance; that is,

$$(I_t/I_0)_{sca} = \exp(-\tau\Delta x) \tag{92}$$

where the subscript *sca* stands for scattering. Examination of Equations (90) and (92) shows how the two complement each other: The first describes absorption without scattering; the second describes scattering without absorption. For a system that displays these two optical effects simultaneously, the following composite relationship applies:

$$(I_t/I_0) = \exp[-(\varepsilon_{abs} + \tau)\Delta x] \tag{93}$$

The experimental extinction in a system that displays these two effects equals the sum of ε_{abs} plus τ.

5.7a.3 Absorption Cross Section and Efficiency

The literature on scattering with absorption often factors absorbance and turbidity into a product of several terms. Both the nomenclature and notation used in this area vary from author to author, but the principal breakdown of both ε_{abs} and τ is as follows:

$$\varepsilon_{abs} = c_N\pi R_s^2\chi_{abs} \tag{94}$$

and

$$\tau = c_N\pi R_s^2\chi_{sca} \tag{95}$$

where c_N is concentration of the particles (in terms of the number of dispersed particles per unit volume), R_s is their radius, and the χ terms are known as the *efficiency factors* for absorption and scattering. The quantity πR_s^2 in these equations gives the geometrical cross section of the dispersed spheres. This area times the efficiency factor defines a quantity known as the *cross section* C_{abs} and C_{sca} for absorption and scattering, respectively:

$$C_{abs} = \pi R_s^2\chi_{abs} \tag{96}$$

$$C_{sca} = \pi R_s^2\chi_{sca} \tag{97}$$

Dimensionally, ε_{abs} and τ describe the attenuation of light per unit optical path and have the units length^{-1}. The area πR_s^2 has units of length^2 per particle, and c_N has units of particles

length $^{-3}$. The efficiency factors are dimensionless. The cross sections have units of area per particle. These should not be taken literally as cross-sectional areas but, rather, as the "blocking power" of a particle as far as the transmission of incident light is concerned.

As stated above, Maxwell's equations have been solved for spherical particles of arbitrary size and arbitrary refractive index. Furthermore, the refractive index may be complex, so the general theory applies to both nonabsorbing and absorbing particles. The solutions to Maxwell's equations are generally given in terms of the efficiency factors χ, the magnitudes of which depend on the wavelength of light, the size of the particles, the real and imaginary parts of the refractive index, and the angle of observation:

$$\chi = f(\lambda, R_s, n, k, \theta) \tag{98}$$

For nonscattering particles, χ_{sca} is zero, and for nonabsorbing particles χ_{abs} is zero. The results of such an analysis are generally reported by giving values of χ_{sca} and χ_{abs} as a function of a size parameter β, where β is defined as

$$\beta = 2\pi\frac{R_s}{\lambda} \tag{99}$$

The calculations involved in computing these results are formidable, and the results available prior to the widespread utilization of computers were very limited. As noted in the introduction to this chapter, the advent of the computer has broadened the applicability of light scattering enormously.

It is difficult to make any generalizations about the variation of the efficiencies with β since the functions are so complicated and vary greatly with the refractive index. About the best that can be done along these lines are the following:

1. At any given angle of observation, χ tends to be an oscillating function of β.
2. For nonabsorbing particles (n real), the amount of oscillation in the χ_{sca} versus β curve is more pronounced the larger n is—if the particles absorb, the amount of oscillation in the curves decreases with increasing k.
3. For any given value of β, χ varies with the angle of observation θ. The number of oscillations in this curve is greater for larger values of β and n.

These generalizations are consistent with the approximations discussed above in the chapter: Small particles with low values of n are the simplest to describe.

In the following two sections we consider two specific systems that have been studied extensively, aqueous dispersions of colloidal gold (Section 5.7b) and aqueous dispersions of colloidal sulfur (Section 5.7c). These illustrate some of the statements made above, as well as show the sort of information that may be obtained from light scattering experiments in this very general case. No attempt has been made to be comprehensive in this presentation. Much more has been done with the systems to be discussed, and cross sections for many other systems have been calculated. The book by Kerker (1969) contains a good bibliography of scattering functions published up to 1969.

5.7b The Mie Theory: Gold Sols

The preceding section indicated some of the complications that arise when the optical properties of dispersions are calculated without placing narrow limitations on the size and refractive index of the particles. The difficulties are still large but somewhat more manageable if only the refractive index is given full range, the particle dimensions being somewhat restricted. This is the situation that was treated by Mie in 1908.

Mie wrote the scattering and absorption cross sections as power series in the size parameter β, restricting the series to the first few terms. This truncation of the series restricts the Mie theory to particles with dimensions less than the wavelength of light but, unlike the Rayleigh and Debye approximations, applies to absorbing and nonabsorbing particles.

The following equations give some indication of the nature of these expansions:

$$\chi_{abs} = A_1\beta + A_2\beta^3 + A_3\beta^4 + \ldots \tag{100}$$

and

$$\chi_{sca} = A_4\beta^4 + \ldots \tag{101}$$

where the values of the coefficients A_1, A_2, A_3, and A_4 are listed in Table 5.6. In this presentation, we have intentionally limited the series to include no terms higher than fourth order in β. Thus only the leading term in χ_{sca} is represented, even though the first three terms in χ_{abs} are included. The neglect of higher order terms permits us to apply this discussion rigorously only to particles that are sufficiently small that terms in β^5 or higher would make a negligible contribution. Even with only these terms retained, it is possible to draw several informative conclusions about χ_{abs} and χ_{sca}.

1. The absorption and scattering efficiencies do not show the same dependence on the particle size parameter β.
2. Both numerical coefficients of the complex refractive index, n and k, appear in both χ_{abs} and χ_{sca} (Table 5.6).
3. The coefficients A_1, A_2, and A_3 equal zero, and A_4 reduces to Rayleigh's law if $k = 0$.
4. The efficiencies are functions of the dimensionless variable alone.
5. For dispersions of uniform spheres, the entire wavelength dependence of the extinction is given by Equations (100) and (101).
6. The "wavelength dependence of the extinction" is simply the spectrum of the dispersion, which is therefore predicted theoretically by the general equations.

In the remainder of this section we see how the theoretical calculations of Mie account for the observed spectrum of colloidal gold. In the next section we consider the inverse problem for a simpler system: how to interpret the experimental spectrum of sulfur sols in terms of the size and concentration of the particles. Both of these example systems consist of relatively monodisperse particles. Polydispersity complicates the spectrum of a colloid since the same χ value will occur at different λ values for spheres of different radii according to Equations (99)–(101).

Colloidal gold is of considerable historic importance in colloid chemistry since many of the scientists who led the early development of the field conducted experiments on this system.

TABLE 5.6 Values for the Constants A_1 to A_4 in Equations (100) and (101)

Coefficient	General case	Special case of $k = 0$
A_1	$\dfrac{24nk}{(n^2 + k^2)^2 + 4(n^2 - k^2) + 4}$	0
A_2	$\dfrac{4nk}{15} + \dfrac{20nk}{3[4(n^2 + k^2)^2 + 12(n^2 - k^2) + 9]} + \dfrac{4.8nk[7(n^2 + k^2)^2 + 4(n^2 - k^2 - 5)]^2}{[(n^2 + k^2)^2 + 4(n^2) - k^2) + 4]^2}$	0
A_3	$\dfrac{-192n^2k^2}{[(n^2 + k^2)^2 + 4(n^2 - k^2) + 4]^2}$	0
A_4	$\dfrac{\frac{8}{3}[(n^2 + k^2)^2 + n^2 - k^2 - 2]^2 + 36n^2k^2}{[(n^2 + k^2)^2 + 4(n^2 - k^2) + 4]^2}$	$\dfrac{\frac{8}{3}(n^2 - 1)^2}{(n^2 + 2)^2}$

Source: From P. B. Penndorf, *J. Opt. Soc. Am.*, **52**, 896 (1962).

Mie set out specifically to account for the brilliant colors displayed by sols of gold and other metals.

Chloroauric acid, $HAuCl_4$, is easily reduced to metallic gold by a wide variety of reducing agents. However, characteristics of the resulting gold are widely different for different reducing agents. Thus, if phosphorus is used, a polydisperse system containing very small particles forms rapidly. If the resulting colloid is used to seed a reaction in which hydrogen peroxide is the reducing agent, the following reduction takes place slowly without additional nucleation:

$$2AuCl_4^- + 3H_2O_2 \rightarrow 2Au + 8\,Cl^- + 6H^+ + 3O_2 \qquad \text{(A)}$$

The gold particles grow to a larger size by this process, with considerable sharpening of the particle size distribution. The particle size may be regulated to some extent by varying the amount of reagents used. Thus approximately monodisperse colloids of several different particle sizes may be prepared and compared. They are found to display different colors, depending on the particle size: The smaller particles produce a red dispersion; somewhat larger particles impart a blue color to the dispersion.

In calculating efficiency factors for absorption and scattering, the wavelength dependence of both the real and imaginary parts of the refractive index must be considered.

Figure 5.14 shows the real and imaginary parts of the complex refractive index of gold plotted against the wavelength of light in air and in water. These values of the refractive index were used to calculate the values of C_{abs} and C_{sca} for three different size spherical gold particles. The results are shown in Figure 5.15 as a function of the wavelength in air. The figure also includes the sum of the two cross sections C_{ext}, which describes the total extinction.

The magnitude and location of the maximum for each of the cross section curves are informative. It will be observed that the height and the location of the absorption curve in Figure 5.15 change relatively little with particle size (note the different scales for the ordinates), as we might expect. The cross section for scattering, on the other hand, is practically negligible for the smallest of the particles and increases to be roughly 50 and 100% larger than

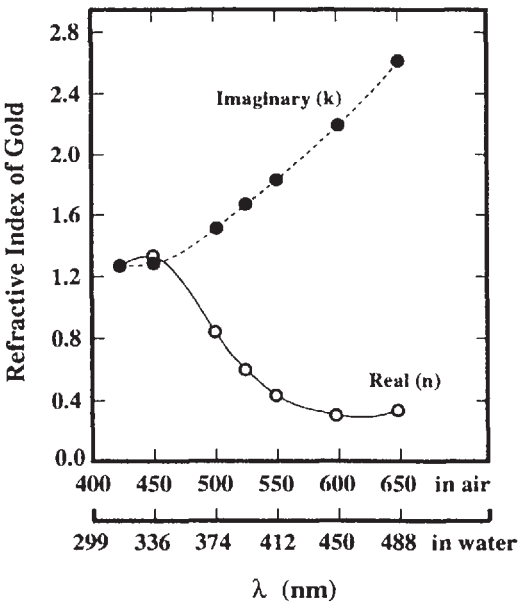


FIG. 5.14 The real and imaginary parts of the complex refractive index of gold versus wavelength in air and in water. (Data from Van de Hulst 1957.)

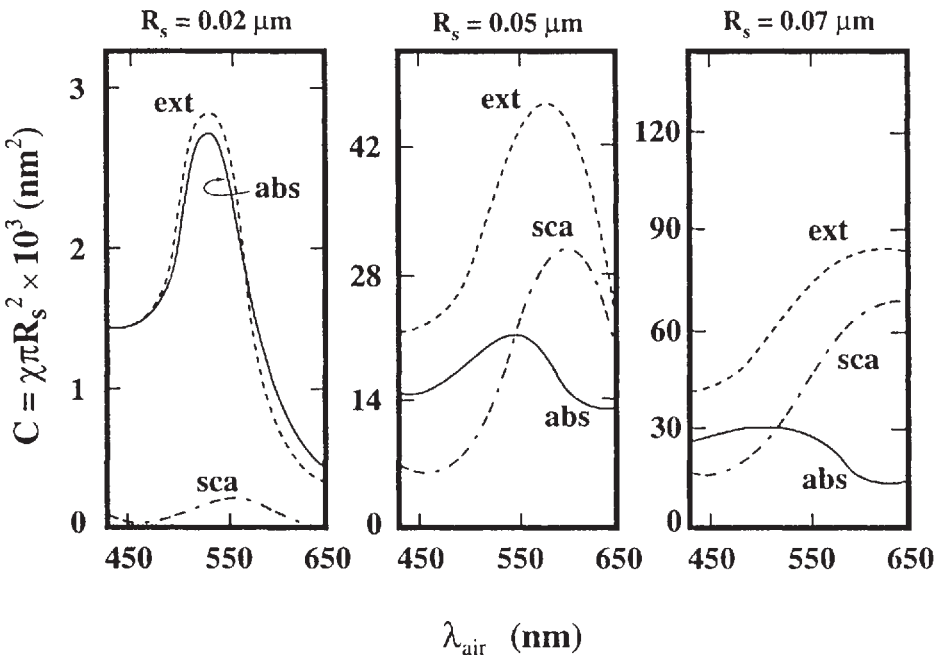


FIG. 5.15 Scattering coefficients versus wavelength for spheres of colloidal gold having three different radii. Note the different scales for the ordinates in each figure. (Data from Van de Hulst 1957.)

absorption for the larger particles. Also noteworthy is the fact that the wavelength location of the scattering maximum shifts to longer wavelengths as the particle size increases. This is a consequence of the fact that the efficiency depends on β rather than the separate values of R_s and λ. The wavelengths at which maximum total extinction occurs lie in the green, yellow, and red portions of the spectrum, respectively, for the particles with R_s = 0.02, 0.05, and 0.07 μm. On the basis of color complementarity, these would appear red, violet, and blue, respectively. These are the hues displayed by actual dispersions. Thus the Mie theory not only accounts for the color of the dispersions but also shows how the color displayed may be used to characterize the particle size of the dispersed phase.

Note that the index of refraction of the continuous phase and that of the dispersed particles enter the evaluation of the various efficiencies. The light that actually strikes the particles is used in the determination of χ. This light differs from that under vacuum by the refractive index of the medium. This effect enters the calculation of the χ values in that it is the ratio of the refractive index of the particle relative to that of the medium that determines the extinction.

5.7c Higher Order Tyndall Spectra: Monodisperse Sulfur Sols

Another colloidal system with light scattering characteristics that have been widely studied is the so-called monodisperse sulfur sol. Although not actually monodisperse, the particle size distribution in this preparation is narrow enough to make it an ideal system for the study of optical phenomena.

The colloid is prepared by rapidly mixing dilute solutions of sodium thiosulfate and hydrochloric acid so that the final concentration of each is about 0.002 M. The following reaction then occurs so slowly that the sulfur precipitates only on those particles that nucleate first:

$$H^+ + S_2O_3^{2-} \rightarrow HSO_3^- + S \qquad (B)$$

Slow growth on the original nuclei is how the narrow distribution of particle sizes is obtained, just as with the colloidal gold described in the preceding section. The formation of sulfur may be terminated at any time by adding I_2 to react with the remaining thiosulfate. These monodisperse sulfur sols have been studied extensively, notably by V. K. LaMer and coworkers. Since these particles are nonabsorbing in the visible spectrum, the range of particle sizes that may be conveniently dealt with is broader than for absorbing particles such as gold. Using the Mie theory, one can evaluate the scattering efficiency as a function of R, for particles having a refractive index relative to the medium of 1.50, which describes the sulfur-water system.

These "monodisperse" sulfur sols are good examples of another light scattering phenomenon: the *higher order Tyndall spectrum*. We observed in Section 5.7a that the scattering cross section is an irregularly oscillating function of θ, at least above a certain threshold value of β. Here it should be recalled that the complete theory reduces to the Rayleigh approximation for very small particles and to the Debye approximation for somewhat larger particles, provided the refractive index values are in the proper range.

The full solution of the Mie theory provides quantitative information about the dependence of the efficiency factors on θ and λ. For uniform spheres over some range of refractive index and size, different colors of light will be scattered in different directions. The sulfur sols described here have the required properties to display this effect. Therefore, if a beam of white light is shown through a sample of the dispersion, various colors will be seen at different θ values. The resulting array of colors is known as the *higher order Tyndall spectrum* (HOTS). Red and green bands are most evident in the sulfur sols, and the number of times these bands repeat increases with the size of the sulfur particles. Therefore the number and angular positions of the colored bands provide a unique characterization of the particle size. In the monodisperse sulfur sols, for example, particles having a radius of 0.30 μm are expected to show red bands at about 60, 100, and 140°. Particles with a radius of 0.40 μm, on the other hand, show red bands at about 42, 66, 105, 132, and 160°. Particle size determinations based on observations of this sort agree well with those determined from electron microscopy. These sulfur sols are quite easy to prepare, and it is interesting to observe the development of higher orders in the Tyndall spectrum as the thiosulfate decomposition reaction progresses.

It was once thought that the appearance of HOTS was evidence in itself for the presence of a monodisperse system. The argument was that one particle size would scatter, say, red light, at a particular angle, whereas another particle size would have the same R, value and therefore the same scattering behavior for light of a complementary color. The resultant would be the obliteration of any distinct color: The scattered light would appear white. Although there may indeed be fortuitous cancellations of this sort at certain angles, it is also possible for certain bands to reinforce each other. In general, then, it is best to say that polydisperse systems may show HOTS, but in this case the angular distribution of bands is a characteristic of the particle size distribution. The angular location and number of bands as determined theoretically for uniform particles may not be used to interpret the HOTS of a polydisperse system correctly.

5.8 DYNAMIC LIGHT SCATTERING

As mentioned in Section 5.1, so far we have focused on what are known as static scattering experiments; that is, the intensities used in the methods discussed until now are time-averaged intensities at any given angle θ. In general, however, the intensity accessible in a scattering experiment depends on time t as well, since the scattering centers are in constant random motion due to their kinetic energy. The variation of the intensity with time, therefore, contains information on the random motion of the particles and can be used to measure the diffusion coefficient of the particles. The measured diffusion coefficient can then be used to determine the size of the particle. The class of light scattering methods based on the time dependence of the scattered light intensity is known as *dynamic light scattering* (DLS), and a large body of work on various aspects of this technique has appeared in the literature in recent years (Brown

1993; Chu 1991; Schmitz 1990). For example, the dynamic version of the diffusing wave spectroscopy described in Vignette V is a form of DLS, although in diffusing wave spectroscopy the method of analysis is different in view of multiple scattering. Most of the advanced developments are beyond the scope of this book. However, DLS is currently a routine laboratory technique for measuring diffusion coefficients, particle size, and particle size distributions in colloidal dispersions, and our objective in this section is to present the most essential ideas behind the method and show how they are used for particle size and size distribution measurements.

5.8a Intensity Fluctuations and the Siegert Relation

In a typical scattering experiment, a detector measures the intensity of the scattered radiation over a period of time, say, t_n, in discrete steps of Δt (see Figure 5.16a). As shown in the figure, the intensity $i(s,t)$ fluctuates around an *average* value because of the random motions of the scatterers. Until now, we have denoted the average by simply $i(s)$ for convenience. To be precise, however, the average should be denoted by, say, $\bar{i}(s)$, and this is an average *over time*, t, defined as

$$\bar{i}(s) = \lim_{t_n \to \infty} \frac{1}{t_n} \int_0^{t_n} i(s,t)\, dt \approx \lim_{n \to \infty} \frac{1}{n} \sum_{j=1}^{n} i(s, j\Delta t) \tag{102}$$

where the limit $t_n \to \infty$ reminds us that the measurement should be made over a sufficiently large time for the average to be accurate.

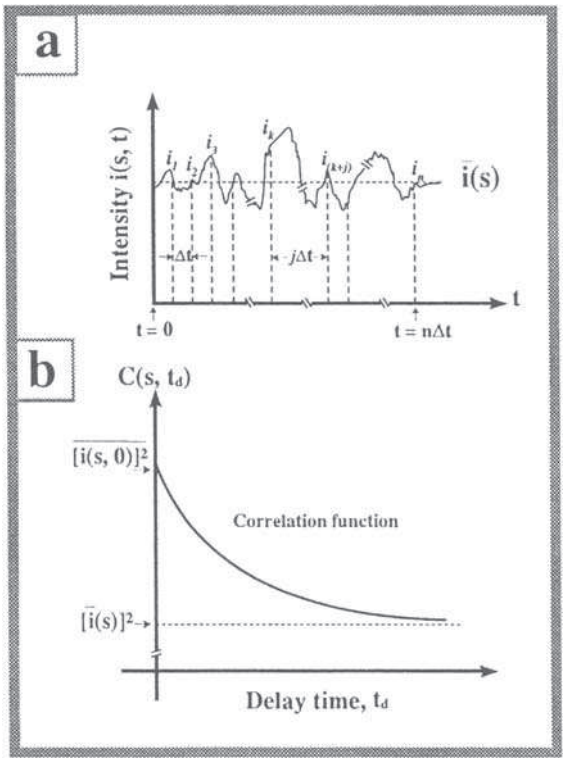


FIG. 5.16 Schematic illustration of intensity measurement and the corresponding autocorrelation function in dynamic light scattering: (a) variation of the intensity of the scattered light with time; (b) the variation of the autocorrelation function $C(s,t_d)$ with the delay time t_d.

Equation (102) also shows how the above time average is measured experimentally. Typically, the intensity is measured in a set of discrete time intervals, $t = \Delta t, 2\Delta t, 3\Delta t, \ldots$, etc., as illustrated in Figure 5.16a, and the arithmetic average shown in Equation (102) is an approximation of the average intensity over time $t_n = n\Delta t$.

In order to be able to use the *fluctuation* of the intensity around the average value, we need to find a way to represent the fluctuations in a convenient manner. In Section 5.3b in our discussion of Rayleigh scattering applied to solutions, we came across the concept of fluctuations of polarizabilities and concentration of scatterers and the role they play in light scattering experiments. In the present section, what we are interested in is the *time dependence* of such fluctuations. In general, it is not convenient to deal with detailed records of the fluctuations of a measured quantity as a function of time. Instead, one reduces the details of the fluctuations to what is known as the *autocorrelation function* $C(s,t_d)$, as defined below:

$$C(s,t_d) = \lim_{t_n \to \infty} \frac{1}{t_n} \int_0^{t_n} i(s,t)i(s,t + t_d)dt = \overline{i(s,0)i(s,t_d)}$$
$$\approx \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n} i(s,k\Delta t)i(s,(k + j)\Delta t) \tag{103}$$

where $t_d = j\Delta t$. The last part of the equation shows how the autocorrelation function is calculated experimentally when the intensity is measured in discrete time steps as illustrated in Figure 5.16a. The time t_d is known as the *delay time* since it represents the delay in time between the two signals $i(s,k\Delta t)$ and $i(s,(k + j)\Delta t)$ and is equal to $j\Delta t$ (see Figure 5.16a). The function $C(s,t_d)$ is obtained for a series of values of t_d by taking $j = 0, 1, 2, 3, \ldots$, etc. The autocorrelation function, as the name implies, is a measure of the correlation between the intensity $i(s,t_1)$ at any time t_1 and the intensity $i(s,t_1 + t_d)$ after a time *delay* of t_d. The correlation function obtained from Equation (103) is shown schematically in Figure 5.16b. Modern dynamic light scattering instruments consist of hardware "correlators" that have a number of channels or registers that keep track of $i(s,k\Delta t)$ for a large number of k's and automatically compute the products and the average in the summation term in Equation (103); see, for example, the schematic representation of a light scattering instrument shown in Figure 5.5.

The autocorrelation function has its highest value $\overline{[i(s,0)]^2}$ at $t_d = 0$. For $t_d \to \infty$, $i(s,t)$ and $i(s,t + t_d)$ become uncorrelated, and it can be shown that $C(s,t_d)$ is again independent of t_d and that it is given by $[\bar{i}(s)]^2$, where $\bar{i}(s)$ is the average value defined in Equation (102). For nonperiodic $i(s,t)$, $C(s,t_d)$ decreases monotonically from $\overline{[i(s,0)]^2}$ to $[\bar{i}(s)]^2$. Therefore, the ratio of the autocorrelation function to its asymptotic value $[\bar{i}(s)]^2$ can be written as

$$\frac{C(s,t_d)}{[\bar{i}(s)]^2} = g_2(s,t_d) = 1 + \xi|g_1(s,t_d)|^2 \tag{104}$$

an equation known as the *Siegert relation*, in which ξ is an instrumental constant approximately equal to unity.

The Siegert relation is valid except in the case of scattering volumes with a very small number of scatterers or when the motion of the scatterers is limited. We ignore the exceptions, which are rare in common uses of DLS, and consider only autocorrelations of the type shown in Equation (104). As mentioned above, modern DLS instruments use computer-controlled correlators to calculate the intensity autocorrelation function automatically and to obtain the results in terms of the function $g_1(s,t_d)$; therefore we only need to concern ourselves here with the interpretation of $g_1(s,t_d)$.

5.8b Dilute (Noninteracting) Dispersions

5.8b.1 Monosize Spherical Particles: Measuring Diffusion Coefficient and Particle Size

For dilute dispersions, i.e., those in which the interparticle spacing is so large that there are no particle-particle interactions, DLS simply measures the intensity fluctuations due to single-

particle motion. For monosize, spherical particles, one can show rigorously that $g_1(s,t_d)$ decays exponentially as follows:

$$g_1(s,t_d) = \exp(-s^2 D t_d) \tag{105}$$

Here, D, which is the quantity we seek from $g_1(s,t_d)$, is the diffusion coefficient of the particle (and s is the magnitude of the "scattering vector" defined in Equation (57)). We can now use the *Stokes-Einstein equation* (see Equation (2.32) and the accompanying comment) to obtain the particle radius R_H from D:

$$D = \frac{k_B T}{6\pi\eta R_H} \tag{106}$$

where η is the viscosity of the fluid, k_B is the Boltzmann constant, and T is the absolute temperature of the dispersion. The radius R_H measured in this manner is usually known as the *hydrodynamic radius* (hence the subscript H) since it relies on the Stokes coefficient, $6\pi\eta R_H$— a result from fluid (or, hydro) dynamics.

The measurement of the diffusion coefficient (and the hydrodynamic radius from D) is one of the most common uses of DLS, but the method can also be used as a nonintrusive technique for measuring the viscosity of a fluid. In this case, one uses "probe" particles with a known radius so that the unknown quantity in the Stokes-Einstein equation is η. More sophisticated uses of the DLS technique using essentially the above concept are discussed in specialized monographs (Brown 1993; Pecora 1985; Schmitz 1990). The diffusing wave spectroscopy mentioned in Vignette V also measures $g_1(t_d)$, but $g_1(t_d)$ in DWS is no longer a function of the angle θ since multiple scattering smears out the angle dependence of the intensity. As a result, the theoretical formalism needed for the analysis of the correlation function differs from what we have presented above and in Section 5.8b.2.

Example 5.5 illustrates one use of the DLS data.

* * *

EXAMPLE 5.5 *Determination of the Effective Diameter of an Enzyme Using Dynamic Light Scattering.* DLS analysis of a dilute solution of the enzyme phosphofructokinase in water at T = 293K leads to the following data for the correlation function $g_2(s,t_d)$:

$s^2 t_d \times 10^{-10}$ (m^{-2} s)	0.4	0.8	1.2	1.6	2.0	2.4
$g_2(s,t_d)$	1.75	1.6	1.47	1.375	1.298	1.236
$s^2 t_d \times 10^{-10}$ (m^{-2} s)	2.8	3.2	3.6	4.0	5.0	10.0
$g_2(s,t_d)$	1.188	1.148	1.119	1.093	1.052	1.003

Assume that the enzyme is roughly spherical and that the instrument constant ξ in the Siegert relation is unity and determine the hydrodynamic radius R_H of the enzyme. Given that the partial molar volume $\bar{V}$ of the enzyme is $0.74 \cdot 10^{-3}$ m^3/kg and the molecular weight M is $4.78 \cdot 10^2$ kg/mol, determine the "dry radius" R_s for the enzyme and obtain the ratio (R_H/R_s). Can the difference between R_H and R_s be attributed to the bound water on the enzyme? The viscosity η of water at 293K may be taken as 0.001 kg/m s.

Solution: The given DLS data can be used to obtain the intensity autocorrelation function $g_1(s,t_d)$ by rewriting the Siegert relation as follows:

$$\ln g_1(s,t_d) = \ln [g_2(s,t_d) - 1]^{1/2} - 1/2 \ln \xi$$

This leads to

$s^2 t_d \times 10^{-10}$ (m^{-2} s)	0.4	0.8	1.2	1.6	2.0	2.4
$\ln g_1(s,t_d)$	−0.118	−0.23	−0.35	−0.464	−0.58	−0.696
$s^2 t_d \times 10^{-10}$ (m^{-2} s)	2.8	3.2	3.6	4.0	5.0	10.0
$\ln g_1(s,t_d)$	−0.81	−0.93	−1.04	−1.16	−2.45	−2.9

Equation (105) shows that the plot of $\ln g_1(s,t_d)$ versus $s^2 t_d$ should give a straight line with a slope equal to the negative of the diffusion coefficient D of the enzyme. A plot based on the above data gives a straight line with a slope of $-2.88 \cdot 10^{-11}$ m^2/s and an intercept of −0.026. Therefore, the diffusion coefficient based on the data is

$$D = 2.88 \cdot 10^{-11}\ \mathrm{m^2/s}$$

(The magnitude of the intercept implies that the instrumental constant ξ is roughly 0.95.)

From the Stokes-Einstein relation, Equation (106), we then get

$$\begin{aligned} R_H &= k_B T/(6\pi\eta D) \\ &= [1.38\cdot 10^{-23}\,(\mathrm{J/K})\cdot 293\mathrm{K}] \div [(6\pi)0.001\ (\mathrm{kg/m\ s})\ 2.88\cdot 10^{-11}(\mathrm{m^2/s})] \\ &= 74.4\cdot 10^{-9}\mathrm{m} \end{aligned}$$

The "dry radius" R_s can be calculated from

$$\begin{aligned} R_s &= [(3/4\pi)\bar{V}M/N_A]^{1/3} \\ &= [3/4\pi)[0.74\cdot 10^{-3}(\mathrm{m^3/g})4.78\cdot 10^2(\mathrm{kg/mol})/6.02\cdot 10^{23}\mathrm{mol}^{-1}]^{1/3} \\ &= 52\cdot 10^{-9}\ \mathrm{m} \end{aligned}$$

The ratio (R_H/R_s) is therefore 1.43.

The source of this 43% difference between the "dry radius" and the hydrodynamic radius is unlikely to be the increase in diameter due to bound water. It is more likely that the shape asymmetry of the enzyme (i.e., the approximation that the enzyme is effectively spherical) is the source of the above difference. ■

* * *

5.8b.2 Effect of Polydispersity: Measuring Size Distribution

The DLS measurements can also be used in more complicated situations, for example, (a) when interparticle interactions are important, (b) for dispersions with particles of other shapes, (c) for monitoring coagulation, and (d) when the dispersion is polydisperse. In all these cases, a significant amount of modeling is often necessary to interpret the measured autocorrelation function, and we do not consider them here. Instead, we restrict our attention to a brief discussion of item (d) above, namely, polydisperse systems, since it is concerned with a problem of more routine interest.

In the case of a polydisperse system, the overall decay of the function $g_1(s,t_d)$ is determined collectively by the decay rate (i.e., s^2D) corresponding to *each particle* (notice that s^2D varies with the particle size as evident from the Stokes-Einstein relation). In principle, the decay function in this case can be written formally in a simple manner as a weighted average of all possible decays:

$$g_1(s,t_d) = \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} w_j(s^2D_j)\exp(-s^2D_j t_d) \tag{107}$$

where $w_j(s^2D_j)$ is a weighting function determined by the amount of particles in size range j; that is, the decay of $g_1(s,t_d)$ for a polydisperse system is an appropriately averaged function of the monodisperse case given in Equation (105). A number of methods are available for determining the size distribution from the experimentally determined $g_1(s,t_d)$. One of the simplest is based on what is known as the cumulant expansion, i.e., a series expansion of ln $g_1(s,t_d)$, given by

$$\begin{aligned} \ln g_1(s,t_d) &= \sum_{n=1}^{\infty} k_n(s)\,\frac{(-t_d)^n}{n!} \qquad \text{in the limit } \tau_d \to 0 \\ &= -\bar{D}\,s^2\,t_d + \sigma^2 s^4\,\frac{t_d^2}{2!} + \text{higher order terms} \end{aligned} \tag{108}$$

where k_n is known as the *nth cumulant*. Equation (108) also shows that the first-order cumulant is related to the average diffusion coefficients of particles of all sizes (denoted here by $\bar{D}$) and the second-order cumulant to the standard deviation σ of the distribution of diffusion coefficients (see Chapter 1 and Appendix C for a discussion of standard deviation and some of the related statistical concepts).

Equation (108) is accurate only for small delay times and, in fact, the higher order terms obtained experimentally are not usually very reliable because of the "noise" in the data.

However, it does illustrate how one can determine an average particle size and a measure of the breadth of the distribution function from experimental data. A plot of $[\ln g_1(s,t_d)/(s^2 t_d)]$ against $(s^2 t_d)$ will lead to a straight line for small t_d's, and $\bar{D}$ and s^2 can be obtained from the intercept and the slope of the straight line.

The logic of the above form of $g_1(s,t_d)$ and additional details are available in advanced books on DLS, and the above description is meant only to illustrate the basic ideas and one data-analysis approach. The cumulant analysis is often used as a first step before more advanced analytical procedures (each of which has its own advantages and disadvantages) are attempted. Most DLS instruments come with computer programs for the analysis of the size distribution, but we should bear in mind that each analysis technique has specific, and often restrictive, assumptions and none is "exact." As a consequence, the results of size distributions from DLS are best interpreted as semiquantitative indicators of polydispersity rather than a true representation of the distribution.

* * *

EXAMPLE 5.6 *Cumulant Analysis of Dynamic Light Scattering Data.* A polystyrene latex dispersion supplied by a manufacturer is claimed to have a "very narrow" size distribution with an average particle diameter of 62 nm. An analysis of the dispersion using DLS leads to the following data for $\ln g_1(s,t_d)$. The DLS experiments are conducted at 20°C using a dispersion in water at a particle volume fraction of 0.005. The wavelength of the laser used and the angle at which the experiments are conducted correspond to $6.5345 \cdot 10^6\ \mathrm{m}^{-1}$ for the magnitude of the scattering vector s. The viscosity of water at 20°C may be taken as 0.001 kg/m s.

$t_d \times 10^3$ (s)	0.05	0.1	0.15	0.2	0.25
$\ln g_1(s,t_d)$	−0.015	−0.0305	−0.0457	−0.061	−0.076
$t_d \times 10^3$ (s)	0.3	0.35	0.4	0.45	0.5
$\ln g_1(s,t_d)$	−0.091	−0.107	−0.122	−0.137	−0.152

Check if the specifications supplied by the manufacturer are correct. State any assumptions you make in your evaluation of the data.

Solution: Assume that the interparticle forces are negligible. Further, since the volume fraction of the dispersion used in the DLS experiments is very low, we may assume that the dispersion is sufficiently dilute and that multiple scattering is negligible.

Equation (108) shows that the cumulant expansion for $\ln g_1(s,t_d)$ may be rearranged to give

$$y = \frac{\ln g_1(s,t_d)}{x} \approx -\bar{D} + \frac{\sigma^2}{2} x \qquad \text{with } x = s^2 t_d$$

A plot of y versus x can now be used to obtain $\bar{D}$ from the intercept and σ^2 of the diffusion coefficient from the slope. Using the given data and the given value of s, we prepare a table of y versus x:

$x \times 10^{-10}$ (m^{-2} s)	0.2135	0.427	0.6405	0.854	1.0675
$y \times 10^{12}$ (m^2/s)	−7.03	−7.14	−7.135	−7.14	−7.15
$x \times 10^{-10}$ (m^{-2} s)	1.281	1.4945	1.708	1.9125	2.135
$y \times 10^{12}$ (m^2 s)	−7.10	−7.16	−7.14	−7.13	−7.12

It is clear from the table that the value of y is essentially constant over the entire range of given delay time. The slope, hence σ, is clearly negligible, and the intercept is approximately

$$\bar{D} = 7.14 \cdot 10^{-12}\ \mathrm{m}^2/\mathrm{s}$$

The use of the Stokes-Einstein relation with the above value of the average diffusion coefficient leads to a hydrodynamic radius of roughly 30 nm, which is consistent with the specification of the manufacturer. ■

* * *

5.8c Dispersions of Interacting Particles: Mutual and Self-Diffusion Coefficients

Even in the case of monodisperse systems, the observed decay rate of $g_1(s,t_d)$ (and hence the diffusion coefficient) in general depends on the angle at which the decay is measured if interparticle interference effects exist. In the case of *dilute* dispersions, in which interactions

of all sorts among the particles may be neglected, the interference is negligible, and the diffusion coefficient measured is independent of the angle and is given by the Stokes-Einstein equation (for spherical particles). This diffusion coefficient is often called the *self-diffusion coefficient* (or *probe diffusion coefficient*) since it represents the unhindered Brownian motion of a typical particle.

Although the analysis becomes complex for more concentrated dispersions (or even for dilute dispersions of charged particles, which can interact over very large distances), some general observations on two limiting cases are useful:

1. Measurements made at large enough values of Q ($= sR_s$) and for $t_d \rightarrow 0$: For $t_d \rightarrow 0$, the particles have very little time to wander far from their positions, i.e., the encounters with neighboring particles are negligible. Moreover, for large values of Q (e.g., large s) the interparticle interference is negligible since the range of interference in the observed intensity, represented by the magnitude of s^{-1}, is small. As a result, the measurements correspond again to the self-diffusion of the particles.
2. Measurements at low Q's: At low Q's, because of the large magnitudes of s^{-1}, the measured intensity and its autocorrelation function are dominated by the *cumulative* diffusion of the particles. The measured decay rate thus represents the *cumulative* or *mutual diffusion coefficient* D_m given by

$$D_m = \frac{1}{6 \pi \eta R_s} \left(\frac{\partial \pi_{osm}}{\partial c_N} \right)_T \tag{109}$$

where π_{osm} is the osmotic pressure of the dispersion and c_N is the concentration of the particles in "number of particles/volume of dispersion." It is the cumulative diffusion coefficient that appears in the Fick's laws discussed in Chapter 2, and the diffusion experiments described in Chapter 2 measure this diffusion coefficient. Thus we have identified another method for measuring mutual diffusion coefficients for (at least spherical) solutes. For dilute dispersions, D_m in Equation (109) reduces to the self-diffusion coefficient. (Note that, for dilute systems, $\pi_{osm} = c_N k_B T$ and $(\partial \pi_{osm}/\partial c_N)_T = k_B T$.)

Moreover, the influence of the motions of the particles on each other (i.e., when the motion of a particle affects those of the others because of communication of stress through the suspending fluid) can also influence the measured diffusion coefficients. Such effects are called "hydrodynamic interactions" and must be accounted for in dispersions deviating from the dilute limit. Corrections need to be applied to the above expressions for D and D_m when particles interact hydrodynamically. These are beyond the scope of this book, but are discussed in Pecora (1985), Schmitz (1990), and Brown (1993).

We have made a note of the hydrodynamic interactions and other interactions to draw attention to an important fact. That is, the analysis of the DLS data is often quite complex, and a simple use of the single-exponential decay function and the Stokes-Einstein relation is not always sufficient, although many instruments available on the market use such an analysis and report an "effective size" for the particles in the dispersion.

REVIEW QUESTIONS

1. Describe briefly what is meant by light *scattering* and the mechanism by which molecules scatter light.
2. Explain what is meant by each of the following terms: (a) electric field, (b) intensity of light, (c) polarization of light.
3. What is *Coulomb's law* and what are the units of the quantities that appear in Coulomb's law?
4. Why is light scattering an important tool in colloid science? What is the range of dimensions of colloidal particles that can be probed by light scattering? Why?
5. What is the difference between *static* light scattering and *dynamic* light scattering?
6. How does light scattering differ from *x-ray* scattering and *neutron* scattering in terms of mechanisms as well as the range of interactions and structure that can be probed by each?
7. What is meant by *Rayleigh scattering*? What are the important assumptions and limitations of the Rayleigh theory?

8. What is the *Rayleigh ratio*? How is it related to the intensity of the scattered light?
9. What is the relation between the Rayleigh ratio and the molecular weight and second virial coefficient of a dispersion of colloids or macromolecules?
10. What type of molecular weight average is measured by light scattering in the case of polydisperse solution?
11. What is meant by *turbidity* of a dispersion? How is the turbidity related to the Rayleigh ratio?
12. Explain what is meant by *interference of scattered light*. When is it important?
13. Why do liquids scatter light? Why isn't there a destructive influence of the lights scattered by all the molecules?
14. At what angle is the intensity of scattered light largest for large particles? Why?
15. How does the interference of scattered light affect the Rayleigh ratio?
16. What is the *Debye theory* of light scattering? What are its assumptions and limitations?
17. What is the definition of the *form factor* (or *intraparticle structure factor*)?
18. What is the magnitude of the scattering vector? How is it useful?
19. Discuss some of the limiting forms of the form factor. Explain what is meant by (a) the *center-of-mass region*, (b) the *Guinier region*, (c) the *fractal region*, and (d) the *Porod region*.
20. Discuss what is meant by *small-angle scattering*. How small is small?
21. What is the *Zimm plot*? Describe its use.
22. What is the *dissymmetry ratio*, and how is it used?
23. What is meant by *static structure factor* or *inter*particle structure factor? Describe its physical significance and use.
24. How is the light scattering theory formulated for large, absorbing particles? What is meant by "absorbing" particles?
25. What are absorbing and scattering cross sections?
26. What is the *Mie theory* of light scattering? What are its assumptions and limitations?
27. What is meant by *Tyndall spectra*?
28. What is meant by *dynamic light scattering* (DLS)? What is measured in DLS?
29. Why does the intensity of the light scattered by a dispersion fluctuate?
30. What is an *autocorrelation function*? Sketch qualitatively the autocorrelation function of the intensity of scattered light from a dispersion for a number of angles.
31. What is the *Siegert relation*? How is it related to the decay of intensity fluctuations measured in DLS experiments?
32. What is the relation between the diffusion coefficient of a monodispersed suspension of spherical particles and the decay of the intensity correlation function?
33. How is the particle size measured in DLS experiments? How is the size distribution measured in such experiments in the case of polydispersed colloids? Comment on the method(s) critically. How do interparticle interactions affect the above measurements?
34. What is the difference between the *self-diffusion coefficient* and the *mutual diffusion coefficient*? Which of these two is described by the Stokes-Einstein equation?

REFERENCES

General References (with Annotations)

Chu, B., *Laser Light Scattering: Basic Principles and Practice*, 2nd ed., Academic Press, Boston, 1991. (Advanced level. A detailed "manual" of light scattering. Covers basic theory as well as hardware and data-analysis procedures.)

Johnson, C. S., Jr., and Gabriel, D. A., *Laser Light Scattering*, Dover, New York, 1981. (Advanced level. An affordable, paperback edition of a long review article originally published as Chapter 5 in *Spectroscopy in Biochemistry*, Vol. 2 (T. E. Bell, Ed.), CRC Press, Boca Raton, FL, 1981. Covers classical as well as dynamic light scattering, including directed-flow light scattering [e.g., electrophoretic light scattering]. Illustrative applications use biological examples.)

Kerker, M., *The Scattering of Light and Other Electromagnetic Radiation*, Academic Press, New York, 1969. (Advanced level. A classic on the theory of light scattering.)

Schmitz, K. S., *An Introduction to Dynamic Light Scattering by Macromolecules*, Academic Press, Boston, 1990. (Advanced level. A textbook on DLS. Numerous examples from the literature.)

Zare, R. N., Spencer, B. H., Springer, D. S., and Jacobson, M. P., *Laser: Experiments for*

Beginners, University Science Books, Sausalito, CA, 1995. (Undergraduate level. A collection of experiments, most at the undergraduate level, on the use of lasers, light scattering, diffraction, and refraction. Chapter 2 describes a number of experiments highly relevant to the focus of the present book. The examples include particle size determination, studying aggregation, formation of microemulsion from a milky macroemulsion, and vesicle aggregation, among others. Chapter 3 describes diffraction and study of colloid crystals, discussed in Chapter 13 of the present book. Highly recommended.)

Other References

Abramowitz, M., and Stegun, I. A., Eds., *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, U.S. Government Printing Office, Washington, DC, 1964.

Atkins, P. W., *Physical Chemistry*, 5th ed., W. H. Freeman, New York, 1994.

Brown, W. (Ed.), *Dynamic Light Scattering: The Method and Some Applications*, Clarendon Press, Oxford, England, 1993.

Huglin, M. B. (Ed.), *Light Scattering from Polymer Solutions*, Academic Press, New York, 1972.

Jackson, C., Nilsson, L. M., and Wyatt, P. J., *J. Appl. Polym. Sci.*, **43**, 99 (1989).

McIntyre, D., and Gormick, F. (Eds.), *Light Scattering from Dilute Polymer Solutions*, Gordon and Breach, New York, 1964.

Pecora, R. (Ed.), *Dynamic Light Scattering: Applications of Photon Correlation Spectroscopy*, Plenum, New York, 1985.

Schaefer, D. W., Martin, J. E., Wiltzius, P., and Cannel, D. S., *Phys. Rev. Lett.*, **52**, 2371 (1984).

Stacey, K. A., *Light Scattering in Physical Chemistry*, Butterworths, London, 1956.

Tanford, C., *Physical Chemistry of Macromolecules*, Wiley, New York, 1961.

Van de Hulst, H. C., *Light Scattering by Small Particles*, Wiley, New York, 1957.

Weitz, D. A., and Pine, D. J., Diffusing-Wave Spectroscopy. In *Dynamic Light Scattering: The Method and Some Applications* (W. Brown, Ed.), Clarendon Press, Oxford, England, 1993.

PROBLEMS

1. Use Figure 5.4a to show that $\sin^2 \phi_y + \sin^2 \phi_z = 1 + \cos^2 \phi_x$. Verify that the angle ϕ is one-half the tetrahedral angle when $\phi_x = \phi_y = \phi_z = \phi$.

2. The intensity of the light scattered by air was measured in the atmosphere at Mt. Wilson, California, in 1913. The following is a selection of the results obtained:*

λ_0 (nm)	350	360	371	384	397
$i_s/I_{o,u}$	0.459	0.423	0.377	0.338	0.285
λ_0 (nm)	413	431	452	475	
$i_s/I_{o,u}$	0.245	0.213	0.174	0.147	

   Show that these results are in agreement with the predictions of the Rayleigh theory, given that $(n - 1)$ is constant at $2.97 \cdot 10^{-4}$ over this range of wavelengths. Instead of $(N_A\rho/M)$, we could have multiplied the equation in Step 6 in Table 5.1 by N_0, the number of molecules per unit volume under appropriate conditions—taken to be STP in this example for simplicity—and use light scattering to measure this quantity. Under the conditions of this research, the numerical and geometrical factors work out to give $i_s/I_{0,u} = 0.1911\lambda_0^4 N_0$ in SI units. Use a representative observation to evaluate N_0 and, from this, Avogadro's number.

3. The turbidity of solutions of sodium silicate containing SiO_2 and Na_2O in a 3.75 ratio has been studied as a function of time.† When the total solute content is 0.02 g cm^{-3}, the following data are obtained:

Time (days)	0	1	3	15	45	88	166	199	351	455
$\tau \times 10^4$ (cm^{-1})	0.81	2.32	3.80	5.47	6.70	8.06	9.53	10.47	13.00	14.43

   (During the same time, the pH of the solution changes from 10.85 to 11.02.) Suggest a

*Fowle, F. E., *Astrophys. J.*, **40**, 435 (1914).

†Debye, P., and Hauman, R. V., *J. Phys. Chem.*, **65**, 5 (1961).

qualitative explanation for these observations in terms of the chemical behavior of silicates (be sure to include references to whatever sources you consult). Why is it important to the study of light scattering to realize that solutions such as these show a variation of turbidity with time?

4. Equation (2.63) suggests that the quantity $\overline{\delta c^2}$ in Table 5.2 is evaluated by solving

$$\overline{\delta c^2} = \int_0^\infty (\delta c)^2 P(\delta c) d(\delta c)$$

Argue that the appropriate form for $P(\delta c)$ is

$$P(\delta c) = A \exp\left(- \frac{G - G_0}{k_B T} \right)$$

Use the value of $G - G_0$ from Table 5.2 [$G - G_0 = \frac{1}{2}(\partial^2 G/\partial c^2)_0 (\delta c)^2$] to verify the value for $\overline{\delta c^2}$.

5. The refractive indices of NaCl solutions have been measured at 20°C as a function of concentration by means of a differential refractometer* with the following results:

$c \times 10^3$ (g NaCl cm^{-3})	3.749	5.468	7.498	7.920
$(n - n_0) \times 10^4$	6.73	9.83	13.18	14.06

Use these data to evaluate the quantity dn/dc for NaCl solutions in this concentration range. On the basis of these data, what is the apparent uncertainty introduced in a light scattering experiment through this quantity?

6. Above a certain concentration, the turbidity of sodium dodecyl sulfate (molecular weight = 288) solutions increases with concentration as if particles in the colloidal size range were present. Use the following data† to evaluate the apparent molecular weight of the species responsible for the scattering. For this system $H = 3.99 \times 10^{-6}$.

$c \times 10^3$ (g cm^{-3})	2.7	4.2	7.7	9.7	13.2	17.7	22.2
$\tau \times 10^4$ (cm^{-1})	1.10	1.29	1.71	1.98	2.02	2.14	2.33

Assuming the scattering centers are aggregates of sodium dodecyl sulfate molecules, estimate the number of these units in the aggregate.

7. The turbidity of Ludox (a colloidal silica manufactured by DuPont) has been studied as a function of concentration with the following results:‡

$c \times 10^2$ (g cm^{-3})	0.57	1.14	1.70	2.30
$\tau \times 10^2$ (cm^{-1})	1.56	2.97	4.25	5.36

Evaluate the molecular weight of the Ludox particles, using a value of $H = 4.08 \times 10^{-7}$ for the system. Calculate the characteristic diameter for these particles, assuming the particles to be uniform spheres of density 2.2 g cm^{-3}.

8. Criticize or defend the following proposition: The accompanying data for the turbidity of dodecylamine hydrochloride solutions§ suggest that at concentrations exceeding about 0.003 g cm^{-3} the solute associates into aggregates of colloidal dimensions.

$c \times 10^3$ (g cm^{-3})	0.77	1.73	3.10	3.31	6.15	8.31
$\tau \times 10^4$ (cm^{-1})	0	0	0	0.95	1.91	2.55

For this system H is approximately 7.7×10^{-6}.

---

*Debye, P., and Hauman, R. V., *J. Phys. Chem.*, **65**, 8 (1961).

†Tartar, H. V., and Lelong, A. L. M., *J. Phys. Chem.*, **59**, 1185 (1955).

‡Deželic, G., and Kratohvil, J. P., *J. Phys. Chem.*, **66**, 1377 (1962).

§Debye, P., *J. Phys. Chem.*, **53**, 1 (1949).

9. Krasna* has measured the turbidity of calf thymus DNA in aqueous solutions. The accompanying table gives $R_\theta \times 10^5$ (in cm^{-1}) for this system, measured at 546 nm:

	$c \times 10^6$ (g cm^{-3})			
θ (deg)	20.6	41.4	62.0	82.5
26	2.47	4.80	6.94	8.75
30	2.06	4.02	5.83	7.67
34	1.77	3.39	4.84	6.44
38	1.75	2.98	4.28	5.61
42	1.38	2.57	3.84	4.87
50	1.01	1.90	2.91	3.71
60	0.76	1.45	2.23	2.89

Prepare a Zimm plot of these results (using $K = 3.63 \times 10^{-7}$) and evaluate M, B, and the radius of gyration of the DNA in this preparation.

10. The following table gives $R_\theta/K \times 10^{-3}$ for different values of c and θ in the system polystyrene-decalin at 30°C, measured with the mercury 435.8-nm line:†

	$c \times 10^3$ (g cm^{-3})		
θ (deg)	0.50	0.99	1.49
30	0.735	1.34	—
45	0.685	1.27	2.04
60	0.625	1.17	1.62
75	0.562	1.05	1.49
90	0.510	0.96	1.37
105	0.467	0.88	1.25

Prepare a Zimm plot of these date and evaluate M, B, and the radius of gyration of the polymer under these conditions.

11. The following table shows the angular location of the green bands in the HOTS of various size spheres of relative refractive index 1.46:‡

	Radius (μm)					
Number of green bands	0.2	0.3	0.4	0.5	0.6	0.8
First	25°	7.5°				
Second	140°	77.5°	57.5°	42.5°	32.5°	17.5°
Third		150°	95°	72.5°	60°	40°
Fourth			140°	117.5°	85°	62.5°
Fifth				115°	110°	87.5°
Sixth				(blue)	130°	110°
Seventh					157.5°	127.5°
Eighth						142.5°

Formulate a generalization correlating the observed number of green bands in the HOTS for this system with the approximate particle size. Briefly describe how this information could be used to "grow" a monodisperse sulfur sol with a dimension that corresponds aproximately to a predetermined size.

*Krasna, A. I., *J. Colloid Interface Sci.*, **39**, 632 (1972).

†Lechner, M. D., and Schulz, G. V., *J. Colloid Interface Sci.*, **39**, 469 (1972).

‡Kerker, M., *The Scattering of Light and Other Electromagnetic Radiation*, Academic Press, New York, 1969, p. 409.

12. General solutions of light scattering equations generate oscillating curves when χ is plotted versus β for a particular value of the relative refractive index of the dispersed phase compared to the continuous phase. Such a curve can be described by the expression $\chi = K\beta^{-n}$, where $-n$ is the local value of the slope of the ln χ versus ln β curve at specific values of β. Suppose this parameter n is known as a function of β for an experimental system. Describe the kind of experimental data and the analysis required thereof to yield a size parameter for the dispersed particles. What are the limitations of this method?

13. Suppose your employer intends to develop a new laboratory to characterize particles in the colloidal size range. Your assignment is to prepare a list of the equipment that should be purchased for such a facility. A brief justification for each major item should be included along with a priority ranking based on the versatility of the method. Assume that your laboratory is already well stocked with such nonspecialized items as laboratory glassware, balances, and the like.

14. Show that the average intensity $\bar{i}(s)$ given in Equation (102) is simply $(1/t_n)$ times the area under the function $i(s,t)$ in Figure 5.16a between $t = 0$ and $t = t_n$.

15. The DLS data for a dilute biological dispersion are given below. The experiment was conducted at 293K in water ($\eta = 0.001$ kg/m s). The scattering intensity was measured at a fixed angle of 30° using a laser with $\lambda = 480$ nm. Determine the average diffusion coefficient and the diameter of the particles based on this diffusion coefficient. What can you say about the polydispersity of the sample studied?

$t_d \times 10^3$(s)	0.05	0.1	0.15	0.2	0.25
ln $[g_2(s,t_d) - 1]^{1/2}$	−0.25	−0.499	−0.747	−0.995	−1.242
$t_d \times 10^3$(s)	0.3	0.35	0.4	0.45	0.5
ln $[g_2(s,t_d) - 1]^{1/2}$	−1.489	−1.735	−1.98	−2.225	−2.469

6

Surface Tension and Contact Angle

Application to Pure Substances

I don't think I said anything about the Third Dimension; and I am sure I did not say one word about "Upward, not Northward," for that would be nonsense, you know. How could a thing move Upward, and not Northward? Upward and not Northward! . . . How silly it is!

From Abbott's *Flatland*

6.1 INTRODUCTION

6.1a What Are Surface Tension and Contact Angle?

Why is it that insects like beetles can walk on water? Why do the bristles of a brush immersed in water cling together as the brush is pulled out? Phenomena such as these arise because of a special property of interfaces that separate two phases. Let us consider another example first. Everyone has had the experience of pouring more beverage into a cup or glass than that container could hold. In addition to the spills this causes, such an experience provides an opportunity to observe surface tension. Most liquids can be added to a vessel until the liquid surface bulges above the rim of the container. The liquid behaves as if it had a "skin" that prevents it—up to a point—from overflowing. Stated technically, a contractile force, which tends to shrink the surface, operates around the perimeter of the surface. This is what we mean when we talk about the *surface tension* of a liquid. All phase boundaries behave this way, not just liquid surfaces; however, the evidence for this is more apparent for deformable liquid surfaces.

The liquid "skin" described above is anchored to the solid walls of the container around the edges of the surface. The angle the liquid surface makes with the solid support is called the *contact angle*. The tendency of most liquids to climb walls—think of a meniscus in a capillary—is a manifestation of the existence of these angles.

Surface tension and contact angle are bulk properties, but they are both a consequence of intermolecular interactions that are short range in macroscopic terms. In the case of water, for example, the molecules like to be in the interior where they can form up to four hydrogen bonds; that is, the hydrogen atoms in a water molecule are weakly bonded to the oxygens of the neighboring molecules. At a water-air interface, however, the molecules have fewer neighbors, and water tries to minimize the number of "broken" bonds by minimizing the surface area. The bristles of the brush try to cling together as a consequence. The beetle does not sink because that would require the stretching of the surface. Unless an insect is heavy enough to counteract the force of surface tension, the insect will not sink and can walk on the surface. We have more to say about the molecular origin of both surface tension and contact angle in this chapter and in Chapter 10.

Surface tension and contact angle, however, are two different things, although they are

closely related. Surface tension is a property of the *interface* between two phases, whereas the contact angle describes the edge of the two-phase boundary where it ends at a third phase. Two phases must be specified to describe surface tension; three are needed to describe contact angle. (In this chapter, we designate solid, liquid, and vapor phases by S, L, and V, respectively; if two liquids are involved, we call them L_1 and L_2 or A and B.)

6.1b Why Are They Important?

Surface tension and contact angle phenomena play a major role in many practical things in life. Whether a liquid will spread on a surface or will break up into small droplets depends on the above properties of interfaces and determines well-known operations such as detergency and coating processes and others that are, perhaps, not so well known, for example, preparation of thin films for resist lithography in microelectronic applications. The challenge for the colloid scientist is to relate the macroscopic effects to the interfacial properties of the materials involved and to learn how to manipulate the latter to achieve the desired effects. Vignette VI provides an example.

VIGNETTE VI SURFACE CHEMISTRY AND MATERIALS SCIENCE: Wettability of Surfaces and Fabrication of Microstructured Materials

Why is understanding surface tension, contact angle, and their role in wettability of surfaces important? Waterproof fabrics, detergency, and coating processes (e.g., preparation of coated surfaces using spin coating) are standard examples that, perhaps, readily come to mind, but wettability of surfaces has a strong influence on many other phenomena of engineering significance as well. For example, the rate of evaporative heat transfer from surfaces (and, as a result, the effectiveness of heat-transfer equipment in which condensation and evaporation play a major role in the overall heat transfer) is influenced strongly by whether the liquid phase wets the surfaces or not.

Instead of considering wettability from the above traditional perspective, let us look at it from the opposite side: What if we can control wettability of a surface at microscopic length scales, perhaps *actively*? New developments in surface chemistry and fabrication of self-assembled monolayers (SAMs; see Chapter 7) have opened up precisely such a possibility. How can we take advantage of such opportunities?

As we see in this chapter, chemical affinities between a surface and a liquid at the molecular level (i.e., preferential molecular interactions) determine the wettability of a surface and the resulting shapes of the liquid drops. The chemical affinities can be controlled by allowing the formation of monolayers of surfactant molecules with desired properties (lyophilicity or lyophobicity) on a surface. Such a procedure, in combination with micromachining (for example, using a scanning tunneling microscope; see Vignette I.8 in Chapter 1) can be used to prepare surfaces with prechosen geometric patterns of wettability.

Figure 6.1 illustrates such an example (Abbott et al. 1992). Here, a gold surface is first coated with a hydrophilic SAM of alkanethiolate. The desired geometric pattern is then formed on the surface through micromachining (in this case, using a surgical scalpel), and the resulting features are covered with a hydrophobic SAM (of dialkyl disulfide) (see Fig. 6.1a). This procedure can be used to construct micrometer-scale hydrophobic lines on the surface so that the resulting shapes and distribution of liquid drops can be controlled (see Fig. 6.1b).

Similar techniques can also be used to control the wettability of a surface *actively* (in time scales of the order of seconds), for example, by electrochemically controlling the interconversion of less polar (less wettable) groups in the SAMs to more polar (more wettable) groups (Abbott and Whitesides 1994).

What benefits do we derive from such control? Control of wettability on small length scales using simple "wet chemistry" provides interesting opportunities, e.g., a simpler and more easily accessible alternative to optical lithography and new methods for the fabrication

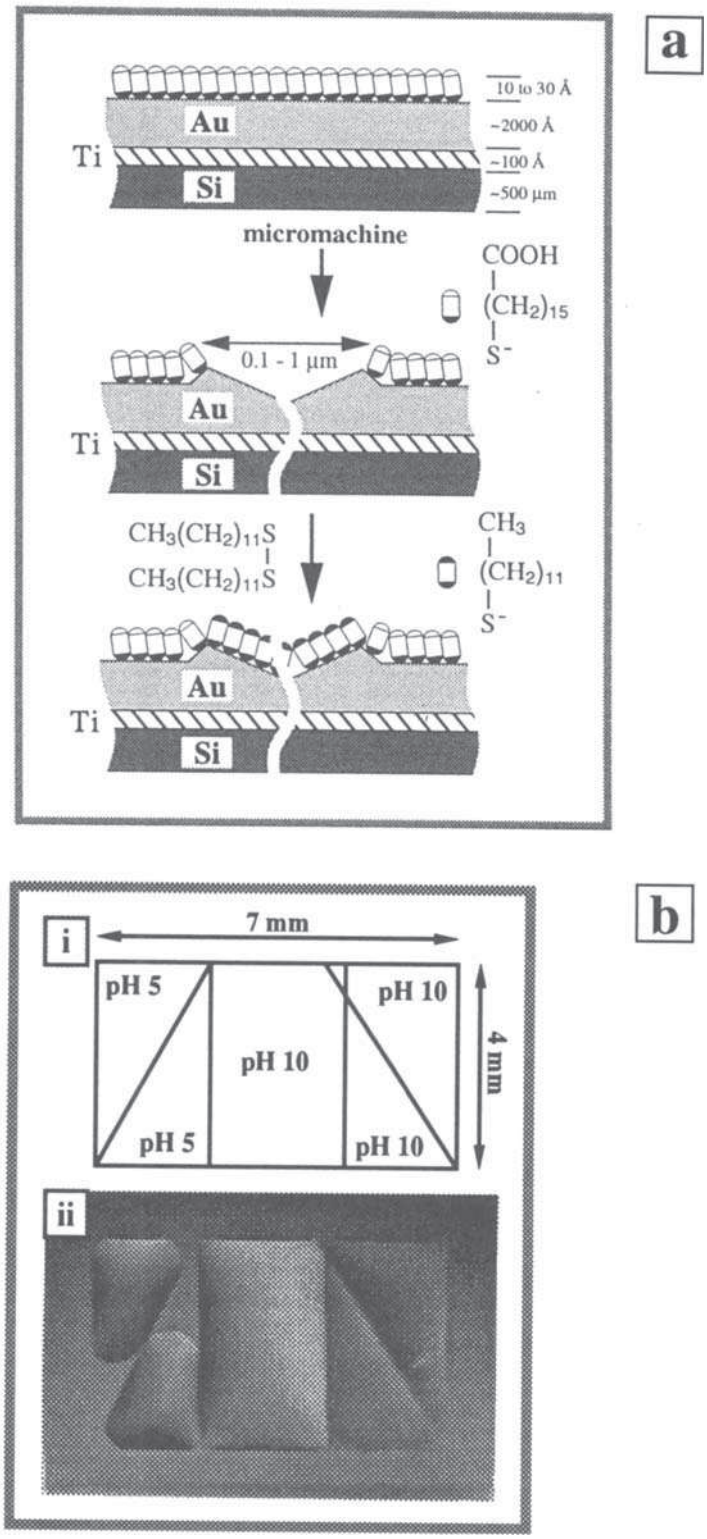


FIG. 6.1 Control of wettability of surfaces through chemistry: (a) schematic illustration of formation of hydrophobic lines on a hydrophilic surface with self-assembled monolayers (SAMs) and micromachining; (b) top view of the shapes and confinement of water drops on an engineered surface. (Reproduced with permission of Abbott et al. 1992.)

of optical switches, electrochemical valves, and pumps for the preparation and operation of diagnostic assays. The technique described above can also be used to generate droplets of controlled shapes so that analyses of such shapes can be used to measure surface excess free energies (Abbott et al. 1994).

Active control of wettability is a subject of current research activity and is beyond our scope here. However, the first step in this process is a study of interfacial tension and contact angle, their basis in thermodynamics, and methods to measure these properties. This is the objective of this chapter.

6.1b Focus of This Chapter

Both this chapter and Chapter 7 are primarily concerned with the equilibrium behavior of surfaces. In this chapter the emphasis is on the surfaces between "pure" phases, whereas in Chapter 7 we consider the effects of solutes on the behavior of surfaces. Most of the phenomena we describe in this chapter will continue to apply to solutions and carry over into the next chapter.

Our major objectives in this chapter are to define surface tension and contact angle and describe how they are measured and what effects they have on the equilibrium behavior of materials.

1. We begin with the definitions and a preliminary look at some measurement techniques in Section 6.2 and show that surface tension can also be thought of as the energy needed to create an interface (Section 6.3).
2. The presence of surface tension has an important implication for the pressures across a curved interface and, as a consequence, for phase equilibria involving curved interphase boundaries. The equation that relates the pressure difference across an interface to the radii of curvature, known as the *Laplace equation*, is derived in Section 6.4, and the implications for phase equilibria are considered for some specific cases in Section 6.5.
3. As mentioned in Section 6.1a, surface tension and contact angle determine wetting phenomena; we examine this in Section 6.6. We take a closer look at the definition of contact angle and some complications associated with it in Section 6.7.
4. Now we are better equipped to consider measurements of surface tension in detail. Section 6.8 describes the use of shapes of menisci, drops, and bubbles for such measurements, and Section 6.9 considers the practically important case of contact of liquids with porous solids and powders.
5. Finally, we close the chapter with a discussion of the relation between molecular interactions and surface tension and contact angle.

6.2 SURFACE TENSION AND CONTACT ANGLE: A FIRST LOOK

6.2a Surface Tension as a Force

As the name implies, surface tension (denoted by γ) is a force that operates on a surface and acts perpendicular and inward from the boundaries of the surface, tending to decrease the area of the interface. In order to illustrate this, we use a simple apparatus based on this notion that might be used to measure surface tension (see Fig. 6.2a). In fact, other methods are used, but the arrangement of Figure 6.2a has the advantage of simplicity and serves to illustrate how the tension in the liquid surface is indeed measured by γ. The figure represents a loop of wire with one movable side on which a film could be formed by dipping the frame into a liquid. The surface tension of a stretched film in the loop will cause the slide wire to move in the direction of decreasing film area unless an opposing force F is applied. In an actual apparatus, the friction of the slide wire might be sufficient for this. In an idealized, frictionless apparatus like that in Figure 6.2a, the force opposing γ could be measured. The force evidently operates along the entire edge of the film and will vary with the length ℓ of the slide wire. Therefore it

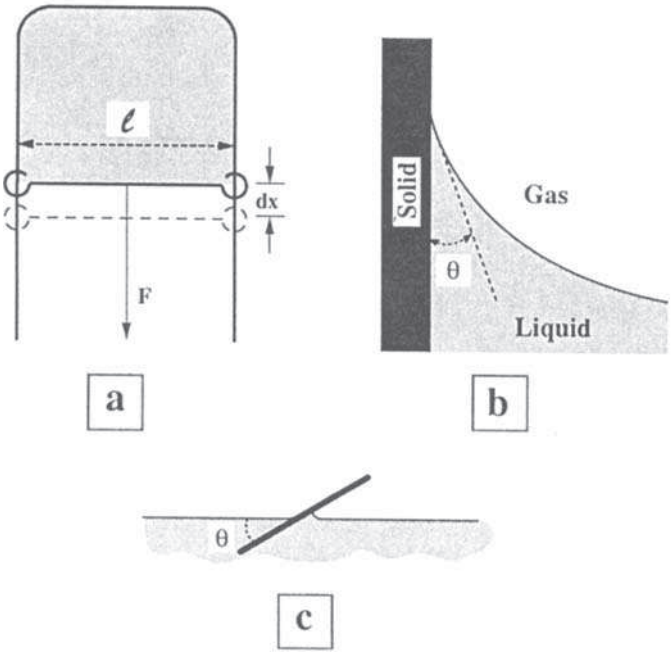


FIG. 6.2 Illustrations of liquid film formation, contact angle, and measurement of contact angle: (a) a wire loop with a slide wire on which a liquid film might be formed and stretched by an applied force **F**. (b) profile of a three-phase (solid, liquid, gas) boundary that defines the contact angle θ. (c) the tilted plate method for measuring contact angles.

is the force per unit length of edge that is the intrinsic property of the liquid surface. Since the film in Figure 6.2a has two sides, the surface tension as measured by this apparatus equals

$$\gamma = F/2\ell \tag{1}$$

Several points should be noted before proceeding any further:

1. Equation (1) defines the units of surface tension to be those of force per length or Nm^{-1} in SI or dynes cm^{-1} in the cgs system. We see presently that these are not the only units used for γ.
2. The apparatus shown in Figure 6.2a resembles a two-dimensional cylinder/piston arrangement. With this similarity in mind, the suggestion that surface tension is analogous to a two-dimensional pressure seems plausible. With certain refinements, this notion will prove very useful in Chapter 7.
3. A gas in the frictionless, three-dimensional equivalent to the apparatus of the figure would tend to expand spontaneously. For a film, however, the direction of spontaneous change is contraction.

6.2b Contact Angle

A quantity that is closely related to surface tension is the *contact angle*. The contact angle θ is defined as the angle (measured in the liquid) that is formed at the junction of three phases, for example, at the solid-liquid-gas junction as shown in Figure 6.2b. Although the surface tension is a property of the two phases that form the interface, θ requires that three phases be specified for its characterization, as mentioned above. The above definition of contact angle is, however, highly simplified, and we take a more in-depth look at the concept later in this chapter.

6.2c Measuring Surface Tension and Contact Angle: Round One

In this section we take an initial look at the experimental determination of γ and θ. The methods we discuss show the complementarity between these two parameters while introducing some important phenomena. As we proceed through the chapter, some additional complications are encountered. We return to the topic of measuring γ and θ in Section 6.8, in which some refinements can be discussed with more meaning.

Contact angle seems like a more straightforward quantity to measure than surface tension. Junctions such as that illustrated in Figure 6.2b are easy to observe and even photograph. This being the case, it seems easy enough to construct a tangent to the liquid surface at the point where it contacts the support and measure the enclosed angle. Actually, this is considerably easier to describe than to carry out. Although the foregoing is impractical as an experimental strategy, it does make the contact angle very accessible conceptually.

Figure 6.2c shows a variation of Figure 6.2b that has been used extensively to measure contact angles. This technique, known as the *tilted plate method*, varies the angle of inclination between a smooth solid penetrating the liquid surface until a position is found at which the liquid makes a horizontal contact with the solid. The horizontal is easier to judge than the tangent to a curve, so this method is preferable to that based on direct measurement of a junction like the one shown Figure 6.2b.

The situation shown in Figure 6.2b is one in which surface tension and contact angle considerations pull a liquid upward in opposition to gravity. A mass of liquid is drawn up as if it were suspended by the surface from the supporting walls. At equilibrium the upward pull of the surface and the downward pull of gravity on the elevated mass must balance. This elementary statement of force balance applies to two techniques by which γ can be measured if θ is known: the *Wilhelmy plate* and *capillary rise*.

6.2c.1 The Wilhelmy Plate Technique

Figure 6.3a represents a thin vertical plate suspended at a liquid surface from the arm of a tared balance. For simplicity, the plate is positioned so that the lower edge is in the same plane as the horizontal surface of the liquid away from the plate as shown in the figure. The manifestation of surface tension and contact angle in this situation is the entrainment of a

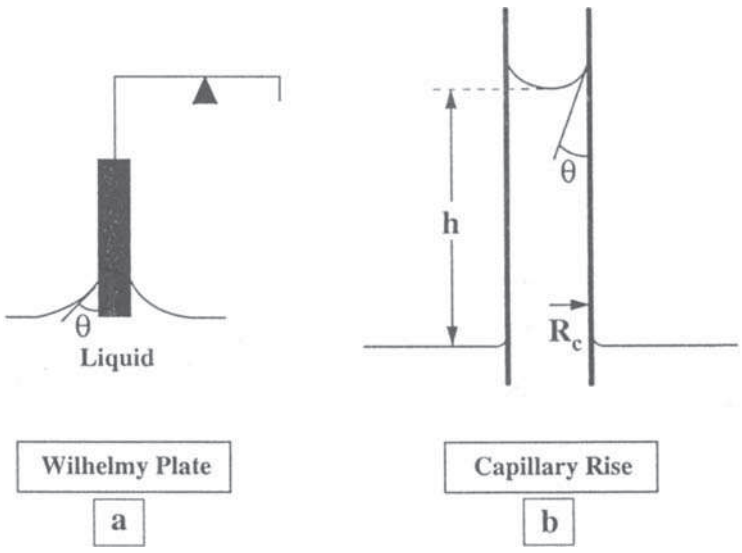


FIG. 6.3 Surface tension and capillary rise: (a) the Wilhelmy plate method for measuring γ; (b) schematic illustration of capillary rise in a cylindrical tube of radius R_c.

meniscus around the perimeter of the suspended plate. Assuming the apparatus is balanced before the liquid surface is raised to the contact position, the imbalance that occurs on contact is due to the weight of the entrained meniscus. Since the meniscus is held up by the tension on the liquid surface, the weight measured by the apparatus can be analyzed to yield a value for γ.

The observed weight w of the meniscus must equal the upward force provided by the surface. This, in turn, equals the vertical component of $\gamma - \gamma \cos \theta$, where θ is the contact angle—times the perimeter of the plate P since surface tension is a force per unit length of surface edge. Therefore, at equilibrium,

$$w = P\gamma \cos \theta \tag{2}$$

For a plate of rectangular cross section having length ℓ and thickness t, $P = 2(\ell + t)$; these dimensions can be accurately measured. By suspending the plate from a sensitive balance, we can also measure w with considerable accuracy. The apparatus is called a *Wilhelmy balance*, and the technique the *Wilhelmy plate method*. Thus, if the contact angle is known from an independent determination by, say, the tilted-plate method, then γ can be evaluated by Equation (2).

Strictly speaking, Equation (2) allows the vertical component of surface tension to be measured. Since this equals $\gamma \cos \theta$, we are actually making a single measurement that involves two parameters. If γ were independently known, the Wilhelmy plate method could also be used to determine θ. Whether we seek to evaluate γ, θ, or both, two experiments are needed, and these may not both involve the factor $\cos \theta$. In Section 6.8a we discuss a second type of measurement that can be made with the Wilhelmy apparatus that supplies a complementary observation so both γ and θ can be determined on a single instrument.

6.2c.2 The Capillary Rise Technique

Capillary rise is also a measure of the vertical component of surface tension, so it is an alternative to the Wilhelmy balance and not its complement. Figure 6.3b shows how a capillary rise experiment is carried out. Conventionally, the height of a liquid column in a capillary above the reference level in a large dish is measured. It is important that the dish be large enough in diameter so that the reference level has a well-defined horizontal surface. In the capillary the liquid will have a curved meniscus, and, as usual, it is the height of the bottom of the meniscus above the horizontal that is measured. It should be noted that capillary depression is also observed, as with mercury, for example. In this case the capillary "rise" is a negative quantity. Our objective is to relate the equilibrium liquid column height h to the surface tension of the liquid.

A simple—but incorrect—relationship between the height of capillary rise, capillary radius, contact angle, and surface tension is easily derived. At equilibrium the vertical component of the surface tension ($2\pi R_c \gamma \cos \theta$) equals the weight of the liquid column, approximated as the weight of a cylinder of height h and radius R_c. This leads to the approximation

$$2\pi R_c \gamma \cos \theta \approx \pi R_c^2 h \Delta\rho g \tag{3}$$

Here $\Delta\rho$ is the difference in the density of the liquid and its surroundings and is used in the above formulation since the surrounding fluid could be a second liquid that has a buoyant effect on the liquid column. Rearranging Equation (3), we obtain

$$\frac{R_c h}{\cos\theta} \approx \frac{2\gamma}{\Delta\rho g} \tag{4}$$

where the cluster of constants on the right-hand side of the equation is called the *capillary constant* and is given the symbol a^2:

$$a^2 = \frac{2\gamma}{\Delta\rho g} \tag{5}$$

Since $a^2 \propto R_c h$, a has units of length. Note that the height to which a liquid climbs in a capillary (assuming $\theta < 90°$) increases as R_c decreases. As would be expected, h is larger for large γ and small $\Delta\rho$. By measuring h for a capillary of known radius, Equation (5) permits an approximate value of γ to be determined if θ is known.

The approximation that limits this analysis of capillary rise originates from neglecting the weight of the liquid in the "crown" of the curved meniscus. We see in Section 6.8b that the height of capillary rise can be related to surface tension without making this approximation, although the connection is somewhat unwieldy. A more detailed description of the experimental aspects of the capillary rise method can be obtained from advanced textbooks (e.g., Adamson 1990).

There are numerous other methods for measuring surface tension that we do not discuss here. These include (a) the measurement of the maximum pressure beyond which an inert gas bubble formed at the tip of a capillary immersed in a liquid breaks away from the tip (the so-called *maximum bubble-pressure method*); (b) the so-called *drop-weight method*, in which drops of a liquid (in a gas or in another liquid) formed at the tip of a capillary are collected and weighed; and (c) the *ring method*, in which the force required to detach a ring or a loop of wire is measured. In all these cases, the measured quantities can be related to the surface tension of the liquid through simple equations. The basic concepts involved in these methods do not differ significantly from what we cover in this chapter. The experimental details may be obtained from Adamson (1990).

It is impossible to complete a discussion of the measurement of surface tension without saying something about the need for extreme cleanliness in any determination of γ. Any precision chemical measurement requires attention to this consideration, but surfaces are exceptionally sensitive to impurities. It is often noted that touching the surface of 100 cm^2 of water with a fingertip deposits enough contamination on the water to introduce a 10% error in the value of γ. Not only must all pieces of equipment be clean, but also the experiments must be performed within enclosures or in very clean environments to prevent outside contamination. In addition, both surface tension and contact angle should be measured under constant temperature conditions.

Both the Wilhelmy and capillary rise methods for determining γ have been based on the concept of surface tension as a force. While this point of view is useful for describing the experimental methods we have discussed, it is only one way of interpreting γ. An energetic interpretation is also possible that makes surface tension amenable to the powerful methods of thermodynamics.

6.3 THERMODYNAMICS OF SURFACES: SURFACE TENSION AS SURFACE FREE ENERGY

Application of a force infinitesimally larger than the equilibrium force to the slide wire in Figure 6.2a will displace the wire through a distance dx. The product of force and distance equals energy, in this case the energy spent in increasing the area of the film by the amount $dA = 2\ell\, dx$. Therefore the work done on the system is given by

$$\text{Work} = Fdx = \gamma 2\ell dx = \gamma dA \tag{6}$$

This supplies a second definition of surface tension: It equals the work per unit area required to produce a new surface. In terms of this definition, the units of γ are energy per area – J m^{-2} in SI or erg cm^{-2} in the cgs system.

We see, therefore, that there are two equivalent interpretations of γ: force per unit length of boundary of the surface and energy per unit area of the surface. The dimensional equivalency of the two is evident if the numerator and denominator of force length $^{-1}$ are multiplied by length.

Equation (6) relates γ to the work required to increase the area of a surface. From thermodynamics, it will be recalled that work is a path-dependent process: How much work is done depends on how it is done. Based on this realization, then, it seems desirable to examine

Equation (6) a little more fully. As already noted, there is a tendency for mobile surfaces to decrease spontaneously in area. Therefore it is convenient to shift our emphasis from work done on the system to work done by the system in such a reduction of area. If the quantity $\delta w'$ is defined as the work done by the system when its area is changed, then Equation (6) becomes

$$\delta w' = -\gamma dA \tag{7}$$

According to Equation (7), a decrease in area (negative dA) corresponds to work done by the system, whereas an increase in area requires work to be done on the system (positive dA and negative $\delta w'$). This sign convention is consistent with the idea expressed in Chapter 1 that energy is stored in surfaces.

We are now in a position to relate the quantity $\delta w'$ to other thermodynamic variables. To do this a brief review of some basic thermodynamics is useful.

According to the first law of thermodynamics, the change in the energy E of a system equals

$$dE = \delta q - \delta w \tag{8}$$

in which δw is the work done by the system and δq is the heat absorbed by the system. The quantity δw is conveniently divided into a pressure-volume term and a non-pressure-volume term:

$$\delta w = \delta w_{pV} + \delta w_{non\text{-}pV} = pdV + \delta w_{non\text{-}pV} \tag{9}$$

It will be recalled from physical chemistry that chemical work is the usual substitution for $\delta w_{non\text{-}pV}$. However, the work defined by Equation (7) may also be classified as non-pressure-volume work.

The second law of thermodynamics tells us that for reversible processes

$$\delta q_{rev} = TdS \tag{10}$$

Substituting Equations (9) and (10) into Equation (8), with the stipulation of reversibility as required by Equation (10), enables us to write

$$dE_{rev} = TdS - pdV - \delta w_{non\text{-}pV} \tag{11}$$

Next we recall the definition of the Gibbs free energy G:

$$G = H - TS = E + pV - TS \tag{12}$$

which may be differentiated to give

$$dG = dE + pdV + Vdp - TdS - SdT \tag{13}$$

Substituting Equation (11) into (13) gives

$$dG_{rev} = TdS - pdV - \delta w_{non\text{-}pV} + pdV + Vdp - TdS - SdT \tag{14}$$

This is a fundamental equation of physical chemistry because it enables us to assign a physical significance to G as defined by Equation (12). Equation (14) shows that for a constant temperature, constant pressure, and reversible process

$$dG = -\delta w_{non\text{-}pV} \tag{15}$$

that is, dG equals the maximum non-pressure-volume work derivable from such a process since maximum work is associated with reversible processes.

We have already seen by Equation (7) that changes in surface area entail non-pressure-volume work. Therefore we identify $\delta w'$ from Equation (7) with $\delta w_{non\text{-}pV}$ in Equation (15) and write

$$dG = \gamma dA \tag{16}$$

Even better, in view of the stipulations made going from Equation (14) to Equation (15), we write

TABLE 6.1 Several Representative Values of γ, S^s, and H^s for a Variety of Liquids Near Room Temperature

Substance	γ at 20°C (mJ m^{-2})	$d\gamma/dT = -S^s$ (mJ m^{-2} K^{-1})	$H^s = \gamma - T(d\gamma/dT)$ (mJ m^{-2})
n-Hexane	18.4	−0.105	49.2
Ethyl ether	17.0	−0.116	51.0
n-Octane	21.8	−0.096	49.9
Carbon tetrachloride	26.9	−0.092	53.9
m-Xylene	28.9	−0.077	51.4
Toluene	28.5	−0.081	52.2
Benzene	29.0	−0.099	58.0
Chloroform	28.5	−0.135	68.3
1,2-Dichloroethane	32.2	−0.139	72.9
Carbon disulfide	32.3	−0.138	72.7
Water	72.8	−0.152	117.3
Mercury	484	−0.220	548

Source: D. H. Kaelbe, *Physical Chemistry of Adhesion*, Wiley, New York, 1971.

$$\gamma = \left(\frac{\partial G}{\partial A}\right)_{T,p} \tag{17}$$

Several things should be noted about Equation (17). This relationship identifies the surface tension as the increment in Gibbs free energy per unit increment in area. *The path-dependent variable $\delta w'$ is replaced by a state variable as a result of this analysis.* Another notation that is often encountered that emphasizes the fact that γ is identical to the excess Gibbs free energy per unit area arising from the surface is to write it as G^s. The energy interpretation of γ, then, has been carried to the point at which it has been identified with a specific thermodynamic function. Many of the general relationships that apply to G apply equally to γ. For example,

$$G^s = \gamma = H^s - TS^s \tag{18}$$

and

$$\left(\frac{\partial G^s}{\partial T}\right)_p = \left(\frac{\partial \gamma}{\partial T}\right)_p = -S^s \tag{19}$$

Equations (18) and (19) may be combined to give

$$\gamma = H^s + T\left(\frac{\partial \gamma}{\partial T}\right)_p \tag{20}$$

For water at 20°C, $\gamma = 7.28 \cdot 10^{-2}$ J m^{-2}; it is convenient to write this in millijoules since 72.8 mJ m^{-2} is numerically equal to the cgs value. For water $d\gamma/dT = -0.152$ mJ m^{-2} deg^{-1}; therefore $H^s = 117$ mJ m^{-2} by Equation (20). Additional values of G^s, S^s, and H^s for various substances are listed in Table 6.1.

6.4 SURFACE TENSION: IMPLICATIONS FOR CURVED INTERFACES AND CAPILLARITY

6.4a Pressure Difference across a Curved Interface: The Laplace Equation

In Section 6.2 we discussed the Wilhelmy and capillary rise experiments as if the supported liquid were hanging from a surface skin. While this is a convenient device for generating

formulas, it is not an adequate description of the physical situation. Now that we have established the connection between surfaces and thermodynamics, we can remedy this situation. The insight that is central to this development is the realization that a pressure difference operates across a curved interface. The pressure difference is such that the greater pressure is on the concave side. Our objective in this section is to relate this pressure difference to the curvature of the surface.

Figure 6.4 shows a portion $ABCD$ of a curved surface. The surface has been cut by two planes perpendicular to one another. Each of the planes therefore contains a portion of arc where it intersects the curved surface. In the figure the radii of curvature are designated R_1 and R_2, and the lengths are designated x and y, respectively, for these two intercepted arcs. Now suppose the curved surface is moved outward by a small amount dz to a new position $A'B'C'D'$. Since the corners of the surface continue to lie along extensions of the diverging radial lines, this move increases the arc lengths to $x + dx$ and $y + dy$. Obviously the area of surface must also increase. The work required to accomplish this must be supplied by a pressure difference Δp across the element of surface area.

The increase in area when the surface is displaced is given by

$$dA = (x + dx)(y + dy) - xy = xdy + ydx + dxdy \approx xdy + ydx \tag{21}$$

where the approximation arises from neglecting second-order differential quantities. The increase in free energy associated with this increase in area is given by γdA:

$$dG = \gamma(xdy + ydx) \tag{22}$$

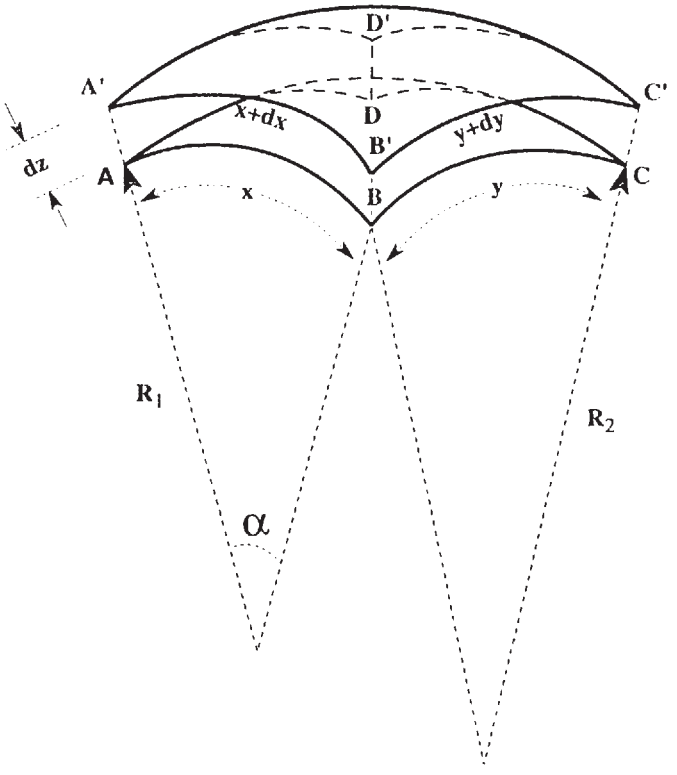


FIG. 6.4 Definition of coordinates describing the displacement of an element of curved surface $ABCD$ to $A'B'C'D'$.

If ordinary pressure-volume work is responsible for the expansion of this surface, then the work equals $\Delta p dV$, where dV is the volume swept by the moving surface. In terms of Figure 6.4, this equals

$$dw = \Delta pxydz \tag{23}$$

Setting Equations (22) and (23) equal to one another gives

$$\gamma(xdy + ydx) = \Delta pxydz \tag{24}$$

Notice that the arc lengths x and $(x + \Delta x)$ in Figure 6.4 are related to the angle α by $x = R_1\alpha$ and $(x + \Delta x) = (R_1 + dz)\alpha$. Therefore, we may set up the following proportions:

$$\frac{x + dx}{R_1 + dz} = \frac{x}{R_1} \tag{25}$$

Similarly,

$$\frac{y + dy}{R_2 + dz} = \frac{y}{R_2} \tag{26}$$

Equations (25) and (26) simplify to

$$\frac{dx}{xdz} = \frac{1}{R_1} \tag{27}$$

and

$$\frac{dy}{ydz} = \frac{1}{R_2} \tag{28}$$

Substituting Equations (27) and (28) into Equation (24) enables us to write the relationship of Δp to R_1, R_2, and γ:

$$\Delta p = \gamma\left(\frac{1}{R_1} + \frac{1}{R_2}\right) \tag{29}$$

This expression is known as the *Laplace equation* and was derived in 1805.

Until now we have not been specific as to the location of the curved surface under consideration. Since this is the case, Equation (29) is general and applies equally well to geometrical bodies with radii of curvature that are constant over the entire surface or to more intricate shapes for which the R_i's change from place to place on the surface. For the former category, there are several special cases of Equation (29) that are worthy of note:

1. For a spherical surface, $R_1 = R_2 = R_s$; therefore

$$\Delta p = 2\gamma/R_s \tag{30}$$

2. For a cylindrical surface, $R_1 \to \infty$; therefore

$$\Delta p = \gamma/R_2 \tag{31}$$

3. For a planar surface, $R_1 = R_2 \to \infty$; therefore

$$\Delta p = 0 \tag{32}$$

It is also possible for a portion of a surface to be locally saddle shaped; in such a case the two radii of curvature lie on opposite sides of the surface and have different signs. It is possible for p to be zero in this situation also.

The Laplace equation applied specifically to spherical surfaces can be derived in a variety of ways. Example 6.1 considers an alternative derivation that points out the thermodynamic character of the result quite clearly.

* * *

EXAMPLE 6.1 *Laplace Equation for Spherical Surfaces: A Thermodynamic Derivation.* The Maxwell relations play an important role in thermodynamics. By including the term dA in the usual differential form for dG, show that $(\partial V/\partial A)_{p,T} = (\partial\gamma/\partial p)_{A,T}$. Evaluate $(\partial V/\partial A)_{p,T}$ assuming a spherical surface and, from this, derive the Laplace equation for this geometry.

Solution: The derivation of the Maxwell relations treats dG as an exact differential and expands it as

$$dG = (\partial G/\partial p)_T dp + (\partial G/\partial T)_p dT$$

The coefficients are then matched with their counterparts in an alternative expression for dG: $dG = Vdp - SdT$. When surface effects are included, these two expressions become

$$dG = (\partial G/\partial p)_{T,A} dp + (\partial G/\partial T)_{p,A} dT + (\partial G/\partial A)_{p,T} dA$$

and

$$dG = Vdp - SdT + \gamma dA$$

or, at constant temperature,

$$dG = (\partial G/\partial p)_A dp + (\partial G/\partial A)_p dA \quad \text{and} \quad dG = Vdp + \gamma dA$$

Since the order of differentiation is immaterial for exact differentials, it follows that

$$\partial/\partial A[(\partial G/\partial p)_A]_p = \partial/\partial p[(\partial G/\partial A)_p]_A$$

or

$$(\partial V/\partial A)_p = (\partial\gamma/\partial p)_A$$

Since dV/dA can be written $(dV/dR_s)(dR_s/dA)$, which is easily evaluated for a sphere of radius R_s as $(4\pi R_s^2)(8\pi R_s)^{-1} = R_s/2$. Therefore $(\partial\gamma/\partial p)_{A,T} = R_s/2$ or $d\gamma = (1/2)\, R_s dp$. This result may be integrated to give $\gamma = (1/2)\, R_s \Delta p$ where the constant of integration has been set equal to zero since there would be no pressure difference if $\gamma = 0$. This is the Laplace equation for spherical particles. ■

* * *

6.4b Laplace Equation and Capillary Rise

As noted above, it is possible that a different pair of radii of curvature applies at different locations on a surface. In this case the Laplace equation shows that Δp also varies with location. This is the reason for the variation of the pressure with z in the meniscus shown in the capillary in Figure 6.3b. As is often true of pressures, it is convenient to define pressure variations relative to some reference plane.

With this idea in mind, the horizontal surface in Figure 6.3b can be taken as a reference level at which $\Delta p = 0$. Just under the meniscus in the capillary the pressure is less than it would be on the other side of the surface owing to the curvature of the surface. The fact that the pressure is less in the liquid in the capillary just under the curved surface than it is at the reference plane causes the liquid to rise in the capillary until the liquid column generates a compensating hydrostatic pressure. The capillary possesses an axis of symmetry; therefore at the bottom of the meniscus the radius of curvature is the same in the two perpendicular planes that include the axis. If we identify this radius of curvature by b, then the Laplace equation applied to the meniscus is $\Delta p = 2\gamma/b$. Equating this to the hydrostatic pressure gives

$$2\gamma/b = \Delta\rho g h \tag{33}$$

or

$$bh = 2\gamma/\Delta\rho g = a^2 \tag{34}$$

Comparison of this result with Equation (4) shows the two relations to be of similar form, with the approximate form becoming exact under specific conditions. The condition under which the approximate result is exact is when the meniscus is a perfectly hemispherical depres-

sion at the liquid surface. When this is the case, $\theta = 0$ and $R_{cap} = R_{men} = b$. Since the meniscus usually only approximates a hemisphere, the use of the capillary radius in Equation (4) remains an approximate relationship. Although Equation (34) is exact, it is not particularly useful since b is not readily measurable. Note that if b were available, Equation (34) would permit the evaluation of γ without requiring that θ be known.

6.4c Arbitrary Variations in Radii of Curvature: Generalization of the Laplace Equation

A variety of drop, bubble, and meniscus shapes have axial symmetry. As is the case in the capillary in Figure 6.3b, p varies with z and, in general, the two radii of curvature may vary from position to position on the surface also. With these ideas in mind, the Laplace equation becomes

$$\Delta p(z) = \gamma[R_1^{-1}(x,y,z) + R_2^{-1}(x,y,z)] \tag{35}$$

In Equation (35) we have added the notion that Δp, R_1^{-1}, and R_2^{-1} may be functions of location in space for any given surface.

The following expressions from analytical geometry are general functions for R_1^{-1} and R_2^{-1} for surfaces with an axis of symmetry:

$$R_1^{-1} = \frac{d^2z/dx^2}{[1 + (dz/dx)^2]^{3/2}} \tag{36}$$

and

$$R_2^{-1} = \frac{dz/dx}{x[1 + (dz/dx)^2]^{1/2}} \tag{37}$$

Substitution of Equations (36) and (37) into Equation (35) generates a complicated differential equation with a solution that relates the shape of an axially symmetrical interface to γ. In principle, then, Equation (35) permits us to understand the shapes assumed by mobile interfaces and suggests that γ might be measurable through a study of these shapes. We do not pursue this any further at this point, but return to the question of the shape of deformable surfaces in Section 6.8b. In the next section we examine another consequence of the fact that curved surfaces experience an extra pressure because of the tension in the surface. We know from experience that many thermodynamic phenomena are pressure sensitive. Next we examine the effect of the increment in pressure small particles experience due to surface curvature on their thermodynamic properties.

6.5 EFFECTS OF CURVED INTERFACES ON PHASE EQUILIBRIA AND NUCLEATION: THE KELVIN EQUATION

In addition to capillarity, another important consequence of the pressure associated with surface curvature is the effect it has on the thermodynamic activity of substances. As a consequence, phase equilibria (including dissolution of chemical species in the different phases) are affected by the presence of interfaces. In this section we consider a few such cases.

6.5a Effect on Vapor-Liquid Equilibria

The influence of curvature on phase equilibria is most readily understood for liquids, for which the activity is measured by the vapor pressure of the liquid. Accordingly, suppose we consider the process of transferring molecules of a liquid from a bulk phase with a vast horizontal surface to a small spherical drop of radius R_s.

According to Equation (32), no pressure difference exists across a plane surface; the pressure is simply p_0, the normal vapor pressure. However, a pressure difference given by Equation (30) exists across a spherical surface. Therefore for liquid-vapor equilibrium at a

spherical surface, both the liquid and the vapor must be brought to the same pressure $p_0 + \Delta p$. If we assume the liquid to be incompressible and the vapor to be ideal, ΔG for the process of increasing the pressure from p_0 to $p_0 + \Delta p$ is given by the following:

1. For the liquid,

$$\Delta G = \int_{p_0}^{p_0 + \Delta p} \overline{V}_L \, dp = \overline{V}_L \, \Delta p = \frac{2 \overline{V}_L \gamma}{R_s} \tag{38}$$

where $\overline{V}_L$ is the molar volume of the liquid.

2. For the vapor, since $\overline{V} = RT/p$ (in view of the ideal gas assumption),

$$\Delta G = RT \ln\left(\frac{p_0 + \Delta p}{p_0}\right) = RT \ln\left(\frac{p}{p_0}\right) \tag{39}$$

When liquid and vapor are at equilibrium, these two values of ΔG are equal:

$$RT \ln\left(\frac{p}{p_0}\right) = \frac{2 \overline{V}_L \gamma}{R_s} = \frac{2 M \gamma}{\rho R_s} \tag{40}$$

since the volume per mole equals M/ρ, where M and ρ are the molecular weight and density of the liquid, respectively. In either of these forms, this expression is known as the *Kelvin equation*. The Kelvin equation enables us to evaluate the actual pressure above a spherical surface and not just the pressure difference across the interface, as is the case with the Laplace equation.

In our derivation of the Kelvin equation, the radius is measured in the liquid. For a gas bubble in a liquid, the same equilibrium is involved, but the bubble radius is measured on the opposite side of the surface. As a consequence, a minus sign enters Equation (40) when it is applied to bubbles. Since γ is on the order of millijoules while RT is on the order of joules, Equation (40) predicts that ln (p/p_0) is very small. It is important to realize, however, that γ/R is also divided by the radius of the spherical particle and therefore becomes more important as R_s decreases. For water at 20°C, the Kelvin equation predicts values of p/p_0 equal to 1.0011, 1.0184, 1.1139, and 2.94 for drops of radius 10^{-6}, 10^{-7}, 10^{-8}, and 10^{-9} m, respectively. For bubbles of the same sizes in liquid water, p/p_0 equals 0.9989, 0.9893, 0.8976, and 0.339. These calculations show that the effect of surface curvature, while relatively unimportant even for particles in the micrometer range, becomes appreciable for very small particles.

The Kelvin effect is not limited to spheres. For example, the "neck" of liquid between two supports is described by a paraboloid of revolution. In this case the radius of curvature of the concave surface is outside the liquid and therefore introduces a minus sign into Equation (40). In addition, the factor of 2 is not required to describe the Kelvin effect for this geometry. Example 6.2 illustrates a test of the Kelvin equation based on this kind of liquid neck.

* * *

EXAMPLE 6.2 *Use of the Kelvin Equation for Determining Surface Tension.* Figure 6.5 shows a plot of experimental data that demonstrates the validity of the Kelvin effect. Necks of liquid cyclohexane were formed between mica surfaces at 20°C, and the radius of curvature was measured by interferometry. Vapor pressures were measured for surfaces with different curvature. Use these data to evaluate γ for cyclohexane. Comment on the significance of the fact that the linearity of Figure 6.5 extends all the way to a p/p_0 value of 0.77.

Solution: For the paraboloid of revolution, the Kelvin equation becomes ln $(p/p_0) = -(M\gamma/\rho RTR_s)$. Therefore a plot of ln (p/p_0) versus $-1/R_s$ is expected to be a straight line of slope $M\gamma/\rho RT$. The slope of the line in Figure 6.5 is $1.13 \cdot 10^{-9}$ m. Equating this with $M\gamma/\rho RT$ and solving for γ, we obtain $\gamma = (1.13 \cdot 10^{-9})(779)(8.314)(293)/0.084 = 0.0255$ J m^{-2} $= 25.5$ mJ m^{-2}, which agrees with the value determined by conventional methods.

In addition to proving the Kelvin effect, the data in Figure 6.5 show that the surface tension of cyclohexane may be regarded as constant down to a radius of curvature given by the reciprocal of the abscissa corresponding to $p/p_0 = 0.77$, which is 4 nm. It is remarkable that the

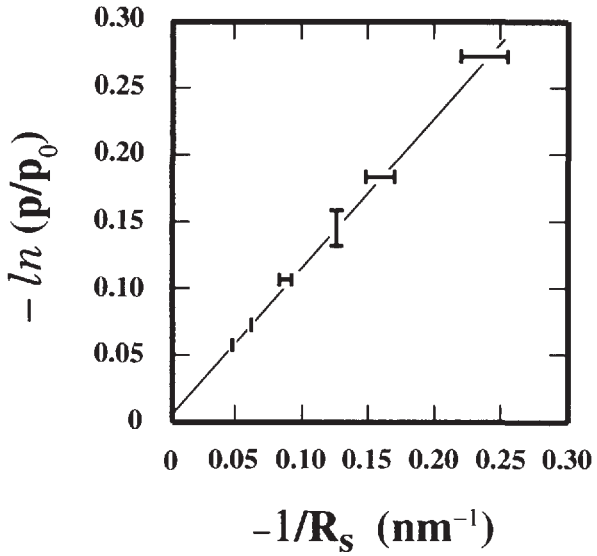


FIG. 6.5 Verification of the Kelvin equation for cyclohexane "necks" of different curvature between mica surfaces (discussed in Example 6.2). (Redrawn from L. R. Fisher and J. N. Israelachvili, *Nature*, 277, 548 (1979).)

same value of γ as determined on macroscopic surfaces continues to apply to within an order of magnitude of molecular dimensions. ■

* * *

6.5b Effect on the Solubility of Solids in Liquids

The Kelvin equation may also be applied to the equilibrium solubility of a solid in a liquid. In this case the ratio p/p_0 in Equation (40) is replaced by the ratio a/a_0, where a_0 is the activity of dissolved solute in equilibrium with a flat surface, and a is the analogous quantity for a spherical surface. For an ionic compound having the general formula M_mX_n, the activity of a dilute solution is related to the molar solubility S as follows:

$$a = (mS)^m(nS)^n \tag{41}$$

Therefore for a solid sphere

$$\frac{2M\gamma}{\rho R_s} = RT\ln\left(\frac{a}{a_0}\right) = (m + n)RT\ln\left(\frac{S}{S_0}\right) \tag{42}$$

where S and S_0 are the solubilities of the spherical and flat particles, respectively. In principle, Equation (42) provides a thermodynamically valid way to determine γ for an interface involving a solid. The thermodynamic approach makes it clear that curvature has an effect on activity for any curved surface. The surface free energy interpretation of γ is more plausible for solids than the surface tension interpretation, which is so useful for liquid surfaces. Either interpretation is valid in both cases, and there are situations in which both are useful. From solubility studies on a particle of known size, γ_S can be determined by the method of Example 6.2.

Although the increase in solubility of small particles is unquestionably a real effect, using it quantitatively as a means of evaluating γ_{SL} is fraught with difficulties:

1. The difference in solubility between a small particle and a larger one will probably be less than 10%. Since a phase boundary exists at all, the solubility is probably low to begin, so there may be some difficulty in determining the experimental solubilities accurately.

2. Solid particles are not likely to be uniform spheres, even if the sample is carefully fractionated; rather, they will be irregularly shaped and polydisperse, although the particle size distribution may be narrow. The smallest particles will have the largest effect on the solubility, but they may be the hardest to measure.
3. The radius of curvature of sharp points or protuberances on the particles has a larger effect on the solubility of irregular particles than the equivalent radius of the particles themselves.

6.5c Effect on Nucleation

The Kelvin equation helps explain an assortment of supersaturation phenomena. All of these—supercooled vapors, supersaturated solutions, supercooled melts—involve the onset of phase separation. In each case the difficulty is the nucleation of the new phase: Chemists are familiar with the use of seed crystals and the effectiveness of foreign nuclei to initiate the formation of the second phase.

The Kelvin equation shows that the ratio S/S_0 or p/p_0 increases rapidly as R_s decreases toward zero. Applying Equation (40) rigorously (and incorrectly) down to $R_s = 0$ would imply infinite supersaturation and make the appearance of a new phase impossible. Equation (40) is derived on the basis of two phases already in existence. To arrive at an understanding of the emergence of a new phase, we must consider what is going on at the molecular level at the threshold of phase separation. A highly purified vapor, for example, may remain entirely as a gas even though its pressure exceeds the normal vapor pressure of the liquid for the temperature in question. At such a point the vapor state is thermodynamically unstable with respect to the formation of a liquid phase with flat surfaces. Whatever stability the vapor has is kinetic stability, arising from the high activation energy required to start the formation of the second phase. A crude kinetic picture of the processes occurring at the molecular level may be informative.

Although a supersaturated vapor is still a gas, it is a very nonideal gas indeed! Clusters of molecules are continually forming and disintegrating; these are the embryonic nuclei of the new phase. Some of these clusters will be dimers, some trimers, and in general n-mers. Each will have its own characteristic radius $R_{s,n}$. An abbreviated derivation of the rate of formation of n-mers proceeds as follows. The rate law will contain both a frequency factor and a Boltzmann factor. The energy term in the latter may be estimated by Equation (38) with volume of the cluster taken to be $(4/3)\pi R_{s,n}^3$. The frequency factor will involve the probability of additional molecules adding by collision; that is, it will depend on the surface area of the cluster ($4\pi R_{s,n}^2$) and the frequency of collisions with a wall as given by kinetic molecular theory. Therefore the rate law may be approximated as

$$\text{Rate} \approx Z(4\pi R_{s,n}^2)\exp\left(-\frac{2\gamma}{R_{s,n}}\frac{4\pi R_{s,n}^3}{3k_B T}\right) = c_1 p R_{s,n}^2 \exp(-c_2 R_{s,n}^2) \tag{43}$$

where Z is the collision frequency. Since the collision frequency certainly increases with p, the pressure has been factored out of the second expression in Equation (43), where c_1 and c_2 are constants.

Two aspects of Equation (43) are especially informative. First, we observe that the pre-exponential term increases with increasing $R_{s,n}$, whereas the exponential term decreases. This means that there exists some critical radius for which the rate law shows a maximum. Clusters with this critical radius may be compared with reaction intermediates or transition states in ordinary chemical reactions. Those clusters that manage to overcome the energy barrier associated with this critical size are capable of further growth, leading to the appearance of the new phase, and smaller clusters disintegrate. In addition, Equation (43) also shows that the rate of cluster growth increases as the pressure increases. All clusters, including those of the critical size, form more rapidly at higher pressures.

The preceding considerations suggest that a point is ultimately reached in the course of increasing supersaturation at which the liquefaction process becomes kinetically as well as

thermodynamically favorable. It must be remembered here that the initial state of the system is one of instability, or, more correctly, metastability. Once liquefaction begins, the pressure drops until the radius-pressure combination that satisfies Equation (40) is reached.

A great deal of work has been done on the kinetics of phase formation; the arguments presented here are intended merely to suggest the direction taken in more detailed treatments. Many aspects of nucleation are of extreme interest in colloid and surface chemistry. The monodisperse colloids are formed by carefully controlling the formation of the solid phase. In the case of the monodisperse sulfur sols, for example, the decomposition of $S_2O_3^{2-}$ (Chapter 5, Section 5.7c) proceeds slowly to quite a high level of supersaturation (the solution must be free of foreign nuclei). Ultimately, clusters of sulfur atoms exceeding the critical size form, nucleate precipitation, and grow until the supersaturation is relaxed. Any additional sulfur that is formed beyond this point will deposit on these particles without forming a new "crop" of nuclei. Relatively monodisperse colloids formed by such condensation processes depend on the fact that supersaturation, and therefore nucleation, occur only once during the formation of the colloid. If the rate of crystallization were too slow, then a second stage of supersaturation and nucleation might be reached, and a polydisperse colloid would result.

Of course, nucleation may also be accomplished by seeding or adding externally formed nuclei. The monodisperse gold sols described in Chapter 5, Section 5.7b, are prepared by this method. The use of AgI crystals and other materials as nuclei in cloud seeding has also been studied extensively. This is especially interesting in view of possible applications to weather modification.

6.6 SURFACE TENSION AND CONTACT ANGLE: THEIR RELATION TO WETTING AND SPREADING PHENOMENA

In the sections above, a variety of important related concepts have been introduced. Among these are the equivalence of surface tension and surface free energy, the applicability of these concepts to solids as well as liquids, and the notion of the contact angle. All these concepts are brought together in discussing the physical situation sketched in Figure 6.6, which is important for understanding wetting, spreading, and related phenomena.

6.6a Relationship Between Surface Tension and Contact Angle: The Young Equation

Suppose a drop of liquid is placed on a perfectly smooth solid surface, and these phases are allowed to come to equilibrium with the surrounding vapor phase. Viewing the surface tensions as forces acting along the perimeter of the drop enables us to write immediately an equation that describes the equilibrium force balance in the horizontal direction:

$$\gamma_{LV} \cos\theta = \gamma_{SV} - \gamma_{SL} \tag{44}$$

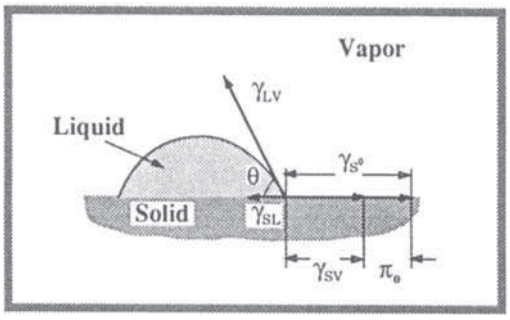


FIG. 6.6 Components of interfacial tension needed to derive Young's equation.

This result was qualitatively proposed by Thomas Young in 1805 and is generally known as *Young's equation*. It is also known as the *Young-Dupré equation* or the *Dupré equation* (MacRitchie 1990).

Young's equation is a plausible, widely used result, but its apparent simplicity is highly deceptive. The two terms that involve the interface between the solid and other phases cannot be measured independently, so experimental verification is difficult, although a variety of experiments have been directed along these lines. There are also a number of objections to Young's equation. These objections may be classified into two categories:

1. Those based on the noncompliance of the experimental system to the assumptions in the derivation
2. Those critical of the assumption of thermodynamic equilibrium in the solid

We discuss these two classes of objections separately.

6.6a.1 Effects of Surface Heterogeneities

Real solid surfaces may be quite different from the idealized one in the above derivation. Actual solid surfaces are apt to be rough and even chemically heterogeneous. This statement is true on a fine scale, even for carefully prepared surfaces. We discuss these complications in more detail in Section 6.7. In Chapter 9 we examine metal surfaces specifically to see the extreme conditions that must exist for these surfaces to be uniform down to an atomic scale.

The model surface of the derivation is thus the exception rather than the rule for solid surfaces. In principle, both roughness and heterogeneity can be incorporated into Young's equation in the form of empirical corrections. For example, if a surface is rough, a correction factor β is traditionally introduced as a weighting factor for $\cos\theta$, for which $\beta > 1$. The logic underlying this correction goes as follows. The factor $\cos\theta$ enters Equation (44) because of the projection of γ_{LV} onto the solid surface in the force balance. If the solid surface is rough, the corresponding surface area will be larger, but much of it will be "overshadowed" (i.e., the surface area will be underestimated) by the projection since Equation (44) assumes that the surface is smooth. The *roughness factor* β corrects for this effect (that is why $\beta > 1$). With the empirical correction factor for roughness included, Young's equation becomes

$$\beta\gamma_{LV}\cos\theta = \gamma_{SV} - \gamma_{SL} \tag{45}$$

A surface may also be chemically heterogeneous. Assuming, for simplicity, that the surface is divided into fractions f_1 and f_2 of chemical types 1 and 2, we may write

$$\gamma_{LV}\cos\theta = f_1(\gamma_{S_1V} - \gamma_{S_1L}) + f_2(\gamma_{S_2V} - \gamma_{S_2L}) \tag{46}$$

where $f_1 + f_2 = 1$.

Both roughness and heterogeneity may be present in real surfaces. In such a case, the correction factors defined by Equations (45) and (46) are both present. Although such modifications adapt Young's equation to nonideal surfaces, they introduce additional terms that are difficult to evaluate independently. Therefore the validity of Equation (44) continues to be questioned.

6.6a.2 Effect of the State of Equilibrium at the Solid Surface

More fundamental objections to Young's equation center on the issue of whether the surface is in a true state of thermodynamic equilibrium. In short, it may be argued that the liquid surface exerts a force perpendicular to the solid surface, $\gamma_{LV}\sin\theta$. On deformable solids a ridge is produced at the perimeter of a drop; on harder solids the stress is not sufficient to cause deformation of the surface. This is the heart of the objection. Is it correct to assume that a surface under this stress is thermodynamically the same as the idealized surface that is free from stress? Clearly, the troublesome stress component is absent only when $\theta = 0$, in which case the liquid spreads freely over the surface, and Figure 6.6 becomes meaningless.

In answer to this, the following argument has been suggested based on the fact that it is the difference $(\gamma_{SV} - \gamma_{SL})$ that appears in Young's equation. Since the same solid is common to both terms, it is really only the local difference at the surface between an adjacent phase

that is liquid and one that is vapor that is being measured. According to this point of view, the nonequilibrium state of the solid is immaterial: The same solid is involved at both interfaces.

Another approach to resolving this theoretical objection to Young's equation is to eliminate the difference $(\gamma_{SV} - \gamma_{SL})$ from the equation entirely, replacing it by some equivalent quantity, thereby shifting attention away from the notion of solid surface tension. An example of this is the use of the heat of immersion to test Young's equation. We return to this in Section 6.6c.

In summary, then, Young's equation is still controversial despite the fact that it has been in existence since the beginning of the 19th century. The reader will appreciate that any relationship that has been around so long and has eluded definitive empirical verification has been the center of much research. Accordingly, the relationship is very widely encountered in the literature of surface chemistry.

6.6b Young's Equation and Equilibrium Film Pressure

To explore Young's equation still further, suppose we distinguish between γ_{SV} and γ_{S°, where the former describes the surface of a solid in equilibrium with the vapor of a liquid and the latter a solid in equilibrium with its own vapor. Since Young's equation describes the three-phase equilibrium, it is proper to use γ_{SV} in Equation (44). The question arises, however, what difference, if any, exists between these two γ's. In order to account for the difference between the two, we must introduce the notion of adsorption. In the present context adsorption describes the attachment of molecules from the vapor phase onto the solid surface. All of Chapter 9 is devoted to this topic, so it is unnecessary to go into much detail at this point. The extent of this attachment depends on the nature of the molecules in the vapor phase, the nature of the solid, and the temperature and the pressure.

For now we may anticipate a result from Chapter 7 to note that adsorption always leads to a decrease in γ. In the present context, therefore, we write

$$\gamma_{S^\circ} \geq \gamma_{SV} \tag{47}$$

We shall use the symbol π_e to signify the difference

$$\gamma_{S^\circ} - \gamma_{SV} = \pi_e \tag{48}$$

and call this quantity the *equilibrium film pressure*. The word *equilibrium* in this designation refers explicitly to the fact that the adsorbed molecules are in equilibrium with a drop of bulk liquid. The molecules adsorbed at an interface may be regarded as repelling one another or as rebounding off one another, thereby relieving some of the tension in the surface. This interpretation makes it sensible to call the reduction of γ due to adsorption π_e a "pressure." Note that π_e is a two-dimensional pressure, measuring the force exerted per unit length of perimeter (Newtons per meter) by the adsorbed molecules. We shall have a good deal more to say about this quantity in the following chapter. With these ideas in mind, Equations (44) and (48) may be combined to give

$$\gamma_{LV} \cos\theta = \gamma_{S^\circ} - \pi_e - \gamma_{SL} \tag{49}$$

Figure 6.6 shows the relationship among γ_{S°, γ_{SV}, and π_e. It is apparent that the value of θ might be quite different between equilibrium and nonequilibrium situations, depending on the value of π_e. There are several concepts that will assist us in anticipating the range of π_e values:

1. Spontaneously occurring processes are characterized by negative values of ΔG.
2. Surface tension is the surface excess free energy; therefore the lowering of γ with adsorption is consistent with the fact that adsorption occurs spontaneously.
3. Surfaces that initially possess the higher free energies have the most to gain in terms of decreasing the free energy of their surface by adsorption.
4. High-energy surfaces bind enough adsorbed molecules to make π_e significant. On the other hand, π_e is negligible for a solid that possesses a low-energy surface.

5. A surface energy in the neighborhood of 100 mJ m^{-2} is generally considered the cutoff value between high- and low-energy surfaces.
6. Silica, glass, metals, metal oxides, metal sulfides, and inorganic salts are examples of high-energy surfaces. Most solid organic compounds, including polymers, have low-energy surfaces.
7. Hard and soft solids are generally classified as having high- and low-energy surfaces, respectively. These criteria must be used cautiously since the mechanical properties of a solid depend on the concentration of defects and dislocations in the bulk.

Because of the simplification that results from $\pi_e = 0$ for low-energy surfaces, they are often chosen as model systems in fundamental research. Even when neglecting π_e is of questionable validity, γ_{SV} and γ_{S° are often used interchangeably for lack of suitable data. We see in Chapter 9—Equation (9.7), for example—how π_e may be determined from experimental adsorption data. Otherwise, we generally assume $\pi_e = 0$.

6.6c Young's Equation and Heat of Immersion

6.6c.1 Enthalpy of Wetting

Small quantities of heat are generally evolved when a dry solid is immersed in a liquid. This can be measured calorimetrically and is called the *heat of immersion*. The physical process with which this heat is associated may be represented by the following equation:

$$\text{Dry solid} \xrightarrow{\text{liquid}} \text{wet solid} \tag{A}$$

Following the usual thermochemical convention, the heat of immersion ΔH_{im} may be written $H_{wet} - H_{dry}$, with these enthalpies expressed per unit area. Since heat is released by this process, ΔH_{im} is negative. The "wet" surface clearly describes the *SL* interface. For the "dry" surface, we ignore π_e and describe the surface by the S° notation. Therefore the heat of immersion may also be written

$$\Delta H_{im} = H^s_{SL} - H^s_{S^\circ} \tag{50}$$

Next we recall Equation (20), which gives us an expression for H^s:

$$H^s = \gamma - T\frac{\partial \gamma}{\partial T} \tag{51}$$

Applying Equation (51) to the right-hand side of Equation (44)—with $\gamma_{SV} = \gamma_{S^\circ}$—gives

$$\Delta H_{im} = \left(\gamma_{SL} - T\frac{\partial \gamma_{SL}}{\partial T}\right) - \left(\gamma_{S^\circ} - T\frac{\partial \gamma_{S^\circ}}{\partial T}\right) \tag{52}$$

Using $\gamma_{LV}\cos\theta$ as a replacement for $\gamma_{S^\circ} - \gamma_{SL}$ we obtain

$$-\Delta H_{im} = \gamma_{LV}\cos\theta - T\frac{d}{dT}(\gamma_{LV}\cos\theta) \tag{53}$$

Carrying out the indicated differentiation yields

$$-\Delta H_{im} = \gamma_{LV}\cos\theta - T\cos\theta\frac{d\gamma_{LV}}{dT} - T\gamma_{LV}\frac{d\cos\theta}{dT} \tag{54}$$

which shows how Young's equation may be tested by comparing experimental heats of immersion with calculated values. Example 6.3 gives an idea of the magnitude of some of these quantities.

* * *

EXAMPLE 6.3 *Determination of Heat of Immersion from Surface Tension and Contact Angle.* Estimate the heat of immersion for the system for which γ and θ are 22 mJ m^{-2} and 30°,

respectively, at 20°C. The temperature variations of γ and $\cos\theta$ are -0.10 mJ m^{-2} K^{-1} and 0.0010 K^{-1}, respectively. These values are close to those observed for liquid alkanes on Teflon. Comment on the implications of this estimate for the ease or difficulty in measuring ΔH_{im}.

Solution: We can estimate ΔH_{im} by substitution into Equation (54):

$$\begin{aligned}\Delta H_{im} &= -\gamma_{LV}\cos\theta + T\cos\theta\,(d\gamma_{LV}/dT) + T\gamma_{LV}\,(d\cos\theta/dT)\\ &= -(22)\cos 30 + 293\cos 30\,(-0.10) + 293\,(22)\,(0.0010)\\ &= -38\text{ mJ m}^{-2}\end{aligned}$$

To get an idea of the problems associated with this kind of experiment, we estimate the temperature change in the liquid as a result of absorbing this heat. Using 2.4 J g^{-1} K^{-1} for the heat capacity (the value for *n*-octane) and taking $T = 1.6$ K as an arbitrary but convenient temperature change, we calculate

$$2.4\text{ J K}^{-1}/\text{g} \cdot 1.6\text{ K} \cdot (\text{m}^2/0.038\text{ J}) \approx 100\text{ m}^2\text{ g}^{-1}$$

That is, for each gram of liquid, about 100 m^2 of surface must be wet. This overestimates the temperature change since it ignores the fact that the solid will absorb some heat. If the heat capacity of the solid is the same as the liquid, 100 m^2 of surface must be wet for each gram of solid-liquid mixture. Using a smaller temperature change would decrease the required area proportionately, but the fact remains that large areas are required to obtain measurable effects. Practically, this means work with powdered solids of small particle size. ■

* * *

6.6c.2 Comments on the Measurement of Heats of Immersion

Although heats of immersion are small, this quantity is measurable. For systems in which both the heat of immersion and the necessary information about γ and θ have been measurable, the prediction of Equation (54) has been verified. Figure 6.7 shows some experimental results for *n*-alkanes wetting Teflon (polytetrafluoroethylene) surfaces. The open circles were determined calorimetrically; the closed ones were calculated from Equation (54). Even though the two sets of values diverge for alkanes larger than *n*-decane, the overall picture is quite acceptable. Incidentally, the value calculated in the example is close to the actual values, even though the numbers used in Example 6.3 were rounded off.

Several aspects of Figure 6.7 and Example 6.3 deserve further comment:

1. Calorimetry requires large areas of interface, which virtually demands powdered solids. We have not considered (yet) the problem of measuring θ and $d\theta/dT$ for powders. Solids that are available as large specimens with smooth surfaces (suitable,

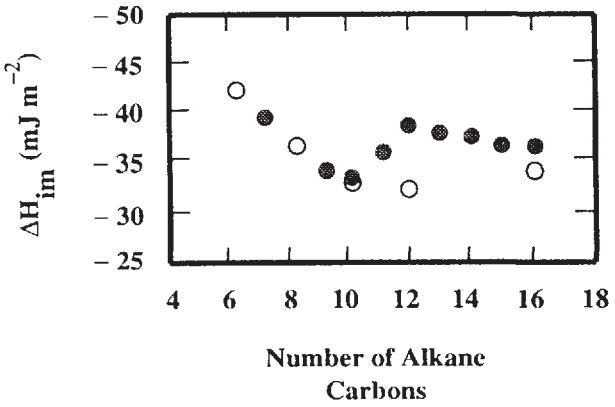


FIG. 6.7 Heats of immersion determined by calorimetry and calculated by Equation (54) for Teflon in various alkanes. (Redrawn from A. W. Neumann, *Adv. Colloid Interface Sci.*, **4**, 105 (1974).)

say, for the tilted plate determination of θ) can be pulverized to increase their surface. The reverse process is often not possible for powders, so a method for measuring powders is important.

2. Results like those shown in Figure 6.7 may be considered an experimental verification of Young's equation. In light of the first item, however, there may still be objections that the surfaces used for calorimetric studies and those used to study γ and θ are not identical even though they are nominally the same.
3. Accepting Equation (54) and Young's equation, on which it is based, suggests calorimetry as a method for measuring contact angles. At this time this is not practical, but the implication that contact angle is a thermodynamic property is a very important realization.

Although we established the thermodynamic significance of γ early in the chapter, θ has been allowed to drift. Its role is clear when we think of surface tension as a force: We use θ to project γ in a specified direction. In thermodynamic terms, contact angle has been an outsider in our presentation. Young's equation is the remedy to this. Rewriting Equation (44), we observe

$$\cos\theta = \frac{\gamma_{SV} - \gamma_{SL}}{\gamma_{LV}} \tag{55}$$

and $\cos\theta$ is fully described by various free energy terms. In view of this new-found (for us) significance, we return shortly to some additional experimental methods for the determination of θ. First, however, we consider the notions of adhesion, cohesion, and spreading, which will enhance our ideas of γ and θ as thermodynamic quantities.

6.6d Surface Tension and Cohesion, Adhesion, and Spreading

In this section we consider some hypothetical processes that provide us with additional ways of thinking about γ and θ. They are related to process (A) stated in Section 6.6c.1 but are not studied calorimetrically as is the case with immersion. In defining adhesion, cohesion, and spreading, we designate the phases A and B without specifying their physical state. Their surface with each other is designated AB; their individual surfaces with their own vapor or air (we make no distinction here) are designated by either A or B. With this notation in mind, we consider the following processes as they affect a unit area:

1. Cohesion:

No surface $\rightarrow$ $2A$ (or B) surfaces (B)

2. Adhesion:

$1AB$ surface $\rightarrow$ $1A + 1B$ surface (C)

3. Spreading (B on A):

$1A$ surface $\rightarrow$ $1AB + 1B$ surface (D)

Schematic illustrations of these processes are shown in Figure 6.8. Two things must be remembered about these sketches: One unit of surface is affected by the processes, and the shape of the affected area is immaterial. It is understood that these are elements of volume and area that are portions of macroscopic samples. Our interest is in the free energy change accompanying each process.

6.6d.1 Work of Cohesion

In Figure 6.8a—which applies to a pure liquid—the process consists of producing two new interfaces, each of unit cross section. Therefore, for the separation process,

$$\Delta G = 2\gamma_A = W_{AA} \tag{56}$$

The quantity W_{AA} is known as the *work of cohesion* since it equals the work required to pull apart a column of liquid A apart. It measures the attraction between the molecules of the two

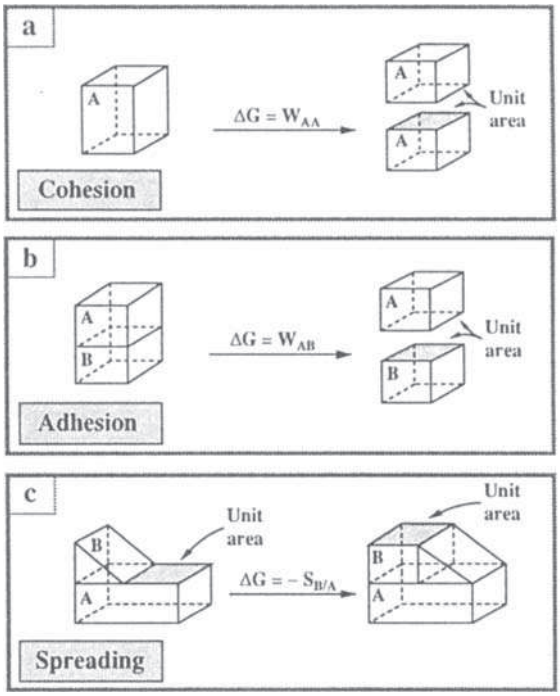


FIG. 6.8 Work of cohesion, adhesion, and spreading. Schematic illustrations of the processes for which ΔG equals (a) the work of cohesion; (b) the work of adhesion; and (c) the work of spreading.

portions. Recall the concept of cohesive energy density in terms of which we discussed in Section 3.4b interactions between pairs of identical molecules. Interpreting γ as half the work of cohesion shows that *surface tension measures the free energy change involved when molecules from the bulk of a sample are moved to the surface.*

6.6d.2 Work of Adhesion

Now let us consider the value of ΔG for the separation of A and B as represented in process (C) and Figure 6.8b. Taking the difference between the final and the initial free energies for this process yields

$$\Delta G = W_{AB} = \gamma_{final} - \gamma_{initial} = \gamma_A + \gamma_B - \gamma_{AB} \tag{57}$$

This quantity is known as the *work of adhesion* and measures the attraction between the two different phases.

The work of adhesion between a solid and a liquid phase may be defined by analogy with Equation (57):

$$W_{SL} = \gamma_{S^\circ} + \gamma_{L^\circ} - \gamma_{SL} \tag{58}$$

By means of Equation (48), γ_{S° may be eliminated from this expression to give

$$W_{SL} = \gamma_{SV} + \pi_e + \gamma_{L^\circ} - \gamma_{SL} \tag{59}$$

Finally, Young's equation may be used to eliminate the difference $(\gamma_{SV} - \gamma_{SL})$:

$$W_{SL} = \gamma_{L^\circ}(1 + \cos\theta) + \pi_e \tag{60}$$

Neglecting π_e as we did in the last section, Equation (59) shows how γ and θ combined measure the work of adhesion between a solid and a liquid.

It is informative to apply Equation (60) to low-energy surfaces for two extreme values of θ, 0° and 180°, for which cos θ is 1 and -1, respectively. For $\theta = 0°$, $W_{SL} = 2\gamma_{LV} = W_{AA}$; the work of solid-liquid adhesion is identical to the work of cohesion for the liquid. In this case interactions between solid and solid, liquid and liquid, and solid and liquid molecules are all equivalent. At the other extreme, with $\theta = 180°$, $W_{SL} = 0$. In this case the liquid is tangent to the solid; there is no interaction between the phases.

6.6d.3 Spreading Revisited: The Spreading Coefficient

Last, if we take the difference between the final and initial states for the process of spreading B over A (symbolized B/A) we obtain

$$\Delta G = \gamma_{AB} + \gamma_B - \gamma_A \tag{61}$$

As usual with free energies, a negative value for $\Delta G_{B/A}$ means that the process represented by process (D) and shown in Figure 6.8c occurs spontaneously. The negative of $\Delta G_{B/A}$ is called the *spreading coefficient* $S_{B/A}$; because of the sign change, a positive spreading coefficient means B spreads freely over A and wets it. The concept of wetting is very important in numerous applications: With lubricants and adhesives it is desirable, and in waterproofing it is undesirable. Since additives are frequently mixed into liquids to affect this property, it is appropriate to postpone any discussion of applications of the spreading coefficient until the next chapter.

By combining Equations (56), (57), and (61), we note that the spreading coefficient can also be written

$$S_{B/A} = W_{AB} - W_{BB} \tag{62}$$

If $W_{AB} > W_{BB}$, the A-B interaction is sufficiently strong to promote the wetting of A by B. This is the significance of a positive spreading coefficient. Conversely, no wetting occurs if $W_{BB} > W_{AB}$ since the work required to overcome the attraction between two B molecules is not compensated by the attraction between A and B. Thus a negative spreading coefficient means that B will not spread over A.

6.7 CONTACT ANGLES: SOME COMPLICATIONS

6.7a Advancing and Receding Contact Angles

In Section 6.2 we saw how the tilted plate method could be used to measure θ; we also noted that it could be determined by the Wilhelmy method if γ were measured independently. For that matter, the three-phase junction might be examined and θ determined by direct observation. These and many other methods—some of which we discuss later—have been used to measure θ. Even if the most careful experimental techniques are employed on carefully prepared surfaces, contact angle data are frequently confusing. The situation is best introduced by referring to Figure 6.9, which shows a drop on a tilted plane. "Teardrop" shapes such as this are familiar to everyone—just look at a raindrop on a window pane. The problem, of course, is that the contact angle is different at different points of contact with the support. It is conventional to call the larger value the *advancing angle* θ_a, and the smaller one the *receding angle* θ_r. The two may be quite different. The presence of contamination is definitely a contributing factor, but it is by no means the only one. Therefore, even with carefully purified materials, both advancing and receding contact angles should be measured.

All the techniques described here are easily conducted, so that both θ_a and θ_r may be observed. When the tilted plate method is used to evaluate the contact angle, θ_r values are obtained if the plate has been pulled out (emersion) from the liquid; θ_a results if the plate is pushed into the liquid (immersion). Likewise, both values of θ may be obtained from the Wilhelmy method, depending on whether the liquid is making an initial contact (θ_a) with the plate or is draining from it (θ_r).

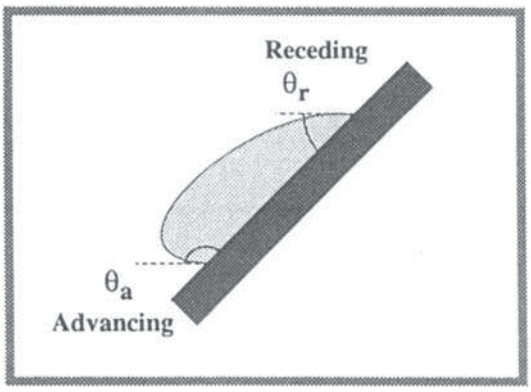


FIG. 6.9 A drop on a tilted plane, showing advancing and receding contact angles. (Adapted from Johnson and Dettree 1969.)

6.7b Contact Angle Hysteresis

A rather interesting example of the difference between advancing and receding contact angles is obtained from the Wilhelmy plate method when the contact angle has a nonzero value. Suppose the Wilhelmy plate is allowed to dip beneath the horizontal liquid surface, as shown in Figure 6.10a. In this case the weight of the meniscus as given by Equation (2) will be decreased by a term w', the buoyant force on the submerged plate. This buoyant force will clearly be proportional to the depth of immersion d. Therefore we write

$$w = \gamma P \cos\theta - w' = \gamma P \cos\theta - kd \tag{63}$$

where k is a suitable proportionality constant. This equation shows that the apparent weight w of the meniscus should give a straight line when plotted against the depth of immersion, and that the intercept should be proportional to $\cos\theta$. If a single value of θ applies to both the immersion and emersion steps, then the line shown in Figure 6.10b would result. Because of the difference between θ_a and θ_r, a curve like that shown in Figure 6.10c is obtained instead. When the immersion-emersion cycle is repeated, a hysteresis loop is obtained that may be as reproducible as the analogous loops observed in magnetization-demagnetization cycles. The difference $(\theta_a - \theta_r)$ is called the *hysteresis of a contact angle*.

The general requirement for hysteresis is the existence of a large number of metastable states that differ slightly in energy and are separated from each other by small energy barriers. The situation is shown schematically in Figure 6.11. The equilibrium contact angle corresponds to the free energy minimum. However, systems may be "frozen" in metastable states of somewhat higher energy by lacking sufficient energy to overcome the energy barrier separating them from equilibrium. An interesting experimental observation is that advancing and receding values of θ converge to a common value when the surface is vibrated. Presumably, the mechanical energy imparted to the liquid by the vibration assists it in passing over the energy barrier and reaching equilibrium. In this sense maximum vibration rather than vibration-free conditions may appear to be the ideal conditions for measuring θ, but "vibration" is a difficult parameter to control reproducibly, so this concept is of little practical help.

Now let us briefly consider the origin of these metastable states. If we exclude the effect of impurities, the metastable states are generally attributed to the roughness of the solid surface, its chemical heterogeneity, or both. Of course, a well-prepared laboratory sample will be fabricated and cleaned as effectively as possible to eliminate gross roughness and chemical heterogeneity. What we are talking about are microscopic irregularities that cannot be eliminated. In size and distribution, these will follow a random pattern on actual surfaces.

An informative model for contact angle hysteresis is obtained by postulating the surface

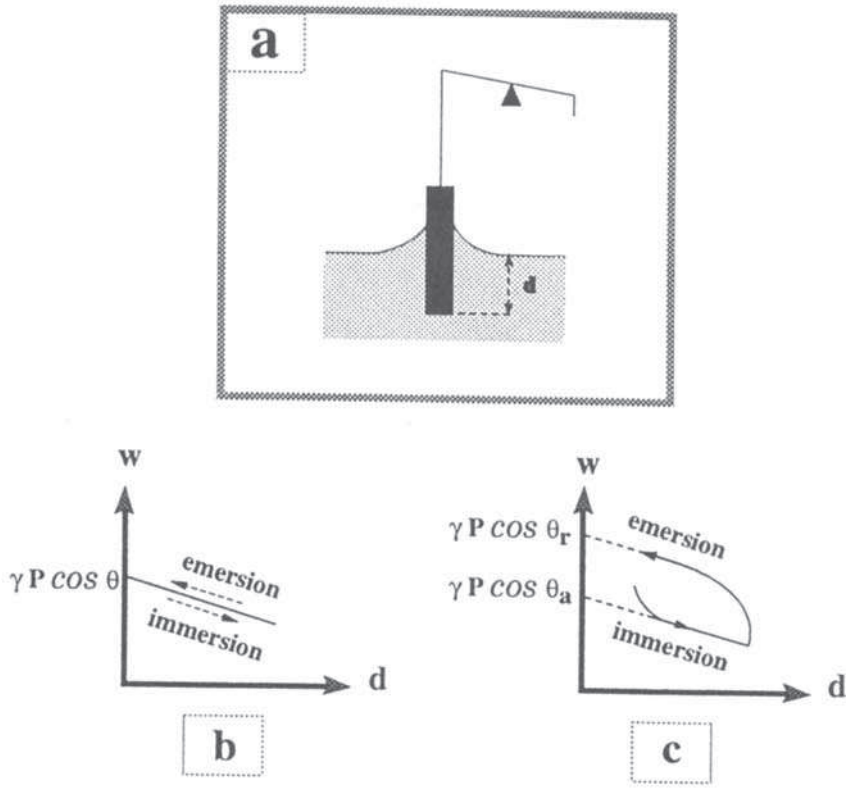


FIG. 6.10 Contact angle hysteresis: (a) weighing a meniscus in a Wilhelmy plate experiment versus the depth of immersion of the plate; (b) both the advancing and receding contact angles are equal; (c) $\theta_a > \theta_r$.

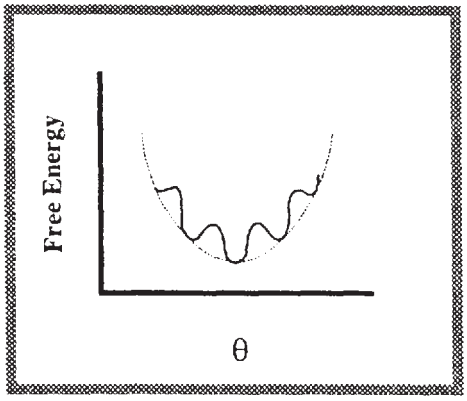


FIG. 6.11 Schematic energy diagram for metastable states corresponding to different contact angles.

to contain a set of concentric grooves on which a drop rests. Figure 6.12 represents the profile of two different drops on such a surface. In both of the profiles shown, the angles of contact between the liquid-vapor interface and the solid are identical, θ_0. With respect to the horizontal, however, two very different apparent contact angles are observed. The two extremes are identified as θ_a and θ_r in this model.

The two drop configurations in Figure 6.12 differ in surface area and in the elevation of their centers of gravity, thus they possess different energies. The change from one configuration to the other involves the distortion of the shape of the drop, which accounts for the energy barrier between the two configurations. Thus the model qualitatively accounts for the kind of metastable states shown in Figure 6.11 that are required for hysteresis. According to this model, contact angle hysteresis arises when a three-phase boundary becomes trapped in transit, lacking sufficient energy to surmount the energy barrier to a lower energy state. The teardrop profile of Figure 6.9 corresponds to the situation in which one edge of the drop has one configuration while the other edge has the second configuration.

This model can also be applied to hysteresis that arises from chemical heterogeneity; however, this time the surface is assumed to be smooth and to contain concentric rings of different chemical composition and hence different θ's. Actual heterogeneity may arise from impurities concentrated at the surface, from crystal imperfections, or from differences in the properties of different crystal faces. The distribution of such heterogeneities on an actual surface will obviously be more complex than the model considers, but the qualitative features of hysteresis are explained by the model nevertheless. Johnson and Dettree (1969) present additional details of model experiments of this sort.

If a surface consists of a patchwork array of high- and low-energy sites, it is generally assumed that the larger, advancing contact angle measures the contact angle that would be obtained for a smooth, homogeneous sample of the low-energy component:

$$\theta_a \approx \theta_{low\ E} \tag{64}$$

Conversely, the receding angle is taken as a measure of θ appropriate to the high-energy component:

$$\theta_r \approx \theta_{high\ E} \tag{65}$$

Both of these conclusions are consistent with Young's equation. All other things being equal in Equation (55), the larger $\gamma_{S\ell} \approx \gamma_{S^\circ}$ is, the larger $\cos\theta$ will be. Since the cosine increases as θ decreases, large values of γ_{S° result in smaller contact angles, and vice versa. We noted

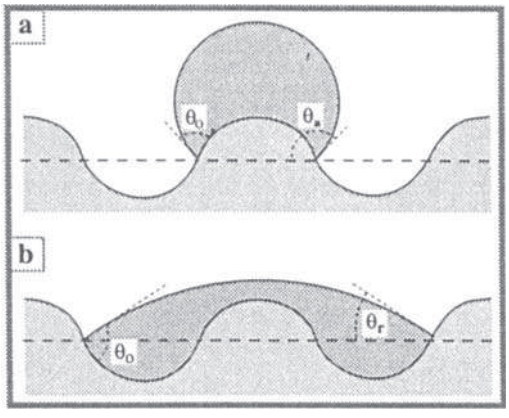


FIG. 6.12 Cross section of a drop resting on a surface containing a set of concentric grooves. For both profiles the contact angles are microscopically identical although macroscopically different. (Adapted from Johnson and Dettree 1969.)

previously that high-energy surfaces have more to gain by adsorption than low-energy surfaces. This helps explain why receding contact angles are less reproducible than advancing angles; adsorption by impurities affects this measurement the most. Correlations with surface roughness reveal that receding angles are also more sensitive to roughness than advancing angles. For these reasons θ_a is more reliable than θ_r as a characteristic of the solid. Obviously, the lower the hysteresis is, the more we can be sure that this quantity truly represents the solid.

In addition to roughness and heterogeneity, time-dependent effects also influence contact angle hysteresis. Dynamic studies have been conducted with liquid fronts moving across a solid substrate at variable speeds. Even the extrapolation to zero rate is unsatisfactory since the limiting dynamic values differ from equilibrium values. This suggests that insufficient mobility at the molecular level may also contribute to hysteresis. Viscous effects operating in the interface may contribute other forces besides those shown in Figure 6.6 that influence the value of the contact angle. This type of effect is harder to analyze than roughness and heterogeneity, but future research may help clarify this third cause of hysteresis.

With this background, let us now return to experiments that yield values for γ and θ.

6.8 MEASURING SURFACE TENSION AND CONTACT ANGLE: ROUND TWO

We have now established that both γ and θ have thermodynamic significance and have seen that their values as well as their temperature coefficients are of interest. In addition, we have seen that the measurement of contact angles presents some complications of its own. All this adds up to a need for more reliable and more accurate methods for the measurement of these parameters than those presented in Section 6.2. One of the most powerful strategies for this involves a second measurement made with the Wilhelmy plate.

6.8a Height of a Meniscus at a Wall

We saw in Section 6.2 that the Wilhelmy plate offers an accurate method for measuring $\gamma \cos \theta$. We thus have one experiment with two unknowns. The Wilhelmy balance measures the weight of the meniscus; in this section we examine the height to which the meniscus climbs on the same surface. We shall see that this distance—which may be accurately measured with a traveling microscope or cathetometer—also depends on γ and θ. The functional relationship between these parameters and the experimental variables is different from the case of the meniscus weight. Therefore we have two experiments with two unknowns that can be solved for γ and θ.

Since both of these measurements can be made with the same interface, difficulties arising from the pairing of mismatched data are eliminated. Of course, the complications associated with surface roughness and chemical heterogeneity persist, and the method applies only to solids that can be prepared as plates. Despite these limitations, this technique is the best general method available for determining γ and θ. In the next section we describe some additional ways of obtaining γ values, and in Section 6.9 we discuss a method for measuring θ on powdered solids.

In Section 6.4 we discussed the pressure difference that exists across a curved surface. The Laplace equation, in the form provided by Equation (35), gives a general description of this pressure difference Δp. Our objective is to apply this relationship to the meniscus formed by a liquid surface at a flat solid wall. The first thing to notice about this is that one of the R's in Equation (35) becomes infinite since the support is planar; hence this R^{-1} term disappears from the Laplace equation.

Figure 6.13a shows a cross-sectional view of this meniscus and defines some pertinent quantities for the experiment. The meniscus is formed by liquid A displacing B. The weight of the displaced B exerts a buoyant force on the meniscus; therefore the density difference must be used. The pressure difference is zero at the flat surface well removed from the wall. We define this to be the $z = 0$ plane. The meniscus makes contact with the wall at $z = h$. Our

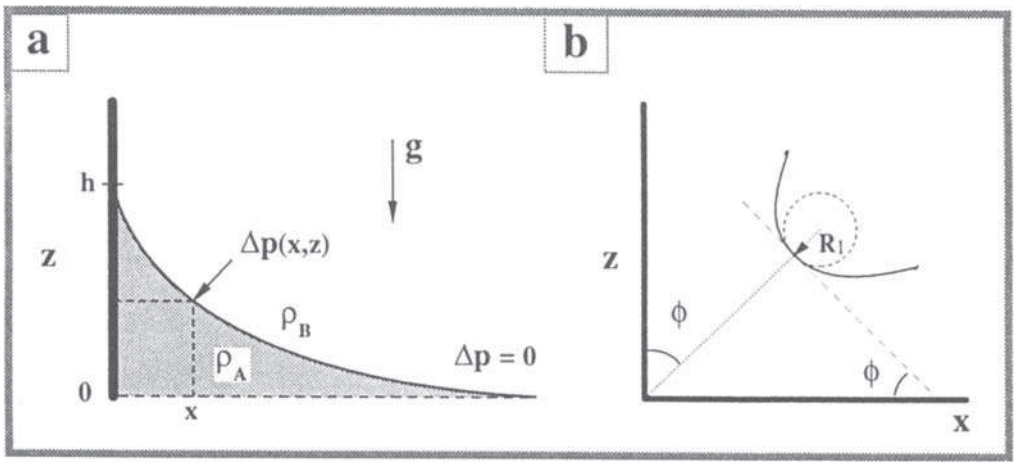


FIG. 6.13 Definitions of variables used to describe a meniscus formed by liquid A displacing liquid B in terms of (a) x and z and (b) ϕ.

interest is in some general point on the surface having the coordinates (x,z). Just beneath the surface at this point the pressure is less by $\Delta p(x,z)$ than in the reference plane. The liquid is elevated at this point by an amount sufficient to produce a compensating hydrostatic pressure. This pressure is given by $[(\rho_A - \rho_B)gz]$, which may be equated to the pertinent form of Equation (35):

$$\Delta\rho g z = \gamma / R_1 \tag{66}$$

where R_1 is the radius of curvature in the plane of Figure 6.13a. Substitution of Equation (36) for $1/R_1$ generates a second-order differential equation with a solution that gives $z = f(x)$. The mathematics are simplified considerably by using some trigonometric relations based on the tangent and the normal to the curve constructed at the point under consideration. Figure 6.13b shows the local radius of curvature of a general curve. The angle ϕ is the angle made by the extension of the normal with the z axis; ϕ also gives the angle between the tangent and the x axis. Inspection of Figure 6.13b reveals that the local slope of the tangent $dz/dx = -\tan\phi = -\sin\phi/\cos\phi$. Substituting this ratio into Equation (36) gives

$$\frac{1}{R_1} = \left| \frac{d}{dx}\left(-\frac{\sin\phi}{\cos\phi} \right) \right| \left| 1 + \left(\frac{\sin\phi}{\cos\phi} \right)^2 \right|^{-3/2} \tag{67}$$

When the indicated differentiations are carried out and the result is simplified, Equation (67) becomes

$$\frac{1}{R_1} = -\cos\phi \frac{d\phi}{dx} \tag{68}$$

which is the same as

$$\frac{1}{R_1} = -\frac{d\sin\phi}{dx} \tag{69}$$

Since $x \propto \sin\phi$ and $z \propto \cos\phi$, Equation (69) is equivalent to

$$\frac{1}{R_1} = -\frac{d\cos\phi}{dz} \tag{70}$$

Substituting this result into Equation (66) gives a relationship that is easily integrated:

$$\frac{\Delta\rho g}{\gamma}\int z dz = -\int d\cos\phi \tag{71}$$

Carrying out the indicated integration and evaluating the constant from the fact that $\phi = 0$ at $z = 0$ (i.e., the tangent is horizontal), we obtain

$$\cos\phi = 1 - \frac{\Delta\rho g}{2\gamma} z^2 \tag{72}$$

At the wall where $z = h$, ϕ is the complementary angle to θ, the contact angle. Therefore at the surface of the solid support

$$\cos\phi = \sin\theta \tag{73}$$

and

$$\sin\theta = 1 - \frac{\Delta\rho g}{2\gamma} h^2 = 1 - \left(\frac{h}{a}\right)^2 \tag{74}$$

where a^2 is given by Equation (5). This result shows that θ may be evaluated by measuring the height to which the meniscus climbs against a wall, provided that γ is known. Example 6.4 gives us an indication of the magnitude of the effect to be determined.

* * *

EXAMPLE 6.4 *Surface Tension and the Height of a Meniscus at a Wall.* Calculate the height to which an *n*-octane surface will climb on a Teflon wall (this is the same system used in Example 6.3) if γ is 22 mJ m^{-2}, $\theta = 30°$, and $\rho = 0.70$ g cm^{-3}. Comment on the ease or difficulty of making this measurement.

Solution: The density of air is insignificant compared to the liquid; hence ρ can be used for $\Delta\rho$. Because the density is given in cgs units, it is convenient to use this system of units throughout; remember, millijoule $meter^{-2}$ and erg cm^{-2} are numerically identical. Substituting numerical values into Equation (74) gives

$$\sin 30 = 1 - [(0.70)(980)\, h^2/2(22)]$$

from which $h^2 = 0.032\ cm^2$ and $h = 0.18$ cm.

Although h is not a large number, it is readily measurable using a cathetometer or low-magnification traveling microscope to measure the vertical distance between the level surface of the liquid and the top of the meniscus. The sighting is done along the liquid surface through a window in the container holding the liquid. Therefore the junction of interest must be viewed through the meniscus at the window. Special care must be taken to establish the right reference plane. If the contact angle is very small, the top of the meniscus might be difficult to see, but good illumination should remedy this. ■

* * *

6.8a.1 Simultaneous Measurement of Surface Tension and Contact Angle

Equation (74) provides a second relationship in addition to Equation (2), which expresses the connection between γ and θ and experimental quantities. The two simultaneous equations can be solved for γ and θ by the following steps:

1. Square Equation (74):

$$\sin^2\theta = [1 - (\Delta\rho g h^2/2\gamma)]^2$$

2. Square Equation (2):

$$\cos^2\theta = (w/P\gamma)^2$$

3. Recognize that $\sin^2\theta + \cos^2\theta$ equals unity:

$$1 = (w/P\gamma)^2 + (1 - \Delta\rho g h^2/2\gamma)^2$$

By multiplying out this last result and simplifying, we obtain

$$\gamma = \frac{\Delta\rho\, g\, h^2}{4} + \frac{w^2}{\Delta\rho\, g\, h^2 P^2} \tag{75}$$

Substituting Equation (75) into Equation (2) gives

$$\cos\theta = \frac{4\,\Delta\rho\, g\, h^2 P\, w}{(\Delta\rho)^2 g^2 h^4 P^2 + 4\, w^2} \tag{76}$$

Equations (75) and (76) allow γ and θ to be evaluated in a perfectly straightforward way from experimental quantities measured on a single system with one apparatus. For systems that can be fabricated into the necessary plates, this is an excellent way to measure these important parameters.

In the next section we examine the shapes of liquid surfaces possessing an axis of symmetry.

6.8b Surface Tension and Shapes of Drops, Bubbles, and Menisci

In the last section we saw that the Laplace equation describes the meniscus formed by a liquid surface at a flat wall. The situation is not essentially different—only more complicated—if the supporting wall is wrapped into a cylinder to generate a capillary meniscus with an axis of symmetry. As a matter of fact, any liquid interface with axial symmetry is described by the same mathematical formalism. For simplicity, we consider a liquid drop resting on a smooth horizontal surface as shown in Figure 6.14. Such a drop, incidentally, is called a *sessile* (sitting) drop. In the absence of gravity, a drop like this would be spherical since this geometry encloses the maximum volume within a minimum surface area. Any departure from a spherical shape increases the area and the surface free energy associated with it. A drop of liquid A in air—or, for that matter, any less dense medium B—is ordinarily acted on by gravitational as well as surface forces. Gravity forces the drop to lower its center of mass and thus flatten. Flattening increases the surface area; the equilibrium shape depends on the balance of the two forces. The same situation holds for the shapes of bubbles and menisci.

6.8b.1 Shape of a Sessile Drop: The Bashforth-Adams Equation

With this physical picture in mind, let us consider how the Laplace equation can be written to describe the profile of a sessile drop. Figure 6.14 represents the profile of a sessile drop; the

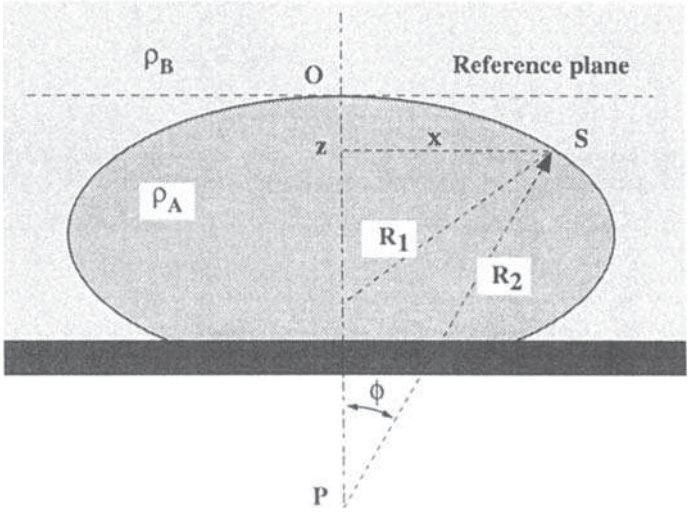


FIG. 6.14 Definition of coordinates for describing surfaces with an axis of symmetry (OP).

actual surface is generated by rotating the profile around the axis of symmetry, the z axis in this representation.

In Figure 6.14 the origin of the coordinate system O is situated at the apex of the surface. The two radii of curvature at point S are defined as follows. The one designated R_1 lies in the plane of the figure and describes the curvature of the profile shown. The radius of rotation of point S around the z axis equals x. The radius R_2 is the radius of curvature of the surface in the plane parallel to the z plane (normal to the plane of the figure) at S. The relationship between x and R_2 is given by

$$x = R_2 \sin \phi \tag{77}$$

since R_2 is defined as a vector originating at the z axis at P and making an angle ϕ with the axis of symmetry as shown in the figure.

Because of the symmetry of the surface, both R_1 and R_2 must be equal at the apex of the drop. The value of the radius of curvature at this location is symbolized b. Therefore at the apex (subscript 0)

$$(\Delta p)_0 = 2\gamma/b \tag{78}$$

according to Equation (30).

Next let us calculate the pressure at point S. At S the value of Δp equals the difference between the pressure at S in each of the phases. These may be expressed relative to the pressure at the reference plane through the apex (subscript 0) as follows:

1. In phase A,

$$p_A = (p_A)_0 + \rho_A g z \tag{79}$$

2. In phase B,

$$p_B = (p_B)_0 + \rho_B g z \tag{80}$$

Therefore Δp at S equals

$$(\Delta p)_S = p_A - p_B = (p_A)_0 - (p_B)_0 + (\rho_A - \rho_B) g z = (\Delta p)_0 + \Delta \rho g z \tag{81}$$

where $\Delta \rho = \rho_A - \rho_B$. Now Equation (78) may be substituted for Δp at the apex to give

$$(\Delta p)_S = 2\gamma/b + \Delta \rho g z \tag{82}$$

The general form of the Laplace equation may be expressed in terms of the coordinates of Figure 6.14 by combining Equations (35), (77), and (82):

$$\gamma\left(\frac{\sin \phi}{x} + \frac{1}{R_1}\right) = \frac{2\gamma}{b} + \Delta \rho\, g\, z \tag{83}$$

In Equation (83) R_1^{-1} is given by Equation (36). Expression (83) is known as the *Bashforth-Adams equation*. It is conventional to express this equation in dimensionless form by expressing all distances relative to the radius at the apex b:

$$\frac{\sin \phi}{x/b} + \frac{1}{R_1/b} = 2 + \frac{\Delta \rho\, g\, b^2}{\gamma} \frac{z}{b} = 2 + \beta \frac{z}{b} \tag{84}$$

The cluster of constants in Equation (84) is defined by the symbol β:

$$\beta = \frac{\Delta \rho\, g\, b^2}{\gamma} \tag{85}$$

The Bashforth-Adams equation—the composite of Equations (36) and (84)—is a differential equation that may be solved numerically with β and ϕ as parameters. Bashforth and Adams solved this equation for a large number of β values between 0.125 and 100 by compiling values of x/b and z/b for $0° < \phi < 180°$. Their tabular results, calculated by hand before the days of computers, were published in 1883. Other workers subsequently extended these

tables. A very useful compilation of these results is found in Padday (1969). Table 6.2 shows a typical result from these tables for $\beta = 25$. The surface profiles sketched in Figure 6.15 were drawn from these tabulated results for (a) $\beta = +10.0$ and (b) $\beta = -0.45$.

It is apparent from inspection of Figure 6.15 that different values of β characterize different drop shapes. This is to be expected since β varies with $\Delta\rho$ compared to γ. According to the convention established in Figure 6.14, g and z are measured in the same direction. Therefore, when the Bashforth-Adams tables are applied to sessile drops, the sign of $\Delta\rho$ determines the sign of β. If $\rho_A > \rho_B$, β will be positive and the drop will be oblate in shape since the weight of the fluid tends to flatten the surface. If $\rho_A < \rho_B$ a prolate drop is formed since the larger buoyant force leads to a surface with much greater vertical elongation. In this case β is negative. A β value of zero corresponds to a spherical drop and, in a gravitational field, is expected only when $\Delta\rho = 0$. Positive values of β correspond to sessile drops of liquid in a gaseous environment. Negative β values correspond to sessile bubbles extending into a liquid. These statements imply that the drop is resting on a supporting surface. If, instead, the drop is suspended from a support (called pendant drops or bubbles), g and z are measured in opposite directions and have different signs; that is, for a pendant drop, z is measured upward from the apex, while g continues to define the downward direction. Because of this sign reversal, the liquid drop will be a prolate ($\beta < 0$) shape, and the gas bubble will be oblate ($\beta > 0$).

6.8b.2 Surface Tension from the Shapes of Sessile Drops

The Bashforth-Adams tables provide an alternate way of evaluating γ by observing the profile of a sessile drop of the liquid under investigation. If, after all, the drop profiles of Figure 6.15 can be drawn using β as a parameter, then it should also be possible to match an experimental drop profile with the β value that characterizes it. Equation (85) then relates γ to β and other measurable quantities. This method is claimed to have an error of only 0.1%, but it is slow and tedious and hence not often the method of choice in practice.

In order to obtain the profile of one of these drops experimentally, it is best to photograph

TABLE 6.2 x/b and z/b for $\beta = 25$ and $0° < \phi \leq 180°$

ϕ (deg)	x/b	z/b	ϕ (deg)	x/b	z/b
5	0.08521	0.00368	95	0.48092	0.28243
10	0.16035	0.01348	100	0.47934	0.29435
15	0.22230	0.02712	105	0.47682	0.30594
20	0.27250	0.04288	110	0.47345	0.31665
25	0.31333	0.05974	115	0.46931	0.32662
30	0.34684	0.07713	120	0.46452	0.33585
35	0.37455	0.09475	125	0.45911	0.34433
40	0.39755	0.11236	130	0.45319	0.35204
45	0.41666	0.12985	135	0.44682	0.35901
50	0.43249	0.14711	140	0.44008	0.36519
55	0.44551	0.16405	145	0.43302	0.37061
60	0.45609	0.18063	150	0.42571	0.37526
65	0.46451	0.19678	155	0.41823	0.37915
70	0.47101	0.21246	160	0.41062	0.38231
75	0.47579	0.22761	165	0.40296	0.38472
80	0.47905	0.24221	170	0.39528	0.38643
85	0.48089	0.25626	175	0.38766	0.38744
90	0.48148	0.26966	180	0.38014	0.38776

Source: F. Bashforth, and J. C. Adams, *An Attempt to Test the Theory of Capillary Action*, Cambridge University Press, London, 1883.

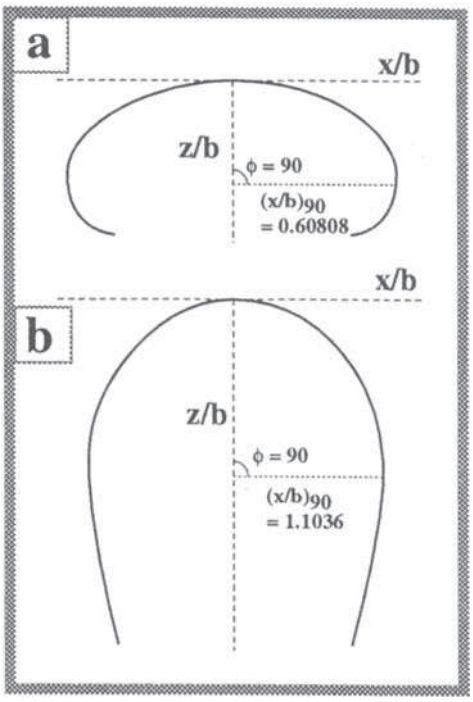


FIG. 6.15 Sessile drop profiles drawn from results of Bashforth and Adams's tables: (a) β = 10.0; and (b) β = −0.45. (Data from Padday 1969.)

the silhouette of the surface. Then experimental and theoretical profiles are compared in an effort to identify the β value that characterizes the experimental surface. With care, it is possible to interpolate between theoretical profiles and arrive at the β value that describes the surface under consideration. Equation (85) shows that knowledge of β alone is not sufficient to permit the evaluation of γ; b and $\Delta\rho$ must also be known. Evaluating the density difference poses no special difficulty. Let us next consider how b is measured.

Once β is known for a particular profile, the Bashforth-Adams tables may be used further to evaluate b:

1. For the appropriate β, the value of x/b at $\phi = 90°$ is read from the tables. This gives the maximum radius of the drop in units of b.
2. From the photographic image of the drop, this radius may be measured since the magnification of the photograph is known.
3. Comparing the actual maximum radius with the value of $(x/b)_{90°}$ permits the evaluation of b.

Figure 6.15a may be used as a numerical example to illustrate this procedure. Suppose an actual experimental drop profile is matched with theoretical profiles and is shown to correspond to a β value of 10.0. Then b is evaluated as follows:

1. The value of $(x/b)_{90°}$ for $\beta = 10$ is found to be 0.60808 from the tables.
2. Assume the radius of the actual drop is 5.00 mm at its widest point.
3. The first two items describe the same point; therefore $b = 5.00/0.608 = 8.22$ mm. This would be the radius at the apex of the drop shown in Figure 6.15a if the maximum radius were 5.00 mm.
4. This numerical example may be completed by using the definition of β to evaluate γ

for the liquid of Figure 6.15a. Assuming $\Delta\rho$ to be $+1.00$ g cm^{-3} and taking $g = 9.8$ m s^{-2} gives

$$\gamma = \Delta\rho g b^2/\beta = (10^3)(9.8)(8.22\cdot 10^{-3})^2/10.0$$
$$= 0.0662 \text{ J m}^{-2} = 66.2 \text{ mJ m}^{-2}$$

Several additional points might be noted about the use of the Bashforth-Adams tables to evaluate γ. If interpolation is necessary to arrive at the proper β value, then interpolation will also be necessary to determine $(x/b)_{90°}$. This results in some loss of accuracy. With pendant drops or sessile bubbles (i.e., negative β values), it is difficult to measure the maximum radius since the curvature is least along the equator of such drops (see Figure 6.15b). The Bashforth-Adams tables have been rearranged to facilitate their use for pendant drops. The interested reader will find tables adapted for pendant drops in the material by Padday (1969). The pendant drop method utilizes an equilibrium drop attached to a support and should not be confused with the drop weight method, which involves drop detachment.

6.8b.3 Other Methods for Determining Surface Tension

Several other methods for determining γ—notably, the maximum bubble pressure, the drop weight, and the DuNouy ring methods (see Section 6.2)—all involve measurements on surfaces with axial symmetry. Although the Bashforth-Adams tables are pertinent to all of these, the data are generally tabulated in more practical forms that deemphasize the surface profile.

A method related to sessile drop and sessile bubble methods is a technique based on the measurement of the profile of a spinning drop. In this method a drop of a liquid is injected into a higher density liquid, and the whole container is rotated at a known angular velocity. The drop elongates because of the imposed centrifugal acceleration against the opposing effect of interfacial tension, which tries to reduce the area. The analysis of the shape of the drop is similar to the sessile drop method in principle, and the centrifugal acceleration can be related to the interfacial tension and the density difference from the liquids through a simple equation (see, for example, Miller and Neogi 1985). In contrast to the sessile drop and sessile bubble methods, the spinning drop technique requires no contact with a solid surface. Moreover, since the drop and the suspending liquid can be made quite thin, the method can be used even when the suspending liquid is moderately turbid. The spinning drop technique is particularly useful when the interfacial tension is very low.

In addition to the methods discussed here and in Section 6.2, there are a few other methods for measuring surface tension that are classified as *dynamic methods* as they involve the *flow* of the liquids involved (e.g., methods based on the dimensions of an oscillating liquid jet or of the ripples on a liquid film). As one might expect, the dynamic methods have their advantages as well as disadvantages. For example, the oscillating jet technique is ill-suited for air-liquid interfaces, but has been found quite useful in the case of surfactant solutions. A discussion of these methods, however, will require advanced fluid dynamics concepts that are beyond our scope here. As our primary objective in this chapter is simply to provide a basic introduction to surface tension and contact angle phenomena, we shall not consider dynamic methods here. Brief discussions of these methods and a comparison of the data obtained from different techniques are available elsewhere (e.g., see Adamson 1990 and references therein).

6.9 CONTACT OF LIQUIDS WITH POROUS SOLIDS AND POWDERS

Any solid can be pulverized into particles of small size; not all can be fabricated into the smooth supports we have discussed in this chapter. This consideration alone—not to mention the many practical applications of powdered solids—encourages us to look for a relationship that describes the junction of a liquid interface with such solids. One thought might be to compress the powder into a pellet and treat the surface of the pellet in the same way as we have treated other solid surfaces. On a fine scale the surface of such a pellet is rough, however, so hysteresis effects should be severe. Therefore we look for some alternative approach.

6.9a Penetration of Liquids Through Pores: Measurement of Contact Angles

Instead of the exterior surface of a powder pellet, let us consider the network of irregular channels that permeate it—the hole instead of the doughnut! Figure 6.16a represents a section through such a pellet for irregularly shaped particles. It is easy to imagine an arrangement such as that shown in Figure 6.16b in which the powder pellet is positioned as a plug in a cylinder such that liquid can be forced through it by a movable piston. For now our main requirement for such an apparatus is that we are able to measure the applied pressure p needed to force the liquid into the plug. In a further development we will also be interested in measuring the depth h into the plug that the liquid penetrates under a pressure p.

Commercial instruments are available that carry out these operations under the control of a computer, which also analyzes the results. Let us consider the penetration of the liquid into the pores of the plug to see the basis for this analysis.

Instead of the actual network of irregular channels, the interpretation of this experiment is based on a model that imagines the plug to consist of a bundle of cylindrical pores of radius R_c. The model is represented by Figure 6.16c. The intrusion of the liquid into the cylindrical pores in response to the applied pressure follows the same mathematical description as the rise of a liquid in a capillary. In view of the approximate nature of the model, it is adequate to use the Laplace equation in the form given by Equation (3) to describe this situation:

$$\frac{2\gamma_{LV}\cos\theta}{R_c} = p \tag{86}$$

In the capillary rise experiment $p = \Delta\rho gh$, whereas p is the applied pressure in the case discussed in the present section.

Now suppose a liquid is pushed through the plug that forms a contact angle of zero with the walls of the channels. Using the subscript 0 to represent this situation, Equation (86) becomes

$$p_0 = \frac{2\gamma_{LV,0}}{R_c} \tag{87}$$

since $\cos\theta = 1.0$ in this case. Equation (87) can be used to evaluate R_c for the plug. The value of R_c thus obtained is a parameter that is characteristic of the experiment; however, it might be very different from any physical distance within the plug. It might be called the radius of an equivalent cylinder, although some researchers feel that this name attaches too much significance to the R_c value thus obtained. The term *tortuosity* is sometimes applied to this

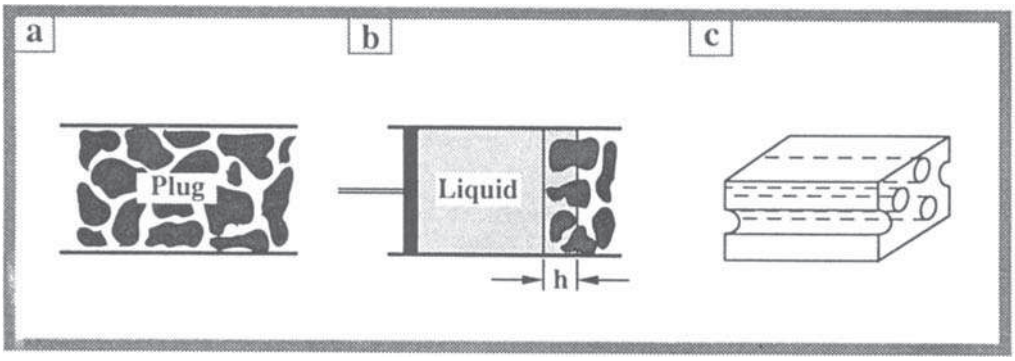


FIG. 6.16 Contact of liquids with pores and powders: (a) schematic view of pores through a plug of particles; (b) liquid intrusion into a plug under pressure; and (c) an idealized plug of cylindrical pores.

empirical R_c value. Whatever terminology is used, this evaluation of R_c amounts to a calibration with a known (with respect to θ) liquid so that R_c can be eliminated from Equation (86) to give

$$\frac{p_0 \gamma_{LV} \cos \theta}{\gamma_{LV,0}} = p \tag{88}$$

or

$$\cos \theta = \frac{p}{p_0} \frac{\gamma_{LV,0}}{\gamma_{LV}} \tag{89}$$

Note that values of γ are assumed to be available when this equation is used. Since numerous methods are available for measuring γ, there is no loss of applicability in assuming this. Since all of the quantities on the right-hand side of Equation (89) are measurable, this approach provides a method for determining θ for powdered solids. The method is not highly reliable, but is preferable to any technique based on the exterior surface of the plug as a liquid support.

6.9b Porosimetry: Structure of Porous Materials

Our approach until now has been to focus attention on the liquid-solid junction and to eliminate the parameter that characterizes the plug. Numerous porous solids exist for which there may be considerable interest in knowing the pore dimensions. The foregoing analysis can also be applied to this problem by assuming that both γ and θ are known. For this application the liquid is chosen for convenience, not because it is part of what is being measured. Mercury is generally used as the penetrating fluid, and the general technique is called *mercury porosimetry*. Contact angles of 130–140° are generally assumed, although it would clearly be preferable to use the θ value that is appropriate for the specific system.

When the pore structure of a solid is being investigated, it is desirable to learn as much as possible from porosimetry. In general, solids will possess a distribution of pore sizes, so the volume (or surface area) associated with pores of various radii is of interest. The significance of this sort of information is apparent for such solids as zeolite catalysts. These are aluminasilicates that can be synthesized with variable Si/Al ratios and variable pore sizes. Pore dimensions in the range of 0.8 to 2.0 nm can be obtained for solids with controllable acidity associated with Al^{3+} in the silica matrix. Zeolites are used on a large scale in the petroleum industry as catalysts for the cracking and isomerization of hydrocarbons (see Vignette 1.9). Part of the specificity of these catalysts arises from the pore size distribution, so porosimetry becomes a valuable tool for the characterization of these materials.

In an apparatus based on Figure 6.16b, the volume of mercury forced into the pores of the solid can be measured as a function of the applied pressure. Equation (86) shows that higher pressures are required for smaller pores. Therefore incremental increases in p will result in the filling of pores of progressively smaller radii. The volume V that has intruded a porous solid at a pressure p gives the cumulative volume of all pores larger than the size associated with p. Since plots of V versus p give information about the cumulative pore distribution, it is the derivative of such data that measures the increment in pore volume associated with an increment in R_c. Written as a formula, $dV/dp \propto dV/(-dR_c)$ since V increases as R_c decreases.

To develop this last relation further, note that Equation (86) can be written as $pR_c =$ constant since γ and θ are independent of pore structure. Therefore $(pdR_c + R_c dp) = 0$, or

$$\frac{dp}{dR_c} = -\frac{p}{R_c} \tag{90}$$

With Equation (90), the slope of the experimental V-p plot can be developed as follows:

$$\frac{dV}{dR_c} = -\frac{dV}{dp}\frac{dp}{dR_c} = \frac{dV}{dp}\frac{p}{R_c} \tag{91}$$

Combining Equations (86) and (91) gives

$$\frac{dV}{dR_c} = \frac{dV}{dp}\frac{p^2}{2\gamma\cos\theta} \tag{92}$$

which shows how the local slope of the experimental data can be converted into the desired information about increments of volume for pores of different radius.

As noted above, much of the data manipulation of commercial porosimeters is computerized, so a pore size distribution is produced automatically by these instruments.

6.9c Contact Angle Measurement and Porosimetry Using the Poiseuille Equation

The rate at which a liquid penetrates a porous solid can also be measured and interpreted at the same level of approximation as used in the above section. The Poiseuille equation, Equation (4.20), gives an expression for the rate at which a liquid flows through cylindrical pores of radius R_c under a pressure p. We shall neglect the gravitational contribution to the driving force, which means that the following is better suited to horizontal than to vertical arrangements. If we write the volume rate of flow $dV/dt = d(\pi R_c^2 h)/dt$ and simplify, Equation (4.20) can be rewritten as

$$\frac{dh}{dt} = \frac{p R_c^2}{8\eta h} \tag{93}$$

where h is the depth of penetration of the intruding liquid and η is its viscosity.

If Equation (86) is used to eliminate p, Equation (93) becomes

$$\frac{dh}{dt} = \frac{2\gamma\cos\theta}{R_c}\frac{R_c^2}{8\eta h} = \frac{\gamma R_c\cos\theta}{4\eta h} \tag{94}$$

As with Equation (86), R_c can be eliminated from Equation (94) by calibration with a liquid for which $\theta = 0$. In that case Equation (94) can be used to determine θ for powders or porous solids. Alternatively, a value of θ may be assumed, and R_c can be evaluated from the rate of mercury intrusion. Equation (94) is called the *Washburn equation*.

Since advancing and receding contact angles are likely to be different in these experiments, mercury permeation curves are expected to be different, depending on whether the mercury is being pushed in or out of the plugs. This type of hysteresis is indeed observed. We encounter another type of hysteresis associated with pore filling in Chapter 9 (Section 9.7a).

6.10 MOLECULAR INTERPRETATION OF SURFACE TENSION

Throughout this chapter we have dealt with surface tension from a phenomenological point of view almost exclusively. From fundamental perspective, however, descriptions from a molecular perspective are often more illuminating than descriptions of phenomena alone. In condensed phases, in which interactions involve many molecules, rigorous derivations based on the cumulative behavior of individual molecules are extremely difficult. We shall not attempt to review any of the efforts directed along these lines for surface tension. Instead, we consider the various types of intermolecular forces that exist and interpret γ for any interface as the summation of contributions arising from the various types of interactions that operate in the materials forming the interface.

6.10a Some Molecular Interactions Important in Interpreting Interfacial Phenomena

To develop the above idea, several broad categories of molecular interactions should be considered.

1. *Hydrogen bonding*: Hydrogen atoms serve as bridges linking together two atoms of high electronegativity. In the present context these atoms are in separate molecules so the molecules themselves are mutually "attracted" by these bonds.

2. *Metallic bonding*: A sea of mobile electrons shared by the atoms of a metal contributes to the attraction between metal atoms in bulk samples.

3. *Permanent dipole interactions*: Polar molecules have relatively positive and negative regions. Regions of opposite charge on different molecules result in an attraction between these molecules. Molecules must possess a permanent dipole moment to display this effect.

4. *London forces*: Deformable electron clouds in adjoining molecules distort one another, resulting in an instantaneous polarity with accompanying attraction between the molecules involved. The polarizability of a molecule (see Chapter 5, Section 5.2b) is a measure of its tendency to display this effect.

Although additional entries could be included in this list, the foregoing entries are sufficient for our purposes. The main thing to recognize about these interactions is that all but London forces are highly specific. London forces require only the presence of electrons and therefore operate between all molecules. The other types of interactions, by contrast, require some specific feature: the metallic state, high electronegativity, or certain molecular geometry. Whether any of the other types of interactions operates or not, London forces are always present.

In Chapter 10 we discuss in considerable detail the attractions between molecules arising from their polarity and polarizability. In that chapter, we also see how surface tension data can be used to obtain information on London forces between macroscopic objects and vice versa, thus providing a direct link between the London force and surface tension. For now, however, it is sufficient to recognize that London forces attract all molecules together, regardless of their specific chemical nature. The other interactions itemized above operate only for systems possessing the requisite special features. What this means for surface tension is best illustrated by considering a specific surface. At the water-mercury interface, for example, the metallic bonds of the mercury end at the interface. Similarly, the hydrogen bonds of the water end at the surface. London attraction exists between water and mercury particles, however, just as it exists for water-water and mercury-mercury pairs. These various London attractions may differ in magnitude as a result of different molecular properties, but they share a common origin.

6.10b Work Needed to Bring a Molecule to an Interface: The Girifalco-Good Equation

Next we apply these concepts to a molecule of A in two different environments: in the interior of a bulk phase and near an interface between two phases. A molecule in the interior of a bulk phase is surrounded on all sides by a homogeneous molecular environment of A. Any movement that would increase its separation from some neighbors would automatically decrease the separation from others, so large deviations from the equilibrium separation are improbable.

For a molecule at an interface between a condensed phase and the gas phase, the environment is quite asymmetrical. Movement toward the bulk phase is impeded by the excluded volume of the A molecules. Movement away from the bulk phase meets no resistance of this sort, although the prevailing attraction between molecules in the condensed phase opposes the A molecules from escaping the condensed phase altogether. Because of this, the equilibrium separation between the molecules at the surface will be larger than between those in the interior. The intermolecular separation has been "stretched" in bringing a molecule to the surface. The contractile force in the interface of the substance is simply the restoring force attempting to return the molecules to their bulk spacing. From an energetic point of view, the difference between the energy of the bulk and surface minimum separations gives the work needed to bring a molecule from the interior to the surface.

Next let us consider the situation in which a second condensed phase B adjoins the reference phase A. The new consideration in this case is the London component of attraction of the molecules in condensed phase B for the A molecules in the interface. This A-B attraction partially overcomes the A-A attraction that opposes the movement of an A molecule to any interface. As a consequence, there is a difference in the energy that must be expended to bring

an A molecule to an interface with a gas and to an interface with another condensed phase. We call this difference ΔE^s.

It is argued that only the London component of intermolecular attractions operates across the interface to decrease the work required to bring a particle to the surface. Suppose we define ϕ as that fraction of the surface tension due to London forces. Next we use the sort of geometric mixing rule that was employed in Chapter 3, Section 3.4b, to estimate ΔE^s as follows:

$$\Delta E^s = \sqrt{(\phi_A \gamma_A)(\phi_B \gamma_B)} \tag{95}$$

The assignment of a geometric mean rather than an arithmetic mean or some other function of the two γ terms is justified primarily on the basis of the successful use of this type of averaging in the theory of nonelectrolyte solubility. Only the London component of γ is used since it is the part of γ that crosses phase boundaries.

According to these ideas, the work required to bring an A molecule to the AB interface is given by

$$(\text{Work})_A = \gamma_A - (\Delta E^s)_A = \gamma_A - \sqrt{(\phi_A \gamma_A)(\phi_B \gamma_B)} \tag{96}$$

when A and B are both in condensed phases. A similar expression applies to the work required to bring a B molecule to the interface. The total work of forming the AB interface is the sum of these contributions:

$$\begin{aligned} \gamma_{AB} &= \left| \gamma_A - \sqrt{(\phi_A \gamma_A)(\phi_B \gamma_B)} \right| + \left| \gamma_B - \sqrt{(\phi_A \gamma_A)(\phi_B \gamma_B)} \right| \\ &= \left| \gamma_A + \gamma_B - 2\sqrt{(\phi_A \gamma_A)(\phi_B \gamma_B)} \right| \end{aligned} \tag{97}$$

Combining the two fractions into the parameter Φ, Equation (97) becomes

$$\gamma_{AB} = \gamma_A + \gamma_B - 2\Phi(\gamma_A \gamma_B)^{1/2} \tag{98}$$

with $\Phi = (\phi_A \phi_B)^{1/2}$. The relationship given by Equation (98) is called the *Girifalco-Good equation*. The parameter Φ may be viewed from two different points of view:

1. We may regard Φ as an empirical constant evaluated by fitting data to Equation (98) or some other form of the basic relationship.
2. We may estimate Φ from first principles by using molecular parameters and relationships from Chapter 10.

With a few additional refinements, the two approaches are found to give satisfactory agreement. The Φ values often lie in the range 0.5 to 1.0, which shows that, in general, half or more of the surface tension can be attributed to London forces.

Equation (98) may be combined with several other relationships of this chapter to generate expressions by which the contributions of London forces can be investigated. For example, if one of the condensed phases is a solid and the other a liquid, Equations (98) and (49) may be combined to give

$$\gamma_{LV} \cos\theta = -\gamma_L + 2\Phi(\gamma_S \gamma_L)^{1/2} - \pi_e \tag{99}$$

Equation (99) can be solved for any one of the factors, depending on the data available. Neglecting π_e as usual (although we might solve for this also), we obtain the following:

1. If γ_S, γ_L, and θ are known,
$$\Phi_{SL} = \gamma_L(1 + \cos\theta)/2(\gamma_S \gamma_L)^{1/2}$$
2. If γ_L, γ_S, and Φ are known,
$$\cos\theta = -1 + 2\Phi(\gamma_S/\gamma_L)^{1/2}$$
3. If γ_L, θ, and Φ are known,

$$\gamma_S = \gamma_L(1 + \cos\theta)^2/4\Phi^2$$

The third item provides a valuable way for evaluating γ_S, a quantity that is otherwise difficult to determine.

6.10c The Fowkes Approximation to the Girifalco-Good Equation

Rather than pursue the fractions ϕ or their composite Φ any further, we turn next to an empirical approach due to Fowkes to estimate the product $\phi\gamma$ for various substances. We use the symbol γ^d to represent the London contributions to γ estimated by his method. The notation is derived from the fact that London forces are also called *dispersion* (superscript *d*) *forces* (see, for example, Chapter 10, Section 10.7b). In this notation Equation (98) can be written

$$\gamma_{AB} = \gamma_A + \gamma_B - 2(\gamma_A^d\gamma_B^d)^{1/2} \tag{100}$$

In this version the relationship is called the *Girifalco-Good-Fowkes equation*. (We use a similar approach again in Chapter 10, e.g., see Equations (10.77) and (10.78), to determine the Hamaker constant for the van der Waals interaction forces.) Although we use $\phi\gamma$ and γ^d interchangeably, it is important to recognize that γ^d values are determined by a particular strategy as illustrated in Example 6.5.

* * *

EXAMPLE 6.5 *Estimation of Interfacial Tensions Using the Girifalco-Good-Fowkes Equation.* The following are the interfacial tensions for the various two-phase surfaces formed by *n*-octane (*O*), water (*W*), and mercury (*Hg*): for *n*-octane-water, $\gamma = 50.8$ mJ m^{-2}; for *n*-octane-mercury, $\gamma = 375$ mJ m^{-2}; and for water-mercury, $\gamma = 426$ mJ m^{-2}. Assuming that only London forces operate between molecules of the hydrocarbon, use Equation (100) to estimate γ^d for water and mercury. Do the values thus obtained make sense? Take γ values from Table 6.1 for the interfaces with air of these liquids.

Solution: In general, Equation (100) contains two unknowns: the dispersion components of γ for *A* and *B*. If γ^d for one of these is known, the other γ^d value can be calculated from experimental results. If it is assumed that $\gamma^d = \gamma$ for the hydrocarbon, then the experimental γ for an interface involving octane can be interpreted by Equation (100) to give the other γ^d value. Thus $375 = 484 + 21.8 - 2(\gamma_{Hg}^d \cdot 21.8)^{1/2}$; therefore $\gamma_{Hg}^d = 196$ mJ m^{-2}. Also $50.8 = 72.8 + 21.8 - 2(\gamma_w^d \cdot 21.8)^{1/2}$; therefore $\gamma_w^d = 22.0$ mJ m^{-2}. The two γ^d values are 30 and 40%, respectively, of the γ's for water and mercury. These fractions are somewhat less than expected in terms of the preceding discussion, but are not totally out of line. A more meaningful test is to use these values to estimate γ for the water-mercury interface. Repeating the above procedure, we obtain

$$\gamma_{Hg\text{-}W} = 484 + 72.8 - 2[(196)(22)]^{1/2} = 425 \text{ mJ m}^{-2}$$

which compares quite favorably with the experimental value of 426 mJ m^{-2}. ■

* * *

Table 6.3 lists the data and calculated results for a number of hydrocarbons forming interfaces with water and mercury. The relative constancy of the calculated γ^d values adds to their plausibility.

The average values of γ^d for water and mercury are clearly successful in their ability to calculate γ correctly for the water-mercury interface. Individually, however, they are the averages of slightly divergent values measured for several different interfaces with hydrocarbons. In this sense the values of γ^d we have considered are analogous to mean bond energies in physical chemistry, which also are averages obtained from a variety of compounds. Although mean bond energies are very useful, they are by nature insensitive to unique, specific effects. With both mean bond energies and values of γ^d, the user must be careful that no such special interactions are present, otherwise quite serious errors could arise. Furthermore, it is important to realize that any errors are perpetuated and compounded by this scheme for evaluating γ^d.

TABLE 6.3 Experimental Values for γ_H, $\gamma_{H\text{-}W}$, and $\gamma_{H\text{-}Hg}$, with H = Hydrocarbon, at 20°C (All Values in mJ m^{-2})[a]

		Mercury (γ_{Hg} = 484)		Water (γ_W = 72.8)	
Hydrocarbon	γ	$\gamma_{H\text{-}Hg}$	γ_{Hg}^d	$\gamma_{H\text{-}W}$	γ_{HW}^d
n-Hexane	18.4	378	210	51.1	21.8
n-Heptane	18.4	–	–	50.2	22.6
n-Octane	21.8	375	199	50.8	22.0
n-Nonane	22.8	372	199	–	–
n-Decane	23.9	–	–	51.2	21.6
n-Tetradecane	25.6	–	–	52.2	20.8
Cyclohexane	25.5	–	–	50.2	22.7
Decalin	29.9	–	–	51.4	22.0
Benzene	28.85	363	194	–	–
Toluene	28.5	359	208	–	–
o-Xylene	30.1	359	200	–	–
m-Xylene	28.9	357	211	–	–
p-Xylene	28.4	361	203	–	–
n-Propylbenzene	29.0	363	194	–	–
n-Butylbenzene	29.2	363	193	–	–
Average			≈201		≈21.9

Source: Data from F. M. Fowkes, *Ind. Eng. Chem.*, **56**, 40 (1964).
[a] γ_{Hg}^d and γ_{HW}^d are calculated as in Example 6.5.

6.10d London Components of Interfacial Tensions of Solid Surfaces

It is not difficult to apply the concept of the dispersion component of γ to solid surfaces. In doing this, it is necessary to treat high- and low-energy surfaces differently. We shall not consider solid interfaces in detail; our treatment is limited to the following observations:

1. For low-energy surfaces, $\pi_e \approx 0$. Manipulation of Young's equation (Equation (44)) generates a relationship that expresses γ_s^d in terms of θ and other experimental quantities.
2. For high-energy surfaces, $\pi_e > 0$ owing to adsorption. Relationships have been derived that express γ_s^d in terms of gas adsorption.

Values of γ_s^d that have been determined for high- and low-energy surfaces by these two methods are listed in Table 6.4.

Examination of the values of γ^d for high- and low-energy solids in Table 6.4 reveals several interesting points. To begin, there is a slight overlap in the data inasmuch as polypropylene was studied by the gas adsorption technique, even though it might reasonably be expected to resemble paraffin wax or polyethylene in γ^d value. As the table shows, the value of γ^d for this substance lies in the same neighborhood as the values for these other compounds, even though very widely different procedures were used to arrive at the different values. The multiple values of γ^d listed for some high-energy solids correspond to evaluations based on the adsorption of different gases at widely different temperatures. For example, the γ^d values for TiO_2 are obtained from adsorption studies conducted with butane at 0°C, heptane at 25°C, and N_2 at −195°C. With these ideas in mind, the observed variation between ostensibly duplicate values becomes more acceptable. The evaluation of γ_S^d also depends on the accuracy of the value of γ for a substance other than the solid. Again we see that errors may be propagated in the analyses that led to the results presented in Table 6.4.

TABLE 6.4 Values of γ_s^d (in mJ m^{-2}) for a Variety of Solids as Determined From Contact Angles and From Gas Adsorption

	γ_s^d	
Material	By measurements of θ	By measurements of π_e
Dodecanoic acid on Pt	10.4, 13.1	–
Polyhexafluoropropylene	11.7,* 18.0	–
Polytetrafluoropropylene	19.5	–
n-$C_{36}H_{74}$	21.0	–
n-Octadecylamine on Pt	22.0,* 22.1	–
Paraffin wax	23.2,* 25.5	–
Polypropylene	–	26, 28.5
Polytrifluoromonochloroethylene (Kel F)	30.8	–
Nylon-6,6	33.6*	–
Polyethylene	31.3,* 35.0	–
Polyethyleneterephthalate	36.6*	–
Polystyrene	38.4,* 44.0	–
$BaSO_4$	–	76
Silica	–	78
Anatase (TiO_2)	–	89, 92, 141
Iron	–	89, 106, 108
Graphite	–	115, 120, 123, 132

Source: Most data from F. M. Fowkes, *Ind. Eng. Chem.*, **56**, 40 (1964); the data with an asterisk are from D. H. Kaelbe, *Physical Chemistry of Adhesion*, Wiley, New York, 1971.

REVIEW QUESTIONS

1. What is *surface tension*? What are its *units*? Explain the physical significance of surface tension.
2. What is *contact angle*? What is the *range* of magnitudes it can take?
3. Describe a few methods to measure surface tension and contact angle of a liquid on a solid substrate and discuss the assumptions and approximations involved in the measurement procedure.
4. Why are the surface tension and contact angle thermodynamic properties of a substance? What does it mean?
5. What is the *Laplace equation*? What is the *Kelvin equation*? What are the differences between the two?
6. What assumptions have been made in the derivation of the Kelvin equation in the text? How restrictive are these assumptions?
7. What is the *Young-Dupré equation*? What are the approximations made in its derivation? Discuss the merits of those approximations.
8. Describe the terms *wetting*, *cohesion*, *adhesion*, and *spreading*.
9. Define the following terms and their relation to surface energies: (a) work of adhesion, (b) work of cohesion, and (c) spreading coefficient.
10. What is the physical significance of the spreading coefficient?
11. What is *equilibrium film pressure*? What is its physical significance?
12. What is the *Bashforth-Adams equation*? How would you use it to determine the surface tension of a liquid and the contact angle of a liquid with a surface?
13. List a few methods for the measurement of surface tension and contact angle. Discuss the basic principles involved in each method. What are the experimental advantages and disadvantages in each case?
14. What are some of the *dynamic methods* for measuring surface tension? What are the differences between these and the *static methods*?

15. Explain the basic principle behind *porosimetry*. How would you use a porosimetry experiment to measure the contact angle of a liquid with a powdered solid? What are the limitations of this approach?
16. Present a molecular interpretation of surface tension.
17. What is the *Girifalco-Good equation*? What is the *Fowkes approximation* to the Girifalco-Good equation?

REFERENCES

General References (with Annotations)

Adamson, A. W., *Physical Chemistry of Surfaces*, 5th ed., Wiley, New York, 1990. (Primarily graduate level. More details on surface tension measurement may be found in this volume.)

Israelachvili, J. N., *Intermolecular and Surface Forces*, 2d ed., Academic Press, London, 1991. (Undergraduate and graduate levels. Contains good discussions and illustrations of the relation among surface tension, wetting, and contact angle phenomena and van der Waals forces.)

Miller, C. A., and Neogi, P., *Interfacial Phenomena: Equilibrium and Dynamic Effects*, Marcel Dekker, New York, 1985. (Graduate level. An advanced-level textbook with a focus on interfaces between liquids, interfacial tension, and dynamics of interfaces.)

Other References

Abbott, N. L., Folkers, J. P., and Whitesides, G. M., *Science*, **257**, 1380 (1992).

Abbott, N. L., and Whitesides, G. M., *Langmuir*, **10**, 1493 (1994).

Abbott, N. L., Whitesides, G. M., Racz, L. M. and Szekely, J., *J. Am. Chem. Soc.*, **116**, 290 (1994).

Adam, N. K., *The Physics and Chemistry of Surfaces*, Dover, New York, 1968.

Good, R. J. In *Surface and Colloid Science*, Vol. 11 (R. J. Good and R. R. Stromberg, Eds.), Plenum, New York, 1979.

Johnson, R. E., Jr., and Dettree, R. H. In *Surface and Colloid Science*, Vol. 2 (E. Matijević, Ed.), Wiley, New York, 1969.

MacRitchie, F., *Chemistry at Interfaces*, Academic Press, San Diego, CA, 1990.

Neumann, A. W., and Good, R. J. In *Surface and Colloid Science*, Vol. 11 (R. J. Good and R. R. Stromberg, Eds.), Plenum, New York, 1979.

Padday, J. F. In *Surface and Colloid Science*, Vol. 1 (E. Matijević, Ed.), Wiley, New York, 1969.

PROBLEMS

1. The frictional force was measured on a strand of viscous rayon fiber moving through a wad of identical fibers as a function of the water content of the wadded fiber.* It was found that the friction incresed from 59 to 133 mN cm^{-1} when the water content decreased to the point at which capillary "necking" between the fibers occurred. A model for this situation may be visualized by considering two parallel, tangent cylinders connected by a "neck" of water held in the neighborhood of the contact by capillary forces. Sketch the situation represented by this model and explain why the force of fiber-fiber attraction increases with decreasing water content. Use this model to discuss (a) the behavior of a wet paintbrush, (b) the practice of wetting the tip of a thread before threading a needle, and (c) the dewatering of cellulose fibers to form paper.

2. The accompanying data give experimental values of the capillary rise for various liquids:†

Liquid	$\Delta\rho$ (g cm^{-3})	h (cm)	R_c (cm)
Water	0.9972	1.4343	0.100990
Benzene	0.8775	1.5425	0.043135
$CHCl_3$	1.4869	1.9210	0.019320

*Skelton, J., *Science*, **190**, 15 (1975).

†Richards, T. W., and Carver, E. K., *J. Am. Chem. Soc.*, **43**, 827 (1921).

Use values from the following table (interpolated from Padday 1969) to evaluate a^2 (and γ) by successive approximation:

β	0	0.02	0.04	0.06	0.08	0.10	0.12	0.14	0.16	0.18	0.20
$(x/b)_{90°}$	1.00	0.997	0.994	0.991	0.978	0.984	0.981	0.978	0.975	0.972	0.970

Compare the values of γ calculated by this procedure with those obtained by the approximation given by Equation (4).

3. Finely dispersed sodium chloride particles were prepared, their specific area was measured, and their solubility in ethanol at 25°C was studied.* It was found that a preparation with a specific area of 4.25×10^5 cm^2 g^{-1} showed a supersaturation of 6.71%. Estimate the radius of the NaCl ($\rho = 2.17$ gm cm^{-3}) particles, assuming uniform spheres. Calculate γ for the NaCl-alcohol interface from the solubility behavior of this sample.

4. Enüstün and Turkevich† prepared $SrSO_4$ ($\rho = 3.96$ g cm^{-3}) precipitates under conditions that resulted in different particle sizes. Particle sizes were characterized by electron microscopy, and solubilities were determined at 25°C by a radiotracer technique. In the following data the supersaturation ratios are presented for different preparations, each of which is characterized by an average particle width and a minimum particle width:

x_{mean} (Å)	x_{min} (Å)	S/S_0
247	96	1.43
269	130	1.35
388	155	1.28
541	168	1.29
629	252	1.16
1260	378	1.10
1660	500	1.07

Which size parameter gives the best agreement with the Kelvin equation? Explain. Use the best fitting data to evaluate γ for the $SrSO_4$-H_2O interface.

5. Bartell and Osterhof‡ describe an experimental procedure for measuring the work of adhesion between liquids and solids. With carbon (lampblack) as the solid, the following values for the work of adhesion were obtained:

Liquid	Benzene	Toluene	CCl_4	CS_2	Ethyl ether	H_2O
W_{AB} (erg cm^{-2})	109.3	110.2	112.4	122.1	76.4	126.8

Use these data together with the surface tensions of the pure liquids from Table 6.1 to calculate the spreading coefficients for the various liquids on carbon black. Use your results to interpret the authors' observations: "About equal quantities of water and organic liquid were put into a test tube with a small amount of the finely divided solid and shaken. It was noted that the carbon went exclusively to the organic liquid phase."

6. The effect of mutual saturation on the L-V and L-L interfacial tensions is effectively illustrated by considering the spreading coefficient of one liquid on another using both the initial (unsaturated) and equilibrium values of γ. Use the following data§ to calculate $S_{B/A}$ (equilibrium) and $S'_{B/A}$ (nonequilibrium):

*Van Zeggeren, F., and Benson, G. C., *Can. J. Chem.*, **35**, 1150 (1957).

†Enüstün, B. V., and Turkevich, J., *J. Am. Chem. Soc.*, **82**, 4502 (1960).

‡Bartell, F. E., and Osterhof, H. J., *J. Phys. Chem.*, **37**, 543 (1933).

§Harkins, W. D., *The Physical Chemistry of Surface Films*, Van Nostrand-Reinhold, Princeton, NJ, 1952.

	H_2O/air	H_2O/IAA	IAA/air	H_2O/air	H_2O/CS_2	CS_2/air
γ' (erg cm^{-2})	72.8	5.0	23.7	72.8	47.4	32.4
γ (erg cm^{-2})	25.9	5.0	23.6	70.3	48.4	31.8

Describe what happens when a drop of pure isoamyl alcohol (IAA) is placed on the surface of pure water. What happens with the passage of time? Repeat this description for the case of pure CS_2 on pure water.

7. Water drops were formed at the mercury-benzene interface by means of a syringe, and the contact angle (measured in the water) was recorded as a function of time.* For the interface between Hg and benzene saturated with water, γ was measured independently as a function of time. The following table summarizes these data (all measured at 25°C):

Time (hr)	$\gamma_{Hg\text{-}benzene}$ (erg cm^{-2})	θ_{obs} (deg)
0.10	363.0	118
0.42	359.5	119
1.0	358.0	122
2.5	354.5	138
5.0	350.0	144
13	336.0	–
23	–	180

Using 379.5 and 34.0 erg cm^{-2}, respectively, as the values of γ for the mercury-water and benzene-water interfaces, compare the observed contact angles with the predictions of Young's equation. Comment on the fact that constant values are used for $\gamma_{Hg\text{-}W}$ and $\gamma_{benzene\text{-}W}$.

8. A cylindrical rod may be used instead of a rectangular plate in a slight variation of the Wilhelmy method. Derive an expression equivalent to Equation (63) for a suspended solid of cylindrical geometry. Prepare semiquantitative plots analogous to Figure 6.10c based on the following data, assuming cylindrical rods 1.0 mm in diameter at the air-water interface:†

Metal	ρ_s (g cm^{-3})	θ_a (deg)	θ_r (deg)
Au	19.3	70	40
Pt	21.5	63	28

In both cases the metal surfaces were carefully polished, washed, steamed, and then heated in an oven for 1 hr at 100°C.

9. The vertical rod method of the preceding problem was used to study the contact angle of water at the gold-water-air junction at 25°C. The following data show how the value of θ_a depends on the prior history of the metal surface.‡ For a gold surface polished, washed, and heat treated for 1 hr at T°C, then allowed to stand in air for t hr we obtain

T (°C)	100	200	300	400	500	600	600	600	600	600	600
t (h)	$< \frac{1}{4}$	$< \frac{1}{4}$	$< \frac{1}{4}$	$< \frac{1}{4}$	$< \frac{1}{4}$	$\frac{1}{4}$	1	5	10	24	120
θ_a (deg)	68	57	45	36	25	13	22	38	47	53	55

Calculate the work of adhesion between water and gold for each of these cases on the assumption that $\pi_e = 0$. Is the variation in W_{SL} consistent (qualitatively? quantitatively?) with the expected validity of the assumption concerning π_e?

10. The tendency of spilled mercury to disperse as small drops that roll freely on most surfaces is a well-known characteristic of this liquid. Discuss this behavior in terms of (a) the work of adhesion and the spreading coefficient for mercury on various substrates, (b) surface tension

*Bartell, F. E., and Bjorkland, C. W., *J. Phys. Chem.*, **56**, 453 (1952).

†Bartell, F. E., Culbertson, J. A., and Miller, M. A., *J. Phys. Chem.*, **40**, 881 (1936).

‡Bartell, F. E., and Miller, M. A., *J. Phys. Chem.*, **40**, 889 (1936).

versus contact angle (measured in Hg) as causes of this behavior, and (c) the implications of the Kelvin equation on the health hazards associated with mercury spills.

11. Use the data of Table 6.2 to plot the profile of a drop with $\beta = 25$. Measure (in cm) the radius of the drop you have drawn at its widest point. By comparing this value with the value of $(x/b)_{90°}$ from the table, evaluate b (in cm) for the drop as you have drawn it. Suppose an actual drop is characterized by this value of β. If the actual radius at the widest point is 0.25 cm and $\Delta\rho = 0.50$ g cm^{-3}, what is γ for the interface of the drop?

12. Suppose the drop profile shown in Figure 6.15b describes an actual drop for which the radius at the widest point equals 0.135 cm. Use the value of $(x/b)_{90°}$ from the figure to calculate γ for each of the following situations:

System	$\lvert\Delta\rho\rvert$ (g cm^{-3})
(a) Oil in water	0.20
(b) Water in oil	0.20
(c) Oil in air	0.80
(d) Air in water	1.00
(e) Water in air	1.00

State whether the drop is pendant or sessile in each case.

13. Sometimes it is difficult to locate the bottom of the meniscus of a colorless solution in a buret. If this is the case, it is surely that much more difficult to try to estimate contact angles by direct observation of a meniscus. In view of this, criticize or defend the following proposition: Before attempting to read contact angles directly by viewing the surface through a low-power microscope, a new worker should practice with simulated drops (recommended by Neumann and Good 1979). By masking off the lower part of a sphere such as a ball bearing and looking at its silhouette, a practice junction is obtained. The angle estimated can be compared with the true value by calculating the true value from the height of the apex and the width of the base. In fact, values for actual drops can be determined by this method if the profile of the entire drop is visible in the microscope.

14. The following combinations of θ, γ_L, and p values were reported by Bartell and Whitney* for the wetting of silica plugs by various liquids:

Liquid	p (g cm^{-2})	γ_L (dyne cm^{-1})	θ
Nitrobenzene	125	25.3	41°25′
Chloroform	192	31.6	22°11′
Benzene	215	34.7	19°16′
Toluene	221	36.5	—
CCl_4	265	44.5	—
Hexane	333	51.0	—

Use the first three sets of data to determine a value of R_c for the plug; then use this value to determine θ for the remaining liquids against silica. Are large or small contact angles more sensitive to, say, a 5% error in R_c?

15. Drake† has reported the accompanying data for the porosimetry analysis of a catalyst preparation; the cumulative pore volume occupied by mercury is given for the applied pressures indicated:

Pressure (psi)	Volume (cm^3 g^{-1})	Pressure (psi)	Volume (cm^3 g^{-1})
1,000	0.082	20,000	0.228
1,500	0.115	25,000	0.249

*Bartell, F. E., and Whitney, C. E., *J. Phys. Chem.*, **36**, 3115 (1932).
†Drake, L. C., *Ind. Eng. Chem.*, **41**, 781 (1949).

2,000	0.132	30,000	0.276
3,000	0.150	35,000	0.307
4,000	0.160	40,000	0.336
6,000	0.177	45,000	0.358
10,000	0.190	50,000	0.363
15,000	0.213		

Plot these data and from the tangents to the curve estimate dV/dR_c at 2000, 10,000, 30,000, and 45,000 psi. To what radii do these pressures correspond? Use $\gamma = 484$ mJ m^{-2} and $\theta = 140°$ for these calculations.

16. By a suitable combination of Equations (49) and (97) show that

$$\gamma_S^d = \frac{\gamma_{LV}^2}{4\gamma_l^d}(1 + \cos\theta)^2$$

for low-energy surfaces. Use the following data (see Table 6.4 for reference) to evaluate either γ_S^d or γ_l^d as appropriate:

S	L	γ_{LV} (erg cm^{-2})	γ^d (erg cm^{-2})	θ (deg) in liquid
Dodecanoic acid on Pt	α-Bromonaphthalene	44.6	10.4 for S	92
Kel F	α-Bromonaphthalene	44.6	30.8 for S	48
Paraffin wax	Glycerol	63.4	36 for L	97
Paraffin wax	Fluorolube	20.2	13.5 for L	31

17. The equation derived in the preceding problem suggests that a plot of cos θ (as ordinate) versus $\sqrt{\gamma_l^d/\gamma_L}$ (as abscissa) should be linear with a slope of $2\sqrt{\gamma_S^d}$ and an intercept of -1. Describe how this result can be used to evaluate γ_S^d when contact angle measurements are made on a particular solid with a variety of liquids for which γ_L and γ_l^d are known. Describe how the same graphing procedure can be used to evaluate $\sqrt{\gamma_l^d/\gamma_L}$ when contact angles are measured in a liquid of unknown γ_l^d on different solids of known γ_S^d. Use the data of the preceding problem to illustrate these two graphical interpretations. Include the additional datum that α-bromonaphthalene forms a contact angle of 58.5° with paraffin (see Table 6.4 for reference).

18. If γ_L and θ are measured for a homologous series of liquids on a given low-energy solid, a plot of cos θ versus γ_L results in a straight line. Verify that this is the case for the following data, determined for alkanes on Teflon at 20°C:*

Compound	γ_L (mJ m^{-2})	θ (deg)	Compound	γ_L (mJ m^{-2})	θ (deg)
Hexadecane	27.6	46	Nonane	22.9	32
Tetradecane	26.7	44	Hexane	21.8	26
Dodecane	25.4	42	Heptane	20.3	21
Undecane	24.7	39	Octane	18.4	12
Decane	23.9	35	Pentane	16.0	spreads

The intercept at $\theta = 0$ is viewed as a kind of critical state for these systems, and the corresponding surface tension is represented by γ_c. What is the value of γ_c for Teflon?

19. Use Equation (99) to show that γ_c as defined in the last problem is equal to γ_S^d for those systems where $\gamma_L = \gamma_l^d$. What else must be assumed to prove this?

*Fox, H. W., and Zisman, W. A., *J. Colloid Sci.*, **5**, 514 (1950).

7

Adsorption from Solution and Monolayer Formation

You are living on a plane. What you style Flatland is the vast level surface of what I may call a fluid, on, or in, the top of which you and your countrymen move about, without rising above it or falling below it.

From Abbott's *Flatland*

7.1 INTRODUCTION

The interest of scientists on monolayers of oil on a water surface is frequently traced to Benjamin Franklin's now-famous experiment on stilling the waves with oil in the pond at Clapham Common in England on a windy day,* but the striking properties of such monolayers are brought to light by an equally simple experiment most students have seen at home or high school. Shake some pepper or flour on a pan of water and touch the surface of the water with a bar of soap; you can see the pepper or flour race to the edges. Why does an oil film dampen the surface? Why does the pepper retreat so rapidly to the edge of the water's surface? Investigations of such questions are what lead us to the material we discuss in this chapter.

7.1a Surfactant Layers: Langmuir and Gibbs Layers and Langmuir-Blodgett Films

Monolayers formed by a substance that is insoluble in the liquid subphase are labeled *Langmuir layers*, whereas those formed by substances soluble in the bulk but adsorbed preferentially at the interface are known as *Gibbs layers*. Langmuir layers transferred to solid substrates are known as *Langmuir-Blodgett* (LB) layers (see Vignette VII below). Our objective in this chapter is to lay the basic foundations necessary for understanding adsorption and formation of such layers and their structures. The Langmuir-Blodgett films usually consist of multiple layers; however, we restrict ourselves to monolayers.

Until now, we have avoided considering the above topics by intentionally excluding solutes of variable concentration from our consideration of surfaces. Now, the effects of such solutes are our specific interest. We are especially concerned with a particular class of solutes that show dramatic effects on surface tension. These are said to be *surface active* and are often simply called *surfactants*. Our primary emphasis here is on the relationship between adsorp-

*An engaging discussion of the history of Benjamin Franklin's experiment and a relatively nontechnical treatment of monolayers and bilayers of surfactants and their implications to biochemistry and biology are presented by Tanford, a pioneer of what is known as the *hydrophobic effect* and the biological applications of mono- and multilayers (Tanford 1989). Almost all of the material discussed in this highly readable volume is relevant to the focus of this chapter.

tion phenomena and surface tension/surface energy. In the process of describing experimental methods, results, and interpretations, however, a variety of related concepts enters the picture, as explained in Section 7.1c.

7.1b Why Are Monolayers and Multilayers Important?

Traditionally, monolayer and multilayer adsorption have been used in detergency, mineral processing, flotation, stability of food and pharmaceutical emulsions, and the like, and, as a consequence, the topics of this chapter have been a central part of colloid science. In recent years, however, research on monolayer and multilayer deposition has mushroomed rapidly because of significant new opportunities.

At the most fundamental level, monolayers of surfactants at an air-liquid interface serve as model systems to examine condensed matter phenomena. As we see briefly in Section 7.4, a rich variety of phases and structures occurs in such films, and phenomena such as nucleation, dendritic growth, and crystallization can be studied by a number of methods. Moreover, monolayers and bilayers of lipids can be used to model biological membranes and to produce vesicles and liposomes for potential applications in artificial blood substitutes and drug delivery systems (see, for example, Vignette I.3 on liposomes in Chapter 1).

Equally important are the new developments in electronic and optical materials and other "advanced materials." Developments in chemical synthesis and in experimental techniques such as atomic force microscopy, scanning tunneling microscopy, and optical imaging techniques have opened the way for manipulating the structure and properties of ultrathin films. As a result, it might be possible to "engineer" surface coatings with special optical, electrical, and magnetic properties, and the monolayer and multilayer structures engineered in this manner can be used for fabricating microelectronic devices, chemical and biochemical sensors, and optical switches and storage devices, to name a few.

Vignette VII presents a historical snapshot of the early experiments of Agnes Pockels that led to the above exciting possibilities. A couple of specific examples along these lines and some additional information follow in Sections 7.10b and 7.10c.

VIGNETTE VII MOLECULAR MONOLAYERS AND MULTILAYERS: Langmuir-Blodgett Films

When Agnes Pockels (1862–1935) of Brunswick, Germany, sent a registered letter on January 12, 1891, to Lord Rayleigh (John William Strutt, Third Baron Rayleigh, 1842–1919) describing her simple apparatus for studying properties of oil films and oil-contaminated water, little did she realize that her work and the essential design of her apparatus would pave the way for fundamental research on molecular films for possible use in molecular electronics, bio- and chemical sensors, tribology, and catalysis, among others.

Pockels described in the letter her design of a rectangular tin trough with a thin tin strip laid across it. The trough was filled to the brim with water, with a thin layer of oil covering the surface of the water on one side of the tin strip and clean water on the other side. The tin strip served to vary the area of the oil-contaminated surface, and a balance measured the force necessary to lift a small disk (a button) from the surface. Pockels used this setup to study the surface tension of the oil-contaminated layer.

This groundbreaking work from a woman with a "domestic background" impressed Lord Rayleigh immensely, and he sent it to *Nature*, a prestigious British journal, with his strong recommendation for publication (Pockels 1891). Additional papers from Pockels and Lord Rayleigh followed, and Lord Rayleigh proposed that layers a single molecule in thickness could be studied using the Pockels technique.

While the experiments of Pockels and Rayleigh generated some activity on the properties of thin films, the area was essentially dormant until the work of Irving Langmuir (1881–1957; Nobel Prize, 1932, for surface chemistry) at the General Electric Corporate Research Laboratories in Schenectady, New York. Langmuir developed a number of new techniques

(including the well-known surface film balance named after him) to study the chemistry and physics of monolayers on water surfaces. Such "floating" monolayers are now known as *Langmuir films*, in honor of Langmuir's pioneering work.

The scientific and industrial importance of monolayers and multilayers of organic and polymeric materials would not have been possible without the work of another pioneer, Katherine Blodgett (1898–1979), the first woman scientist to join G.E. Research Laboratories, and the first woman to obtain a doctorate from the famed Cavendish Laboratory in Cambridge, England. Blodgett devised an apparatus to transfer fatty acid monolayers to solid substrates from water surfaces. The films thus produced are now known as *Langmuir-Blodgett films* and can be fabricated using a number of different substances (see Fig. 7.1).

The Langmuir-Blodgett films are believed to hold considerable promise in numerous high-technology applications such as molecular electronics, piezoelectric organic films, waveguides, nonlinear optics, and optical information storage, in addition to classical applications such as adhesion, prevention of corrosion, catalysis, and solubilization (see Section 7.10 and Roberts 1990). The challenge to the surface chemists, physicists, and engineers is to devise and fabricate films with special chemical and physical properties (i.e., use molecular engineering to manipulate the architecture of the films to provide the desired properties). This field is thus a very active area of current research and development in colloid and surface science.

7.1c Focus of This Chapter

The material in this chapter is organized broadly in two segments. The topics on monolayers (e.g., basic definitions, experimental techniques for measurement of surface tension and surface-pressure-versus-area isotherms, phase equilibria and morphology of the monolayers, formulation of equation of state, interfacial viscosity, and some standard applications of monolayers) are presented first in Sections 7.2–7.6. This is followed by the theories and experimental aspects of adsorption (adsorption from solution and Gibbs equation for the relation between

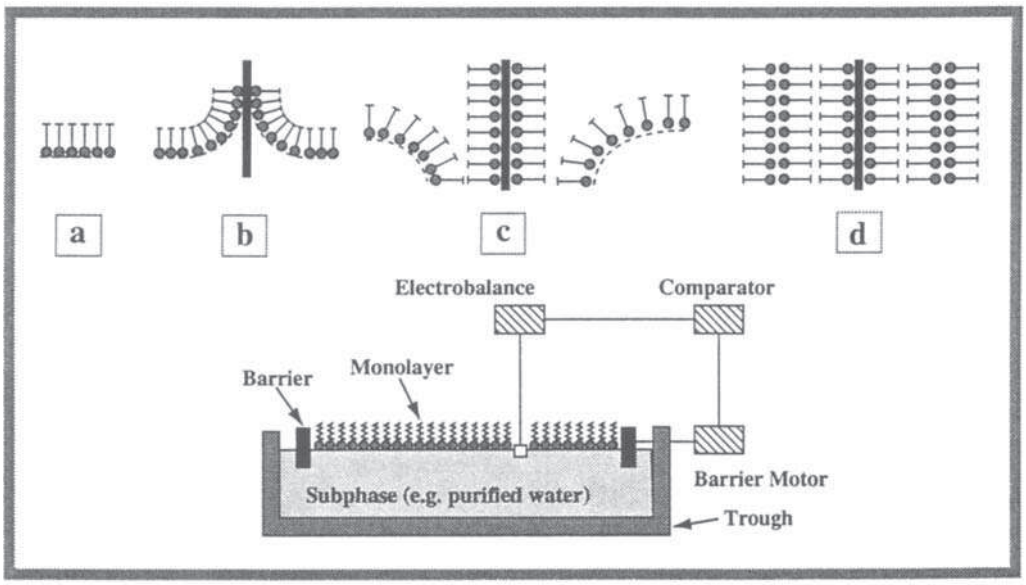


FIG. 7.1 Deposition of multilayers using the Langmuir-Blodgett technique: (a and b) monolayer deposition; and (c and d) multilayer deposition.

surface tension and surface excess concentration, experimental aspects of Gibbs equation, adsorption on solid substrates, the Langmuir equation and its application, and examples of traditional and some modern applications of adsorbed layers) in Sections 7.7–7.10. We close the chapter with a brief discussion of adsorption in the presence of applied potentials as this is an important topic in many aspects of colloid science.

Despite the above organization, in studying the material of this chapter it may be helpful to realize that the topics covered may be grouped in several different ways. Let us enumerate what these various ways of looking at the material are.

First, we may focus our attention on the solubility of the adsorbed species in one or both of the adjacent phases. In this way two broad categories of phenomena emerge: insoluble and soluble surface layers.

A second way of classifying the material is on the basis of the experimental methods involved. For mobile interfaces, surface tension is easily measured. For these it is easiest to examine the surface tension-adsorption relationship starting with surface tension data. When insoluble surface films are involved, we shall see how the difference in γ between a clean surface and one with an adsorbed film may be measured directly. For solid surfaces, surface tension is not readily available from experiments. In this case adsorption may be measurable directly, and the relationship between adsorption and surface tension may be examined from the reverse perspective.

Third, the material of this chapter is a mix of descriptive and theoretical concepts. The most important descriptive observation is the existence of two-dimensional phases. The theoretical content of the chapter is mostly thermodynamic in origin. Three major results are the equations named after Gibbs, Langmuir, and Lippmann. We are mostly concerned with uncharged surfaces, except for a brief discussion of electrolyte adsorption at a polarizable mercury electrode in Section 7.11.

Finally, the material may also be regarded as a mixture of fundamentals and applications. Although the entire book stresses principles, applications are considered from time to time as examples of more abstract ideas. This is also the intent of the sections on applications in this chapter. In addition, however, many applications of adsorption phenomena are the basis of large and important areas of technology. To omit mention of them would lead to a very incomplete picture of these fields. As it is, many important applications must be omitted for lack of space, and those mentioned are sketched in only a superficial way.

7.2 INSOLUBLE MONOLAYERS: SPREADING OF SURFACTANTS ON AQUEOUS SURFACES

7.2a Spread or Insoluble Monolayers

Suppose a dilute solution is prepared from an aliphatic solvent and an organic solute RX in which R is a long-chain alkyl group and X is a polar group. Then, a small amount of this solution is placed on a large volume of water with a horizontal surface. The components of this system were chosen because they are assumed to meet the following experimental criteria:

1. The solubility in water of both components of the organic phase is negligible at room temperature.
2. The likelihood of any complex being formed between the organic solvent and solute is exceedingly low.
3. The volatility of the organic solvent is high and that of the solute is low.

With these facts in mind, let us examine the fate of the drop of solution placed on the surface of water. The initial spreading coefficient $S_{O/W}$ (Equation (6.61)) for the organic layer on water is positive. This is primarily because $\gamma_{O/W}$ is unusually low and γ_W is high, even with an adsorbed layer of the organic solvent. After spreading, we allow sufficient time to elapse for all the solvent to evaporate from the spread layer. At this point the surface will contain a

layer of the organic solute similar to that which would result from the spreading of a sessile drop of pure liquid solute or from the adsorption of vapors of the solute component from the gas phase. Using a solution with a volatile solvent to form such a layer is a very common technique and has the advantage of permitting very small amounts of solute to be quantitatively deposited on a surface.

The nature of the layer that remains after the solvent has evaporated depends on the amount of solute deposited and the area available to it. It is convenient to distinguish among three situations in this regard. If the amount of added material and the area are such that the water surface is covered uniformly to a depth of one molecule with the solute, the resulting film is called a *monolayer*. On the other hand, submonolayer coverage and multilayer coverage result when the amount of added material per area is less or more, respectively, than that which produces the monolayer. In this chapter we are concerned mostly with degrees of coverage up to and including the monolayer. If a large excess of spread material (beyond the amount needed for monolayer coverage) is used, the excess collects into droplets of a bulk phase. The equilibrium situation is then identical to what would be produced by the spreading of a sessile drop of the solute material. Films of the sort described here are called either *spread monolayers* (when the method of their preparation is stressed) or *insoluble monolayers* (when the chemical nature of the solute is emphasized). We use these terms interchangeably.

7.2b Some Properties of Spread Monolayers

Now let us examine some of the properties of the spread monolayer that we have described. It was seen in the preceding chapter (e.g., Equation (6.48)) that the presence of an adsorbed layer lowers the surface tension of an interface. The phenomenon is quite general, so we redefine π (no subscript) in the following symbols:

$$\pi = \gamma_0 - \gamma \tag{1}$$

where γ_0 refers to the surface tension of any phase in the absence of an adsorbed layer, and γ refers to the tension of the same surface with an adsorbed layer. Specific subscripts are used only when the problem clearly involves more than one interface.

The spread monolayer just described may be discussed from two points of view. First, there are those aspects of the film that pertain explicitly to the chemical nature of the components: water and the organic solute. Second, there are certain properties of the monolayer that depend on physical variables such as temperature, area of the water surface, and number of molecules of RX present. Let us briefly discuss both of these viewpoints.

7.2b.1 Structural Aspects of Monolayers

The organic solute RX is a prototype of an important array of surface-active materials. Many surface-active substances are composed of what are known as *amphipathic* molecules. This term simply means that the molecule consists of two parts, each of which has an affinity for a different phase. We are concerned mostly with surfaces in which one of the phases is aqueous, so the surfactants that we consider will contain polar or ionic groups, or "heads," and nonpolar organic residues, or "tails." In the compound RX, for example, R is an alkyl group, generally containing 10 or more carbon atoms. The literature of surface chemistry contains many references to these organic groups by both their International Union of Pure and Applied Chemistry (IUPAC) and their common names. Table 7.1 lists some of the more commonly encountered examples. In RX, the polar X group may be $-OH$, $-COOH$, $-CN$, $-CONH_2$, or $-COOR'$, or an ionic group such as $-SO_3^-$ $-OSO_3^-$, or $-NR_3^+$.

With the foregoing ideas in mind, one characteristic of the adsorbed monolayer becomes apparent: molecular orientation at surfaces. For a film of RX on water, the picture that emerges is one in which the polar groups are incorporated into the aqueous phase with the hydrocarbon part of the molecule oriented away from the water. Such details as the depth of immersion of the tail and the configuration of the alkyl group are best approached by considering how the properties of the monolayer depend on the physical variables.

TABLE 7.1 IUPAC and Common Names for a Variety of Normal Saturated and Unsaturated Surface Active Compounds

	Normal, saturated compounds		
	Carboxylic acids		Alcohols, amines, sulfates, etc.
n (number of C atoms)	IUPAC name	Common name	Common name
12	Dodecanoic	Lauric	Lauryl
14	Tetradecanoic	Myristic	Myristyl
16	Hexadecanoic	Palmitic	Cetyl
17	Heptadecanoic	Magaric	Heptadecyl
18	Octadecanoic	Stearic	Stearyl
20	Eicosanoic	Arachidic	Eicosyl, arachic
22	Docosanoic	Behenic	Docosyl
	Normal, unsaturated carboxylic acids		
	IUPAC name		Common name
18	cis-9-Octadecenoic		Oleic
18	cis,cis-6,9-Octadecenoic		Linoleic
18	cis,cis,cis-3,6,9-Octadecenoic		Linolenic
18	trans-9-Octadecenoic		Elaidic
22	cis-9-Docosenoic		Erucic
22	trans-9-Docosenoic		Brassidic

7.2b.2 Surface Pressure Versus Area Isotherms

Next let us consider some of the physical properties of the spread monolayer we have described. Equation (1) states that the surface tension of the covered surface will be less than that of pure water. It is quite clear, however, that the magnitude of γ must depend on both the amount of material adsorbed and the area over which it is distributed. The spreading technique already described enables us to control the quantity of solute added, but so far we have been vague about the area over which it spreads. Fortunately, once the material is deposited on the surface, it stays there—it has been specified as insoluble and nonvolatile for precisely this reason. This means that some sort of barrier resting on the surface of the water may be used to "corral" the adsorbed molecules. Furthermore, moving such a barrier permits the area accessible to the surface film to be varied systematically. In the laboratory this adjustment of area is quite easy to do in principle. As we see below, the actual experiments must be performed with great care to prevent contamination.

Suppose that the initial film is spread on water that fills to the brim a shallow tray made of some inert material. Rods with low-energy surfaces may then be drawn across the water to adjust the area accessible to the molecules of the monolayer. Figure 7.2a indicates schematically how such an arrangement might appear. In practice, several barriers would be used, first to sweep the surface free of insoluble contaminants and then to confine the monolayer.

7.2b.3 Wilhelmy-Plate Technique for Surface Tension Measurement

Next, an experiment such as that shown in Figure 7.2b could be conducted. The apparatus consists of a pair of *Wilhelmy plates* attached to two arms of a balance. One plate contacts the clean surface and the other the surface with the monolayer. Note that the barrier separates the two portions of surface. The surface tension will be different in the two regions, and the weight (and volume) of the meniscus entrained by the plate will be larger for the clean surface

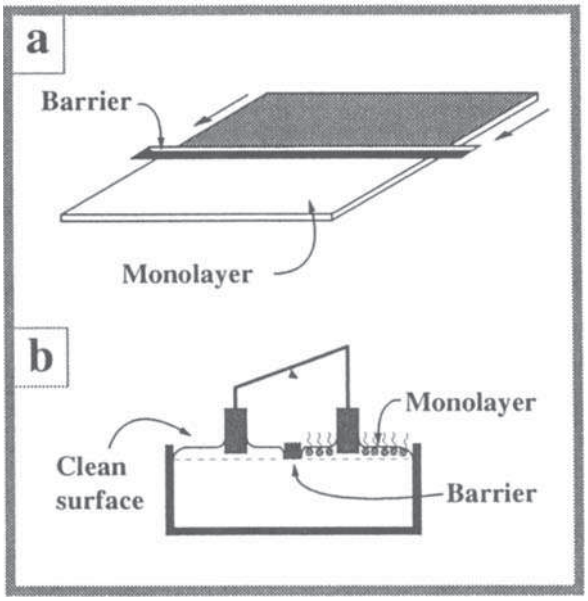


FIG. 7.2 Schematic illustrations of a monolayer and a Wilhelmy plate arrangement for surface tension measurement: (a) schematic illustration of a barrier delineating the area of a monolayer; and (b) a Wilhelmy plate arrangement for measuring the difference in γ on opposite sides of barrier.

because of its higher surface tension (Equation (6.2)). Ideally, the two plates are identical in weight, perimeter, and contact angle, although the last may be difficult to achieve in practice. If these conditions are met, however, the additional weight needed to bring the apparatus to balance measures the difference in γ for the two surfaces. By Equation (6.2), this is given by

$$w_{clean} - w_{film} = P \cos \theta \, (\gamma_0 - \gamma) \tag{2}$$

Of course, there is no necessity to measure both γ_0 and γ on the same apparatus; they may be determined independently by any of the methods of Chapter 6. The experiment represented by Figure 7.2b is intended mainly to emphasize that the surface tension of the two areas will be different. Furthermore, as the barrier is moved in such a way as to compress the area of the spread monolayer, the value of π will increase.

Although π and the area A of the surface vary inversely, the precise functional form by which they are related is more difficult to describe. For very large areas, π and A show a simple inverse proportionality such as pressure and volume for an ideal gas. As the area is decreased, a more complex relationship is needed to connect these variables, just as the equation of state becomes more complex for nonideal gases and condensed phases. This analogy of π and A with p and V turns out to be a very profitable way of thinking about insoluble monolayers. For one thing, it suggests an alternative to the difference between two values of surface tension as a means of measuring π. In addition, the analogy to three-dimensional states suggests models for understanding monolayers. In succeeding sections each of these points is developed in greater detail.

Identification of area as the two-dimensional equivalent of volume is a straightforward geometrical concept. That π should be interpreted as the two-dimensional equivalent of pressure is not so evident, however, even though the notion was introduced without discussion in Chapter 6, Section 6.6. Figure 7.3 helps to clarify this equivalency as well as suggest how to compare quantitatively two- and three-dimensional pressures. The figure sketches a possible profile of the air-water surface with an adsorbed layer of amphipathic molecules present. In

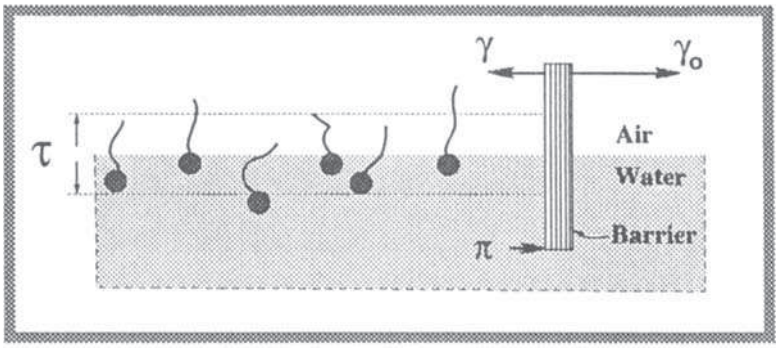


FIG. 7.3 Schematic profile of the air-water interface at a barrier that separates a monolayer from the clean surface.

general, we must allow for the fact that different configurations might exist among the adsorbed molecules; nevertheless, the surface layer has some mean thickness τ.

If the barrier represents the limit of the monolayer, then it is clear that the contractile force exerted by the surface is different on opposite sides of the barrier. Since γ is less than γ_0, it is as if the film were exerting a force on the barrier along the perimeter of the film equal to π. Force per unit length—the units of γ—is the two-dimensional equivalent of force per unit area, the units of pressure in the bulk.

The surface layer does not have zero thickness, of course, even though it is conceptually convenient to think of it as two-dimensional matter. If we assume that the film pressure π extends over the entire thickness of the film, then it is an easy problem to convert the two-dimensional pressure to its three-dimensional equivalent. Taking 10 mN m^{-1} as a typical value for π and 1.0 nm as a typical value for τ enables us to write

$$p = \frac{\pi}{\tau} = \frac{10^{-2}\ \mathrm{N\ m^{-1}}}{10^{-9}\ \mathrm{m}} = 10^{7}\ \mathrm{N\ m^{-2}} \tag{3}$$

or

$$p = 10^{7}\ \mathrm{N\ m^{-2}} \cdot \frac{1\ \mathrm{atm}}{1.013 \cdot 10^{5}\ \mathrm{N\ m^{-2}}} \approx 100\ \mathrm{atm} \tag{4}$$

In view of this calculation, it is not too surprising that insoluble monolayers do not usually display a simple inverse proportionality between π and A. At pressures this high, three-dimensional matter is not likely to obey the ideal gas law either.

7.3 EXPERIMENTAL MEASUREMENT OF FILM PRESSURE

7.3a LANGMUIR FILM BALANCE

The considerations of the preceding section suggest a second way to study spread monolayers. This technique involves measuring the film pressure directly rather than calculating it from surface tension differences by Equation (1). Figure 7.4a is a schematic representation of an apparatus called the Langmuir film balance after Irving Langmuir, a pioneer in this field as mentioned in Vignette VII. Its base is a shallow tray or trough of some inert material. As was the case in Figure 7.2a, the surface must be swept by barriers both to clean the surface and to compress monolayers. In the Langmuir balance, however, one of the barriers is attached to a pivoted arm, arranged in such a way that a torque balance around point P can be measured, for example, by adding weights to the pan. As the figure shows, an insoluble monolayer may be confined to a portion of the surface adjacent to the pressure-sensing float. The shaded area

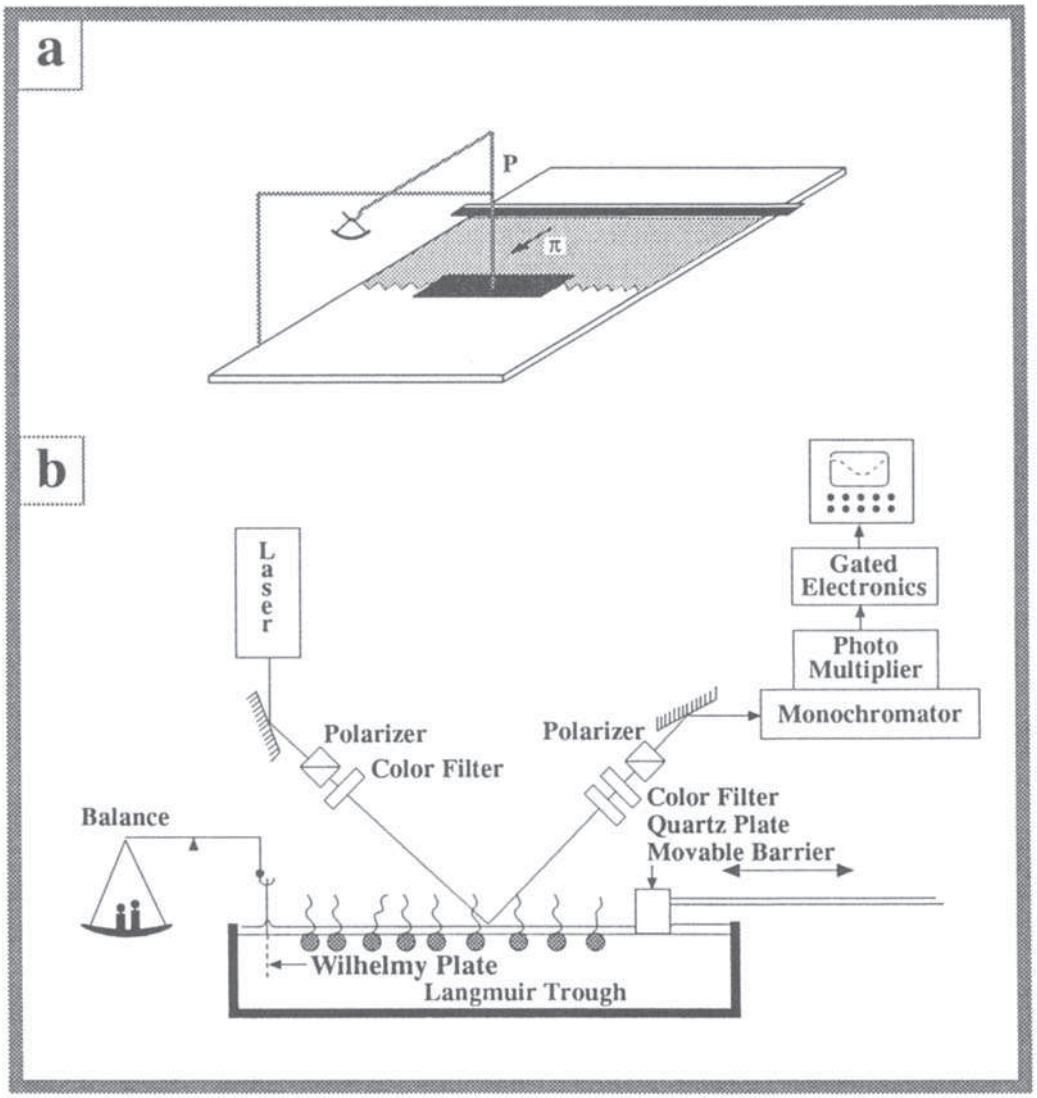


FIG. 7.4 Langmuir film balance: (a) a schematic representation of a Langmuir balance; and (b) a Langmuir trough with a laser optics instrument to measure the orientations of the hydrocarbon tails of the surfactant molecules. The apparatus shown monitors the orientation of the tails through the second harmonic signals generated at various angles of incident light beam. (Redrawn with permission of G. A. Somarjai, *Introduction to Surface Chemistry and Catalysis*, Wiley, New York, 1994.)

in the figure corresponds to the area of the film. It is obviously adjustable by moving the other barrier. To prevent the film from leaking past the edges of the float, flexible barriers connect the ends of the float to the edges of the tray.

By means of this apparatus, it is possible to vary the area of a spread monolayer and measure the corresponding film pressure directly. Many different variations of the film balance are available, and a number of instrumentational techniques can be combined with the Langmuir balance to obtain information on the microstructure of the films and the properties of the films. Figure 7.4b illustrates, for example, a laser optics arrangement to monitor the molecular orientation of the hydrocarbon tails of the surfactant molecules. Below in this

chapter we come across another technique, based on fluorescence microscopy, for observing the different "surface phases" that can develop in the film. Figure 7.5 is a photograph of a commercial apparatus, shown without the peripherals such as illustrated in Figure 7.4b.

7.3b Using the Langmuir Balance

Although the Langmuir balance is quite simple conceptually, obtaining unambiguous results by this technique is far from simple. As we have done in discussing other experimental techniques in this book, we only touch on those aspects of the method that will somehow contribute to our fundamental understanding of insoluble monolayers. Anyone considering experiments of this sort should consult more detailed discussions, such as the book by Gaines (1966).

A convenient way to discuss the Langmuir balance is to examine the difficulties involved in measuring each of the two-dimensional state variables π, A, T, and, of course, the number of moles n of material in the insoluble layer.

The float-torsion wire assembly is the pressure-sensing system in the film balance. The flexible barriers that connect the float to the edges of the tray must be considered part of this mechanism. As with gas pressure determinations, it is essential that the system be leakproof. For this reason the float and flexible barriers must always be hydrophobic, that is, not wetted by the aqueous substrate. If these surfaces were wet by water, the possibility of surfactant transferring to them—that is, a leak in the system—would be enhanced. Thin pieces of mica

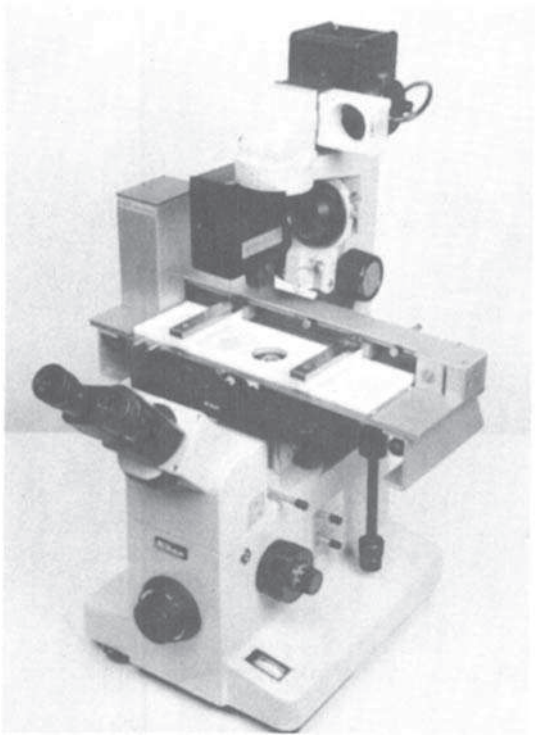


FIG. 7.5 Photograph of a commercial film balance. The photograph shows a "minitrough" with microscopy setup. (Courtesy of KSV Instruments, Ltd., Höyläämötie 11B, SF-00380 Helsinki, Finland.)

are commonly used as float material, and platinum ribbons or threads of silk or nylon are often used for flexible barriers. All of these are waxed to give them suitably hydrophobic characteristics. The sweeping barriers and the tray must also be hydrophobic for the same reasons. These are usually waxed metal, although Teflon is also quite popular because of its inertness. The barriers must make intimate contact with the edges of the tray, also to prevent leaks. Therefore both tray and barriers must be carefully machined to assure good contact. Plastics are generally unsuitable as barrier materials because they are too light to make good contacts.

To convert the measured torque into a two-dimensional pressure, it is necessary to know both the length of the float and the distance between the float and the torsion wire. The distance measurement requires that the water level be controlled quite accurately. As far as the length of the float is concerned, the flexible connectors must be included in this figure. Since they are anchored at one end, only part of their length may be considered a part of the pressure-sensing system. Some approximations are required here, but if the length of the connector is small compared to the total length of the float, the error is negligible.

The float is effectively a two-dimensional manometer, and, like its open-ended counterpart, it measures the film pressure difference between the two sides of the float. This is another reason why it is imperative that no leakage occur past the float assembly: Leakage would increase the pressure on the reference side of the float. For the same reason, the side of the float opposite the monolayer must be carefully checked for any possible source of contamination, not just misplaced surfactant. One way of doing this is to slide a barrier toward the float from that side to verify that no displacement of the float occurs. In all aspects of film pressure measurement, the torque must be measured with sufficient sensitivity to yield meaningful results.

Measuring the area of the film is less troublesome. If the edges of the tray are parallel and the barriers perpendicular to them, the area of the rectangular surface is easily determined. The curvature of the surface at the hydrophobic boundaries introduces a small error but—since the total area is of the order of magnitude of 10^{-2} m^2—this is generally negligible.

The results obtained in π versus A experiments may be sensitive to the rate at which the film area is changed. We shall not discuss the factors responsible for this, but merely note that the same film pressures should be obtained on compression and expansion if true equilibrium values are being measured.

Temperature is an important variable in any equation of state. The experiments we are describing are isothermal; therefore it is important that both the water substrate and the adjoining vapor be thermostated.

Next let us consider those difficulties associated with the determination of the amount of material deposited on the surface. We have already noted that the method of depositing insoluble monolayers by spreading permits the accurate determination of n. Since the spreading technique requires solvent volatility, care must be exercised to prevent the stock solutions from changing concentration due to evaporation prior to their application to the surface. Also, precise microvolumetric methods must be used to dispense the solution on the aqueous surface since the quantity used is small. The solvent (as well as the solute) must be free from contaminants. There is also the possibility that the solvent will extract spreadable contaminants from the waxed surfaces of the float, barriers, and tray. Some workers advocate addition and evaporation of one drop at a time to minimize this. Oily contaminants may also reach the water surface from the fingers and from the atmosphere. These last sources are particularly hard to control: Tests for reproducibility and blank compressions (i.e., moving the barrier toward the float on a "clean" surface) are the best evidence of their absence.

Not all solvents are equally suitable for spreading monolayers. The requirement that the solvent evaporate completely is self-evident. It has been suggested that if the organic solvent dissolves much water, the properties of the monolayer will be different from those in which no water is trapped. Verification that no artifacts are entering the observations from the solvent may be accomplished by conducting duplicate experiments with different solvents.

Until now we have concentrated on those difficulties in using the Langmuir balance that

arise from the determination of π, A, T, and n. A few remarks are also in order about considerations that may affect the monolayer and originate in the adjacent phases. We have already discussed contaminants originating in the gaseous phase. For some monolayer materials, air oxidation may also be a problem. The aqueous substrate is the source of a wide assortment of contaminants in addition to spreadable oily matter. Ionic impurities, including those that affect the pH, are quite troublesome. The charge state of many amphipathic molecules, for example, amines and carboxylic acids, is obviously pH dependent. Salts or complexes formed between amphipathic molecules and ions in the aqueous phase will have different monolayer properties from those of the unreacted surfactant molecule.

An extensive discussion of the Langmuir balance technique and a comparison with the Wilhelmy plate method are given by MacRitchie (1990). This book also discusses modifications that are possible to the techniques and other experimental details.

7.4 RESULTS OF FILM BALANCE STUDIES

The preceding section shows that it is possible to determine π-A isotherms for surfaces just as p-V isotherms may be measured for bulk matter. The results that are obtained for surfaces are analogous to bulk observations also, although some caution must be expressed about an overly literal correlation between bulk and surface phenomena. We return to a discussion of these reservations below. There can be no doubt, however, that analogies with bulk behavior supply a familiar framework within which to consider π-A isotherms.

The curve sketched in Figure 7.6, which is drawn with grossly distorted coordinates to encompass all features, contains several similarities to p-V isotherms. Not all the features

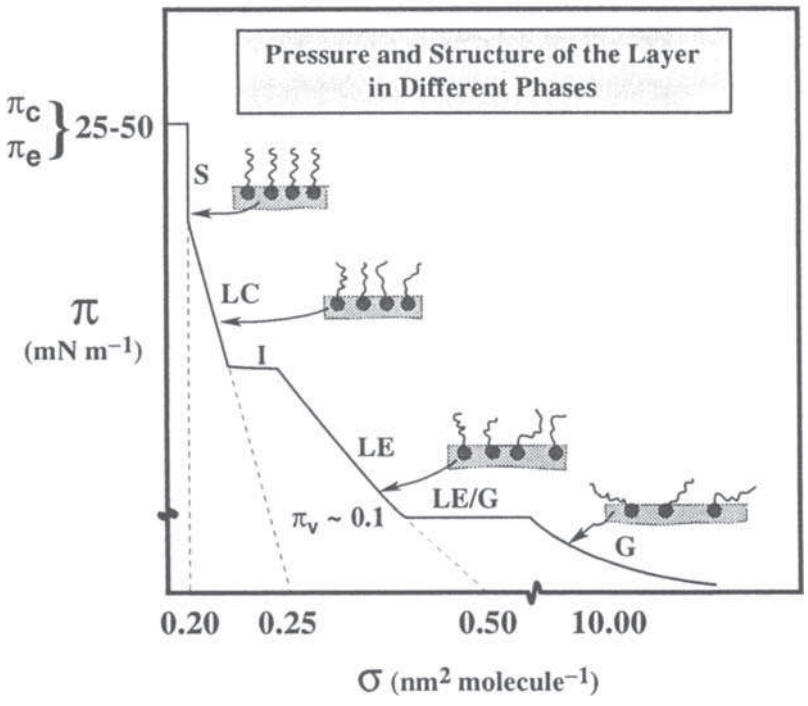


FIG. 7.6 Composite two-dimensional pressure π versus area σ isotherm, which includes a wide assortment of monolayer phenomena. Note that the scale of the figure is not uniform so that all features may be included on one set of coordinates. The sketches of the surfactants show the orientations of the molecules in each phase at various stages of compression.

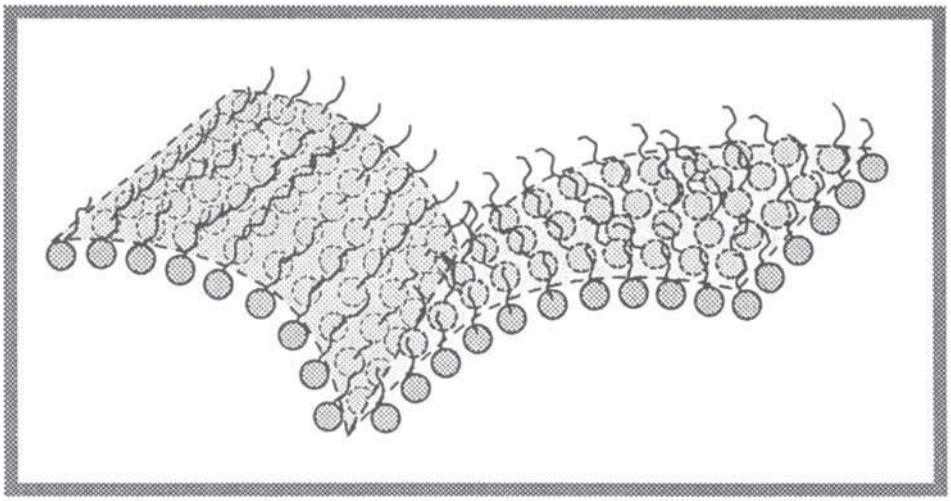


FIG. 7.7 Schematic illustration showing the collapse of the film.

shown here are always observed, nor are all known idiosyncrasies of π-A isotherms represented. The presence or absence of various features and their π-A coordinates vary with temperature for a particular amphipathic molecule and from one amphipathic substance to another. Last, there is some diversity in the terminology used to describe various monolayer phenomena. In short, Figure 7.6 is a composite isotherm that will introduce and summarize a variety of observations.

In this section we discuss in turn the various two-dimensional phases and phase equilibria represented in Figure 7.6, progressing from low values of π to high ones. The existence of these two-dimensional states and the properties they possess are presumably unfamiliar to most readers. Therefore it is important to keep the following ideas in mind in reading this section:

1. We are concerned with two-dimensional matter situated at the boundary between two bulk phases.
2. The properties of the two-dimensional phases are relatively independent of the properties of the bulk phases of the same material.
3. Many surface states are two-dimensional analogs of three-dimensional states. As with any analogy, however, there are points of similarity and points of difference between the surface and bulk states. For most of the states we discuss, we consider the phenomenological behavior as represented by Figure 7.6 and Figure 7.7.

7.4a Microstructural Phases in Monolayers

Let us first take a look at the types of microstructural phases one typically finds in a monolayer. For this consider the isotherm shown in Figure 7.6. The Langmuir layers have a (two-dimensional) "gaslike" distribution of surfactants when the area per molecule is large compared to the dimensions of the molecule. In this phase, denoted as G in Figure 7.6, the hydrocarbon tails of the molecules make significant contact with water. As the concentration of the surfactant at the surface increases, the gas phase begins condensing to a liquidlike phase known as the *liquid-expanded* (LE) phase. This two-phase region corresponds to the plateau region in Figure 7.6 (analogous to the three-dimensional gas-liquid coexistence region on the pressure-volume isotherm). At higher surfactant concentrations the gas phase completely condenses to the LE phase. In the LE phase, the hydrocarbon tails of the surfactants lift from the surface of the water but remain disordered. At further increases in surfactant concentra-

tion, the LE phase is transformed to another "liquid" phase known as the *liquid-condensed* (LC) phase. Current research suggests that the LC phase is not a liquid (Knobler 1990b), as one observes a higher degree of alignment of the hydrocarbon chains (and possibly some long-range order). The LC phase becomes a solidlike (S) phase with further compression as shown in Figure 7.6. The area per molecule in the solid phase is close to what one would expect for close packing of the hydrocarbon chains. The phase diagram shown in Figure 7.6 can also be plotted in terms of a temperature-versus-area diagram (similar to the temperature-volume diagram in three dimensions) as shown in Figure 7.8, in which some sample tie lines (the horizontal lines connecting coexisting phases) are also marked. The diagram shown is only a schematic representation, as emphasized in the legend.

It is instructive to examine the morphology of a monolayer as the above phase changes occur. The different phases can be visualized to a limited extent by using fluorescence microscopy. In this technique, a monolayer is prepared with a small amount of solute with a fluorescence label. One can then excite the fluorescence markers with a laser and observe the morphology using a high-sensitivity camera. A sequence of such results is shown in Figure 7.9 (Knobler 1990b). Figure 7.9a corresponds to a point in the LE/G coexistence region in Figure 7.8 (at a density of 1 molecule per 61 $Å^2$ in this case). The dark regions in the figure are the gas "bubbles", which appear dark because of the difference in the density of the phases and (possibly) because of the quenching of the fluorescence of the labeled molecules in the gas phase. The white regions are the LE phase.

As the layer is compressed at constant temperature (i.e., as one moves horizontally from right to left in Fig. 7.8), one sees the sequence of morphologies shown in Figures 7.9b–7.9e. The layer in Figure 7.9b is still the LE/G coexistence region, but the amount of gas phase has clearly decreased. On further increase in density one enters the single-phase LE region (Fig. 7.9c), as indicated by the completely white layer. Figure 7.9d may be misleading as it appears similar to Figures 7.9a and b, but in this case the dark circles correspond to the LC phase in which the (fluorescence-labeled) probe molecules have low solubility (hence the contrast). Figure 7.9e shows that the LC phase increases in quantity at a higher solute concentration. Although the fluorescence technique described here provides only black-and-white pictures and it may appear that the technique is limited to two-phase regions, detection of more than two phases can be achieved by analyzing the evolution of the features in the morphology.

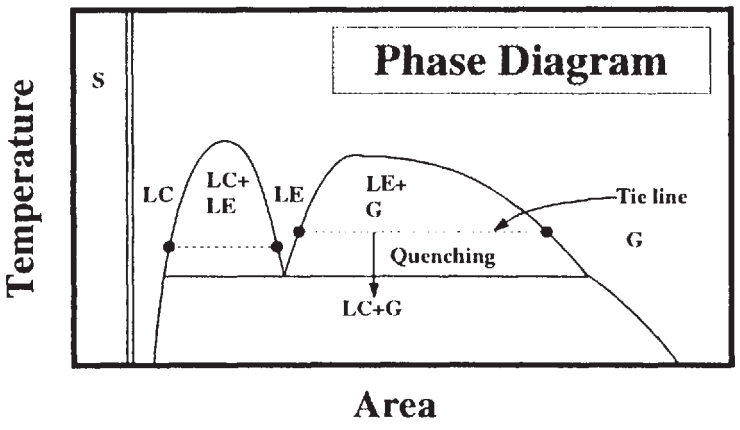


FIG. 7.8 A schematic representation of the temperature-versus-area diagram for a Langmuir layer (a two-dimensional phase diagram). The coexistence regions are exaggerated for clarity. The horizontal lines shown are the tie lines. The arrow marked "quenching" starts at the LE + G coexistence region and is used in the text to illustrate the morphological changes when a two-phase liquid-expanded/gaslike (LE + G) mixture is quenched. (Redrawn with permission of Knobler 1990b.)

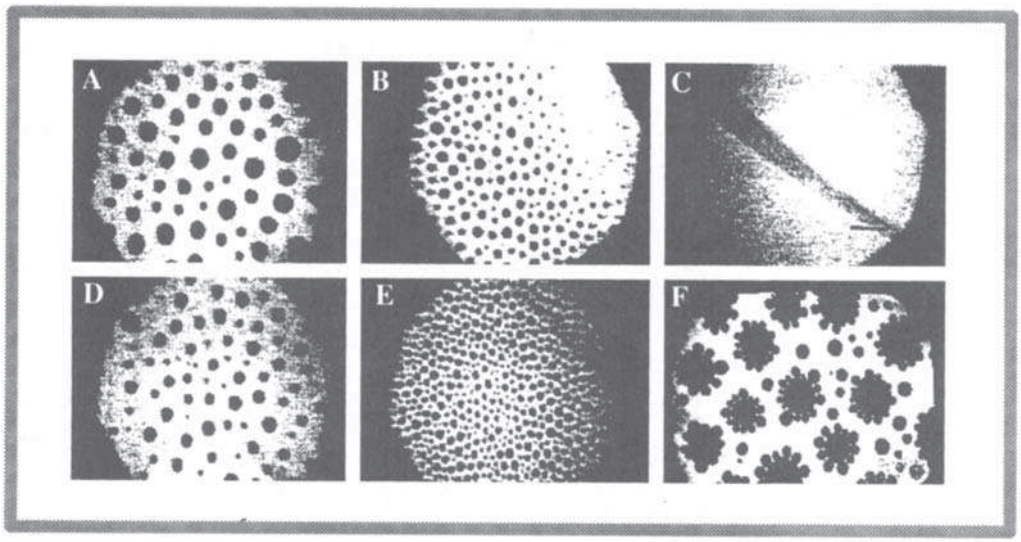


FIG. 7.9 Fluorescence microscope pictures of a monolayer of pentadeconoic acid (PDA) containing 1% fluorescent probe (4-(hexadecylamino)-7-nitrobenz-2-oxa-1,3-diazole, i.e., NBD-hexadecylamine): (A) 1 molecule per 61 Å^2 at 25°C with G (dark) and LE (white) phases; (B) 1 molecule per 50 Å^2 at 25°C with G (dark) and LE (white) phases; (C) 1 molecule per 36 Å^2 at 25°C with a single LE phase; (D) 1 molecule per 27 Å^2 at 25°C with LC (dark) and LE (white) phases; (E) 1 molecule per 24 Å^2 at 25°C with LC (dark) and LE (white) phases; (F) temperature quench starting at the LE/G coexistence region (overall density = 1 molecule per 51 Å^2). The final point is a three-phase region consisting of LC, LE, and G phases. See the text for details. (Redrawn with permission of Knobler 1990b.)

One such case is shown in Figure 7.9f, which is obtained by quenching a state in the LE/G coexistence region to a point in the LC/LE/G three-phase region (the arrow marked "quenching" in Fig. 7.8). The large circles in the centers of the flower-shaped domains correspond to the LC phase. Circling the LC "drops" are small circular domains of the gas phase. These grow at the expense of the LE phase, and with further quenches one moves into an LC/G mixture (see the arrow marked "quenching" in Fig. 7.8). A considerable number of studies have appeared in the literature in recent years on the morphological changes in Langmuir monolayers (see Knobler 1990b) because of the potential and the projected uses of such monolayers in advanced technological applications (see Sections 7.6 and 7.10).

7.4b The "Gas" Phase and Its Models

As evident from the above discussion, if measurements can be made at sufficiently low pressures, all monolayers will display gaseous behavior, represented by region G in Figure 7.6. The gaseous region is characterized by an asymptotic limit as $n \rightarrow 0$. In the limit of very low film pressures, a two-dimensional equivalent to the ideal gas law applies:

$$\pi A = nRT \tag{5}$$

where R is the gas constant, usually in SI units. This is a convenient place to define another quantity, the area occupied per molecule in the interface σ. Since R equals Avogadro's number times the Boltzmann constant, and nN_A equals the total number of surface molecules, we may write

$$\pi\left(\frac{A}{nN_A}\right) = \pi\sigma = k_B T \tag{6}$$

As a model for this highly expanded state we may choose a situation like that depicted in one of the insets of Figure 7.6 corresponding to very low surface coverage; the inset corresponding to the gas phase shows the hydrocarbon chain lying flat on the surface, blocking an area $\pi\ell^2$, where ℓ is the length of the "tail." We may then use Equation (6) to calculate the value of π corresponding to this area per molecule. At 25°C and with $\ell = 1.0$ nm, we obtain

$$\pi = \frac{k_B T}{\sigma} = \frac{(1.38 \cdot 10^{-23})(298)}{(3.14)(10^{-9})^2} = 1.31 \text{ mN m}^{-1} \tag{7}$$

Allowing the above amount of area per molecule is equivalent to taking the distance of closest approach for the typical distance between the molecules. Therefore ideal behavior will be expected only if the areas per molecule are larger than this, on the order of 10 nm^2 perhaps. Correspondingly lower film pressures will be involved also. If we recall that the pressure of this example is equivalent to a bulk gas pressure of about 13 atm, it is less surprising that such low film pressures are needed to observe gaseous monolayer behavior. The repulsion between particles in a charged monolayer increases the effective area these molecules occupy at the surface. The effect of this is to increase the pressure of charged films, making them more accessible to measurement. Since ionic surfactants are soluble, techniques other than the film balance must be used to study their π-σ isotherms.

It is not difficult to propose and develop a model for the gaseous state of insoluble monolayers. The arguments parallel those developed in kinetic molecular theory for three-dimensional gases and lead to equally appealing results. The problem, however, is that many assumptions of the model are far less plausible for monolayers than for bulk gases. To see this, a brief review of the derivation seems necessary.

7.4b.1 Equation of State from Kinetic Theory

Suppose we imagine a single molecule bouncing back and forth across a surface between two restraining barriers. If we define the direction of this motion to be the x direction and the velocity of the molecule to be v_x, then the change in momentum at each collision (if we assume they are elastic) is

$$\Delta(\text{momentum})/\text{collision} = mv_x - (-mv_x) = 2\,m\,v_x \tag{8}$$

where m is the mass of the molecule. The time interval between two successive collisions at the same wall is given by

$$\text{elapsed time/collision} = 2\ell/v_x \tag{9}$$

if the distance between barriers is ℓ since the molecule must cross the distance between the barriers twice before returning to the same spot. The force exerted by the molecule on impact equals the rate of change in momentum, which, in turn, equals the ratio of Equation (8) to Equation (9):

$$\Delta(\text{momentum})/\Delta(\text{time}) = F_x = m\,v_x^2/\ell \tag{10}$$

This force is converted to two-dimensional pressure by dividing it by the length of the edge to which the force is applied. Assuming the accessible surface area to be a square means that the length of this edge is also ℓ; the pressure contribution of this one collision equals

$$\pi = F_x/\ell = mv_x^2/\ell^2 \tag{11}$$

The quantity ℓ^2 in this equation clearly describes the area accessible to the molecule.

Since pressure is isotropic, we assume that the forces on the perpendicular barriers are identical. Therefore, if the surface contains N molecules, they behave as if $N/2$ were exerting a pressure given by Equation (11) on the barriers perpendicular to the x direction, with the other $N/2$ exerting an identical pressure in the other direction. That is, for a surface containing N molecules

$$\pi A = \frac{N}{2} m\overline{v^2} = N(\overline{KE}) \tag{12}$$

The average value of the square velocity has been used in Equation (12) to allow for the fact that a distribution of molecular velocities exists. The nature of the averaging procedure to be used in this case is well established from physical chemistry. We also know from physical chemistry that the average kinetic energy per molecule ($\overline{KE}$) per degree of freedom is

$$\frac{\overline{KE}}{\text{degree of freedom}} = \frac{1}{2} k_B T \tag{13}$$

Since the molecules on the surface have two translational degrees of freedom, Equations (12) and (13) may be combined to give

$$\pi \frac{A}{N} = \pi\sigma = k_B T \tag{14}$$

which is identical to Equation (6).

The foregoing derivation is a straightforward two-dimensional analog of the three-dimensional case and leads to a result that describes the experimental facts. From a pragmatic point of view, it is a great success. One of the theoretical assumptions underlying Equation (13), however, is that translational quantum states are sufficiently close together to justify treating them as continuous rather than discrete. This is unquestionably true for gases. For an amphipathic molecule with a polar head that contacts—and interacts with—the aqueous substrate, it is somewhat harder to justify. We see below that there is a totally different way of looking at Equation (14) that is free from this objection.

Like its three-dimensional counterpart, Equation (5) is a limiting law, which means that deviations may be expected at higher pressures, lower temperatures, or with more strongly interacting molecules. Figure 7.10 is a plot of $\pi\sigma/k_B T$ versus π for several members of the carboxylic acid homologous series. The film balance was used to collect the data only for the C_{12} acid. Shorter chain compounds are soluble and were investigated by surface tension measurements and interpreted by the Gibbs equation, which we discuss in Section 7.7.

The main point to note about Figure 7.10 is the strong resemblance it bears to similar plots for three-dimensional gases. Negative deviations occur at low pressures, becoming more pronounced as the length of the alkyl chain increases. As with gases, this may be attributed to attraction between molecules, an effect that increases with chain length. At higher pressures, deviations tend to be positive. This is analogous to the excluded volume effect for gases, except that it becomes an excluded area in two dimensions.

* * *

EXAMPLE 7.1 *Determination of the Molecular Weight of a Solute from π Versus A Isotherms.* The molecular weight of a gas may be determined by measuring the mass of the sample as well as a set of p, V, and T values for the same sample. Use this fact to criticize or defend the

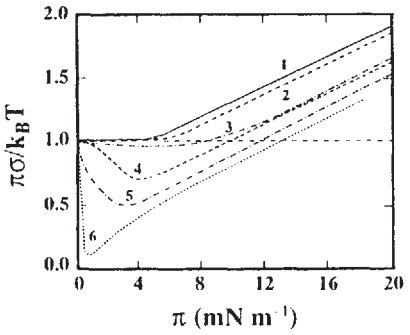


FIG. 7.10 Plots of $\pi\sigma/k_B T$ versus π for n-alkyl carboxylic acids: (1) C_4, (2) C_5, (3) C_6, (4) C_8, (5) C_{10}, and (6) C_{12}. (Data from N. K. Adam, *Chem. Rev.*, **3**, 172 (1926)).

following proposition: The molecular weight of a solute in a monolayer may be calculated by the formula $M = mRT/\pi A$, where A/m gives the area mass $^{-1}$ of the monolayer under a pressure π and at a temperature T.

Solution: The formula given is the two-dimensional analog of the ideal gas law; therefore two conditions must be met to justify its use. First, the monolayer must be in the gaseous state, and, second, the gas pressure must approach zero. The first point may appear trivial since no one would apply $pV = nRT$ to a bulk liquid sample for which p, V, and T had been measured. The two-dimensional state of a monolayer is not directly perceptible, however, and π-σ data must be evaluated to verify that the monolayer is indeed in the G state. With π-A data measured over a range of (low) π's, Figure 7.10 shows that the limiting value of $\pi A/RT$ is the number of moles in the sample. Dividing the mass of the sample by this value of n yields the correct molecular weight. The proposition needs to be qualified by adding "in the limit of $\pi \rightarrow 0$." ■

* * *

7.4b.2 Van der Waals Equation of State for Monolayers

We noted above that the applicability of Equation (14) to insoluble monolayers is severely restricted to very low values of π. Figure 7.10 shows that the deviations from Equation (14) with increases in π are very similar to what is observed for nonideal gases. Specifically, the positive deviations associated with excluded volume effects in bulk gases and the negative deviations associated with intermolecular attractions are observed. It is tempting to try to correct Equation (14) for these two causes of nonideality in a manner analogous to that used in the van der Waals equation:

$$\left(\pi + \frac{a}{\sigma^2}\right)(\sigma - b) = k_B T \tag{15}$$

where a and b are the two-dimensional analogs of the van der Waals constants. Note that b, the excluded area per molecule, is conceptually equivalent to σ^0, although which of the values $(\sigma^0_{LE}, \sigma^0_{LC}, \sigma^0_S)$ best fits the data cannot be predicted a priori.

The temperature at which the van der Waals equation goes from one having three real roots to one having one real root is generally identified with the critical temperature. In the same way, Equation (15) may be considered to connect both the gaseous (G) and liquid-expanded (LE) states in monolayers, the transition between which also displays a critical point. Statistical mechanics shows, however, that the van der Waals constant a explicitly ignores orientation effects as contributing anything to the energy in gases; it is hard to imagine the properties of insoluble monolayers as being independent of orientation. Therefore any attempt to correct Equation (14) in such a way as to extend its range encounters difficulties. Ultimately, all objections to two-dimensional equations of state seem to center on their neglect or unsatisfactory inclusion of the substrate.

7.4b.3 Monolayers as Two-Dimensional Binary Solutions

An alternative way of looking at monolayers is to consider them as two-dimensional binary solutions rather than two-dimensional phases of a single component. The advantage of this approach is that it does acknowledge the presence of the substrate and the fact that it plays a role in the overall properties of the monolayer. Although quite an extensive body of thermodynamics applied to two-dimensional solutions has been developed, we consider only one aspect of this. We examine the film pressure as the two-dimensional equivalent of osmotic pressure. It will be recalled that, at least for low osmotic pressures, the relationship among π_{osm}, V, n, and T is identical to the ideal gas law (Equation (3.25)). Perhaps the interpretation of film pressure in these terms is not too farfetched after all!

As we saw in Chapter 3, the heart of any osmotic pressure experiment is a semipermeable membrane that allows the solvent but not the solute to pass. The float of a Langmuir balance accomplishes this. That portion of surface with the monolayer is considered to be the two-dimensional solution; the clean surface is the two-dimensional solvent. The solvent can certainly pass from one region to another (remember that the mechanism of the partitioning has

nothing to do with the equilibrium osmotic pressure) through the bulk substrate. However, the insoluble solute is restrained by the float.

For osmotic equilibrium, the chemical potential of the solvent must be the same on both sides of the membrane. In the two-dimensional analog also the chemical potential must be the same for the water on both sides of the float. The presence of the solute lowers the chemical potential of the solvent, but the excess pressure compensates for this. Therefore, by analogy with Equation (3.19), we write

$$\mu_{1s}^{0} = \mu_{1s}^{0} + RT \ln a_{1s} + \int_{0}^{\pi} \overline{A}_1 \, dp \tag{16}$$

where the subscript s indicates the surface and the subscript 1 identifies the solvent. If the partial molal area of the solvent $\overline{A}_1$ is assumed to be independent of π, Equation (16) may be integrated to give

$$-RT \ln a_{1s} = \pi \overline{A}_1 \tag{17}$$

For ideal (dilute) solutions the activity is replaced by the mole fraction x_{1s}, which, for a two-component surface solution, equals $(1 - x_{2s})$. With the customary expansion of the logarithm as a power series (see Appendix A), these substitutions yield

$$RTx_{2s} = \pi \overline{A}_1 \tag{18}$$

the two-dimensional equivalent of Equation (3.23). Finally, it is necessary to relate the surface mole fraction and the molar area of the solvent to more familiar variables.

The surface mole fraction is entirely analogous to the bulk value of this quantity,

$$x_{2s} = \frac{n_{2s}}{n_{1s} + n_{2s}} = \frac{N_{2s}}{N_{1s} + N_{2s}} \tag{19}$$

where the n terms are the numbers of moles, and the N terms are the numbers of molecules. For dilute surface solutions—that is, expanded monolayers—$N_{1s} \gg N_{2s}$; therefore

$$x_{2s} \cong N_{2s}/N_{1s} \tag{20}$$

The total area of the surface may be written

$$A_T = n_{1s} \overline{A}_1 + n_{2s} \overline{A}_2 = N_{1s}\sigma_1^0 + N_{2s}\sigma_2^0 \tag{21}$$

Applying these various relationships to Equation (18) leads to the following result for low film osmotic pressures:

$$\pi(A_T - N_{2s}\sigma_2^0) = n_{2s} RT = N_{2s}k_B T \tag{22}$$

Dividing through by N_{2s} to express the total area as area per solute molecule gives

$$\pi(\sigma - \sigma_2^0) = k_B T \tag{23}$$

Equation (23) obviously gives the two-dimensional ideal gas law when $\sigma \gg \sigma_2^0$ and with the σ_2^0 term included represents part of the correction included in Equation (15). This model for surfaces is, of course, no more successful than the one-component gas model used in the kinetic approach; however, it does call attention to the role of the substrate as part of the entire picture of monolayers. We saw in Chapter 3 that solution nonideality may also be considered in osmotic equilibrium. Pursuing this approach still further results in the concept of phase separation to form two immiscible surface solutions, which returns us to the phase transitions described above.

* * *

EXAMPLE 7.2 *Use of the van't Hoff Equation for Monolayers.* A monolayer of egg albumin was spread on a concentrated aqueous solution of ammonium sulfate and π-A data were collected at 25°C by Bull (1945). Use the two-dimensional van't Hoff equation to evaluate the molecular weight of the albumin if $(\pi/c')_0 = 5.54 \cdot 10^5$ erg g^{-1}. In this expression c' is the

two-dimensional concentration in practical units, g cm^{-2}. How does this interpretation of π-A data compare with the interpretation given in Example 7.1?

Solution: According to the van't Hoff equation,

$$1/M = (1/RT)(\pi/c')_0 = (5.54 \cdot 10^5)/(8.314 \cdot 10^7)(298)$$
$$= 2.24 \cdot 10^{-5} \text{ mole g}^{-1}$$

therefore M = 44,700 g $mole^{-1}$. This is completely equivalent to the interpretation given in Example 7.1 since $\pi \to 0$ as $c' \to 0$. Therefore the same limiting value is obtained whether the limit is taken in terms of π or c'. ■

* * *

In summary, we see that insoluble monolayers may be viewed either as examples of two-dimensional phases of one component or as two-dimensional solutions with two components. The former model is somewhat simpler and is often adequate. The latter, although more complex, is more realistic. In spite of our interest in the monolayer, we must not neglect the fact that none of the monolayer phenomena would exist without the aqueous phase as the substrate.

7.4c The Liquid-Expanded Phase

Next, let us return to Figure 7.6, discussing the features labeled LE and LE/G. As mentioned above, the horizontal line of the LE/G region in Figure 7.6 is analogous in every way to the corresponding feature in bulk matter. At a given temperature there is a constant-pressure region over which a significant compression occurs. The film pressures at which LE/G equilibrium occurs are known as film vapor pressures π_v. Like the gaseous state itself, the LE/G equilibrium occurs at very low pressures. Tetradecanol, for example, has a two-dimensional vapor pressure of $1.1 \cdot 10^{-4}$ N m^{-1} at 15°C.

It is important to remember the significance of π_v. It refers specifically to the equilibrium between two surface states. There is a danger of confusing π_v with the equilibrium spreading pressure π_e, introduced in Chapter 6. The latter is the pressure of the equilibrium film that exists in the presence of excess bulk material on the surface. It is the equilibrium spreading pressure that is involved in the modification of Young's equation (Equation (6.49)), for which a bulk phase is present on the substrate. For tetradecanol at 15°C, the equilibrium spreading pressure is about $4.5 \cdot 10^{-2}$ N m^{-1}, so π_e and π_v are very different from one another.

The temperature variation of π_v may be analyzed by a relationship analogous to the Clapeyron equation to yield the two-dimensional equivalent to the heat of vaporization. The numerical values obtained for this quantity more nearly resemble the bulk values for hydrocarbons than those for polar molecules. This suggests that most of the change in the surface transition involves the hydrocarbon tail of the molecule rather than the polar head.

Finally, note that the LE/G equilibrium region disappears above a certain temperature that is the two-dimensional equivalent of the critical temperature for liquid-vapor equilibrium (see Fig. 7.8).

As described in Section 7.4a, the liquid-expanded state is the first of several condensed states. Because it is bound on the low-pressure side by a two-phase region with a critical temperature, the LE state is easily compared to a bulk liquid state. Since gaseous behavior is observed only at very low pressures, it is easy to extrapolate the isotherm for the LE state to the $\pi = 0$ value, as the dashed line in Figure 7.6 shows. For amphipathic molecules with saturated unbranched R groups, this intercept is in the range 0.45 to 0.55 nm^2. We shall identify the limiting area per molecule (superscript zero) for this state (subscript LE) by the symbol σ^0_{LE}. The presence of branched chains or double bonds—particularly in the cis configuration—increases the value of this limiting area.

The precise structural details of the liquid-expanded state at the molecular level are not fully understood, but several generalizations do appear to be justified. The value of σ^0_{LE} is several times the actual cross-sectional area of the amphipathic molecule, when oriented perpendicular to the surface. At the same time the area per molecule is considerably less than

could be permitted if the entire tail were free to move in the surface. The inset corresponding to the LE phase in Figure 7.6 represents a model of the surface at the molecular level. Here, part of the hydrocarbon chain lies in the surface, and some has been lifted out of the surface plane. That portion of the tail in the surface defines the effective area per molecule. Neither this area nor the length of the chain out of the surface will be the same for all molecules, so experimental values of σ^0_{LE} correspond to average values. Furthermore, there will be considerable lateral interaction between those segments that are not in contact with the substrate.

The compressibility of the LE state is expected to be far less than that of the gaseous state, but not yet incompressible, because the average area per molecule may be altered by squeezing additional CH_2 groups out of contact with the water.

7.4d The Intermediate Liquid Phase

With sufficient compression the isotherm of the LE state shows a sharp break to enter a situation variously known as the intermediate or transition state, indicated by I in Figure 7.6. Recent results suggest that the region I does represent a first-order phase transition (Knobler 1990a).

For a number of years, no satisfactory molecular interpretation existed for the I state. As its name implies, the I state was thought of as an intermediate transition state between the LE and LC states. If, however, the I state is a transition state that somehow becomes trapped in transit, it should be described as a metastable rather than a stable state. We saw in Chapter 6, Section 6.7 that metastable states are associated with nonreversible behavior, and I states seem to be reversible. A considerable number of studies have emerged in recent years on the evolution of the LC phase from the LE phase. It is known that very complex patterns (e.g., fractal dendritic structures) develop during this transition. In some cases, additional phases have also been identified (see Knobler 1990a). We shall not go into these here and merely emphasize that the isotherm shown in Figure 7.6 and the phase diagram Figure 7.8 are highly simplified.

7.4e The Liquid-Condensed and the Solid Phases

If the area of an insoluble monolayer is isothermally reduced still further, the compressibility eventually becomes very low. Because of the low compressibility, the states observed at these low values of σ are called *condensed* states. In general, the isotherm is essentially linear, although it may display a well-defined change in slope as π is increased, as shown in Figure 7.6. As mentioned above, the (relatively) more expanded of these two linear portions is the *liquid-condensed* state LC, and the less expanded is the *solid* state S. It is clear from the low compressibility of these states that both the LC and S states are held together by strong intermolecular forces so as to be relatively independent of the film pressure.

Because of the near linearity of these portions of the isotherm, it is easy to extrapolate both regions to their value at $\pi = 0$. The intercepts for the solid and liquid-condensed regions, σ^0_S and σ^0_{LC}, respectively, differ only slightly. Values of σ^0_{LC} for alcohols are about 0.22 nm^2, and for carboxylic acids about 0.25 nm^2, more or less independent of the length of the hydrocarbon chain. The intercept σ^0_S has a value of about 0.20 nm^2, independent of both the length of the chain and the nature of the head. The film pressures in the condensed states (LC or S) are of the same magnitude as the equilibrium spreading pressure for amphipathic molecules.

For the condensed LC and S states a molecular interpretation is again possible. In both the values of σ^0 are close to actual molecular cross sections when the molecules are oriented perpendicular to the surface. The difference between these two regions seems to involve the polar part of the molecule more than the hydrocarbon chain, which was more important for the more expanded states. The difference between σ^0_S and σ^0_{LC} may involve a more efficient packing of the heads or the formation of fairly specific lateral interactions through hydrogen bonds, for example. The values of σ^0 that are observed for monolayers of saturated *n*-alkyl compounds are only slightly larger than the close-packed cross sections obtained for these compounds in the bulk solid state by x-ray diffraction.

Additional compression eventually leads to the collapse of the film. The pressure π_c at which this occurs is somewhere in the vicinity of the equilibrium spreading pressure. Figure 7.7 represents schematically how this film collapse may occur. The mode of film buckling shown in Figure 7.7 is not the only possibility: head-to-head as well as tail-to-tail configurations can be imagined. The second structure strongly resembles that of cell membranes, which we discuss in the next chapter.

Film collapse represents the squeezing out of the surface of highly ordered aggregates that may quite plausibly be regarded as nuclei to bulk phase particles. If collapse marked the appearance of the amphipathic material in a bulk phase, then the collapse pressure and the equilibrium spreading pressure should be identical. Here a significant complication appears. The collapse pressure is highly sensitive to the rate at which the film is compressed. This indicates nonequilibrium conditions, showing that the two-dimensional solid phase resembles three-dimensional solids in this respect also, a difficulty in the attainment of thermodynamic equilibrium.

In summary, it must be emphasized again that there are wide variations in the properties of insoluble monolayers. Extensive compilations of π-A data for a large number of surfactants are available, although not over the full range of densities (Mingotaud et al. 1993). Some of the phenomena reported in the literature are probably artifacts due to impurities or nonequilibrium conditions. Others are probably unique effects that apply only to a very specific system. In the descriptive material of this section both phenomenological and modelistic information were provided for various stages along the isotherm. At the very least, the models serve the pedagogical function of assisting the student in remembering an assortment of probably unfamiliar facts. At best, the models provide the basis for quantitatively understanding these phenomena. In the next section we take a more quantitative look at the model for the gaseous state.

7.5 VISCOUS BEHAVIOR OF TWO-DIMENSIONAL PHASES

Our discussion of two-dimensional phases has drawn heavily on the analogy between bulk and surface behavior. This analogous behavior is not restricted to thermodynamic observations, but extends to other areas also. The viscosity of surface monolayers is an excellent example of this. To illustrate the parallel between bulk and surface viscosity, let us retrace some of the introductory notions of Chapter 4, restricting the flow to the surface region.

7.5a Coefficient of Surface Viscosity

We begin by defining the coefficient of surface viscosity η^s. Equation (4.1) serves to define the *bulk* viscosity; for surfaces we ignore the area extending in the z direction and consider the force per unit edge of the surface ℓ. Assuming a velocity gradient dv/dy exists between two edges of an element of area, we can write the two-dimensional analog of Equation (4.1) as

$$\frac{F}{\ell} = \eta^s \frac{dv}{dy} \tag{24}$$

This gives η^s units of mass time^{-1} in contrast to bulk viscosity, which has dimensions mass length^{-1} time^{-1}. In discussing Figure 7.3 we noted that—from a chemical if not geometrical point of view—surfaces extend over a thickness τ. By analogy with Equation (3), we might expect surface and bulk viscosities to be related by

$$\eta = \eta^s/\tau \tag{25}$$

7.5b Measurement of Surface Viscosity

Next let us consider how surface viscosities can be measured. A variety of methods for measuring η^s exist, including a method based on concentric rings, the two-dimensional equivalent of the concentric cylinder viscometer. We limit our discussion to the analog of the

capillary viscometer. Figure 7.11a shows an arrangement by which a monolayer can be pushed through a narrow channel by a moving barrier. Following a derivation that parallels the development of the Poiseuille equation (Equation (4.20)), the area rate at which the monolayer emerges from the channel under an applied pressure $\gamma^0 - \gamma = \Delta\gamma$ can be obtained. As in Section 4.4, we continue to describe the locations within the channel in terms of the distance r from its center line, with h the distance of the wall from the center line (see Figure 7.11b). The following steps are highlights of the derivation and parallel the presentation in Section 4.4:

1. Viscous and surface pressure forces balance under stationary-state conditions along the edge of the film a distance r from the center line. In terms of Figure 7.11b, this force balance is given by

$$\eta^s \ell \frac{dv}{dr} + \Delta\gamma r = 0 \tag{26}$$

2. Integrating Equation (26) and using the nonslip condition at the wall ($v = 0$ at $r = h$) to evaluate the constant of integration yields

$$\frac{\eta^s \ell}{\Delta\gamma} v = \frac{1}{2}(h^2 - r^2) \tag{27}$$

3. The rate of emergence of area, A/t, of the monolayer from the channel is twice the integral of $v dr$ between $r = 0$ and $r = h$:

$$\frac{A}{t} = \frac{2}{3}\frac{\Delta\gamma}{\eta^s \ell} h^3 \tag{28}$$

This is the two-dimensional equivalent of Poiseuille's equation. All of the other quantities besides η^s in Equation (28) are measurable, so η^s can be evaluated by measuring the rate at which the monolayer flows through the channel. In practice, a second barrier is moved along in front of the advancing interface to maintain a constant film pressure for an insoluble monolayer.

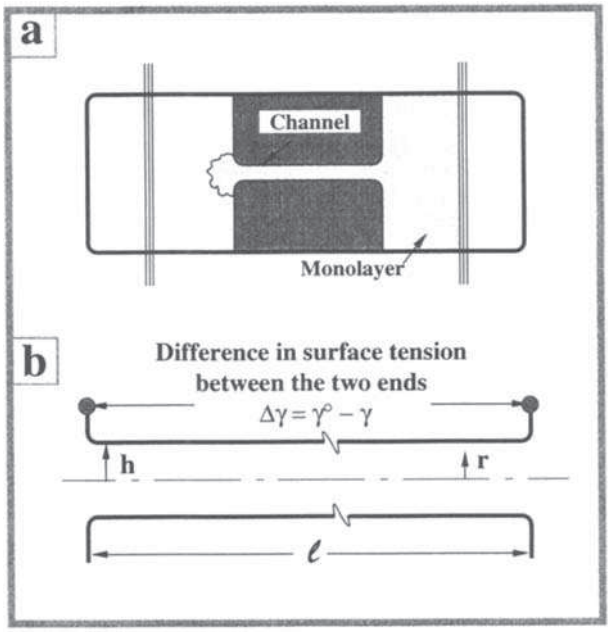


FIG. 7.11 Schematic representation of a surface viscometer: (a) a monolayer is pushed through a narrow channel; and (b) definition of variables for analysis.

In the two-dimensional gaseous state, surface viscosities can be as low as 10^{-8} kg s^{-1}, while for condensed states values range from 10^{-7} to 10^{-5} kg s^{-1}. While it makes sense that the surface viscosity increases as we move from G to L to S states, the numbers themselves mean little to us. To remedy this we repeat the sort of calculation done for two-dimensional pressure and examine what Equation (25) tells us about the equivalent bulk viscosity. As with Equation (3), we assume $\tau = 1.0$ nm; therefore a surface viscosity of 10^{-7} kg s^{-1} is equivalent to a bulk viscosity of

$$\eta = \frac{\eta^s}{\tau} = \frac{10^{-7}\ \text{kg s}^{-1}}{10^{-9}\ \text{m}} = 10^2\ \text{kg s}^{-1}\ \text{m}^{-1} = 10^3\ \text{P} \tag{29}$$

a value that suggests a consistency like that of butter for the molecules of the monolayer in condensed states. This is consistent with a high degree of lateral interaction between the vertically oriented chains. Once again, we see that a highly localized phenomenon translates into a dramatic effect when scaled up to macroscopic dimensions. In the next section we see some applications of monolayers or monolayer concepts that take advantage of the properties we have discussed.

Surface viscosities have been measured for soluble and insoluble monolayers, for charged and uncharged molecules, and at water-air and water-oil interfaces. We will not consider all these possibilities, but note instead that all share as a common feature the presence of polar head groups in the aqueous phase. In general, this means hydrogen bonding will occur between the amphipathic molecules at the surface and the substrate. A certain amount of water is dragged along with the surface molecules as a consequence. This effect was not taken into account in the derivation of Equation (28), although it has been investigated extensively. The effect of the entrained substrate may be incorporated into Equation (28) by multiplying the equation by a correction factor of the form $(1 + 2h\eta/\pi\eta^s)^{-1}$. This shows that the resistance to the motion of the monolayer due to this effect increases as the ratio η/η^s increases. This factor is incorporated into Equation (28) when η^s is determined experimentally by this method.

7.6 APPLICATIONS OF MONOLAYERS AND MONOLAYER CONCEPTS

At the surface of water, amphipathic molecules are oriented in such a way as to interact extensively, at least for ordinary surface concentrations. This results in the formation of the various two-dimensional condensed phases with the attendant effect on surface viscosity. In this section we consider some situations for which monolayers or the concepts involved in their discussion find application.

7.6a Retardation of Evaporation Using Monolayers

One area in which monolayers have been successfully employed is the retardation of evaporation. Particularly in arid regions of the world, evaporation of water from lakes and reservoirs constitutes an enormous loss of a vital resource. Under some conditions the water level of such bodies may change as much as 1 ft per month due to evaporation. The usual unit for water reserves is the acre-foot, a volume of water covering an acre of surface to the depth of 1 ft. It equals about 1/3 million gallons for each acre of water surface. Considerable research has been conducted both in the laboratory and in the field on the effectiveness of insoluble monolayers in reducing evaporation. An American Chemical Society Symposium in 1960 dealt exclusively with this topic; the proceedings of that symposium are given by LaMer (1962).

Laboratory research in this area is conducted by suspending a porous box of desiccant very close to the surface of a film balance. The rate of water uptake is determined by weighing at various times. This way the retardation of evaporation may be measured as a function of film pressure and correlated with other properties of the monolayer determined by the same method. As might be expected, the resistance to evaporation that a monolayer provides is enhanced by those conditions that promote the most coherent films, most notably high film pressures and straight-chain compounds. To see how this is quantified, consider the Example 7.3.

* * *

EXAMPLE 7.3 *Suppression of Evaporation by Monolayers.* The rate of evaporation is quantified by a parameter called the transport resistance r. For water with octadecanol monolayers at surface pressures of 10, 20, 30, and 40 mN m^{-1}, r is about 1, 2, 3, and 4 s cm^{-1}, respectively. This resistance drops off rapidly at lower pressures and approaches $2 \cdot 10^{-3}$ s cm^{-1} for pure water. By considering the rate of water uptake as a diffusion problem, suggest how these r values are calculated from data collected in an experiment like that described above. Use the fact that $1/r$ is dimensionally equivalent to the diffusion coefficient D divided by a length.

Solution: To be collected by the desiccant, molecules evaporating from a surface of area A must diffuse across a gap of width Δx between the water surface and the desiccant. The gap contains the monolayer as well as the air space, so the diffusion coefficient used is an effective value rather than the actual D value for a homogeneous region. According to Equations (2.20) and (2.22), the rate at which the desiccant increases in weight is given by

$$dQ/dt = AD(\Delta c/\Delta x)$$

In this expression Δc describes the difference between the concentration of water vapor at the water surface and that at the desiccant surface: $\Delta c = c_w - c_{des}$. Since $c_{des} \ll c_w$, Δc may be replaced by c_w, which in turn equals pM/RT, where p and M are the vapor pressure and molecular weight, respectively, of water. With this substitution the expression for dQ/dt becomes $A(D/\Delta x)(pM/RT)$. The ratio $D/\Delta x$ has units length time $^{-1}$, so we identify it as the reciprocal of the transport resistance. Note that r increases as the effective value of Δx increases and the effective value of D decreases. Thus $1/r$ is the only unknown in the expression $dQ/dt = A(1/r)(pM/RT)$ and can be calculated from the measured rate of weight increase with different monolayers present. ■

* * *

To be acceptable for use in the field, the monolayer material must have the following properties:

1. It must spread easily, probably as bulk material, so a high value of π_e is desirable.
2. It must be self-healing since surface ripples will disrupt the monolayer. This implies viscous rather than rigid monolayers.
3. It must be inexpensive, which means, effectively, capable of forming good films from naturally occurring mixtures.
4. It must be nontoxic and free from other deleterious effects on aquatic life.

Hexadecyl and octadecyl alcohol have been extensively studied and shown to be highly effective in evaporation retardation. Scattering powdered samples of commercial-grade alcohols by boat on lake surfaces or the continuous addition of alcohol slurries from floating dispensers are two of the methods that have been employed to apply these monolayers. Wind conditions and the activity of aquatic birds have a considerable effect on the stability of the monolayer and therefore on the rate at which the monolayer chemicals must be reapplied. Rates of application rarely exceed 0.5 lb $acre^{-1}$ day^{-1}, however, so that the cost of the materials used is not excessive.

An indication of the effectiveness of such field treatment is seen in Figure 7.12, which compares the amount of water lost by evaporation from two small adjacent lakes in Illinois. Treatment of North Lake (ordinate) with commercial hexadecanol was begun in late June 1957; untreated South Lake was the control (abscissa). Prior to treatment the evaporation losses from the two lakes were identical, as shown by the 45° line in the figure. After treatment was begun, however, the loss of water from North Lake fell considerably behind that from South Lake. By the end of the summer a difference of about 40% in the water loss was observed. This was equivalent to about 7600 gallons of water saved per pound of hexadecanol used. For areas where water is scarce—the southwestern United States, Israel, and western Australia, for example—such conservation of water is highly valued, and research continues to look for methods to improve the efficiency of this technique.

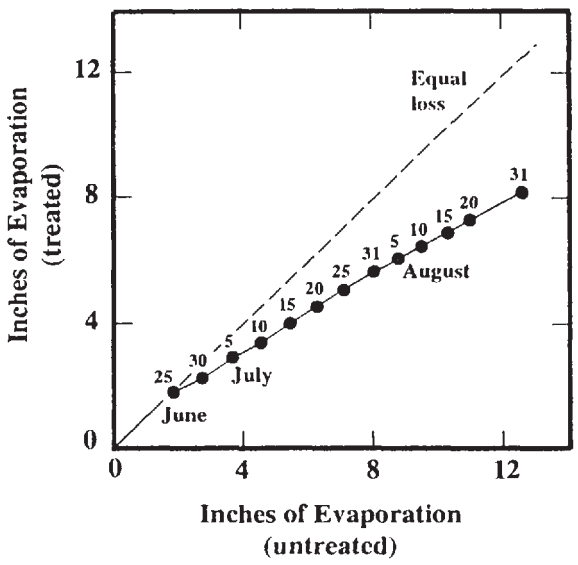


FIG. 7.12 Comparison of the water level in two adjacent lakes during the summer, 1957. The ordinate shows the level in the lake with the monolayer; the abscissa is the level in the untreated lake. (Redrawn with permission of LaMer 1962.)

7.6b Damping of Waves

Surface viscosity also has observable effects on macroscopic bodies of water. As mentioned in the opening paragraph of this chapter, the calming effect on surface turbulence of pouring oil on the sea has been known from antiquity. In terms of the concepts of this chapter, the increase in surface viscosity produced by the film has a damping effect on waves. Such damping has been studied both theoretically and in laboratory situations; the book by Davies and Rideal (1961) contains some interesting photographs of the ripples on a pond before and after the application of hexadecanol to the surface.

7.6c Stabilization of Emulsions and Foams

Emulsions and foams are two other areas in which dynamic and equilibrium film properties play a considerable role. Emulsions are colloidal dispersions in which two immiscible liquids constitute the dispersed and continuous phases. Water is almost always one of the liquids, and amphipathic molecules are usually present as emulsifying agents, components that impart some degree of durability to the preparation. Although we have focused attention on the air-water surface in this chapter, amphipathic molecules behave similarly at oil-water interfaces as well. By their adsorption, such molecules lower the interfacial tension and increase the interfacial viscosity. Emulsifying agents may also be ionic compounds, in which case they impart a charge to the surface, which in turn establishes an ion atmosphere of counterions in the adjacent aqueous phase. These concepts affect the formation and stability of emulsions in various ways:

1. Most emulsions are formed by some sort of comminution process in which large blobs of the dispersed phase are eventually ground down to small drops. This is a complex process, but basically consists of drops becoming elongated under shearing forces, necking, and finally separating into smaller drops. The adsorption of a surface film with the attendant lowering of γ and increase in η^s clearly enters the picture.

2. The first step in the "breaking" of an emulsion is the coming together of the individual drops. If water is the continuous phase and the emulsifier is ionic, then it is the ion atmo-

spheres of the approaching particles that make the first contact. We have already seen in Chapter 4 that the overlap of the electrical double layers affects the viscosity of the dispersion. We see in Chapter 13 that the details of this first encounter can determine whether the drops form an aggregate or go their separate ways.

3. If droplets can aggregate into a single kinetic unit, they might also coalesce into a single geometrical unit. This involves rupture of the thin film of continuous phase that separates them in an aggregate. Again, surface tension and surface viscosity are certainly pertinent to the coalescence process.

The huge variety of emulsions used as food, medicinal, cosmetic, and other industrial products make these colloids important practical systems in which the surface monolayers exert considerable influence. We have already discussed the use of lecithin to control the viscosity and the texture of chocolate in Vignette IV in Chapter 4.

Foams are colloidal systems in which a gas is the dispersed phase. Although a whole range of concentrations is possible, we shall focus on those foams that consist of volume-filling, distorted polyhedra separated by liquid films. For aqueous foams the high area of air-water interface requires adsorption to lower the surface tension sufficiently to make the foam in the first place. Foams drain by losing liquid through the channels that occur at the junction of the planar film surfaces. Such surfaces meet with curved menisci between them. This means that the pressure is lower in the junction than in the flat faces of the film according to the Laplace equation, Equation (6.29). As a consequence, liquid flows from the planar regions into these junctions—called plateau borders—through which the drainage occurs. As the film thickness of the continuous phase decreases, the probability of rupture due to thermal or mechanical fluctuations increases. Surface energetics and viscosity are important here also. As with emulsions, foams occur in many systems familiar to consumers, such as fire-fighting foams, whipped cream, shaving lather, and the head on a glass of beer!

7.6d Preparation of Langmuir-Blodgett Films

As we pointed out in the vignette at the beginning of this chapter, many potential applications emerge when the Langmuir layers are transferred to a solid substrate. We have some more to say about Langmuir-Blodgett films in Section 7.10c, but it is clear, based on what we have discussed so far, that understanding the different structural features of Langmuir layers and how to control the stability of the layers is the first prerequisite for depositing Langmuir-Blodgett layers on solid substrates.

Finally, it is worth noting that the monolayers of the type discussed in the previous sections may serve as good model systems for examining some of the theories of condensed matter physics.

Although we started out this chapter by discussing insoluble monolayers, it is evident that we have slipped into examples for which soluble amphipathics are being considered. In the next section we examine the thermodynamics of adsorption from solution.

7.7 ADSORPTION FROM SOLUTION: THERMODYAMICS

Until now we have discussed only insoluble monolayers. Although their behavior is complex, they have the conceptual simplicity of being localized in the interface. It has been noted, however, that even in the case of insoluble monolayers, the substrate should not be overlooked. The importance of the adjoining bulk phases is thrust into even more prominent view when soluble monolayers are discussed. In this case the adsorbed material has appreciable solubility in one or both of the bulk phases that define the interface.

7.7a The Gibbs Equation: Multicomponent Systems

Gibbs treated this situation as part of his investigations into phase equilibria. Suppose we consider two phases α and β in equilibrium with a surface s dividing them. For the system so constituted, we may write

$$G = G^{\alpha} + G^{\beta} + G^{s} \tag{30}$$

where the superscripts indicate the contribution from each category. For the bulk phases

$$G = E + pV - TS + \sum_i \mu_i n_i \tag{31}$$

where the chemical potential terms are summed for all components i. The superscript has been omitted for convenience. The volume term is replaced by an area term in the corresponding expression for G^s:

$$G^s = E^s + \gamma A - TS^s + \sum_i \mu_i n_i \tag{32}$$

where μ_i's and n_i's here are for the surface phase. Substituting Equations (31) and (32) into Equation (30) and taking the total derivative yields

$$dG = \sum_{\alpha,\beta,s} \left(dE + pdV + Vdp - TdS - SdT + \sum_i \mu_i dn_i + \sum_i n_i d\mu_i \right) + Ad\gamma + \gamma dA \tag{33}$$

For a reversible process

$$dE = \delta q - \delta w = \sum_{\alpha,\beta,s} dE = \sum_{\alpha,\beta,s} [TdS - (pdV + \delta w_{non\text{-}pV})] \tag{34}$$

Substituting this result into Equation (33) gives

$$dG = \sum_{\alpha,\beta,s} \left(Vdp - SdT + \sum_i \mu_i dn_i + \sum_i n_i d\mu_i - \delta w_{non\text{-}pV} \right) + Ad\gamma + \gamma dA \tag{35}$$

As we saw in Chapter 6 (Equation (6.16)), the quantity γdA may be equated to non-pressure-volume work when surface energy is being considered. With this consideration, Equation (35) simplifies still further to become

$$dG = \sum_{\alpha,\beta,s} \left(Vdp - SdT + \sum_i \mu_i dn_i + \sum_i n_i d\mu_i \right) + Ad\gamma \tag{36}$$

Another well-known relationship from thermodynamics may be introduced at this point:

$$dG = Vdp - SdT + \sum_i \mu_i dn_i \tag{37}$$

Applying Equation (37) to the bulk phases and the surface and subtracting the result from Equation (36) gives

$$\sum_i n_i^{\alpha} d\mu_i + \sum_i n_i^{\beta} d\mu_i + \sum_i n_i^{s} d\mu_i + Ad\gamma = 0 \tag{38}$$

When only one phase is under consideration, only one of the terms in Equation (30) is required, and only one of the bulk phase summations in Equation (38) survives. The result in this case is the famous *Gibbs-Dühem equation*:

$$\sum_i n_i d\mu_i = 0 \tag{39}$$

It will be recalled from physical chemistry that this relationship permits the evaluation of the activity of one component from measurements made on the other in binary solutions.

By means of the Gibbs-Dühem equation, we may eliminate the terms in Equation (38) that apply to bulk phases and write

$$\sum_i n_i^{s} d\mu_i + Ad\gamma = 0 \tag{40}$$

This is the *Gibbs adsorption equation* that relates γ to the number of moles and the chemical potentials of the components in the interface.

7.7b Two-Component Systems

In subsequent developments we consider only two-component systems and identify the solvent (usually water) as component 1 and the solute as component 2. In terms of this stipulation, Equation (40) becomes

$$n_1^{\sigma} d\mu_1 + n_2^{\sigma} d\mu_2 + A d\gamma = 0 \tag{41}$$

It is conventional to divide Equation (41) through by A to give

$$d\gamma = -\frac{n_1^{\sigma}}{A} d\mu_1 - \frac{n_2^{\sigma}}{A} d\mu_2 \tag{42}$$

The quantity n_i^{σ}/A is called the *surface excess* of component i and is given the symbol Γ_i

$$\Gamma_i = n_i^{\sigma}/A \tag{43}$$

In this notation, Equation (42) becomes

$$-d\gamma = \Gamma_1 d\mu_1 + \Gamma_2 d\mu_2 \tag{44}$$

We return to further simplifications of this equation after a brief discussion of how to define the position of a surface and how to define surface excess quantities.

7.7c Location of the Surface and the Meaning of Surface Excess Properties

Before proceeding any further, it is necessary to examine just what the concept of a surface excess means. To do this it is convenient to consider the changes that occur in some general property P as we move from phase α to phase β. The situation is represented schematically by Figure 7.13, in which x is the distance measured perpendicular to the interface. The scale of this figure is such that variations at the molecular level are shown. The interface is not a surface in the mathematical sense, but rather a zone of thickness τ across which the properties of the system vary from values that characterize phase α to those characteristic of β. In spite of this, we generally do not assign any volume to the surface, but treat it as if the properties of α and β applied right up to some dividing plane situated at some specific value of x. What is this position x_0 at which we draw such a boundary?

Suppose the solid line in Figure 7.13 represents the actual variation of property P. The

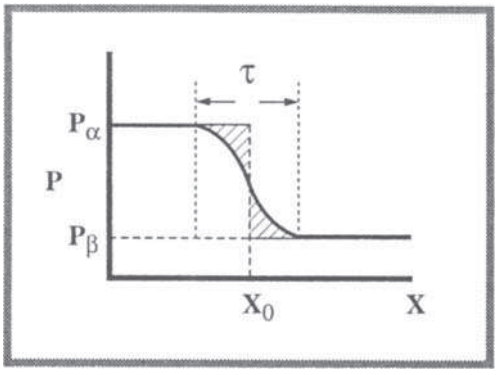


FIG. 7.13 Variation of some general property P with perpendicular distance from the surface in the vicinity of an interface between two phases α and β.

squared-off extensions of the bulk values of this property represent the approximation made in assuming the surface to have zero thickness. Then the shaded area to the left of x_0 shows the amount by which the value of P for the system as a whole has been overestimated by extending P_α. Likewise, the shaded area to the right of x_0 shows how the extension of P_β leads to an underestimation of P for the system as a whole. In principle, the "surface" may be located at an x value such that these two areas compensate for one another; that is, x_0 may be chosen so that the two shaded areas in the figure are equal.

This is where the trouble begins! Generally speaking, the kind of profile sketched in Figure 7.13 will be different for each property considered. Therefore we may choose x_0 to accomplish the compensation discussed herein for one property, but this same line will divide the profiles of other properties differently. The difference between the "overestimated" property and the "underestimated" one accounts for the "surface excess" of this property.

From the point of view of thermodynamics—which is oblivious to details at the molecular level—the dividing boundary may be placed at any value of x in the range τ. The actual placement of x_0 is governed by consideration of which properties of the system are most amenable to thermodynamic evaluation. More accurately, that property that is least convenient to handle mathematically may be eliminated by choosing x_0 so that the difficult quantity has a surface excess of zero.

For example, if the property in Figure 7.13 was G and the dividing surface was placed so that the two shaded regions would be equal, then there would be no surface excess G: The last term in Equation (30) would be zero. The Gibbs free energy is convenient to work with, however, so such a choice for x_0 would not be particularly helpful. Until now we have not had any reason to identify the surface of physical phases with any specific mathematical surface. We had not, that is, until Equation (44) was reached. Now things are somewhat different.

Suppose the property represented in Figure 7.13 is the number of moles of solvent per unit area in a slice of solution at some value of x. This quantity will clearly undergo a transition in the vicinity of an interface. We choose x_0 so that the shaded areas are equal when this is the quantity of interest. This placement of the dividing surface means

$$\Gamma_1 = 0 \tag{45}$$

With this situation, Equation (44) becomes

$$d\gamma = -\Gamma_2 \, d\mu_2 \tag{46}$$

The physical significance of Γ_2 is determined by the arbitrary placement of the mathematical surface that made $\Gamma_1 = 0$; that is, Γ_2 equals the algebraic difference between the "overestimated" and "underestimated" areas of the curve describing moles of solute when this curve is divided at a location x_0 that makes the surface excess of the solvent zero.

It is important to realize that the mathematical dividing surface just discussed is a reference level rather than an actual physical boundary. What is physically represented by this situation may be summarized as follows. Two portions of solution containing an identical number of moles of solvent are compared. One is from the surface region and the other from the bulk solution. The number of moles of solute in the sample from the surface minus the number of moles of solute in the sample from the bulk give the surface excess number of moles of solute according to this convention. This quantity divided by the area of the surface equals Γ_2. To emphasize that the surface excess of component 1 has been chosen to be zero in this determination, the notation Γ_2^1 is generally used.

It should be evident from the foregoing discussion that the property defined to have zero surface excess may be chosen at will, the choice being governed by the experimental or mathematical features of the problem at hand. Choosing the surface excess number of moles of one component to be zero clearly simplifies Equation (44). The same simplification could have been accomplished by defining the mathematical surface so that Γ_2 would be zero, a choice that would obviously deemphasize the solute. If the total number of moles N, the total volume V, or the total weight W had been the property chosen to show a zero surface excess, then in each case both Γ_1 and Γ_2 (which would be identified as Γ^N, Γ^V, or Γ^W for these three conventions) would have nonzero values. Last, note that the surface "excess" is an algebraic

quantity that may be either positive or negative depending on the convention chosen for Γ. A variety of different experimental methods are encountered in the literature to measure "surface excess" quantities; one must be careful to understand clearly what conventions are used in the definition of these quantities.

7.7d Relation Between Surface Tension and Surface Excess Concentration

Equation (46), one form of the Gibbs equation, is an important result because it supplies the connection between the surface excess of solute and the surface tension of an interface. For systems in which γ can be determined, this measurement provides a method for evaluating the surface excess. It might be noted that the finite time required to establish equilibrium adsorption is why dynamic methods (e.g., drop detachment) are not favored for the determination of γ for solutions. At solid interfaces, γ is not directly measurable; however, if the amount of adsorbed material can be determined, this may be related to the reduction of surface free energy through Equation (46). To understand and apply this equation, therefore, it is imperative that the significance of Γ_2 be appreciated.

Now let us return to the development of Equation (46). The chemical potential depends on the activity according to the equation

$$\mu_2 = \mu_2^0 + RT \ln a_2 \tag{47}$$

In applying these results to adsorption from solution, the activity equals the pressure or concentration multiplied by the activity coefficient f. Differentiation of Equation (47) at constant temperature yields

$$d\mu_2 = RT \frac{da_2}{a_2} = RT\, d \ln (fc) \tag{48}$$

This relationship may also be applied to the adsorption of gases by replacing concentration by gas pressure and continuing to use the appropriate activity coefficient. We return to the application of this result to the adsorption of gases in Chapter 9.

For adsorption from dilute solutions the activity coefficient approaches unity, in which case the combination of Equations (46) and (48) leads to the result

$$\Gamma_2^1 = -\frac{c}{RT}\left(\frac{d\gamma}{dc}\right)_T = -\frac{1}{RT}\left(\frac{d\gamma}{d \ln c}\right)_T \tag{49}$$

This form of the Gibbs equation shows that the slope of a plot of γ versus the logarithm of concentration (or activity if the solution is nonideal) measures the surface excess of the solute. It might also be noted that the choice of units for concentration is immaterial at this point.

7.8 THE GIBBS EQUATION: EXPERIMENTAL RESULTS

7.8a Typical Variations of Surface Tension in Aqueous Solutions

Surface tensions for the interface between air and aqueous solutions generally display one of the three forms indicated schematically in Figure 7.14. The type of behavior indicated by curves 1 and 3 indicates positive adsorption of the solute. Since $d\gamma/dc$ and therefore $d\gamma/d \ln c$ are negative, Γ_2^1 must be positive. On the other hand, the positive slope for curve 2 indicates a negative surface excess, or a surface depletion of the solute. Note that the magnitude of negative adsorption is also less than that of positive adsorption.

Curve 1 in Figure 7.14 is the type of behavior characteristic of most un-ionized organic compounds. Curve 2 is typical of inorganic electrolytes and highly hydrated organic compounds. The type of behavior indicated by curve 3 is shown by soluble amphipathic species, especially ionic ones. The break in curve 3 is typical of these compounds; however, this degree of sharpness is observed only for highly purified compounds. If impurities are present, the

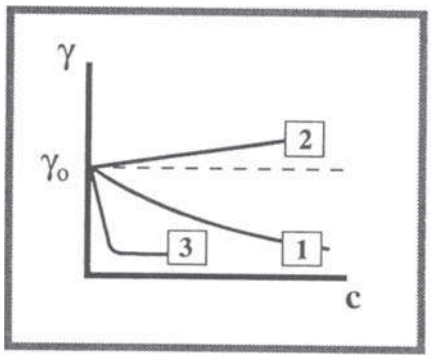


FIG. 7.14 Three types of variation of γ with c for aqueous solutions: (1) simple organic solutes, (2) simple electrolytes, and (3) amphipathic solutes.

curve will display a slight dip at this point. All three of these curves correspond to relatively dilute solutions. At higher concentrations effects other than adsorption may lead to departures from these basic forms. We say a bit more about adsorption from binary solutions over the full range of compositions in Section 7.9c.4.

7.8a.1 Simple Organic Solutes

For the limit as $c \rightarrow 0$, curve 1 may be presented by the equation of a straight line:

$$\gamma = \gamma_0 - mc \tag{50}$$

where m is the initial slope of the line. This is the same as

$$\pi = mc \tag{51}$$

From Equation (50) $d\gamma/dc = -m$, and from Equation (51) $c = \pi/m$; therefore Equation (49) may be written

$$\Gamma_2^1 = \pi/RT \tag{52}$$

Recalling the definition of Γ_2^1 provided by Equation (43), we see that Equation (52) may also be written

$$\pi A = n_2^s RT \tag{53}$$

the two-dimensional ideal gas law again! Those carboxylic acids containing less than 12 carbons in the alkyl chain for which results were presented in Figure 7.10 were investigated by this method. This same analysis also applies to the branch of curve 3 in Figure 7.14 as $c \rightarrow 0$.

7.8a.2 Simple Electrolytes

Curve 2 in Figure 7.14 indicates a negative surface excess of simple electrolytes. This means that portions of solution from both the surface and bulk regions that contain the same number of moles of solvent will have more solute in the bulk region than at the surface. Obviously, the surface is enriched over the bulk in solvent, a fact that is easily understood when the hydration of the ions is considered. Water molecules interact extensively with ions, a fact that accounts in part for the excellent solvent properties of water for ionic compounds. To move an ion directly to the air-water interface would require considerable energy to partially dehydrate the ion. Accordingly, the first couple of molecular diameters into the solution will be a layer of essentially pure water, the ions being effectively excluded from this region. The surface tension is not that of pure water, but is increased slightly due to the small surface deficiency of solute. Other highly solvated solutes such as sucrose also show this effect.

7.8a.3 Amphipathic Solutes

Curve 3 in Figure 7.14 applies primarily to amphipathic species. Most long-chain amphipathic molecules are insoluble unless the hydrophobic alkyl part of the molecule is offset by an ionic head or some other suitably polar head such as a polyethylene oxide chain, $-(CH_2CH_2O)_n-$. Like their insoluble counterparts, these substances form an oriented monolayer even at low concentrations. Figure 7.15 shows some actual experimental plots of type 3 for the ether that consists of a dodecyl chain and a hexaethylene oxide chain ($n = 6$) in the general formula just given. Example 7.4 illustrates the application of the Gibbs equation to these data.

* * *

EXAMPLE 7.4 *Determination of Surface Excess Concentration from Surface Tension Data.* The slope of the 25°C line in Figure 7.15 on the low-concentration side of the break is about -16.7 mN m^{-1}. Calculate the surface excess and the area per molecule for the range of concentrations shown. How would Figure 7.15 be different if accurate measurements could be made over several more decades of concentration in the direction of higher dilution? Could the data still be interpreted by Equation (49) in this case?

Solution: The surface excess is constant over this range of concentrations as indicated by the linearity of Figure 7.15. Equation (49) can be used directly to evaluate Γ_2^1. Since base 10 logarithms are used in the figure, we write

$$\Gamma_2^1 = -(0.0167)/(2.303)(8.314)(298) = 2.93 \cdot 10^{-6} \text{ mole m}^{-2}$$

The reciprocal of this gives the area of surface occupied by a mole of adsorbed molecules. Division by Avogadro's number converts the reciprocal of Γ_2^1 into a value for σ:

$$\sigma = (1 \text{ m}^2/2.93 \cdot 10^{-6} \text{ mole}) \cdot (1 \text{ mole}/6.02 \cdot 10^{23} \text{ molecules}) \cdot (10^{-9} \text{ nm}/1 \text{ m})^2$$
$$= 0.56 \text{ nm}^2$$

If accurate measurements could be made to increasingly lower concentrations, the surface excess would gradually decrease toward zero. This means that the lines in Figure 7.15 must eventually curve until they show a slope of zero at infinite dilution. Curved lines on a semilogarithmic plot of γ versus c are interpreted by drawing tangents at the concentrations of interest and applying the Gibbs equation to the slopes of the tangents to give the corresponding surface excesses. ■

* * *

The polar heads of the solute molecules in Figure 7.15 are much bulkier than those of the simple amphipathic molecules with insoluble monolayers that we discussed above. This is

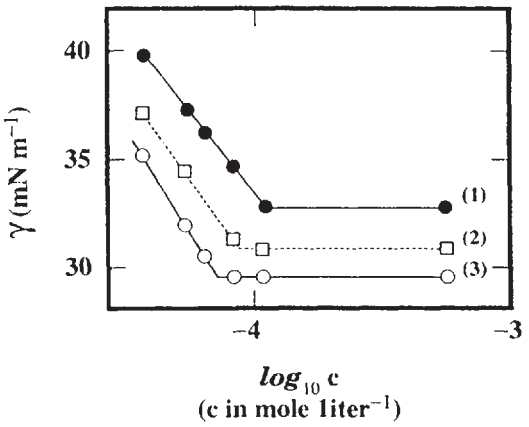


FIG. 7.15 Plot of γ versus $\log_{10} c$ for the dodecyl ether of hexaethylene oxide at three temperatures: (1) 15°C, (2) 25°C, and (3) 35°C. (Redrawn with permission of J. M. Corkill, J. F. Goodman, and R. H. Ottewill, *Trans. Faraday Soc.*, **57**, 1927 (1961).)

especially true when the hydration of the ether oxygens is considered. In view of this, the value calculated in Example 7.4 is probably about as small a value as these molecules can achieve. This suggests that the amphipathic molecules are in a highly condensed surface state at the concentrations investigated in Figure 7.15.

The break in curve 3 in Figure 7.14 is characteristic of this type of plot for soluble amphipathic molecules. Note that it appears in the experimental curves of Figure 7.15 also. The break is understood to indicate the threshold of *micelle* formation (see Chapter 1, Section 1.3a), known as the *critical micelle concentration* (see Chapter 8). We do not discuss this phenomenon any further since the next chapter is devoted entirely to micelles and related structures.

7.8b Effect of Ionic Dissociation on Adsorption

It is instructive to consider the effect of dissociation on the adsorption of amphipathic substances since many of the compounds that behave according to curve 3 are electrolytes. We consider only the case of strong 1 : 1 electrolytes; for weak electrolytes the equilibrium constant for dissociation must be considered.

If an ionic solute is totally dissociated into positive and negative ions, then its activity is given by

$$a_{MR} = a_M a_R \cong c_M c_R \tag{54}$$

where the subscripts M and R refer to the cation and amphipathic anion, respectively. Analogous results would be obtained if the cation were the amphipathic species. The approximation included in Equation (54) applies to the case in which the activity coefficient equals unity. Substituting this result into Equation (49) gives

$$\Gamma^1_{MR} = -\frac{1}{2RT}\left(\frac{d\gamma}{d\ln c}\right)_T \tag{55}$$

The assumption that no other electrolyte is present is implicit in this result. Now let us consider what happens when the system also contains a nonamphipathic electrolyte with a common ion to the surface-active electrolyte.

If a second electrolyte MX is present in addition to MR, then Equation (44) must be written

$$-d\gamma = \Gamma^1_M d\mu_M + \Gamma^1_R d\mu_R + \Gamma^1_X d\mu_X \tag{56}$$

Now the condition of surface neutrality becomes

$$\Gamma^1_M = \Gamma^1_R + \Gamma^1_X \tag{57}$$

so Equation (56) may be written

$$-d\gamma = \Gamma^1_R(d\mu_M + d\mu_R) + \Gamma^1_X(d\mu_M + d\mu_X) \tag{58}$$

This result may now be simplified by invoking some previous results. Recalling curve 2 from Figure 7.14, we know that the surface excess of the X^- ions is likely to be a small negative number that we shall set equal to zero as a first approximation. With this approximation, Equation (58) becomes

$$-d\gamma \approx \Gamma^1_R(d\mu_M + d\mu_R) = \Gamma^1_R RT\left(\frac{dc_M}{c_M} + \frac{dc_R}{c_R}\right) \tag{59}$$

Now let us consider a small change in the concentration of MR while the concentration of MX remains constant and considerably greater than the total MR concentration. Under these conditions, $dc_M = dc_R$ and $c_M \gg c_R$; therefore $d\ln c_M \ll d\ln c_R$ and Equation (59) becomes

$$-d\gamma = \Gamma^1_R(RT\, d\ln c) \tag{60}$$

Equations (55) and (60) are thus seen to describe the adsorption of MR in the absence of electrolyte and in the presence of swamping amounts of electrolyte, respectively. It is clear

from the difference between these two results that extreme care must be taken in the study of charged monolayers if the effect of the charge on the state of the monolayer is to be properly considered in the interpretation of experimental results.

The difference between Equations (55) and (60) may be qualitatively understood by comparing the results with the Donnan equilibrium discussed in Chapter 3. The amphipathic ions may be regarded as restrained at the interface by a hypothetical membrane, which is of course permeable to simple ions. Both the Donnan equilibrium (Equation (3.85)) and the electroneutrality condition (Equation (3.87)) may be combined to give the distribution of simple ions between the bulk and surface regions. As we saw in Chapter 3 (e.g., see Table 3.2), the restrained species behaves more and more as if it was uncharged as the concentration of the simple electrolyte is increased. In Chapter 11 we examine the distribution of ions near a charged surface from a statistical rather than a phenomenological point of view.

7.8c Measuring Surface Excess Concentrations

We have noted previously that measuring γ as a function of concentration is a convenient means of determining the surface excess of a substance at a mobile interface. In view of the complications arising from charge considerations, the need for an independent method for measuring surface excess becomes apparent. Some elaborate techniques have been developed that involve skimming a thin layer off the surface of a solution and comparing its concentration with that of the bulk solution.

A simpler method for verifying the Gibbs equation involves the use of isotopically labeled surfactants. If the isotope emits a low-energy β particle, the range of β in water will be very low. Thus a detector placed just above the surface will count primarily those emissions originating from the surface region. Tritium (^{3}H), for example, emits a 0.0186-MeV β particle with a range in water of only about 17 μm, which means that only a negligible fraction of the β particles can travel farther than this in water. In fact, most are absorbed in an even shorter distance, so any ^{3}H β particles detected above an aqueous solution of tritiated surfactant probably originate within approximately 3 μm of the surface. The contribution of the bulk solution to the "background" of the former measurement is made using the same isotope in a compound that is known not to be adsorbed. By such studies the kinds of effects just described have been investigated and verified.

The surface-active substances we have discussed have been purified, research-quality materials. In practical situations the cost of synthesizing and purifying such surfactants is prohibitive. The materials commercially used, therefore, are inevitably mixtures. Commercial surfactants originate, for example, from the esterification of sugars or the sulfonation of alkyl-aryl mixtures. Such mixtures are marketed under a bewildering variety of trade names, and often as members of number- or letter-designated series that correspond, roughly, to a set of homologs. Table 7.2 lists examples of several specific members of such series, along with a brief description of the general nature of the family to which they belong.

7.9 ADSORPTION ON SOLID SURFACES

7.9a The Langmuir Equation: Theory

Throughout most of this chapter we have been concerned with adsorption at mobile surfaces. In these systems the surface excess may be determined directly from the experimentally accessible surface tension. At solid surfaces this experimental advantage is missing. All we can obtain from the Gibbs equation in reference to adsorption at solid surfaces is a thermodynamic explanation for the driving force underlying adsorption. Whatever information we require about the surface excess must be obtained from other sources.

If a dilute solution of a surface-active substance is brought in contact with a large adsorbing surface, then extensive adsorption will occur with an attendant reduction in the concentration of the solution. To meet the requirement of a large surface available for adsorption, the solid—which is called the *adsorbent*—must be finely subdivided. From the analytical data

TABLE 7.2 Some Familes of Commercial Surfactants and Specific Examples from Each

Name of series	General chemical nature	Specific designation ("____") and chemical nature of example	Example
Igepon "____"	Fatty acid amide of methyltaurine	"TN" R = palmityl	$RCON(CH_3)C_2H_4SO_3^-Na^+$
Aerosol "____"	Alkyl ester of sulfosuccinic acid	"OT" R = octyl	CH_2-COOR $\vert$ CH_2-COOR $\vert$ $SO_3^-Na^+$
Span "____"	Fatty acid esters of anhydrosorbitols	"60" R = stearyl	$HO-CH-C\ HOH$ $\vert\quad\quad\quad

$$K' = \frac{a_2^s a_1^b}{a_1^s a_2^b} \tag{62}$$

where a stands for the activity of the species and the superscripts s and b signify surface and bulk values, respectively. Next let us assume that the two-dimensional surface solution is ideal, an assumption that enables us to replace the activity at the surface by the mole fraction at the surface x^s:

$$K' = \frac{x_2^s a_1^b}{x_1^s a_2^b} \tag{63}$$

Since the surface contains only two components, $x_1^s + x_2^s = 1$ and Equation (63) becomes

$$K' = \frac{x_2^s a_1^b}{(1 - x_2^s) a_2^b} \tag{64}$$

Equation (64) may be rearranged to give

$$x_2^s = \frac{K' a_2^b / a_1^b}{K' a_2^b / a_1^b + 1} \tag{65}$$

In dilute solutions the activity of the solvent is essentially constant, so the ratio K'/a_1^b may be defined to equal a new constant K, in terms of which Equation (65) becomes

$$x_2^s = \frac{K a_2^b}{K a_2^b + 1} \tag{66}$$

This is one form of the Langmuir adsorption isotherm.

An equivalent form of the Langmuir equation expressed in slightly different variables is obtained by writing both x_1^s and x_2^s in Equation (63) in terms of area fractions occupied by '1' and '2'. We have already postulated that both the solvent and solute molecules occupy equal areas on the surface. Therefore, x_i^s equals the fraction of the surface occupied by component i, θ_i. Since $\theta_1 + \theta_2 = 1$, we have

$$\theta_2 = \frac{K a_2^b}{K a_2^b + 1} \tag{67}$$

In this form the Langmuir equation shows how the fraction of surface adsorption sites occupied by solute increases as the solute activity in solution increases. From now on we drop the subscript 2 and the superscript b. Since Equation (67) is written solely in terms of the solute, these designations are redundant.

Two limiting cases are of special interest:

1. At infinite dilution $a \to 0$ and Equation (67) becomes

$$\theta = Ka \tag{68}$$

2. If $Ka \gg 1$, Equation (67) becomes

$$\theta = 1 \tag{69}$$

Equation (68) shows that θ increases linearly with an initial slope that equals K. This slope will be larger the farther to the right the equilibrium represented by Equation (61) lies. At higher concentrations, Equation (69) indicates that saturation of the surface with adsorbed solute is achieved. Figure 7.16a shows how these two limiting conditions affect the appearance of the isotherm.

Experimentally, one does not measure the fraction of sites containing adsorbed solute directly; instead, either the number of moles of solute adsorbed per unit weight of adsorbent n_2^s/w or the number of moles per unit area of adsorbent n_2^s/A is measured. These quantities are related by the equation

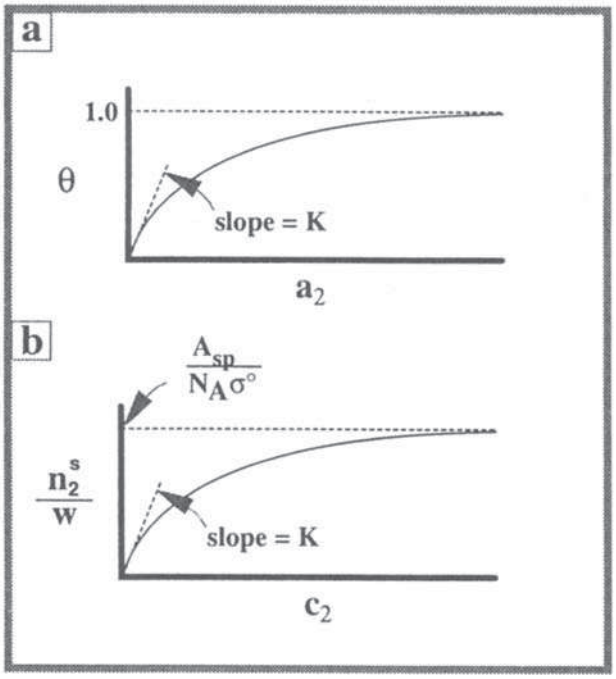


FIG. 7.16 Schematic plots of the Langmuir equation showing the significance of the initial slope and the saturation value of the ordinate: (a) the fraction covered versus solute activity; and (b) the number of moles of solute adsorbed per unit weight of adsorbent versus concentration.

$$\frac{n_2^s}{A} A_{sp} = \frac{n_2^s}{w} \tag{70}$$

where A_{sp} is the *specific area* of the adsorbent (see Chapter 1, Section 1.2). The fraction covered is related to these quantities as follows:

$$\theta = \frac{n_2^s}{A} N_A \sigma^0 = \frac{n_2^s N_A \sigma^0}{w A_{sp}} \tag{71}$$

where N_A is Avogadro's number and σ^0 is the area occupied per molecule. The level at which saturation adsorption occurs may be identified with $\theta = 1$. Therefore Equation (71) shows the saturation values of the usual ordinates to be either

$$\left(\frac{n_2^s}{w}\right)_{sat} = \frac{A_{sp}}{N_A \sigma^0} \tag{72}$$

or

$$\left(\frac{n_2^s}{A}\right)_{sat} = \frac{1}{N_A \sigma^0} \tag{73}$$

Since the entire derivation of the Langmuir isotherm assumes dilute solutions, the concentration c_2 of the solute (denoted by c, for simplicity) rather than the activity is generally used in presenting experimental results. Figure 7.16b shows how actual experimental data might appear.

7.9b The Langmuir Equation: Application

Many systems that definitely do not conform to the Langmuir assumptions—the adsorption of polymers, for example—nevertheless display experimental isotherms that resemble Figure 7.16. Although these can be fitted to Equation (67), the significance of the constants is dubious. Therefore the Langmuir equation is often written as

$$m\left(\frac{n_2^s}{w}\right) = \frac{(m/b)c}{(m/b)c + 1} \tag{74}$$

where m and m/b are regarded simply as empirical constants. A method for obtaining the numerical values for these constants from experimental data is easily seen by rearranging Equation (74) to the form

$$\frac{c}{n_2^s/w} = mc + b \tag{75}$$

This form suggests that a plot of $c/(n_2^s/w)$ versus c will be a straight line of slope m and intercept b.

If the experimental system matches the model, then the values of m and b can be assigned a physical significance by comparing Equation (74) with Equations (67) and (71):

$$m = \frac{N_A \sigma^0}{A_{sp}} \tag{76}$$

and

$$m/b = K \tag{77}$$

If the model does not apply, these constants are treated merely as empirical parameters that describe the adsorption isotherm.

When the model does apply, the experimental value of m permits A_{sp} to be evaluated if σ^0 is known, or σ^0 to be evaluated if A_{sp} is known. It is often difficult to decide what value of σ^0 best characterizes the adsorbed molecules at a solid surface. Sometimes, therefore, this method for determining A_{sp} is calibrated by measuring σ^0 for the adsorbed molecules on a solid of known area, rather than relying on some assumed model for molecular orientation and cross section.

Example 7.5 illustrates how adsorption data can be interpreted if the data conform to the Langmuir model.

* * *

EXAMPLE 7.5 *Use of Langmuir Adsorption Isotherm.* The moles of solute B adsorbed per gram of solid C were determined by measuring concentration changes in the solution. The accompanying results report the adsorption versus the concentration of the equilibrium solution:

c (mole B liter^{-1})	0.75	1.40	2.25	3.00	3.50	4.25
$n/w \cdot 10^4$ [mole B (g C)$^{-1}$]	6.00	8.00	9.57	10.0	10.4	10.8

Plot these data in the form suggested by Equation (75) and evaluate the slope and intercept. If A_{sp} for solid C is known by independent study to be 325 m^2 g^{-1}, what is σ^0 for the adsorbate? Alternatively, suppose σ^0 for the adsorbate is known to be 0.25 nm^2 on this surface. What value of A_{sp} is consistent with the adsorption data?

Solution: The ratio $c/(n/w)$ is evaluated as required to test Equation (75):

$c(n/w)^{-1} \cdot 10^{-3}$ (g C liter^{-1})	1.25	1.75	2.35	3.00	3.35	3.95

A plot of these values against the equilibrium concentrations is shown in Figure 7.17. The slope and intercept of the line drawn are 769 g C (mole B)$^{-1}$ and 0.0700 g C liter^{-1}, respectively.

Equation (76) permits the interpretation of the slope m.

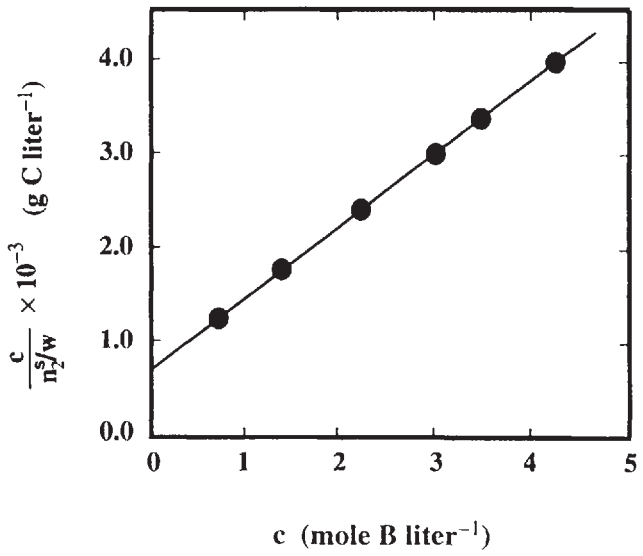


FIG. 7.17 Plot of the Langmuir equation in the form given by Equation (75) for data in Example 7.5.

If A_{sp} is known,

$$\sigma = mA_{sp}/N_A = (769)(325)(10^9)^2/6.02 \cdot 10^{23} = 0.42 \text{ nm}^2$$

and if σ^0 is known,

$$A_{sp} = N_A\sigma^0/m = (6.02 \cdot 10^{23})(0.25)(10^{-9})^2/769 = 196 \text{ m}^2\text{g}^{-1}$$

The value of K is given by the ratio m/b according to Equation (77). For this (hypothetical) system, $K = (769)/(0.0700) = 1.10 \cdot 10^4$ liter (mole B)$^{-1}$. These reciprocal concentration units are appropriate for K of Equation (66) since the activity of the bulk solvent has been absorbed into the definition of K. ■

* * *

The method of Example 7.5 applied to the adsorption of benzene, naphthalene, and anthracene on carbon black from heptane solutions gives values of 0.42, 0.67, and 0.83 nm^2, respectively, for σ^0. The progression of sizes indicates that the molecules lie flat on the surface of the carbon.

It might also be noted that K' (Equation (62)) may be related to ΔG^0 for the adsorption process if the model applies to the experimental system. Therefore, from studies of adsorption at different temperatures, values of ΔH^0 and ΔS^0 may be determined for the process described by Equation (61). It must be emphasized that compliance with the form predicted by the Langmuir isotherm is not a sensitive test of the model; therefore interpretations of this kind must be used cautiously.

In summary, adsorption from dilute solutions frequently displays the qualitative form required by the Langmuir equation. If this form is observed, it may be quantitatively described by Equation (75), in which m and b are empirical constants. Sometimes there may be a justification for further interpretation of these parameters in terms of the theoretical model.

7.9c Limitations of the Langmuir Equation

We should not be too surprised that the Langmuir equation often yields only an empirical isotherm. There are several reasons why real systems are likely to deviate from the theoretical model:

1. The adsorption process described by Equation (61) is a complex one involving several different kinds of interactions: solvent-solute, solvent-adsorbent, and solute-adsorbent.
2. Few solid surfaces are homogeneous at the molecular level.
3. Few monolayers are ideal.
4. Our interest often extends beyond the region of dilute concentrations.

We briefly comment on each of these limitations.

7.9c.1 Multiplicity of Interactions

In discussing adsorption from solution, there is nothing that can be done about the multiplicity of possible interactions, except possibly to avoid systems in which highly specific interactions are to be expected. In Chapter 9 we again discuss the Langmuir isotherm as it applies to the adsorption of gases. In that case there are considerably fewer interactions involved in the adsorption process, making it more amenable to analysis.

7.9c.2 Surface Heterogeneity

The assumption of surface homogeneity is one that was not explicitly stated in deriving the Langmuir equation. It is essential, however; otherwise a different value of K would apply to Equation (61) at various places on the surface. Attempts to deal with surface heterogeneity have been undertaken, but this enterprise seems more likely to be successful for gas adsorption rather than for adsorption from solution because the variety of interactions that must be considered is less in the former than in the latter. There is an equation—known as the *Freundlich isotherm*—that may be derived by assuming a certain distribution function for sites having different ΔG^0 values for the process represented by Equation (61) and assuming Langmuir adsorption at each type of site. The Freundlich isotherm is given by the expression

$$\theta = ac^{1/n} \tag{78}$$

in which a and n are constants with $n > 1$. This equation was in use long before the interpretation of a certain distribution of sites was assigned to it. Therefore it is best regarded as an empirical isotherm, the constants for which may be evaluated from the slope and intercept of a log-log plot of θ versus c. The Freundlich isotherm is no cure-all for surface heterogeneity: Its theoretical derivation depends on a highly specific distribution of site energies. In addition, the Langmuir equation gives adequate results in many cases in which surface heterogeneity is known to be present. Note that with $n = 1$ the Freundlich isotherm is identical to the low-concentration limit of the Langmuir isotherm (Equation (68)), and with $n = \infty$ to the high-concentration limit (Equation (69)).

7.9c.3 Multilayer Adsorption

In the Langmuir derivation the adsorbed molecules are allowed to interact with the adsorbent but not with each other: The adsorbed layer is assumed to be ideal. This necessarily limits adsorption to a monolayer. Once the surface is covered with adsorbed molecules, it has no further influence on the system. The assumption that adsorption is limited to monolayer formation was explicitly made in writing Equations (72) and (73) for the saturation value of the ordinate. It is an experimental fact, however, that adsorption frequently proceeds to an extent that exceeds the monolayer capacity of the surface for any plausible molecular orientation at the surface. That is, if monolayer coverage is postulated, the apparent area per molecule is only a small fraction of any likely projected area of the actual molecules. In this case the assumption that adsorption is limited to the monolayer fails to apply. A model based on multilayer adsorption is indicated in this situation. This is easier to handle in the case of gas adsorption, so we defer until Chapter 9 a discussion of multilayer adsorption.

7.9c.4 Adsorption from Concentrated Solutions

Next let us consider adsorption from solutions that are not infinitely dilute. Suppose, for example, that adsorption is studied over the full range of binary liquid concentrations. Figure 7.18 is an example of such results for the benzene-ethanol system adsorbed on carbon. At

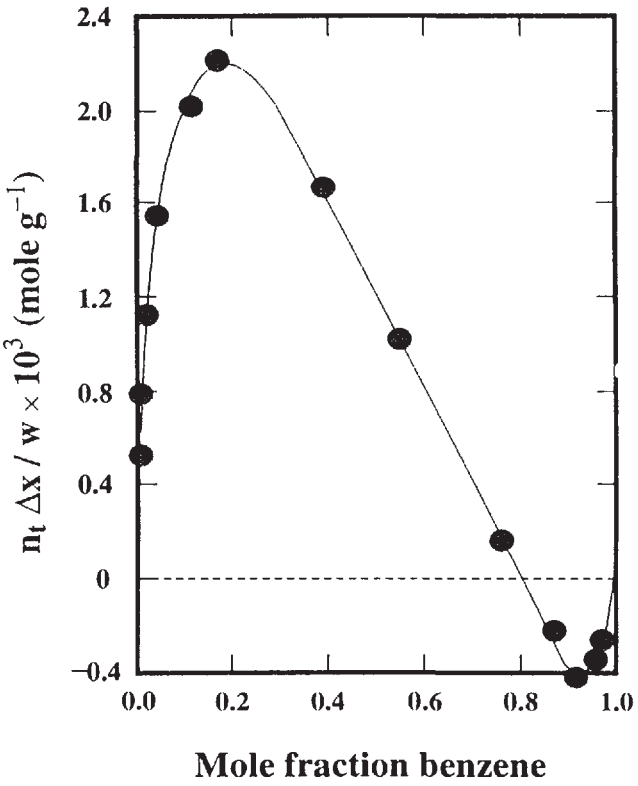


FIG. 7.18 Adsorption on carbon from the ethanol-benzene system. The ordinate equals the total number of moles of solution times the change in solution mole fraction per unit weight of carbon. (Data from F. E. Bartell and C. K. Sloan, *J. Am. Chem. Soc.*, **51**, 1643 (1929).)

first these results appear quite bewildering, displaying maximum, minimum, and negative adsorption. Recall, however, that what is actually measured is an isotherm of concentration change. The observed change in concentration is then expressed as moles of solute adsorbed. In a totally different range of solution concentrations, the solvent rather than the solute may adsorb. The associated change in the solution would then be an increase in solute concentration or the apparent negative adsorption of solute. A curve like that shown in Figure 7.18 should therefore be understood as a composite of two distinctly different isotherms. A good deal of work has been done with composite isotherms, particularly toward separating them into individual isotherms. A summary of this kind of research can be found in Kipling (1965).

7.10 APPLICATIONS OF ADSORPTION FROM SOLUTION

No discussion of adsorption from solution is anywhere near complete unless it includes some indication of its enormous practical applicability. As a matter of fact, the examples we briefly consider—detergency and flotation—encompass a wide variety of concepts from almost all areas of surface and colloid chemistry. We have chosen to stress principles rather than applications, however, so these subjects will receive an amount of attention that belies their actual importance. Following these traditional applications, two examples of new applications that are envisioned for surfactant layers deposited on solid substrates are discussed.

7.10a Detergency and Flotation: Similarities

It is impossible to do justice to the complex phenomena of *detergency* and *flotation* in a few paragraphs. All we can do is point out some of the ways in which the principles of colloid and surface chemistry apply in these areas. There are several ways in which detergency and flotation phenomena resemble one another:

1. Both terms give simple names to processes involving many different steps. The more familiar of the two, detergency, may be defined as the process by which some unwanted foreign matter is removed from a substrate by a combination of chemical treatment, temperature, and mechanical agitation. Flotation is the process by which a specific mineral component of an ore mixture is separated from other components (called "gangue") by being concentrated in the froth of an aerated slurry. Chemical additives and mechanical forces are involved here also.

2. In actual practice, both detergency and flotation deal with systems that are terribly difficult to idealize by any sort of model. In a laundering operation, for example, there will be present a variety of different fabric surfaces (cotton, polyester, etc.), different kinds of foreign matter (particulate, oily, etc.), and different chemical additives (detergent, inorganic phosphate, fluorescent whitening agents, as well as the solvent, water). In flotation, all three states of matter—solid, liquid, and gas—are involved and each of these involves several chemical components. The ore is a complex mixture of minerals (assumed to be crushed to such an extent that each particle is a different phase), the air is a mixture of gases (including chemically reactive oxygen), and the liquid contains at least three deliberately added reagents (known as regulator, collector, and frother), in addition to whatever dissolved minerals are present in the water.

3. A third point of resemblance between detergency and flotation (perhaps redundant in view of what has already been said) is that both have developed largely by empirical research with (partially satisfactory) explanations trailing far behind the actual practice.

7.10a.1 Detergency

With this much general background, let us now consider these two processes separately. In discussing detergency we must first examine the availability of the surfactant. Weak acid soaps form insoluble compounds with Ca^{2+}, for example, and are converted to insoluble molecular acids at low pH levels. One of the reasons for the addition of inorganic phosphate to laundry products is to prevent or minimize these reactions. In this discussion we assume that the impurity has not been imbibed into the interior of the fiber (soaking might help if it has) and that it is a semiliquid soiled spot rather than a solid contaminant with which we are dealing. One advantage of washing this type of soiled material at high temperatures is that the viscosity of the oily spot is lowered so that the shape of these drops is more readily altered.

The process of removing an oily drop from a solid substrate may be described in terms of the work of adhesion, given by Equation (6.57). Applying this idea to the separation of oil (O) from solid (S) gives

$$W_{adhesion} = \gamma_{OW} + \gamma_{SW} - \gamma_{OS} \tag{79}$$

where the subscript W describes the aqueous solution. For the separation to be spontaneous in the thermodynamic sense, this quantity (ΔG) must be negative. Positive surface excesses of surfactant molecules at the interface between the aqueous phase and the oil and/or the solid will lower γ for these surfaces. This change is a favorable one for the process of removing foreign matter.

In addition, the contact angle between an oil spot and a solid surface to be cleaned may be a contributing factor in detergency. For example, Figures 7.19b and 7.19d illustrate schematically two different situations for an oily drop being lifted off a substrate by currents in the adjacent phase. The contact angles θ_1 between the drop and the substrate are assumed to be the same at "lift off" (Figs. 7.19b and 7.19d) as in the quiescent state (Figs. 7.19a and 7.19c). It is evident from the figure, however, the necking of the drop for $\theta_1 < 90°$ is likely to leave a residue, whereas $\theta_1 > 90°$ would lead to a clean detachment. Young's equation (Equation (6.44)) may be applied to this situation:

$$\gamma_{OW} \cos \theta_1 = \gamma_{SW} - \gamma_{OS} \tag{80}$$

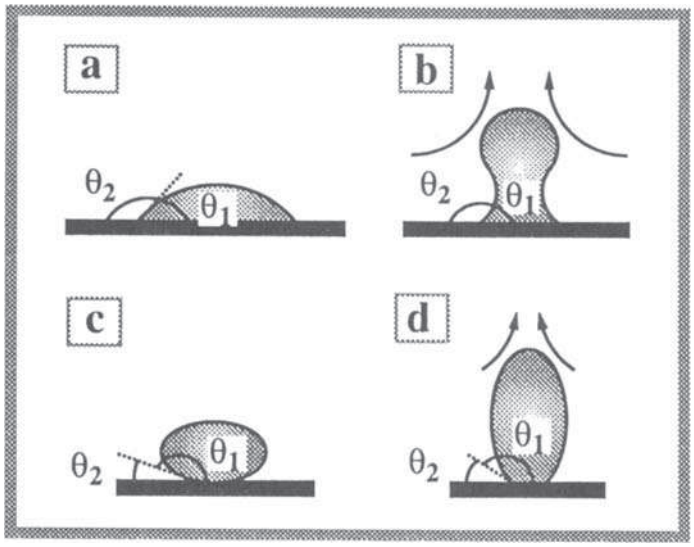


FIG. 7.19 Schematic illustration of several configurations of three phases useful in the discussion of detergency and flotation. The shaded region represents the soiled spot in detergency and θ_1 is the relevant contact angle; the shaded region is an air bubble in flotation, and θ_2 is the appropriate contact angle. The arrows in (b) and (d) indicate flow in the adjacent phase.

where θ_1 is measured in the oil drop as shown in Figure 7.19. Equation (80) shows that $\theta_1 > 90°$ and $\theta_1 < 90°$ correspond to $\gamma_{SW} < \gamma_{OS}$ and $\gamma_{SW} > \gamma_{OS}$, respectively. Any adsorption at the solid-water interface will lower γ_{SW} and therefore be conducive to a contact angle that favors the complete "rollback" of the oily spot.

Once the dirty spot is removed from the substrate being laundered, it is important that it not be redeposited. Solubilization of the detached material in micelles of surfactant has been proposed as one mechanism that contributes to preventing the redeposition of foreign matter. Any process that promotes the stability of the detached dirt particles in the dispersed form will also facilitate this. We see in Chapter 11 how electrostatic effects promote colloidal stability. The adsorption of ions—especially amphipathic surfactant ions—onto the detached matter assists in blocking redeposition by stabilizing the dispersed particles. Materials such as carboxymethylcellulose are often added to washing preparations since these molecules also adsorb on the detached dirt particles and interfere with their redeposition.

7.10a.2 Flotation

Now let us turn to a brief examination of *flotation*. Virtually all nonferrous metallic ores are concentrated by the flotation process. Sulfide ores have been studied particularly extensively, although the method has been used with oxides and carbonates as well as such nonmetallic materials as coal, graphite, sulfur, silica, and clay. Something on the order of a billion tons of ore a year are processed in this way.

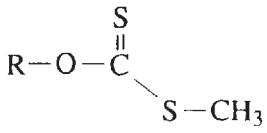
We assume that the ore has been pulverized and mixed with water so that our involvement begins with a slurry known as *pulp*. We consider, in turn, the chemical nature and the effects of each of the three broad classes of chemicals added in the flotation process.

The first class of chemical additives to be considered is the regulator, a compound that affects the adsorption of the collectors. Regulators, like catalysts, may be positive or negative in their role. For the case in which the collector adsorption is enhanced, the regulator is called an activator; when the effect is negative, it is called a depressant. Regulators are frequently compounds that control the pH and sequester metallic cations that would otherwise compete with the mineral particle surfaces for the surface-active collectors. The pH affects not only the availability of certain collectors, but also the charge of the mineral particles (see Chapter 11,

Section 11.2 for a discussion of potential-determining ions). This last consideration plays a role in determining whether the mineral particles will be dispersed as small units (easier to lift by flotation) or whether they will be aggregated. Ammonia, lime, and sources of CN^- and HS^- are commonly used as regulating agents.

Collectors are surface-active additives that adsorb onto the mineral surface and prepare the surface for attachment to an air bubble so that it will float to the surface. Therefore collectors must adsorb selectively if flotation is to result in any fractionation of the crude ore. In addition, the adsorbed collector must impart a hydrophobic character to the particle surface so that an air bubble will attach to the mineral or vice versa.

Amphipathic substances such as we have discussed throughout this chapter are used as collectors. Alkyl compounds with C_8 to C_{18} chains are widely used with carboxylate, sulfate, or amine polar heads. For sulfide minerals, sulfur-containing compounds such as mercaptans, monothiocarbonates, and dithiophosphates are used as collectors. The most important collectors for sulfides are xanthates, the general formula for which is



In the collectors used, R is generally in the C_2 to C_6 range. Xanthates are readily oxidized to dixanthogens, and the extent of this reaction may have a big effect on the efficiency of the collector.

The fundamental role of the collector is to produce a solid surface that is sufficiently hydrophobic so that it will attach to an air bubble when the pulp is aerated. Figure 7.19 may also be used to represent this situation, except that for flotation the shaded region is an air bubble. Since contact angles are measured in the liquid phase, the contact angle in the flotation case will be θ_2. For good bubble adhesion contact angles greater than 90° are preferred. Unlike the parallel situation in detergency, the adhesion of the bubble rather than its detachment is required for the success of the process.

Once again, we may use Young's equation to decide what adsorption situation is most conducive to values of $\theta_2 > 90°$

$$\gamma_{W,A} \cos \theta_2 = \gamma_{AS} - \gamma_{SW} \tag{81}$$

From this, we see the optimum condition corresponds to $\gamma_{SW} \gg \gamma_{AS}$, or extensive adsorption at the air-solid surface and minimum adsorption at the solid-water interface. The hydrophobic nature of the collectors and their chemical affinity for specific solids promote this situation.

The formation of a large bubble that facilitates flotation requires a large area of attachment or, more specifically, a large perimeter of attachment since the three-phase contact boundary occurs along the perimeter. Increasing this perimeter is favored by a positive value of the spreading coefficient (Equation (6.61)). In the notation of this problem (air spreading on solid in water), the spreading coefficient equals

$$S_{A/S} = \gamma_{SW} - (\gamma_{AW} + \gamma_{SA}) \tag{82}$$

The collector lowers γ_{SA}, an effect favorable to a positive spreading coefficient.

Finally, the frothing agents are intended to stabilize the mineral-laden foam at the surface of the aeration tank until it can be scooped off. Alkyl or aryl alcohols in the C_5 to C_{12} range are typical frothers. We have already seen how the long-chain members of this series form monolayers at the air-water surface. This lowers γ_{AW}, which is beneficial to the stability of the foam and also favors a large contact angle (if $\theta_2 > 90°$), positive spreading, and large bubbles for flotation. Neither the collector nor the frother is adsorbed exclusively at the solid-air or the water-air surface where their respective effects would be greatest. To a certain extent these two classes of additives compete with each other for adsorption sites; therefore conditions

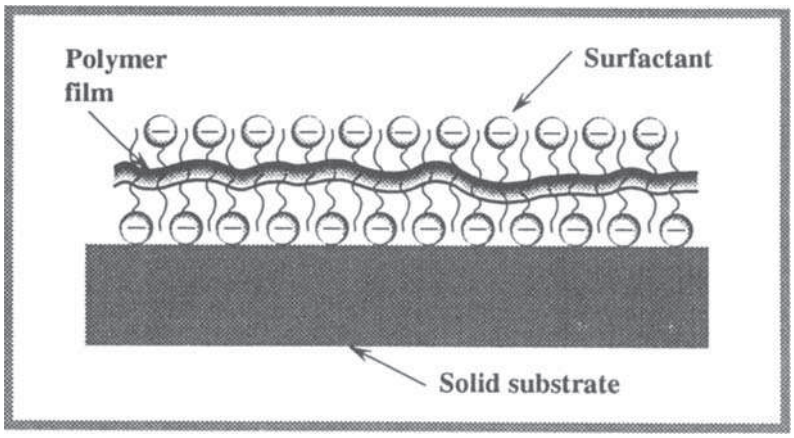


FIG. 7.20 An admicelle (a bilayer adsorbed on a solid substrate) as a two-dimensional solvent for a polymerization reaction. (Redrawn with permission of J. Wu, J. H. Harwell, and E. A. O'Rear, *J. Phys. Chem.*, **91**, 623 (1987).)

under which each produces the maximum effect are difficult to achieve, so compromise conditions in which the net effect is optimized are sought.

Another aspect of the frothers used is the fact that they form fairly condensed and therefore relatively viscous slow-draining films. In addition to thermodynamic considerations, then, kinetic factors are also important in stabilizing the froth.

A number of additional applications of the ideas of this chapter could be profitably considered if space permitted. Included among these are adhesives, lubricants, waterproofing, and the recovery of oil from the pores of rocks. Like detergency and flotation, these topics involve a variety of surface and colloid phenomena. The interested reader will find an introduction to these fields in some of the references listed at the end of this chapter, especially Adamson (1990), Davies and Rideal (1961), and Osipow (1962).

7.10b Surfactant Films as Two-Dimensional Solvents and Reactors

Surfactants adsorbed on solid substrates can also be used as "two-dimensional solvents" or "reactors" (see Fig. 7.20). The objective is to form thin polymer layers on a solid surface for potential use in devices that require ultrathin polymer films with specific properties (e.g., ultrathin photoresists and waveguides for integrated optical systems). The basic idea is to form a layer of surfactants on a suitably treated solid surface so that the hydrocarbon tails of the surfactants project out from the surface. An additional layer of surfactants can then be deposited with the hydrocarbon chains of the second layer in contact with the exposed tails from the first layer (see Fig. 7.20). The resulting film, a bilayer adsorbed on a solid substrate, is sometimes called an *admicelle* (*ad*sorbed *micelle*). This bilayer serves as a two-dimensional solvent for the chosen monomer. The monomer solution is exposed to the surfactant film and, when the appropriate amount of monomers is dissolved in the film, the polymerization reaction is initiated by chemical, thermal, or photochemical methods. The surfactant film serves to localize the polymerization reaction and allows a thin polymer film to form on the solid surface. In principle, such a procedure can be used for fabricating single and multilayer polymer coatings for use in a wide variety of applications such as corrosion protection, solid lubricants, conducting or semiconducting films, and films for controlled release of drugs, in addition to the ones mentioned above.

7.10c Langmuir-Blodgett Films

As mentioned above (Vignette VII and Section 7.6d), Langmuir-Blodgett (LB) films have received considerable attention in recent years as potentially capable of providing a number of

applications. Excellent reviews of such applications are available in the compendium on LB films by Roberts (1990). The projected applications include the use of LB films in molecular electronics, passive thin-film applications such as electron-beam microlithography and lubrication, piezoelectric films, optical devices (e.g., waveguides), semiconductor devices (e.g., metal-insulator-semiconductor [MIS] diodes), and chemical and biological sensors. Some of these applications are speculative at this time, whereas others are closer to reality. Detailed discussions of any of these will require concepts that are outside the area of colloid science, and we restrict ourselves to one brief example (Ball 1994), which will illustrate how the structures and properties of LB films are taken advantage of in devising novel applications. Specialized monographs such as MacRitchie (1990), Roberts (1990), Ulman (1991), and Tredgold (1994) contain other examples.

For instance, optical memories or switches may be fabricated using LB films such as the one illustrated in Figure 7.21. The LB film in this case is made of molecules consisting of 7,7,8,8-tetracyano-p-quinodimethane (TCNQ, for short) at one end and a long hydrocarbon chain on the other. Each hydrocarbon chain contains an azobenzene unit, which can switch between two isomeric forms when exposed to visible and ultraviolet light. The stacking of the TCNQ units is affected by the isomeric state of the azobenzene unit, and this in turn changes the conductivity of the film. One can thus use photoisomerization to change the film from conducting to semiconducting, and it might be possible to use such an LB film in electronic devices such as MIS sandwich structures.

7.11 ADSORPTION IN THE PRESENCE OF AN APPLIED POTENTIAL

We conclude this chapter with a discussion of adsorption at the interface between mercury and a solution (usually aqueous) under the influence of an applied potential. The γ value for this interface is easily measured, and potential and electrolyte concentration can be studied as variables. The Nernst equation provides a familiar reminder that potentials can be dealt with by thermodynamics.

Moving a charge of q coulombs to a surface at which the potential is ψ volts involves a quantity of work (in joules)

$$\delta w = q\psi \tag{83}$$

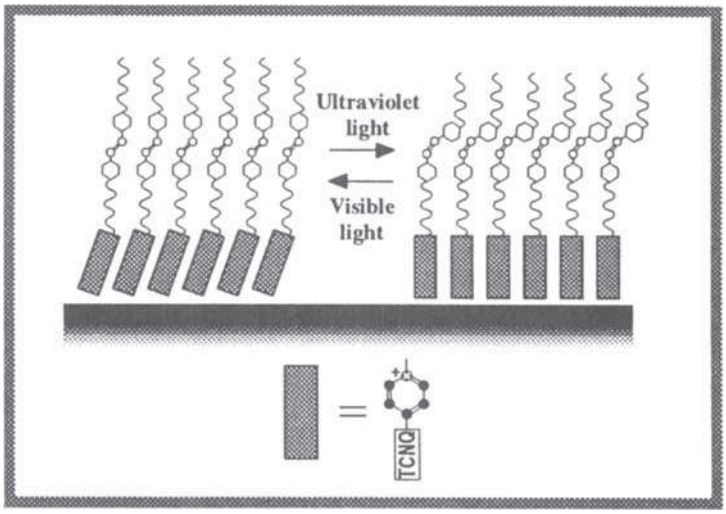


FIG. 7.21 An optical switching device based on Langmuir-Blodgett films. (Redrawn with permission of P. Ball 1994.)

Since this is non-pressure-volume work, it can be identified with a change in Gibbs free energy. In addition, the charge q carried by n moles of ions having a relative charge of $\pm z$ is given by $q = \pm nN_A ze$, where e is the proton charge, $1.60 \cdot 10^{-19}$ C. The product $N_A e$ is called the Faraday constant $\mathfrak{F}$ and equals 96,480 C. In view of these ideas, we can write

$$\Delta G = \pm nN_A ze\psi = \pm zn\mathfrak{F}\psi \tag{84}$$

where the sign depends on whether the charge is moved with or against the potential.

7.11a Electrocapillarity

Figure 7.22 illustrates an apparatus by which the relationship between interfacial tension, applied potential, and electrolyte concentration can be investigated using what is known as the *electrocapillary effect*. Since mercury has a contact angle (measured in the mercury) that is greater than 90°, the mercury-solution interface is depressed (again in reference to the mercury). If an etched mark is placed on the capillary, it is possible to bring the mercury-solution meniscus to that mark by varying the height of the mercury reservoir. It turns out that this is quite sensitive to the potential E between the electrodes. The height of the capillary depression (a negative capillary "rise") can be readily converted to the interfacial tension through Equation (6.4). We see, therefore, that γ for the Hg-solution interface depends on the electrical potential across the system. In addition, the detailed shape of the so-called electrocapillary curve—a plot of γ versus E—depends on the concentration and nature of the electrolyte present. Figures 7.23a and 7.23b show examples of typical electrocapillary curves.

Several generalizations are evident from an inspection of Figure 7.23:

1. The general shape of the curves is roughly parabolic. The coordinates of the maximum in the electrocapillary curve depend on the electrolyte content of the system. Since γ decreases on both sides of the electrocapillary maximum and since reductions in γ are associated with adsorption, we conclude that adsorption increases as we move in either direction from the maximum; that is, the electrocapillary maximum seems to be a point of minimum adsorption.

2. The left-hand branch of the electrocapillary curves (also called the rising or ascending branch) is sensitive to the chemical nature of the anion present. In Figure 7.23a, for example, different potassium salts are used as the electrolytes. Although the ascending branches of these curves differ, the descending branches (to the right of the maximum) lie on a common curve. Figure 7.23b shows that the right-hand branch of the curves is sensitive to the nature (and

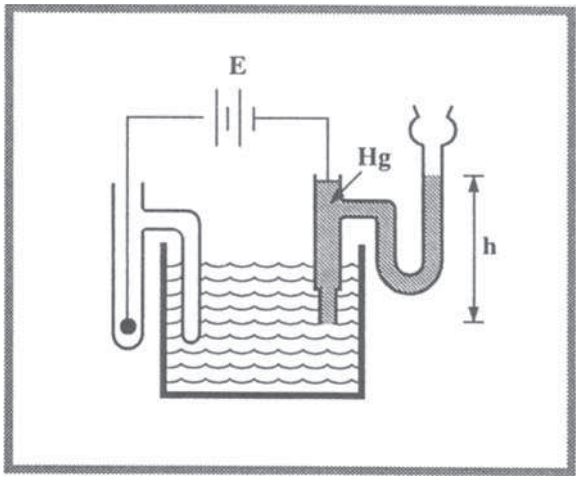


FIG. 7.22 Schematic illustration of an apparatus to measure the electrocapillary effect.

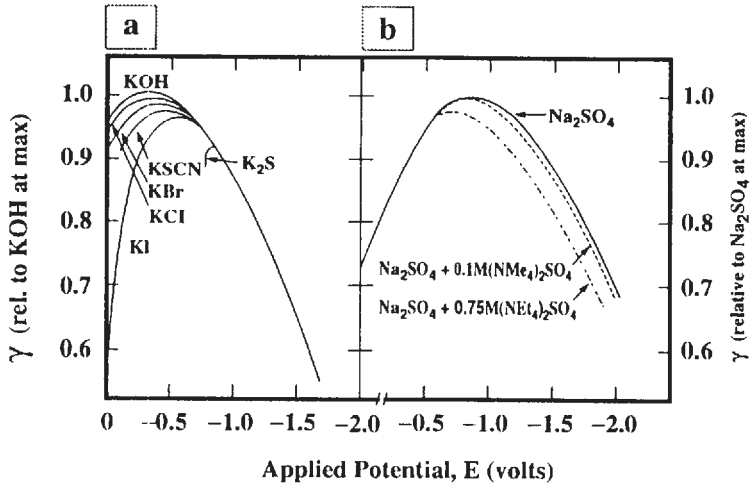


FIG. 7.23 Typical electrocapillary curves: (a) anions are adsorbed; and (b) cations are adsorbed. (Redrawn with permission of N. K. Adam, *The Physics and Chemistry of Surfaces*, Dover, New York, 1968.)

concentration) of the cation. These observations, coupled with the above remarks about adsorption, suggest that the anion is preferentially adsorbed on the ascending branch of the curve, the cation is preferentially adsorbed on the descending branch, and that neither ion is preferentially adsorbed at the maximum, although both may be adsorbed in equivalent amounts.

3. The charge carriers in the system—ions in the solution and electrons in the mercury—come to an equilibrium distribution that is consistent with the applied potential. To the left of the maximum, the mercury surface is a positively charged anode (this branch is also called the anodic branch), toward which anions are attracted. Electrons in the mercury are repelled from the surface by the anions in the water, so the potential on the mercury must become progressively more positive (relative to the maximum) for increased anion adsorption to occur. To the right of the maximum, the mercury is negative (the cathodic branch), attracts cations, and must become progressively more negative (relative to the maximum) for increased cation adsorption.

4. The potential at which the maximum occurs is different for different ions. In light of the third item, the differences in the voltage coordinates of the maxima must reflect differences in the chemical (as opposed to purely electrostatic) affinities of the ions for the interface. Nonionic solutes have also been investigated extensively, but we do not go into this aspect of the subject.

Electrocapillary phenomena have been studied for a long time; the apparatus shown in Figure 7.22 is essentially that used by G. Lippmann in 1875 in his comprehensive studies of electrocapillarity. We do not examine either the experimental or the theoretical aspects of this system in great detail; however, an interpretation of the results that is more quantitative than that just outlined qualitatively is possible with relatively little additional effort.

7.11b The Lippmann Equation

As a starting point, it is convenient to return to the kind of argument that leads to the Gibbs-Dühem equation (Equation (39)) for one-phase systems and to the Gibbs adsorption equation (Equation (44)) for systems containing an interface. In the present context, we are interested not only in the interface, but also in possible charge effects at that interface. Accordingly, it is convenient to distinguish between charged and uncharged components in the

system; we use the subscript i to identify the former and j for the latter. In this notation, Equation (44) may be written

$$-d\gamma = \sum_i \Gamma_i d\bar{\mu}_i + \sum_j \Gamma_j d\mu_j \tag{85}$$

In this expression μ_j is the usual chemical potential for the uncharged species. The quantity $\bar{\mu}_i$ is called the *electrochemical potential* and is related to μ_j as follows:

$$\bar{\mu}_i = \mu_i + z_i e \psi \tag{86}$$

The second term on the right-hand side of Equation (86) arises from Equation (84) since chemical potential is the partial molar free energy. In this expression z_i is the relative charge of the ith species and ψ is the potential of the surface.

In this discussion the mercury-solution interface is assumed to be completely polarizable; that is, there is no transfer of charge across the interface. Therefore each of the charged species (including electrons) occurs in only one of the phases. The requirement of complete polarizability means that the surface excess may be divided between that in the aqueous phase (W) and that in the mercury (Hg). In view of Equations (85) and (86), we write

$$-d\gamma = \sum_i (\Gamma_i d\mu_i)_W + \sum_i (z_i \Gamma_i e \psi)_W + \sum_i (\Gamma_i d\mu_i)_{Hg} + \sum_i (z_i \Gamma_i e \psi)_{Hg} + \sum_j \Gamma_j d\mu_j \tag{87}$$

Since the system as a whole is electrically neutral, the charge density on each of the phases must be equal and opposite; that is,

$$\sum_i (z_i e \Gamma_i)_W = -\sum_i (z_i e \Gamma_i)_{Hg} \tag{88}$$

Substituting Equation (88) into Equation (87) yields

$$-d\gamma = \sum_i (\Gamma_i d\mu_i)_W + \sum_i (\Gamma_i d\mu_i)_{Hg} + \sum_j \Gamma_j d\mu_j + \sum_i (z_i \Gamma_i e)_W (\psi_W - \psi_{Hg}) \tag{89}$$

The difference in potential between the two phases varies as dE, the externally applied potential difference. Therefore Equation (89) may also be written

$$-d\gamma = \sum_i (\Gamma_i d\mu_i)_W + \sum_i (\Gamma_i d\mu_i)_{Hg} + \sum_j \Gamma_j d\mu_j + \sum_i (z_i \Gamma_i e)_W dE \tag{90}$$

For a completely polarizable electrode, the concentration of each of the components in both solutions is constant and therefore so is the chemical potential for each. Therefore we obtain

$$-\left(\frac{\partial \gamma}{\partial E}\right)_\mu = \sum_i (z_i e \Gamma_i)_W \tag{91}$$

This result is known as the *Lippmann equation*. The charge density on the right-hand side of this equation refers exclusively to the solution phase; therefore the subscript is no longer retained.

For monovalent ions, Equation (91) is simply

$$-\left(\frac{\partial \gamma}{\partial E}\right)_\mu = e(\Gamma_+ - \Gamma_-) \tag{92}$$

7.11c Use of Electrocapillary Curves

When the surface excess of anions exceeds that of cations, $d\gamma/dE$ is positive, as is observed in Figure 7.23a. Negative values of $d\gamma/dE$ correspond to larger surface excesses of cations, as shown by Figure 7.23b. Finally, the condition $d\gamma/dE = 0$ corresponds to equal amounts of positive and negative adsorbed charge, that is, surface neutrality. Note that this is not the same as saying no ions are adsorbed. The slope of the electrocapillary curve measures the

difference between the cation and anion surface excesses. This is explored further in Example 7.6.

* * *

EXAMPLE 7.6 *Interpretation of Adsorption Using Electrocapillary Curves.* In electrocapillary curves like those shown in Figure 7.23a, the ascending branches cross for NaCl and $Ca(NO_3)_2$ solutions, with the steeper NaCl curve cutting across the flatter $Ca(NO_3)_2$ curve. Use this observation to criticize or defend the following proposition: Both solutes must show equal adsorption at the crossover point since the lowering of γ depends on the surface excess. Since both show the same value for γ, both must have the same surface excess.

Solution: According to the Lippmann equation, it is not γ per se, but rather $d\gamma/dE$ that is pertinent to these data. Even though the two interfaces show the same value of γ at the point in question, they do not have the same slopes. In fact, $(d\gamma/dE)_{NaCl} > (d\gamma/dE)_{Ca(NO_3)_2}$, which means, according to Equation (91),

$$-e[(+1)\Gamma_{Na^+} + (-1)\Gamma_{Cl^-}] > -e[(+2)\Gamma_{Ca^{2+}} + (-1)\Gamma_{NO_3^-}]$$

or

$$(\Gamma_{Cl^-} - \Gamma_{Na^+}) > (\Gamma_{NO_3^-} - 2\Gamma_{Ca^{2+}})$$

This does not mean that $\Gamma_{Cl^-} > \Gamma_{NO_3^-}$, but rather that the surface excess of the anion compared to the cation is greater for NaCl than for $Ca(NO_3)_2$. ■

* * *

Another interpretation of the electrocapillary curve is easily obtained from Equation (89). We wish to investigate the effect of changes in the concentration of the aqueous phase on the interfacial tension at constant applied potential. Several assumptions are made at this point to simplify the desired result. More comprehensive treatments of this subject may be consulted for additional details (e.g., Overbeek 1952). We assume that (a) the aqueous phase contains only 1 : 1 electrolyte, (b) the solution is sufficiently dilute to neglect activity coefficients, (c) the composition of the metallic phase (and therefore μ_i^{Hg}) is constant, (d) only the potential drop at the mercury-solution interface is affected by the composition of the solution, and (e) the Gibbs dividing surface can be located in such a way as to make the surface excess equal to zero for all uncharged components ($\Gamma_j = 0$). With these assumptions, Equation (89) becomes

$$-d\gamma = RT(\Gamma_+ + \Gamma_-)_w \, d\ln c + e(\Gamma_+ - \Gamma_-)_w \, dE \tag{93}$$

where c is the concentration of the electrolyte. If we specify constant applied potential, Equation (93) becomes

$$-\left(\frac{\partial\gamma}{\partial \ln c}\right)_E = RT(\Gamma_+ + \Gamma_-)_w \tag{94}$$

This result shows that the vertical displacements (at fixed potential) of the electrocapillary curve with changes in electrolyte concentration measure the sum of the surface excesses at the solution surface. Curves such as those in Figure 7.23b may be interpreted by this result. We have already seen that $\Gamma_+ = \Gamma_-$ at the electrocapillary maximum (where $E = E_{max}$); therefore

$$-\left(\frac{\partial\gamma}{\partial \ln c}\right)_{E_{max}} = 2RT\Gamma_{\pm,w} \tag{95}$$

A final result that can be extracted from the Lippmann equation is readily obtained by differentiating Equation (91) with respect to E at constant μ:

$$-\left(\frac{\partial^2\gamma}{\partial E^2}\right)_\mu = \left(\frac{\partial(e\Sigma z_i\Gamma_i)}{\partial E}\right)_\mu = C \tag{96}$$

The quantity ($e\Sigma_i z_i \Gamma_i$) gives the charge density of the surface, and d(charge density)$/dE$ is called the *differential capacitance* C of an interface. Although experimental values of surface capacitance provide valuable information about the distribution of charge at an interface, we shall not pursue this topic. Our interest in Equation (96) is primarily in the suggestion it offers that the distribution of charge at an interface can be modeled as a capacitor. When we discuss the ion atmosphere at a charged surface in Chapter 11, we begin by using the parallel plate capacitor as a preliminary model.

REVIEW QUESTIONS

1. Provide a simple mechanistic description of adsorption.
2. What is your conceptual picture of an interface between two liquids or a liquid and a gas? For example, in the case of a gas-liquid interface, can you imagine an infinitesimally thin surface that separates the gas from the liquid? If not, how would you define a surface that serves as the interface between the gas and the liquid?
3. What is a *surface-active agent* (or a *surfactant*)? Why is it often referred to as an *amphipathic* molecule?
4. What are *Langmuir* and *Gibbs* layers? What is a *Langmuir-Blodgett* film? How do they differ from each other?
5. Describe the formation and the possible structures of a surfactant monolayer at an air-water interface.
6. What is the meaning of "two-dimensional" phases?
7. What types of two-dimensional phases can occur in a monolayer when the surface concentration of the surfactant and the temperature are changed individually?
8. Describe a technique to visualize the structures that form in a Langmuir film.
9. Describe the *Langmuir film balance* and how it is used to measure surface tension.
10. What is the meaning of *surface pressure*? Describe one or two models of pressure-area isotherms for a monolayer.
11. List a few applications of monolayers. List a few applications in which the formation of monolayer at a fluid-fluid interface is used in industry.
12. What is a *Gibbs surface*? What are the definition and meaning of *surface excess* properties?
13. Explain why the air-water interfacial tension decreases with adsorption of a surfactant at the interface.
14. What is the *Gibbs equation*, and how is it derived?
15. Sketch the surface tension of water as a function of solute concentration in the bulk for different types of solutes such as an ionizable surfactant, a nonionic surfactant, and an inorganic electrolyte. Why is there a sharp change in such a plot in the case of ionic surfactants?
16. What is the *Langmuir adsorption isotherm*, and when is it applicable? Describe its limitations.
17. What is a *Freundlich isotherm*?
18. List a few (traditional as well as modern) applications of adsorbed layers of surfactants.
19. What is the *electrocapillary effect*? How is it used to study adsorption of solutes under an applied potential?

REFERENCES

General References (with Annotations)

Adamson, A. W., *Physical Chemistry of Surfaces*, 5th ed., Wiley, New York, 1990. (Graduate level. A more extended and somewhat more advanced treatment of adsorption at liquid interfaces in Chapter 4 and adsorption at solid-liquid interfaces in Chapter 11.)

Ball, P., *Designing the Molecular World: Chemistry at the Frontier*, Princeton University Press, Princeton, NJ, 1994. (Undergraduate level. A *Scientific-American*-style discussion of the role of chemistry in the molecular design of materials. Chapter 7 has a discussion of Langmuir and Langmuir-Blodgett layers. Recommended as an introduction.)

Davies, J. T., and Rideal, E. K., *Interfacial Phenomena*, Academic Press, New York, 1961. (Undergraduate level. A standard reference on interfacial phenomena. Somewhat outdated, but has good discussions on many of the conventional applications of adsorption of surfactants

at interfaces, such as detergency, lubrication, etc. Also contains some interesting photographs of the ripples on a pond before and after the application of hexadecanol to the surface.)

Gaines, G. L., *Insoluble Monolayers at Liquid-Gas Interface*, Wiley, New York, 1966. (Undergraduate level. A classic on the subject.)

MacRitchie, F., *Chemistry at Interfaces*, Academic Press, San Diego, CA, 1990. (Graduate and undergraduate levels. This book is devoted almost entirely to the contents of the present chapter and to surface tension and contact angle phenomena discussed in Chapter 6. A number of applications of monolayers and multilayers, including Langmuir-Blodgett films, are summarized in Chapter 11. Chapter 10 is devoted to the implications of adsorption and monolayer formation in biology and medicine.)

Tanford, C., *Ben Franklin Stilled the Waves*, Duke University Press, Durham, NC, 1989. (A nontechnical treatment. A historical perspective of Langmuir layers is intertwined nicely with a readable, popular introduction to surfactant films.)

Other References

Adam, N. K., *The Physics and Chemistry of Surfaces*, Dover, New York, 1968.

Bull, H. B., *J. Am. Chem. Soc.*, **67**, 4 (1945).

Kipling, J. J., *Adsorption from Solutions of Nonelectrolytes*, Academic, New York, 1965.

Klassen, V. I., and Mokrousov, V. A., *An Introduction to the Theory of Flotation*, Butterworths, London, 1963.

Knobler, C. M., *Adv. Chem. Phys.*, **77**, 397 (1990a).

Knobler, C. M., *Science*, **249**, 870 (1990b).

LaMer, V. K. (Ed.), *Retardation of Evaporation by Monolayers*, Academic Press, New York, 1962.

Mingotaud, A.-F., Mingotaud, C., and Patterson, L. K., *Handbook of Monolayers*, Vols. 1 and 2, Academic Press, San Diego, CA, 1993.

Osipow, L. I., *Surface Chemistry*, Van Nostrand-Reinhold, Princeton, NJ, 1962.

Overbeek, J. Th. G. In *Colloid Science*, Vol. 1 (H. R. Kruyt, Ed.), Elsevier, Amsterdam, Netherlands, 1952.

Pockels, A., *Nature*, **43**, 437 (1891).

Roberts, G. (Ed.), *Langmuir-Blodgett Films*, Plenum Press, New York, 1990.

Tredgold, R. H., *Order in Thin Organic Films*, Cambridge University Press, Cambridge, England, 1994.

Ulman, A., *Ultrathin Organic Films*, Academic Press, Boston, MA, 1991.

PROBLEMS

1. By analogy with the behavior of gases, monolayers are expected to expand to cover an entire surface. When the underlying liquid is flowing, the motion supplies a natural barrier to spreading at the edge of the monolayer. Use this concept to interpret the following bucolic scene:*

   On a calm day an observer who finds the right place on a stream or a river will see an unobtrusive yet startling phenomenon, a line on the surface of the water. The line may lie still, or it may contort itself, one way and another, in response to eddies. Very likely he will think a spider thread has fallen onto the water and try to cut it with his canoe paddle. As the disturbance caused by the cutting fades, the line reappears, mended and whole.

   Use the concepts of this chapter to discuss the existence of the bulge or line at the edge of the monolayer.

2. In some general chemistry laboratory courses the following experiment is done to evaluate Avogadro's number. A watch glass is filled to the brim with water, and then a solution of stearic acid in benzene is slowly deposited dropwise on the surface until such

*McCutchen, C. W., *Science*, **170**, 61 (1970).

time that a drop is added that "will not spread out, but will instead form a thick, lens-shaped layer."* What is the name of the film pressure at the "end point" of this experiment? What would be a reasonable estimate for σ at this pressure? Estimate the number of 0.005-ml drops of stearic acid solution (c = 0.200 g liter^{-1}) needed to form a monolayer on a watch glass with a 14-cm diameter. Outline how data such as these could be interpreted to lead to a value for N_A. Discuss some of the sources of error in this experiment.

3. Isotherms of π versus σ at both 15 and 25°C were studied for oleic acid on a substrate of pH 2.0 using a high-sensitivity film balance.† The following results were obtained:

	π (dyne cm^{-1})	
σ (Å^2 molecule^{-1})	At 15°C	At 25°C
10,000	0.028	0.037
8,000	0.035	0.040
6,000	0.044	0.052
4,000	0.055	0.071
3,000	0.058	0.086
2,000	0.076	0.095
1,000	0.075	0.095
500	0.074	0.095

Prepare a plot of π versus σ from these data. What is the apparent significance of the discontinuity in the curves? What quantity can be evaluated from the temperature variation of this discontinuity? Estimate this quantity from the available data.

4. The accompanying data give σ values corresponding to different film pressures for monolayers of various lecithins spread on 0.1 M NaCl at 22°C:‡

	σ (Å^2 molecule^{-1})				
π (dyne cm^{-1})	Dibehenoyl	Distearoyl	Dipalmitoyl	Dimyristoyl	Dicapryl
2	51.7	53.3	96.7	96.7	99.2
4	50.0	52.5	88.3	90.0	93.8
6	49.2	50.8	82.2	85.0	86.7
8	48.3	50.0	76.7	81.7	82.5
10	47.9	49.5	66.7	77.5	78.3
15	46.7	48.0	53.3	70.8	71.7
20	45.5	46.7	50.0	65.8	65.8
30	45.0	45.0	46.3	58.3	58.3
40	44.7	44.7	45.0	53.8	53.8

Plot π versus σ for these isotherms and label the apparent two-dimensional phase present for various parts of the curves. Write the structural formulas for each of the lecithins and discuss the features of the curves in terms of the structure of the molecules.

5. If gas densities can be measured as a function of pressure, the molecular weight of a gas may be calculated from the expression

*Ifft, J. B., and Roberts, J. L., Jr., *Frantz/Malm's Essentials of Chemistry in the Laboratory*, 3d ed., W. H. Freeman, San Francisco, CA, 1975.

†Pagano, R. E., and Gershfeld, N. L., *J. Colloid Interface Sci.*, **41**, 311 (1972).

‡Phillips, M. C., and Chapman, D., *Biochim. Biophys. Acta*, **163**, 301 (1968).

$$M = RT\left(\frac{d}{p}\right)_{\lim p \to 0}$$

Likewise, if surface concentrations (in weight per area) are measured as a function of π, the molecular weight of a solute that forms an insoluble monolayer may be determined. Romeo and Rosano* obtained the following data for a monolayer of acetyl lipopolysaccharide on 0.2 M NaCl at 20° C:

c (mg m^{-2})	0.06	0.09	0.11	0.14	0.17	0.23
π (millidyne cm^{-1})	10.3	16.4	20.4	25.9	34.3	50.0

What is the molecular weight of the acetyl lipopolysaccharide?

6. Cockbain† measured the interfacial tension of the water–decane surface at various concentrations of sodium dodecyl sulfate (NaDS). The experiments were done at 20°C both in the presence and absence of NaCl. Use the suitable form of the Gibbs equation in each case to calculate Γ_R^1 and σ at γ values of 10 and 20 dyne cm^{-1} from the following data:

Pure H_2O		0.1 M NaCl (swamping)	
c_{NaDS} (mole liter^{-1})	γ (dyne cm^{-1})	c_{NaDS} (mole liter^{-1})	γ (dyne cm^{-1})
0.0079	8.5	0.0014	5.2
0.00694	10.8	0.000694	11.7
0.00521	15.3	0.000347	17.4
0.00347	20.8	0.000173	22.7
0.001735	28.3	0.0000867	27.5

Is the variation of σ with interfacial film pressure qualitatively consistent with the expected behavior of monolayers in general? Of charged monolayers in particular?

7. Blank‡ has reported the permeability (in cm^3 of gas s^{-1} cm^{-2} surface) of various spread monolayers to water vapor at 25°C. For several different RX-type compounds at different π values, the permeabilities are as follows:

R	X	π (dyne cm^{-1})	Permeability × 10^3 (cm^3 s^{-1} cm^{-2})
C_{16}	OH	44	380
C_{18}	OH	44	300
C_{17}	COOH	24	430

Discuss the observed differences in permeability between (a) the two alcohols at the same film pressure and (b) the two 18-carbon surfactants at different pressures. In your comments include comparisons of the molecular structure of the surfactants and the efficiencies of these monolayers in retarding evaporation.

8. The pendant drop method has been used§ to measure the interfacial tension at the surface between mercury and cyclohexane solutions of stearic acid at 30 and 50°C.

*Romeo, D., and Rosano, H. L., *J. Colloid Interface Sci.*, **33**, 84 (1970).

†Cockbain, E. G., *Trans. Faraday Soc.*, **50**, 874 (1954).

‡Blank, M. In *Retardation of Evaporation by Monolayers* (V. K. LaMer, Ed.), Academic Press, New York, 1962.

§Ambwani, D. S., Jao, R. A., and Fort, T., Jr., *J. Colloid Interface Sci.*, **42**, 8 (1973).

Interpret the accompanying data by means of the Gibbs equation to evaluate Γ^1 and σ for stearic acid at these two temperatures when the equilibrium bulk concentrations are 10^{-3}, 10^{-4}, and 10^{-5} M:

T = 30°C		T = 50°C	
γ (dyne cm^{-1})	c_{Eq} (mole liter^{-1})	γ (dyne cm^{-1})	c_{Eq} (mole liter^{-1})
362	4.8×10^{-6}	364	4.8×10^{-6}
355	8.5×10^{-6}	362	9.6×10^{-6}
334	6.6×10^{-5}	354	7.4×10^{-5}
307	2.7×10^{-4}	334	4.4×10^{-4}
286	1.0×10^{-3}	314	1.6×10^{-3}

Estimate the concentrations at which the stearic acid film at the interface reaches a condensed packing at the two temperatures.

9. A scintillation counter is used to measure tritium β particles adjacent to the surfaces of tritiated sodium dodecyl sulfate in 0.115 M aqueous NaCl solution and tritiated dodecanol in dodecanol. The former system is surface active and the latter is not, so the difference between the measured radioactivity above the two indicates the surface excess of sodium dodecyl sulfate. The number of counts per minute arising from the surface excess A_s is related to the surface excess in moles per square centimeter Γ^1 by the relationship $A_s = 4.7 \times 10^{12}\ \Gamma^1$.* Use the following data (25°C) to construct the adsorption isotherm for sodium dodecyl sulfate on 0.115 M NaCl:

	Activity $\times 10^{-3}$ (cpm)		Surfactant concentration $\times 10^3$ (mole kg^{-1})
^{3}H μC (g solution)$^{-1}$	Tritiated nonsurfactant	Tritiated surfactant	
2	–	1.9	0.17
4	–	2.1	0.34
6	–	2.3	0.50
8	0.50	2.5	0.67
10	–	2.6	0.84
15	0.95	2.9	1.20
20	–	3.2	1.65
30	1.85	3.8	2.45

Briefly outline how the proportionality constant between A_s and Γ^1 might be determined experimentally.

10. A quantity called the HLB (for hydrophile-lipophile balance) number has proved to be a useful way to match a surfactant to a particular application. For example, surfactants with HLB numbers in the range 4 to 6 produce water-in-oil emulsions; those in the range 7 to 9 are useful as wetting agents; those ranging between 8 and 18 produce oil-in-water emulsions; and those with values in the range 13 to 15 make good detergents. Use these considerations plus the following specific examples to formulate a generalization about the dependence of the HLB number on the molecular structure of the surfactant:†

*Muramatsu, M., Tajima, K., Iwahashi, M., and Nukina, K., *J. Colloid Interface Sci.*, **43**, 499 (1973).

†Osipow, L. I., *Surface Chemistry*, Van Nostrand-Reinhold, Princeton, NJ, 1962.

Surfactant	HLB number
Sodium dodecyl sulfate	40
Sodium oleate	18
Tween 80 (n = 20, R = oleate)	15
Sorbitan monolaurate	8.6
Span 60	4.7
Sorbitan tristearate	2.1

11. The adsorption of straight-chain fatty acids from n-heptane on Fe_2O_3 has been studied.* In all cases studied the adsorption isotherms conform with the Langmuir equation. The following are values of the amount adsorbed at saturation:

Fatty acid	$\left(\frac{n_2^s}{w}\right)_{sat} \times 10^5$ (mole g^{-1})
Acetic	3.00
Propionic	2.36
n-Butyric	2.11
n-Hexanoic	1.78
n-Heptanoic	1.40
n-Octanoic	1.30
Lauric	1.04
Myristic	0.97
Palmitic	0.91
Magaric	0.82
Stearic	0.81

Calculate the area per molecule (σ) of each on the saturated surface if the Fe_2O_3 is known to have a specific surface of 3.45 $m^2\ g^{-1}$. Do the adsorbed molecules appear to be in the same two-dimensional phase in each of these systems?

12. The adsorption of various aliphatic alcohols from benzene solutions onto silicic acid surfaces has been studied.† The experimental isotherms have an appearance consistent with the Langmuir isotherm. Both the initial slopes of an n/w versus c plot and the saturation value of n/w decrease in the order methanol > ethanol > propanol > butanol. Discuss this order in terms of the molecular structure of the alcohols and the physical significence of the initial slope and the saturation intercept. Which of these two quantities would you expect to be most sensitive to the structure of the adsorbed alcohol molecules? Explain.

13. The Michaelis-Menton equation is an important biochemical rate law. It relates the rate of the reaction v to a substrate concentration [S] in terms of two constants v_{max} and K_M:

$$v = \frac{v_{max}[S]}{K_M + [S]}$$

It will be noted that this equation follows the same functional form as the Langmuir equation, specifically $v \rightarrow v_{max}$ as [S] increases. The biochemical literature contains three different graphical procedures to evaluate the constants v_{max} and K_M from kinetic data:

*Allen, T., and Patel, R. M., *J. Colloid Interface Sci.*, **35**, 647 (1971).

†Hoffman, R. L., McConnell, D. G., List, G. R., Evans, C. D., *Science*, **157**, 550 (1967).

Name of method	Plotted on ordinate	Plotted on abscissa
Lineweaver-Burk	$1/v$	$1/[S]$
Hanes	$[S]/v$	$[S]$
Eadie	v	$v/[S]$

Describe how the three equivalent variations of the Langmuir equation would be plotted. Give the interpretation of the slope and intercept in each case.

14. In laboratory tests of flotation, fluorite (CaF_2) particles (range of particle diameters, 0.074–0.147 mm) with oleic acid as collector were aerated under three different conditions. These conditions and the percent CaF_2 recovery after 10 min are as follows:*

Aeration conditions	Percent recovery after 10 min of aeration
Bubbles precipitated from solution	5
Bubbles produced by mechanical means	20
Combination of both means of bubble production	70

Suggest an interpretation for this variation in the efficiency of fluorite recovery.

15. Grahame† gives the following data for γ_{max} versus log c (corrected for activity) for the interface between mercury and aqueous KI at 18°C:

γ_{max} (dyne cm^{-1})	390	398	407	414	419	422
log c (c in mole liter^{-1})	0.5	0	−0.5	−1.0	−1.5	−2.0

Use these results to estimate Γ, the surface excess of KI at the electrocapillary maximum, for 1.0 and 0.1 M KI. Express your results as moles of KI adsorbed per square centimeter and as total charge q_Γ per square meter.

16. Show by the double integration of Equation (96), that is $-(\partial^2\gamma/\partial E^2)_\mu = C$, that a parabolic relationship between $\gamma_{max} - \gamma$ and $E - E_{max}$ is expected if the capacitance of the double layer is constant. Use the equation you derive to test the assumed constancy of C for the interface between mercury and 1 M NaCl from the following data:

γ (dyne cm^{-1})	340	376	396	410	418	423[a]	421
E (V)	−0.02	−0.08	−0.18	−0.28	−0.38	−0.56[a]	−0.68
γ (dyne cm^{-1})	415	405	396	384	373	358	340
E(V)	−0.78	−0.88	−0.98	−1.08	−1.18	−1.28	−1.38

[a]Maximum.

Double-layer capacitance values are usually expressed as microfarads per square centimeter; remember that practical electrical units, including the farad, are consistent with SI units. Comment on these results in terms of anion and cation adsorption.

*Klassen, V. I., and Mokrousov, V. A., *An Introduction to the Theory of Flotation*, Butterworth, London, England, 1963.

†Grahame, D. C., *Chem. Rev.*, **41**, 441 (1947).

8

Colloidal Structures in Surfactant Solutions

Association Colloids

I for my part have never known an Irregular who was not also what Nature evidently intended him to be— . . . up to the limits of his power, a perpetrator of all manner of mischief.

From Abbott's *Flatland*

8.1 INTRODUCTION

8.1a What Is Self-assembly and What Are Association Colloids?

In the last chapter we examined the tendency of surfactant molecules in aqueous solutions to adsorb at a surface in the form of a monolayer. In this chapter we continue to study surfactant solutions, this time considering a few of the many possible modes of organization they can adopt within a *bulk* phase. This process of organization is thermodynamically driven and is spontaneous, as in the case of Langmuir layers we discussed in Chapter 7. It is therefore often called *self-assembly*, and the resulting aggregates are known as *association colloids*. Both chapters share an interest in amphipathic solutes: Chapter 7 focuses on surface activity, while this chapter is concerned with structures in the colloidal size range formed by these molecules. In both, the head-to-head/tail-to-tail ordering of the amphipathic molecules is observed. The difference is that in Chapter 7 the ordering occurs at the surface, under the influence of the surface, while it occurs in the bulk in this chapter. Quite an assortment of bulk surfactant structures is known, including ordinary and reverse micelles, liquid crystals, bilayers, vesicles, and microemulsions. This chapter is concerned primarily with micelles (Sections 8.2–8.8) and microemulsions (Sections 8.9 and 8.10), although some additional structures are mentioned in passing.

The colloidal structures we examine in this chapter are formed as a result of physical interactions among amphipathic molecules, rather than by covalent bonding. This sort of physical association has been recognized for a long time, although contemporary students may be relatively (or totally!) unaware of it. It is interesting to note that in the early days of polymer chemistry, macromolecules were believed to be physically associated rather than covalently bonded structures. The birth of modern polymer chemistry can be traced to the acceptance of the covalent character of these substances. Associated structures do exist, however, and we see by the end of this chapter that their investigation is a very lively area of chemical research.

8.1b Why Are Association Colloids Important?

In the last couple of decades one important insight has triggered a tremendous upsurge of interest in surfactant structures. This is the recognition that these structures may mimic biological structures in some ways. Enzymes, for example, are protein molecules into which a

reactant molecule somehow "fits" to form a reactive intermediate. The highly efficient and specific catalytic effect of enzymes makes their investigation one of the most fruitful areas of biochemical research. Likewise, cell membranes not only compartmentalize biological systems, but also play a variety of functions in the life of the cell. Surfactant structures can be used as model systems to mimic both enzymes and membranes. A whole new field of mimetic (exhibiting mimicry) chemistry has grown around this concept, and colloidal structures formed by surfactants are at the center of the entire subject. Surfactant aggregates known as liposomes are common in physiological systems, and the use of specially designed liposomes as drug-delivery vehicles was discussed in Chapter 1 in Vignette I.3. Self-assembled structures such as micelles and reverse micelles also play an increasingly important role in separation processes in engineering and environmental science and technology; one such example is described in Vignette VIII.

VIGNETTE VIII SEPARATION PROCESSES BASED ON MICELLAR ENCAPSULATION: Micellar-Enhanced Ultrafiltration

Removal of dissolved organics and polyvalent cations from contaminated groundwater, industrial wastewater, and oil-field-produced water in a cost-effective manner is one of the most important engineering challenges of this decade. Hazardous waste sites contain groundwaters polluted by leaking underground organic storage tanks, and the leaching of the hazardous materials and accidental waste releases contribute further to the water pollution problems. The use of water in chemical industries in separation operations is common, and the water thus used picks up at least small amounts of dissolved organics and other hazardous chemicals.

The current technologies popular for removing such wastes include adsorption on activated carbon, ion-exchange membranes, and bioremediation. Activated carbon adsorption is effective but expensive since regeneration of the carbon is often difficult and expensive. Ion-exchange membranes are restricted to specific ions. An alternative technology that is currently investigated relies on the ability of surfactants to self-assemble in water as *micelles* and trap hydrocarbon contaminants. We have already seen an example of encapsulation of chemical species by surfactant aggregates known as liposomes in Chapter 1 and have seen the use of surfactants in detergency in Chapter 7. Micellar encapsulation works on similar principles and may turn out to be simultaneously effective for both dissolved organics and polyvalent ions since an ionic micelle with an appropriate head group can adsorb ionic contaminants.

"Biologically friendly" ionic surfactants can be added to the wastewater at concentrations above the threshold value beyond which the surfactants self-assemble to form micelles. The resulting micelles can trap the hydrocarbon wastes since the hydrocarbon solutes prefer the hydrocarbon interior of the micelle over the aqueous environment outside. In addition, ionic wastes in the water adsorb to the polar heads of the surfactants (see Fig. 8.1). The resulting waste-laden micelles can then be removed more easily using ultrafiltration methods. Such a process, known as *micellar-enhanced ultrafiltration* (MEUF), can be made continuous, scalable, cost effective, and environmentally friendly (through the use of biodegradable surfactants).

Some of the topics we discuss in this chapter are essential for understanding processes such as MEUF. The same ideas can also be used for other separation processes (e.g., protein separation in reverse micelles) and in genetic engineering, as mentioned in Vignette I.3 in Chapter 1. We also see in this chapter other applications such as using micelles as "microreactors," i.e., using the unique environment inside micelles for catalysis and material synthesis.

8.1c Focus of This Chapter

As mentioned above, the bulk of this chapter is concerned with the simplest of the surfactant aggregates that form in solution, namely, micelles. There is plenty to focus on in this respect, and the material is already very complex.

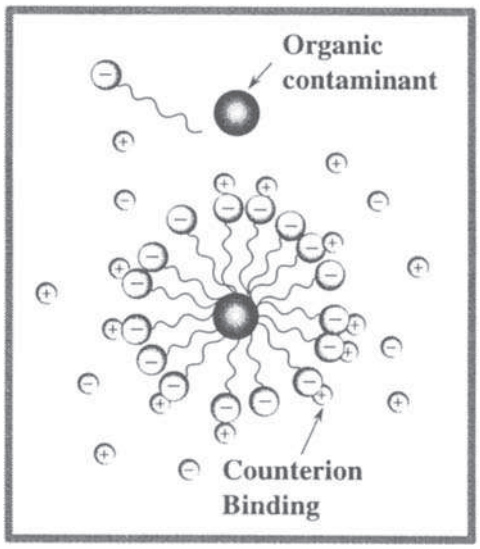


FIG. 8.1 Use of micelles to trap hydrocarbon and ionic wastes.

1. In Section 8.2 we are concerned with the threshold concentration of surfactants at which micellization occurs. This concentration, known as the *critical micelle concentration* (CMC for short), is one of the most important properties of surfactant solutions. We look at two different ways of modeling micellization and discuss briefly when they are appropriate.
2. In Sections 8.3 and 8.4, we examine the structure of the different parts of a micelle and illustrate how the shape of the micelle is determined by the relative sizes of the head group and the tail of the surfactant. This relative measure provides a geometric rationale for the structures of aggregates that are possible, and is simple to use.
3. Following this, the thermodynamic arguments needed for determining CMC are discussed (Section 8.5). Here, we describe two approaches, namely, the *mass action model* (based on treating micellization as a "chemical reaction") and the *phase equilibrium model* (which treats micellization as a phase separation phenomenon). The entropy change due to micellization and the concept of *hydrophobic effect* are also described, along with the definition of thermodynamic standard states.
4. The next two sections focus on solubilization and catalysis in micelles, two important applications of micelles.
5. An overview of other forms of micellar systems follows in the next three sections. Formation of reverse micelles, in nonaqueous media, is discussed briefly in Section 8.8. Sections 8.9 and 8.10 present an introduction to microemulsions (oil, or water, droplets stabilized in water or oil, respectively) and their applications.
6. Finally, a discussion of surfactant self-assembly will not be complete without a mention of surfactant assemblies in biological systems. Although they are outside the scope of our book, we have already drawn attention to such biological applications of colloid science in Chapters 1 and 7 and above in this chapter. Some additional discussion is provided in the last section of this chapter (Section 8.11).

8.2 SURFACTANTS IN SOLUTION: EXPERIMENTAL OBSERVATIONS AND MODELS

8.2a Self-Assembly into Micelles: Representation as Reactions

Curve 3 in Figure 7.14 was presented as a typical illustration of the way surface tension γ varies with concentration for an amphipathic solute in aqueous solution. In discussing that

figure, we noted that the break in the curve marks the concentration above which these solute molecules aggregate to form clusters known as *micelles*. If we represent the amphipathic species by S, then this clustering process can be described by the reaction

$$nS \leftrightarrow S_n \qquad \text{(A)}$$

in which S_n is the micelle with a *degree of aggregation* n. Reaction (A) is the first of a series of reactions of increasing complexity we use to represent the micellization process. Note that Reaction (A) is reversible: Dilution shifts the equilibrium toward the monomeric surfactant. The surfactant species that undergo micellization are essentially the same amphipathic molecules with adsorption behavior that we discussed in the last chapter. They have the general formula RX, in which R is a hydrocarbon chain and X is a polar group. The hydrocarbon chains in R are ordinarily C_8 or greater, may be saturated or unsaturated, may be linear or branched, and may contain an aromatic ring.

The polar X group in the amphipathic molecule may be nonionic or ionic. The nonionic X group most widely used in the study of surfactant structures is relatively short-chain polyoxyethylene moieties, $-(C_2H_4O)_x-$, in which x ranges from 3 or 4 to 20 or more. Several commercial surfactants of this sort (Tween and Triton) are listed in Table 7.2. The polyoxyethylene "heads" of these molecules are essentially short-chain polymers themselves and, as such, are generally polydisperse. The x values that characterize these preparations are average values, and a distribution of chain lengths around the average is typical of commercial nonionic surfactants.

Among numerous possible ionic groups, sulfate ($-SO_4^-$), sulfonate ($-SO_3^-$), and carboxylate ($-CO_2^-$) are the most common anionic groups; various quaternary ammonium groups ($-NR_3^+$) are the most common X substituents in cationic surfactants. Zwitterionic surfactants, which combine both positively and negatively charged groups, have also been investigated. With ionic surfactants, the amphipathic ion is accompanied by a counterion. Although other possibilities exist, we consider only univalent counterions. In general, we write ionic surfactants as M^+S^- and assume these are 100% dissociated into M^+ and S^-. A common surfactant of this type is sodium dodecyl sulfate, often written SDS, in which $M^+ = Na^+$ and $S^- = C_{12}H_{25}SO_4^-$.

The clustering of low molecular weight anionic surfactant molecules to form micelles can also be represented by the following reaction:

$$nS^- + mM^+ \leftrightarrow (S_n^- M_m^+)^{z-} \qquad \text{(B)}$$

in which n is the degree of aggregation and the net charge z of the micelle is given by $z = n - m$. Note that the charge of the micelle can also be expressed as the fraction ionized, $\alpha = (n - m)/n$. This formalism is readily extended to micelles formed from cationic surfactants; for nonionics, m and z are zero.

Reaction (B) is clearly an extension of Reaction (A), with the former admitting the possibility of counterion binding to the micelle. Additional refinements can be introduced into this reaction. The micelle still carries a net charge of $-z$, which means that zM^+ ions must be present in solution to assure electroneutrality. This may be included in the representation of micellization by writing

$$nS^- + mM^+ \leftrightarrow (S_n^- M_m^+)^{z-} + zM^+ \qquad \text{(C)}$$

Unless noted otherwise, we use anionic micelles as prototypes; that is, we use Reaction (B) to represent the formation of micelles.

8.2b Critical Micelle Concentration

The threshold concentration at which micellization begins is known as the *critical micelle concentration*. Surface tension is by no means the only property of the solution that displays a discontinuity in slope when plotted against increasing concentration. Figure 8.2 shows schematically how several different experimental quantities vary with concentration. The figure idealizes things somewhat since the various experimental quantities tend to "see" slightly different thresholds for micellization. However, the differences in CMC values by various methods of determination are not so diverse as to invalidate the notion of a "critical concentration." Below the breaks in the curves in the figure, anionic surfactants behave as expected for strong electrolytes; above the CMC, however, the behavior observed is consistent with the presence of particles in the colloidal size range. That this is true for the curves shown in Figure 8.2 is seen by the following arguments:

1. In the case of variation of osmotic pressure π with concentration c, to a first approximation we know that π/c is proportional to $1/M$ (Equation (3.34)). This is roughly equivalent to saying that the slope of a plot of π versus c is proportional to $1/M$, where M is the molecular weight. (Osmotic pressure measurements generally give the number-average molecular weight $\overline{M}_n$ as illustrated in Chapter 3, Section 3.3c. However, as also noted there, in the case of micellar solutions, osmometry can lead to *weight average* molecular weight $\overline{M}_w$; see Puvvada and Blankschtein 1989.) The decrease in the slope of the osmotic pressure plot at the CMC indicates an increase in the average molecular weight of the solute at this point.
2. In the case of turbidity τ, to a first approximation (Hc/τ) is proportional to $1/M$ (Equation (5.39)). This means that the slope of a plot of τ versus c is roughly proportional to M. The break in the curve again corresponds to an increase in the molecular weight of the solute.
3. The conductivity κ of the solution decreases at the CMC owing to the lower mobility of the larger micelles. Dividing by concentration to convert to equivalent conductivity Λ leads to a sharp reduction in the last quantity at the CMC.

These data can be used quantitatively as well as qualitatively to determine the molecular weight of the micelle. For example, the colloidal particles may be analyzed by light scattering by modifying Equation (5.39) as follows:

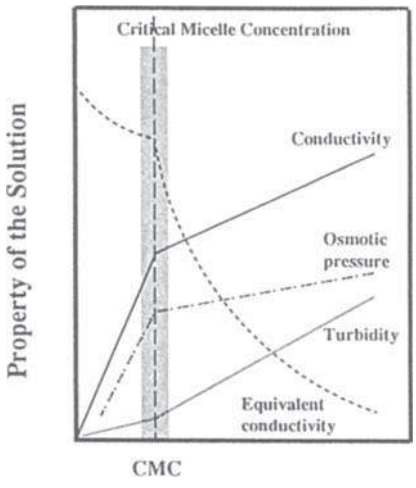


FIG. 8.2 Schematic illustration of variation of properties of a surfactant solution with surfactant concentration. Note the change in behavior at the critical micelle concentration (CMC). The shading emphasizes the fact that the CMC is not necessarily sharply defined.

$$\frac{H(c - c_{CMC})}{\tau - \tau_{CMC}} = \frac{1}{M} + 2B(c - c_{CMC}) \tag{1}$$

since the micelles form in a solution that already has a turbidity τ_{CMC} at the CMC. Turbidity data can then be plotted as in Figure 5.7 to give the molecular weight and the second virial coefficient of the micelles. The second virial coefficient can be analyzed further to give the net charge of the micelle if the amphipathic species are ionic as assumed. Most of the techniques we considered in Chapters 2–5, along with numerous other experimental procedures, have been used to characterize micelles. Of particular interest to us at this point are the average molecular weight $\overline{M}$ and the net charge z. Also of interest are the particle shape and hydration (available from sedimentation and diffusion measurements) and polydispersity (available by combining determinations of average $\overline{M}_n$ and $\overline{M}_w$). For nonionic micelles the degree of association is given by the ratios of the molecular weights of the micelle and of the individual surfactant molecules. For ionic micelles the measured molecular weight includes the bound counterions, but n can be determined if z is known.

Table 8.1 lists experimental CMC values for several ionic surfactants in water and aqueous NaCl solutions, as well as n and α values measured at the CMC. The data in Table 8.1, supplemented by numerous additional studies, allow us to make several generalizations about the state of aggregation and charge of micelles at the CMC:

1. All other things being equal, the aggregation number n increases as the length of the hydrocarbon chain increases.
2. The aggregation number for sodium dodecyl sulfate increases with increasing amounts of the indifferent electrolyte NaCl.
3. There is a general tendency for α to decrease owing to the addition of electrolyte. More extensive data show that α is generally in the range 0.1 to 0.4; a value of about 0.25 is a good choice for an "average" fraction of ionization.

TABLE 8.1 Critical Micelle Concentration, Degree of Aggregation, and Effective Fractional Ionization for Several Surfactants With and Without Added Salt

Surfactant	Solution	Critical micelle concentration (mole liter^{-1})	Aggregation number n	Ratio of charge to aggregation number, z/n
Sodium dodecyl sulfate	Water	0.00810	80	0.18
	0.02 M NaCl	0.00382	94	0.14
	0.03 M NaCl	0.00309	100	0.13
	0.10 M NaCl	0.00139	112	0.12
	0.20 M NaCl	0.00083	118	0.14
	0.40 M NaCl	0.00052	126	0.13
Dodecylamine hydrochloride	Water	0.01310	56	0.14
	0.0157 M NaCl	0.01040	93	0.13
	0.0237 M NaCl	0.00925	101	0.12
	0.0460 M NaCl	0.00723	142	0.09
Decyl trimethyl ammonium bromide	Water	0.06800	36	0.25
	0.013 M NaCl	0.06340	38	0.26
Dodecyl trimethyl ammonium bromide	Water	0.01530	50	0.21
	0.013 M NaCl	0.01070	56	0.17
Tetradecyl trimethyl ammonium bromide	Water	0.00302	75	0.14
	0.013 M NaCl	0.00180	96	0.13

Source: J. N. Phillips, *Trans. Faraday Soc.*, **51**, 561 (1955).

4. As a further refinement of the previous item, additional studies suggest that the order of ion binding to negative micelles is $Cs^+ > Rb^+ > Na^+ > Li^+$ and for positive micelles $I^- > Br^- > Cl^-$; that is, larger ions that are more polarizable and that tend to be less hydrated bind more effectively.
5. For nonionics, increasing the polyoxyethylene chain length for a constant R group decreases n.
6. Those factors that increase n tend to lower the CMC.
7. An assortment of additional evidence suggests that, at the CMC, many micelles are roughly spherical and relatively monodisperse.

The explanations for some of these may be seen from the relation between surfactant packing and the resulting shapes of the micelles described in Section 8.4. Here we merely note that the variables in terms of which micelles have been examined are by no means exhausted by this list. We encounter some additional experimental observations as we proceed. At this point, however, these results indicate that the tendency toward aggregation is increased by those variations that increase the hydrophobic character of the surfactant: increasing the hydrocarbon chain length, decreasing the polyoxyethylene chain length, or increasing counterion binding.

8.2c Micellization: Chemical Reaction or Phase Equilibrium?

The last point we consider in this section is the question of whether micellization should be viewed in terms of chemical reaction equilibrium or phase equilibrium. If we think of micellization as a chemical reaction, then Reaction (A) should surely be written as a sequence of stepwise additions:

$$2S \overset{s}{\leftrightarrow} S_2 \leftrightarrow S_3 \leftrightarrow S_4 \leftrightarrow \cdots \leftrightarrow S_n \qquad \text{(D)}$$

The law of mass action implies that a continuous distribution of species—monomers, dimers, trimers, etc.—should be present. Increasing the surfactant concentration should shift the equilibria in Reaction (D) to the right, but the observed transition from monomers to n-mers with $n > 50$ is not expected. Experimentally, large micelles do form sharply at the CMC in aqueous solutions, and—to a good approximation—the concentration of monomeric surfactant in solution varies little above the CMC. *This type of behavior is typical of phase equilibria*; examples include the (constant) concentration of a saturated solution and the (constant) vapor pressure of a liquid. Energetically, some minimum value of n is apparently necessary before the exclusion of hydrophobic tails from the aqueous medium is effective. Once the solution is concentrated enough for aggregates with this critical n value to be formed, any additional surfactant added to the solution goes into the micelle.

We see in Section 8.8 that surfactants undergo aggregation in nonaqueous solvents also, but the degree of aggregation is very much less ($n < 10$), and the threshold for aggregation is far less sharp than in water. The mass action model for micellization seems preferable for nonaqueous systems.

In summary, whether a reaction equilibrium or a phase equilibrium approach is adopted depends on the size of the micelles formed. In aqueous systems the phase equilibrium model is generally used. In Section 8.5 we see that thermodynamic analyses based on either model merge as n increases. Since a degree of approximation is introduced by using the phase equilibrium model to describe micellization, micelles are sometimes called *pseudophases*.

An illustration of how both the reaction equilibrium and phase equilibrium models can be applied to micellization is provided by Example 8.1.

* * *

EXAMPLE 8.1 *Reaction Equilibrium and Phase Equilibrium Models of Micellization.* Research in which the CMC of an ionic surfactant M^+S^- is studied as a function of added salt, say M^+X^-,

shows that a plot of log CMC versus log ([MX] + CMC) yields a straight line of slope ($\alpha - 1$). Derive this relationship by considering Reaction (B) as describing both a chemical equilibrium and a phase equilibrium.

Solution: As a chemical equilibrium, Reaction (B) is described by an equilibrium constant K, which may be written

$$K = [\text{micelle}]/[S^-]^n[M^+]^m$$

where the brackets signify molar concentrations of the indicated species, activity effects being neglected. Taking logarithms and rearranging gives

$$n \log [S^-] + m \log [M^+] = \log ([\text{micelle}]/K)$$

As a phase equilibrium, the concentration of surfactant in equilibrium with the micellar pseudophase equals the CMC regardless of the apparent concentration as long as micelles are present. Indicating the CMC value by a subscript, we write

$$\log [S^-]_{CMC} = (1/n) \log ([\text{micelle}]/K) - (m/n) \log [M^+]$$

At the CMC the total cation concentration equals the CMC value plus the concentration of the added 1 : 1 electrolyte: $[M^+] = [M^+]_{salt} + [M^+]_{CMC}$. Since $m/n = 1 - \alpha$, these results can be combined to give

$$\log [S^-]_{CMC} = (1/n) \log ([\text{micelle}]/K) - (1 - \alpha) \log \{[M^+]_{salt} + [M^+]_{CMC}\}$$

which is the desired result. The picture of micellar charge as developed here is in reasonable agreement with that determined by other methods. ■

* * *

8.3 STRUCTURE OF MICELLES

In considering the structure of micelles, we continue to base our discussion on aqueous, anionic surfactant solutions as prototypes of amphipathic systems. Cationic micelles are structured no differently from anionics, and nonionics are described parenthetically at appropriate places in the discussion. We summarize present thinking about the structure of micelles at surfactant concentrations equal to or only slightly above the CMC. We see that in nonaqueous systems (Section 8.8) and in concentrated aqueous systems (Section 8.6), the surfactant molecules are organized quite differently from the structure we describe here.

Although McBain suggested over 80 years ago that soap molecules form micellar structures of lamellar and spherical shape (McBain 1913), most of the subsequent work focused on spherical micelles. The earliest concrete model for spherical micelles is attributed to Hartley (1936), whose picture of a liquidlike hydrocarbon core surrounded by a hydrophilic surface layer formed by the head groups, has been essentially verified by modern techniques, and the "Hartley model" still dominates our thinking. We present an overview of the structure of the micelle first and then go on to examine the details a little bit more closely.

8.3a Internal Structure of Micelles: An Overview

We saw in the last section that at the CMC surfactant molecules cluster into roughly spherical aggregates containing 50–100 amphipathic units. For purposes of orientation, let us take an imaginary journey radially outward from the center of a micelle into its aqueous surroundings, identifying some distinctly different regions for subsequent discussion. This first inventory is incomplete; we add refinements as we proceed. Moving outward from the center of the micelle, we find the following:

1. The central core is predominantly hydrocarbon. The expulsion of the hydrophobic tails of the surfactant molecules from the polar medium is an important driving force behind micellization. The amphipathic molecules aggregate with their hydrocarbon tails pointing together toward the center of the sphere and their polar heads in the water at its surface.

For the surfactant this mode of organization competes with monolayer adsorption, which it somewhat resembles.

2. In ionic micelles the hydrocarbon core is surrounded by a shell that more nearly resembles a concentrated electrolyte solution. This consists of ionic surfactant heads and bound counterions in a region called the *Stern layer* (see Chapter 11, Section 11.8). Water is also present in this region, both as free molecules and as water of hydration.

3. In typical nonionic micelles the shell surrounding the hydrocarbon core also resembles a concentrated aqueous solution, this time a solution of polyoxyethylene. The ether oxygens in these chains are heavily hydrated, and the chains are jumbled into coils to the extent that their length and hydration allow.

4. Beyond the Stern layer, the remaining z counterions exist in solution. These ions experience two kinds of force: an electrostatic attraction drawing them toward the micelle and thermal jostling, which tends to disperse them. The equilibrium resultant of these opposing forces is a diffuse ion atmosphere, the second half of a double layer of charge at the surface of the colloid. Chapter 11 provides a more detailed look at the diffuse part of the double layer.

As the first step in refining our ideas about the various regions in and around micelles, it is important to realize that in reality these domains are very different from the static, well-delineated regions described. The various steps represented by Reaction (D) take place very rapidly, so the micelle exists in a state of *dynamic equilibrium.* Furthermore, by simple rotations around carbon-carbon bonds in the hydrocarbon chains, the spatial extension of that chain varies, either pulling the head deeper into the core or allowing it to protrude into the Stern layer. This, coupled with the comings and goings of ions and water molecules, makes the surface of the micelle quite turbulent at the molecular level. These considerations alone argue that the regions cited above are not sharply defined on a molecular scale. We see presently that uncertainties as to the extent of water penetration into the core also blur the distinction between these regions.

8.3b Structure of Micelles: Some Additional Details

Now let us retrace our way, in reverse order, through the various regions of the micelle enumerated above.

8.3b.1 The Exterior of the Micelle

The diffuse part of the double layer is of little concern to us at this point. Chapters 11 and 12 explore in detail various models and phenomena associated with the ion atmosphere. At present it is sufficient for us to note that the extension in space of the ion atmosphere may be considerable, decreasing as the electrolyte content of the solution increases. As micelles approach one another in solution, the diffuse parts of their respective double layers make the first contact. This is the origin of part of the nonideality of the micellar dispersion and is reflected in the second virial coefficient B as measured by osmometry or light scattering. It is through this connection that z can be evaluated from experimental B values.

Ionic micelles will migrate in an electric field, and the ion atmosphere of the colloidal particle is dragged along with it. Interpretation of micellar mobility (conductivity experiments) must take this into account. The same is true, however, of the mobility of simple ions, but the situation is more involved here since the micelle and the ion atmosphere have comparable dimensions. We see in Chapter 12 how particle and double-layer dimensions affect the interpretation of mobility experiments.

In the second item above, the presence of bound and free water molecules was noted. Both bound ions and ionic surfactant groups are hydrated to about the same extent in the micelle as would be observed for the independent ions. The dehydration of these ionic species is an endothermic process, and this would contribute significantly to the ΔH of micellization if ion dehydration occurred. In the next section we discuss the thermodynamics of micellization, but it can be noted for now that there is no evidence of a dehydration contribution to the ΔH of micelle formation. The extent of micellar hydration can be estimated from viscosity

data, assuming spherical particles. From this approach, sodium dodecyl sulfate micelles are estimated to be hydrated to about 39% by weight. For this system, 24 wt% hydration can be explained by assuming that sulfate groups and sodium ions are hydrated with 1 and 4 water molecules, respectively, at $\alpha = 0.3$. Although charge solvation accounts for a considerable part of the hydration of a micelle, these estimates show that an additional 15 wt% hydration remains unaccounted.

One important point to recognize about the Stern layer in ionic micelles is that the bound counterions help overcome the electrostatic repulsion between the charged heads of the surfactant molecules. For nonionics no such repulsion exists. It is incorrect to think that ionic micelles form and then adsorb counterions. The Stern layer is part of the micelle, and the energetics of its formation are part of the thermodynamics of micellization.

8.3b.2 The Core

Next let us return to the hydrocarbon core of the micelle. Of particular interest are the geometrical constraints imposed on the micellar structure by the length of the hydrocarbon tail and the location of the yet unexplained water of hydration. As a first approximation, the core of the micelle can be viewed as a drop of liquid hydrocarbon, the maximum radius of which equals the length of the fully extended hydrocarbon tail. To get an idea of the size of this region, let us consider a numerical example.

* * *

Example 8.2 *Calculating the Geometric Parameters of the Core of a Micelle.* Calculate the radius, volume, and surface area of the core of a micelle formed by the aggregation of dodecyl groups. The fully extended chain has a zigzag structure with angles of 109.5°, and the carbon-carbon bond length is 0.154 nm. The van der Waals radius of the terminal methyl group equals 0.21 nm, and 0.06 nm may be taken as half the length of the bond to the polar head.

Solution: The radius of the spherical core equals the length of the fully extended hydrocarbon tail (see Example 8.3 below). The 12 carbon atoms are connected by 11 bonds, each of length 0.154 nm. What must be added together, however, are the projections of these bond lengths along the direction of the chain. The distance between every other carbon in the fully extended chain—the base of a triangle opposite the tetrahedral angle—is given by the law of cosines:

$$a = (b^2 + c^2 - 2bc\cos\theta)^{1/2} = [(2)(0.154)^2(1 - \cos 109.5°)]^{1/2}$$
$$= 0.252 \text{ nm}$$

Half of this is the projection of each bond along the chain length; the sum of these projections is $(11)(0.252/2) = 1.39$ nm. Adding the contributions of the two ends to this gives the radius of the sphere: $1.39 + 0.21 + 0.06 = 1.66$ nm. (Compare this with the length obtained from Equation (4) for a chain with 12 carbon atoms.)

The volume and surface area of the core are now readily calculated:

$$V = (4/3)\pi(1.66)^3 = 19.1 \text{ nm}^3, \quad \text{and}$$
$$A = 4\pi(1.66)^2 = 34.5 \text{ nm}^2$$ ■

* * *

Dividing the surface area calculated in the example by typical aggregation numbers gives the area per head group at the surface of the core. Using n values of 50 and 100 gives 0.69 and 0.35 nm^2 per group, respectively, quite plausible numbers in view of the behavior of surfactants in monolayers.

The alkyl tails of surfactant molecules are not lines without thickness that can radiate outward from a common center in unlimited number. Instead, the chains themselves occupy a certain volume, represented by the bars in Figure 8.3. In this figure the circles represent head groups and the shaded region is water. Figure 8.3a shows schematically that close packing of the head groups requires unacceptable overlapping of the chain ends if the radius of the core is equal to the length of the fully extended chain. Figure 8.3b continues to use bars for the radius of the sphere, but fans them out in such a way as to avoid overlapping. The hole in the center is an artifact of this pictorial representation, but the wedges that allow water to pene-

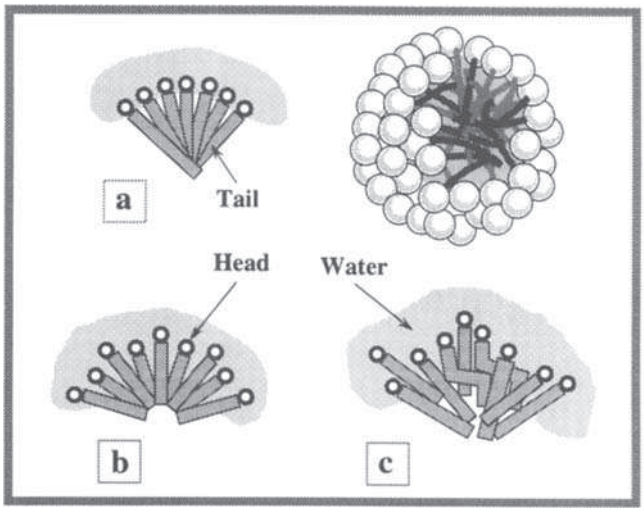


FIG. 8.3 A schematic representation of the structure of an aqueous micelle. Three possibilities are illustrated: (a) tails overlap at the center; (b) water penetrates core; and (c) chain protrusion and bending correct deficiencies of (a) and (b).

trate deeply into the core are a real feature of this model. The representation in Figure 8.3c differs from the preceding models as follows:

1. Some surfactant molecules protrude from the surface of the core farther than others, thereby alleviating crowding at the center of the micelle.
2. Chain flexibility is explicitly acknowledged, allowing some chains to twist and bend in such a way as to fill wedges that would otherwise contain water.
3. These modifications, taken separately or combined, overcome objections to the hypothetical structures shown in Figure 8.3a and 8.3b, which incidentally are whimsically known as *reef* and *fjord* structures, respectively.
4. Because of the protrusion of portions of the hydrocarbon chains into the Stern layer, the core acquires a rough surface. This model seems like a reasonable "snapshot" of what we have already described as a rapidly changing surface region.

One consequence of roughness at the surface of the micellar core is an increased contact between water and hydrocarbon. Figure 8.3b seems unrealistic because the water-hydrocarbon contact is scarcely less than in the bulk solution, a situation that apparently undermines an important part of the driving force for micellization. Figure 8.3c minimizes this effect without eliminating it. At the same time it allows for some water entrapment, which accounts for that part of the micellar hydration that was unexplained by the hydration of ions and charged groups.

8.3b.3 The Palisade Layer

The rough water-hydrocarbon surface of the core introduced in Figure 8.3c suggests that the core of the micelle should really be considered as two distinct regions: an inner core that is essentially water-free and a hydrated shell between the inner core and the polar heads. This partly aqueous shell is sometimes called the *palisade layer*. The extent to which the hydrocarbon chains protrude into the water is problematic, but we can get an idea of the volume of the palisade layer as follows.

Suppose a section of chain three methylenes long defines the thickness of the palisade layer. For a dodecyl chain this corresponds to the outer 3/12 of the radius, meaning that the radius of the inner, anhydrous core is only 3/4 of what we have been using. Cubing this fraction shows that $(3/4)^3 = 0.42$ is the fraction of the original core that is anhydrous, while

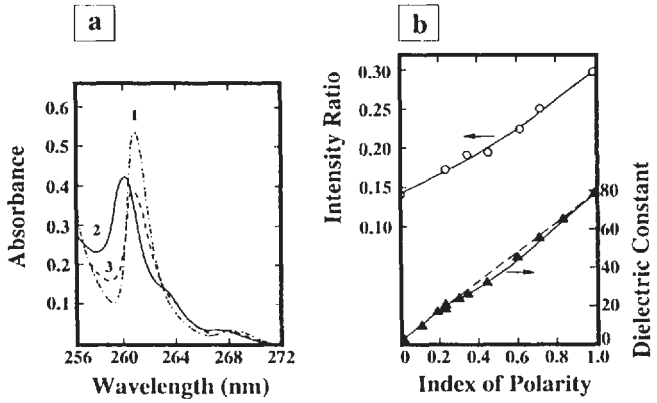


FIG. 8.4 Determination of the microenvironment of a molecule: (a) a portion of the ultraviolet spectrum of benzene in (1) heptane, (2) water, and (3) 0.4 M sodium dodecyl sulfate; and (b) ratio of the intensity of the solvent-induced peak to that of the major peak for benzene versus the index of solvent polarity. The relative dielectric constant is also shown versus the index of polarity. (Redrawn, with permission, from P. Mukerjee, J. R. Cardinal, and N. R. Desai, In *Micellization, Solubilization and Microemulsions*, Vols. 1 and 2 (K. L. Mittal, Ed.), Plenum, New York, 1976.)

58% is hydrated. Other fractions can be calculated by assuming other extents of protrusion. What is significant about this calculation is the fact that hydration of a relatively small fraction of the chain results in the presence of water in a significantly larger fraction of the volume of the micelle. After these idealized calculations, we should ask: Is there any experimental evidence for the presence of water in micelles?

We see below in this chapter that there is a great deal of interest in molecules that are "guests" in micelles. In the present context water molecules are the guests and their abundance and location have been and continue to be the basis for considerable research and controversy. We merely describe one set of pertinent experiments and their interpretation.

Figure 8.4a shows a portion of the ultraviolet spectrum of benzene in three media: (1) heptane, (2) water, and (3) 0.4 M aqueous sodium dodecyl sulfate. It is not the prominent peaks in these spectra that interest us, but rather the small bands located 3.6 nm on the long-wavelength side of the major features. This band is absent in benzene vapor, but is present with variable intensity in solutions. Accordingly, it is described as a solvent-induced band with an intensity that depends on the polarity of the solvent.

This intensity is shown in Figure 8.4b, in which the characteristics of spectra measured in different reference liquids and liquid mixtures are plotted. The abscissa in Figure 8.4b is an index of solvent polarity, specifically the molar concentration of $-OH$ groups in the reference liquids relative to the concentration of such groups in water, namely, 55.5 mole liter^{-1}. Thus abscissa values of 0.0 and 1.0 correspond to hydrocarbon and water, respectively, as solvents; intermediate values describe solvents of intermediate polarity.

The ordinate in Figure 8.4b gives the ratio of the intensity of the solvent-induced band relative to the intensity of the major peak at a 3.6-nm shorter wavelength. With this as a calibration curve, the behavior of the benzene in the sodium dodecyl sulfate micelles can be interpreted in terms of the micellar microenvironment of the benzene molecules. In this case the microenvironment is about 62% of the way between hydrocarbon and water.

Figure 8.4b also shows the relative dielectric constant of the reference liquids plotted against the same polarity index. The spectrum of the benzene suggests that the benzene molecules are located in a microenvironment of relative dielectric constant 46 (ϵ_r values for alkanes and water are about 2 and 78, respectively; $\epsilon_r = 43$ for glycerol).

This experiment introduces the use of a probe molecule to explore the microenvironment within a micelle. The results in this case show the environment to be quite "wet," but this observation alone does not tell us where either the benzene or the water is. Any one of the

models in Figure 8.3 has regions of water-hydrocarbon contact. We examine some additional data in Section 8.6 that support the palisade layer as the location of benzene in these micelles.

8.4 MOLECULAR ARCHITECTURE OF SURFACTANTS, PACKING CONSIDERATIONS, AND SHAPES OF MICELLES

Although one expects the *details* of the molecular architecture of a surfactant molecule to play a prominent role in the shapes and structures of association colloids the surfactant forms in solution, a remarkably useful picture of the shapes of the resulting aggregates can be obtained using packing considerations based on some of the crude, general geometric features of the surfactant. This is somewhat analogous to the simple van der Waals picture of gas-liquid phase transition in atomic and molecular fluids in which one uses two parameters (one for the "size" or "excluded-volume" effects of repulsion and another for attraction) to deduce the general features of phase transition. In the case of surfactant solutions, one can predict with a reasonable accuracy the shapes of the association colloids resulting from self-assembly using three effective geometric parameters of the surfactant: (a) the optimal head group area a_h, (b) the volume v_t of the tail, and (c) the critical chain length $l_{c,t}$ of the tail (see Fig. 8.5).

In addition to its simplicity, such a geometric argument can also be used to predict the changes in the structure of the aggregates as variables such as pH, charge, electrolyte concentration, and chain length of the tail are varied. The importance of the relative effects of the head group area and the size of the tails was first emphasized by Tatar (1955), and the details were developed subsequently by Tanford (1980) and others (see Wennerström and

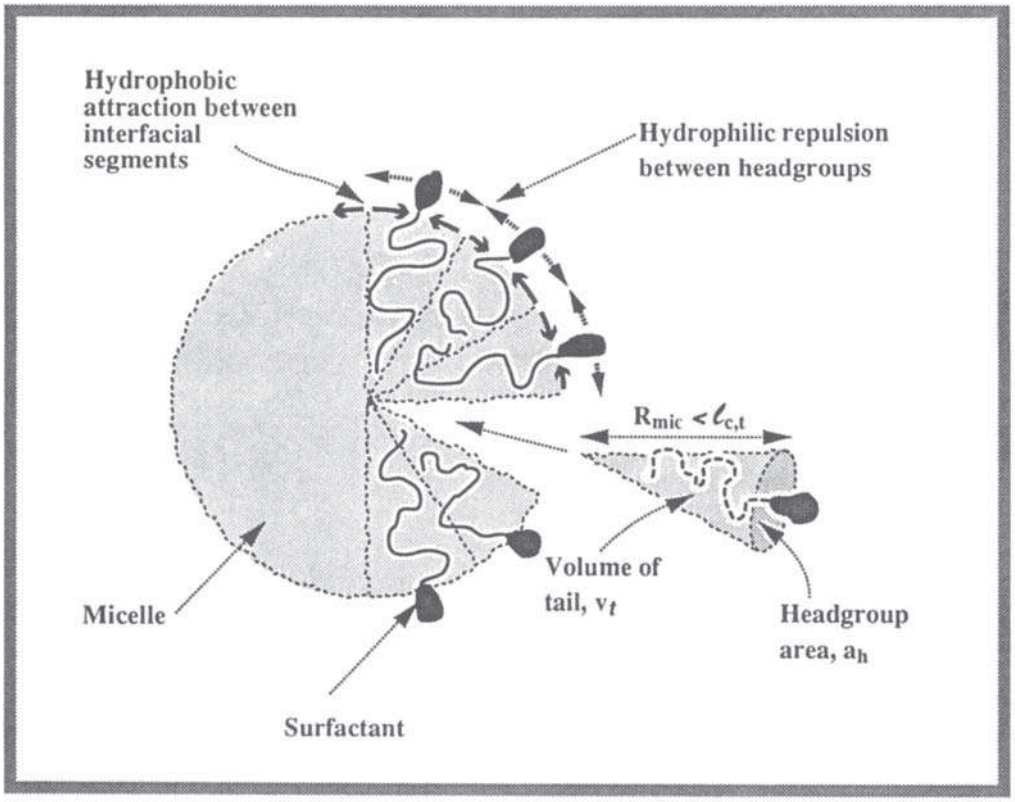


FIG. 8.5 Illustration of the geometric parameters for a surfactant and how they influence the size of the micelle. (Redrawn from Israelachvili 1991.)

Lindman 1979 and Chapter 17 of Israelachvili 1991). In what follows we restrict ourselves to the features essential for illustrating how packing considerations influence the resulting shapes of the surfactant aggregates and ignore the presence of water inside the Stern layer.

We first define the basic geometric parameters of the surfactants and follow this up with an illustration and a discussion of packing considerations that allow us to predict the shapes of the aggregates.

8.4a Optimal Head Group Area

The surface area a taken up by the head group of each surfactant molecule on the surface of a micelle depends on a number of factors, some of which tend to increase the area per head group while others tend to decrease it. This idea of two mutually opposing forces, suggested by Tanford (1980), refers to (a) an attractive force caused by the hydrophobic attraction of the hydrocarbon chain units at the hydrocarbon-water interface, and (b) a repulsive force between adjacent head groups arising from hydrophilic, steric, and ionic (in the case of charged head groups) repulsion. These opposing forces together determine the *optimal* area occupied by the head group, as illustrated in Figure 8.5.

The hydrophobic attraction is due to the preference of the hydrocarbon units near the surface to be close to its counterparts in the adjacent surfactant molecule. This attraction tends to decrease the effective area occupied by the head group. The attractive free energy contribution arising from this force is proportional to a and can be written as γa, where γ is about 20–50 mJ m^{-2}. The repulsive force mentioned in the previous paragraph is more complex and is not well understood presently. We know that this repulsion is caused by the tendency of the hydrophilic head groups to allow as many water molecules as possible in their neighborhood and by simple steric forces. Moreover, additional contributions to this repulsion may also come from electrostatic repulsion if the head groups carry charges. Although exact quantitative details of the repulsion are hard to formulate, the magnitude of the repulsion is expected to vary inversely with area a. The repulsion, of course, tends to increase the effective area occupied by each head group.

The total contribution to the interfacial free energy G may therefore be written as

$$G = \gamma a + (K/a) \tag{2}$$

where K is a constant that can be eliminated by writing it in terms of the optimal head group area a_h i.e., the area that minimizes G (i.e., $\partial G/\partial a = 0$ at $a = a_h$). The minimization of G in Equation (2) implies that $K = \gamma a_h^2$, and the value of G at the minimum is therefore $2\gamma a_h$. One can now rewrite Equation (2) in terms of a_h as

$$G = 2\gamma a_h + (\gamma/a_h)(a - a_h)^2 + \text{term of the order of } (a - a_h)^3 \tag{3}$$

using Taylor's series expansion (see Appendix A). Equation (3) allows us to express G in terms of two measurable parameters. Although the above equations are approximate, they represent the interactions among the surfactants to a first approximation and illustrate the meaning of *optimal* in the term *optimal head group area*. The above arguments can be refined further to account for specific head group interactions (e.g., ionic interactions), effects of curvature, and so on, but these are not important for our purpose here.

8.4b Volume and Critical Chain Length of the Hydrocarbon Tail

We saw in Section 8.3b.2 that the core of the micelle is essentially a hydrocarbon liquid, and we may assume this to be incompressible. This in essence defines the *volume* v_t of the hydrocarbon tail of the surfactant in the core; that is, v_t is simply the volume of the hydrocarbon liquid per hydrocarbon molecule. The critical chain length $\ell_{c,t}$ of the tail is the effective length of the hydrocarbon chain in the liquid state. This length sets a rough upper limit on the effective length of the chain, i.e., large extensions beyond this limit may prevent the collection of hydrocarbon chains from being considered a liquid. The chain length thus defined is a semiempirical parameter, although it is expected to be of the same order as the length of the

fully extended hydrocarbon molecule ℓ_{max}. Following arguments similar to Example 8.2, Tanford (1980) has given the following expressions for $\ell_{c,t}$ (and v_t) for saturated hydrocarbon chains of n carbon atoms:

$$\ell_{c,t} \leq \ell_{max} = (0.154 + 0.1265\, n)\ \text{nm} \tag{4}$$

and

$$v_t = (27.4 + 26.9n) \cdot 10^{-3}\ \text{nm}^3 \tag{5}$$

Since both expressions are linear in n, for large values of n (in fact, even for n larger than 5), the ratio $(v_t/\ell_{c,t})$ approaches 0.21 nm^2, which defines the minimum cross-sectional area a hydrocarbon can have.

Once the estimates for a_h, v_t, and $\ell_{c,t}$ are available, one can determine the preferred shapes of the surfactant aggregates using geometric packing considerations, as discussed in the following subsection and illustrated in Example 8.3.

8.4c Packing Considerations and Shapes of Aggregates

The basic packing considerations that restrict the shape of an aggregate are rather straightforward and simply serve to reconcile the volume-to-surface-area ratio of an aggregate of any shape to the requirements imposed by the optimal head group area and the liquidlike structure of the core. The optimal head group area determines the number of surfactants that can be accommodated in an aggregate of any specified shape and, therefore, the volume of the corresponding hydrocarbon tails. The packing considerations demand that the shape and the size of the core of the aggregate for the volume of the tails thus obtained be such that the aggregate has a *liquidlike* hydrocarbon core. For instance, for a spherical micelle to have a liquidlike core, the ratio $(R_s/\ell_{c,t})$, where R_s is the radius of the micelle, has to be less than or equal to unity. For larger R_s, the chains will extend further or will have more space than needed for a liquidlike core. This implies that the volume-to-surface-area ratio of the micelle specified by the radius R_s must be consistent with the ratio of the volume of the chain of the surfactant in question to its optimal head group area; otherwise, aggregation into a spherical micelle is not possible for the given surfactant (see Example 8.3).

These considerations imply that a dimensionless group, known as the *packing parameter* $\mathcal{P}$ given by

$$\mathcal{P} = v_t/(a_h\ell_{c,t}) \tag{6}$$

can be defined and used as an indicator of the shapes one can expect for the aggregate. We illustrate this using the example below and consider subsequently what this implies for predicting and controlling the aggregate shapes using variables such as pH, electrolyte concentration, and the like.

* * *

Example 8.3 *Packing Parameter for Spherical Micelles.* Show that the packing parameter $\mathcal{P}$ of a surfactant has to be less than 1/3 for it to form spherical micelles.

Solution: For a spherical micelle of radius R_s and aggregation number n_a, one has

$$n_a = [(4/3)\pi R_s^3]/v_t = 4\pi R_s^2/a_h$$

This implies that

$$(v_t/a_h) = (R_s/3)$$

that is,

$$\mathcal{P} = v_t/(a_h\ell_{c,t}) = (R_s/\ell_{c,t})(1/3)$$

Since R_s has to be less than or equal to $\ell_{c,t}$ for the core of the micelle to be liquidlike, $\mathcal{P}$ has to be less than or equal to 1/3 for the micelle to be spherical. ■

* * *

One can extend the above analysis to come up with the values of the packing parameter for which different shapes of aggregates are favored (Israelachvili 1991). These are summarized in Table 8.2. The results shown also serve as "rules of thumb" for what one can expect as one changes the chemical conditions of the solvent or the structure of the surfactant or for controlling the shape or aggregation number. For example,

1. One can change the optimal head group area by (a) decreasing or increasing the electrolyte concentration in the case of ionic surfactants or (b) changing the pH to effect an increase or decrease in the dissociation of the head group. Similarly, in the case of nonionic surfactants (such as polyoxyethylene surfactants), changing the degree of ethoxylation of the head group, or decreasing the temperature (which increases the extent of hydration), will increase the optimal head group area.
2. One can also change the ratio $(v_t/\ell_{c,t})$ by increasing the number of alkyl chains, introducing branching in the tails, using unsaturated hydrocarbon chains, and so on.

Some of the observations summarized at the end of Section 8.2b follow from the above. Although the above considerations provide some guidelines for predicting or controlling the aggregate structures based on the architecture of the surfactant molecules and the packing details, we must remember that the simplicity of the above development could be misleading. For example, the delineation points in the packing parameter do not indicate an abrupt change in the shape of the aggregates. For instance, a value of $\mathcal{P} = 0.35$ does not imply that the micelle is cylindrical. It simply means that the surfactants cannot pack themselves "neatly" into spheres; the shapes are simply expected to be slightly nonspherical. Further, we have not accounted for a number of "second-order" considerations such as the effect of curvature on the interfacial energy, effects of specific head group interactions such as ion bridging, and so on. These and other factors may be found in advanced or more specialized books such as those by Israelachvili (1991) and Clint (1992).

Now we turn our attention to thermodynamic considerations of micellization.

8.5 CRITICAL MICELLE CONCENTRATION AND THE THERMODYNAMICS OF MICELLIZATION

In this section we consider the thermodynamics of micellization from two points of view: the *law of mass action* and *phase equilibrium*. This will reveal the equivalency of the two approaches and the conditions under which this equivalence applies. In addition, we define the thermodynamic standard state, which must be understood if derived parameters are to be meaningful.

8.5a Mass Action Model

In the mass action approach we use Reaction (B) as a prototype for the process of micellization. The equilibrium constant for this reaction is given by

$$K = a_{mic}/a_S^n a_M^m \tag{7}$$

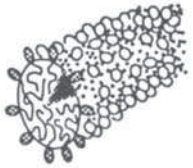
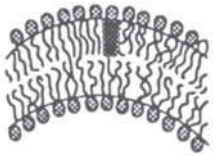
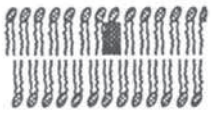
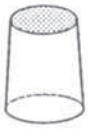
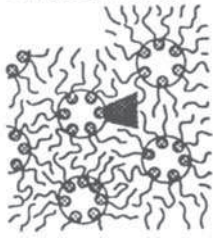
in which the a's are the activities of the indicated species. The well-known thermodynamic result $\Delta G^0 = -RT \ln K$ can be applied to Equation (7) to give ΔG^0 for Reaction (B), that is, the ΔG^0 value for micelle formation:

$$\Delta G^0 = -RT(\ln a_{mic} - n \ln a_S - m \ln a_M) \tag{8}$$

Dividing both sides of Equation (8) by n expresses this free energy change per mole of surfactant; we shall label this ΔG^0_{mic}. At the CMC, $a_M \approx a_S = a_{CMC}$, so per mole of surfactant, Equation (8) becomes

$$\Delta G^0_{mic} = RT\left[\left(1 + \frac{m}{n}\right) \ln a_{CMC} - \frac{1}{n} \ln a_{mic}\right] \tag{9}$$

TABLE 8.2 Packing Parameter and Its Relation to Shapes of Aggregates

Lipid	Critical packing parameter	Critical packing shape	Structures formed
Single-chained lipids (surfactants) with large head group areas: SDS in low salt	$< 1/3$	Cone 	Spherical micelles 
Single-chained lipids with small head group areas: SDS and CTAB in high salt, nonionic lipids	1/3–1/2	Truncated cone 	Cylindrical micelles 
Double-chained lipids with large head group areas, fluid chains: phosphatidyl choline (lecithin), phosphatidyl serine, phosphatidyl glycerol, phosphatidyl inositol, phosphatidic acid, sphingomyelin, DGDG,[a] dihexadecyl phosphate, dialkyl dimethyl ammonium salts	1/2–1	Truncated cone 	Flexible bilayers, vesicles 
Double-chained lipids with small head group areas, anionic lipids in high salt, saturated frozen chains: phosphatidyl ethanolamine phosphatidyl serine + Ca^{2+}	~ 1	Cylinder 	Planar bilayers 
Double-chained lipids with small head group areas, nonionic lipids, poly (cis) unsaturated chains, high T: unsaturated phosphatidyl ethanolamine, cardiolipin + Ca^{2+} phosphatidic acid + Ca^{2+} cholesterol, MGDG[b]	> 1	Inverted truncated cone or wedge 	Inverted micelles 

Source: Adapted with permission from Israelachvili 1991.
[a]DGDG: digalactosyl diglyceride, diglucosyl diglyceride.
[b]MGDG: monogalactosyl diglyceride, monoglucosyl diglyceride.

Since n is large, the second term on the right is smaller than the first at the CMC and can be neglected. Introducing this approximation yields

$$\Delta G^0_{mic} \approx RT\left(1 + \frac{m}{n}\right) \ln a_{CMC} \tag{10}$$

Critical micelle concentrations occur in dilute solutions, so activity may be replaced by the concentration of the surfactant at the CMC:

$$\Delta G^0_{mic} \approx RT\left(1 + \frac{m}{n}\right) \ln c_{CMC} \tag{11}$$

Equation (11) can be used to evaluate ΔG^0_{mic} from readily available CMC values. Note that setting $m = 0$ for ionic micelles is equivalent to reverting from Reaction (B) to (A) for a description of micellization. The ΔG^0_{mic} value calculated in this case describes the contribution of the surfactant alone without including the contribution of counterion binding. Since $m = 0$ for nonionics, the surfactant contribution alone is useful when comparisons between ionic and nonionic micelles are desired.

8.5b Defining the Standard State: The Phase Equilibrium Approach

It is apparent that CMC values can be expressed in a variety of different concentration units. The measured value of c_{CMC} and hence of ΔG^0_{mic} for a particular system depends on the units chosen, so some uniformity must be established. The issue is ultimately a question of defining the standard state to which the superscript on ΔG^0_{mic} refers. When mole fractions are used for concentrations, ΔG^0_{mic} directly measures the free energy difference per mole between surfactant molecules in micelles and in water. To see how this comes about, it is instructive to examine Reaction (A)—this focuses attention on the surfactant and ignores bound counterions—from the point of view of a phase equilibrium. The thermodynamic criterion for a phase equilibrium is that the chemical potential of the surfactant (subscript S) be the same in the micelle (superscript *mic*) and in water (superscript W): $\mu_S^{mic} = \mu_S^W$. In general, $\mu_i = \mu_i^0 + RT \ln a_i$, in which μ_i^0 is the standard state for the chemical potential. We write the activity as the product of the mole fraction and an appropriate *activity coefficient* f_i. We recognize that some of the surfactants are in the water and some are in the micelle and use the labels W and *mic* to differentiate between the two. If x is the mole fraction of surfactant, then

$$\mu_S^W = \mu_S^{0,W} + RT \ln x_W + RT \ln f_W \tag{12}$$

In this equation the standard state corresponds to the state that results from letting $f_W \to 1$ and $x_W \to 1$, in which case $\mu_S^W = \mu_S^{0,W}$. Letting $f_W \to 1$ is equivalent to saying that the surfactant behaves ideally, and letting $x_W \to 1$ is equivalent to having "pure" surfactant possessing the kind of interactions it has when surrounded by water. Physically, this corresponds to an infinitely dilute solution of surfactant in water. Using the primed symbol to represent the chemical potential of surfactant in micelles per mole of micelles, we write

$$(\mu_S^{mic})' = (\mu_S^{0,mic})' + RT \ln\left(\frac{x_{mic}}{n}\right) + RT \ln f_{mic} \tag{13}$$

The mole fraction of micelles is given by x_{mic}/n for particles of aggregation n. By the same logic as used above, $(\mu_S^{0,mic})'$ describes the surfactant in a standard state of pure micelle. We can write Equation (13) per surfactant molecule by dividing it by n:

$$\mu_S^{mic} = \mu_S^{0,mic} + \frac{RT}{n} \ln\left(\frac{x_{mic}}{n}\right) + \frac{RT}{n} \ln f_{mic} \tag{14}$$

The condition for phase equilibrium is given by equating Equations (12) and (14):

$$\mu_S^{0,mic} + \frac{RT}{n}\ln\left(\frac{x_{mic}}{n}\right) + \frac{RT}{n}\ln f_{mic} = \mu_S^{0,W} + RT\ln x_W + RT\ln f_W \tag{15}$$

This last result is seen to be equivalent to Equation (11) with $m = 0$ under the following conditions:

1. Both activity coefficients are set equal to unity, causing the ln f terms in Equation (15) to go to zero.
2. For large n,

$$\frac{RT}{n}\ln\left(\frac{x_{mic}}{n}\right) \ll \frac{RT}{n}\ln x_W$$

   so the smaller term can be neglected.
3. Since the concentration of surfactant in water is essentially constant and roughly equal to the CMC value after micellization, $x_W = c_{CMC}$.
4. This shows that $(\mu_S^{0,mic} - \mu_S^{0,W})$ is identical to ΔG^0_{mic} since it describes the difference per mole between the free energy of surfactant in a micelle and in water.

Although there are some aspects of micellization that we have not taken into account in this analysis—the fact that n actually has a distribution of values rather than a single value, for example—the above discussion shows that CMC values expressed as mole fractions provide an experimentally accessible way to determine the free energy change accompanying the aggregation of surfactant molecules in water. For computational purposes, remember Equation (3.24), which states that $x_2 \approx n_2/n_1$ for dilute solutions. This means that CMC values expressed in molarity units, [CMC], can be converted to mole fractions by dividing [CMC] by the molar concentration of the solvent, [solvent]: $x_2 \approx$ [CMC]/[solvent]; for water, [solvent] = 55.5 mole liter^{-1}.

The Gibbs-Helmholtz equation provides another familiar thermodynamic relationship that is useful in the present context:

$$\Delta H^0_{mic} = \left(\frac{\partial(\Delta G^0/T)}{\partial(1/T)}\right)_p = -T^2\left(\frac{\partial(\Delta G^0/T)}{\partial T}\right)_p \tag{16}$$

Using Equation (11) as the expression for ΔG^0, Equation (16) becomes

$$\Delta H^0_{mic} = R\left(1 + \frac{m}{n}\right)\left(\frac{\partial \ln c_{CMC}}{\partial(1/T)}\right)_p = -RT^2\left(1 + \frac{m}{n}\right)\left(\frac{\partial \ln c_{CMC}}{\partial T}\right)_p \tag{17}$$

The first version of Equation (17) suggests that a plot of ln c_{CMC} versus $1/T$ is a straight line of slope $\Delta H^0_{mic}/[R(1 + m/n)]$ if ΔH^0_{mic} and m/n are independent of T. Backtracking to Equation (9) or (15) shows that n must also be constant with respect to temperature for this to be valid. In fact, n increases with temperature for polyoxyethylene nonionics; any temperature dependence of n is generally assumed to be absent in ionic systems. Even if the ln c_{CMC} versus $1/T$ plot is nonlinear, the second form of Equation (17) allows a value of ΔH^0_{mic} to be evaluated from the slope of a tangent to a plot of ln c_{CMC} versus T at a particular temperature. Once ΔG^0_{mic} and ΔH^0_{mic} are known, the entropy of micellization is readily obtained from $\Delta G = \Delta H - T\Delta S$. Example 8.4 illustrates the use of these relationships.

* * *

Example 8.4 *Gibbs Energy and Entropy Changes Due to Micellization.* At 25°C the CMC for sodium dodecyl sulfate is $8.1 \cdot 10^{-3}$ M. An analysis of the temperature variation of the CMC according to Equation (17) (with $m = 0$) gives $\Delta H^0_{mic} = 2.51$ kJ mole^{-1}. Evaluate ΔG^0_{mic} and ΔS^0_{mic} from these data. Omit counterion binding by taking $m = 0$.

Solution: Convert molarity to mole fraction concentration units:

$$x = 8.1 \cdot 10^{-3} \text{ mole liter}^{-1}/55.5 \text{ mole liter}^{-1} = 1.46 \cdot 10^{-4}$$

Then $\Delta G^0_{mic} = (8.314)(298) \ln (1.46 \cdot 10^{-4}) = -21.9$ kJ mole^{-1};

$\Delta S^0_{mic} = (\Delta H^0_{mic} - \Delta G^0_{mic})/T = 2.51 - (-21.9) = 81.9$ J K^{-1} mole^{-1}

Note that the quantity ($T\Delta S^0_{mic}$) represents about 90% of the value of ΔG^0_{mic}. ■

* * *

Some values of ΔG^0_{mic}, ΔH^0_{mic}, and ΔS^0_{mic} as determined by this method are listed in Table 8.3. Several observations can be made concerning these results:

1. By taking $m = 0$, we have focused on the aggregation of surfactant only, making it easier to compare ionic and nonionic micelles.
2. The ΔG^0_{mic} values are negative, indicating spontaneous micellization.
3. The ΔH^0_{mic} values are both positive and negative. Values of ΔH^0_{mic} calculated by this method generally show poor agreement with those determined calorimetrically, at least for ionic surfactants.
4. The ΔS^0_{mic} values are positive and make a far larger contribution to ΔG^0_{mic} than ΔS^0_{mic}.

Despite controversies and uncertainties as to the best values for ΔH^0_{mic}, there is no question that the principal driving force behind the aggregation of surfactant molecules is a large, positive value of ΔS^0_{mic}. We return to this below.

The variation of ΔG^0_{mic} with changing molecular parameters has been extensively investigated. It is an experimental fact that $\ln c_{CMC}$ varies linearly with the number of carbon atoms in the alkyl chain of the surfactant molecules, the CMC decreasing with increasing chain length. This variation is readily explained by breaking ΔG^0_{mic} into contributions from the terminal methyl group (subscript CH_3), chain methylene groups (subscript CH_2), and the polar head group (subscript *PH*):

$$\Delta G^0_{mic} = \Delta G_{CH_3} + \nu \Delta G_{CH_2} + \Delta G_{PH} \tag{18}$$

In Equation (18), ν is one less than the number of carbons in the alkyl chain. If we assume that neither ΔG_{CH_3} nor ΔG_{CH_2} is affected by the length of the tail, then Equation (18) can be combined with Equation (11) to give

$$\ln c_{CMC} = \frac{\nu \Delta G_{CH_2}}{RT} + \text{constant} \tag{19}$$

which is the equation of a straight line ($\ln c_{CMC}$ versus ν) and explains the observed behavior. The slope of such a plot for an assortment of *n*-alkyl ionic surfactants averages about -0.69, which corresponds to a ΔG_{CH_2} of -1.72 kJ mole^{-1} at 25°C. Since 0.693 equals ln 2, it follows that the addition of a methylene group to a hydrocarbon chain decreases the CMC of these

TABLE 8.3 Some Thermodynamic Properties for the Micellization Process at or Near 25°C for Various Surfactants

Surfactant	ΔG^0_{mic} (kJ mole^{-1})	ΔH^0_{mic} (kJ mole^{-1})	ΔS^0_{mic} (J K^{-1} mole^{-1})
Dodecyl pyridinium bromide	−21.0	−4.06	+56.9
Sodium dodecyl sulfate[a]	−21.9	+2.51	+81.9
N-Dodecyl-*N*,*N*-dimethyl glycine	−25.6	−5.86	+64.9
Polyoxyethylene(6) decanol	−27.3	+15.1	+142.0
N,*N*-Dimethyl dodecyl amine oxide	−25.4	+7.11	+109.0

Source: Data from J. H. Fendler and E. J. Fendler, *Catalysis in Micellar and Macromolecular Systems*, Academic Press, New York, 1975.

[a]Calculated in Example 8.4.

surfactants by almost exactly a factor of 2. In the presence of added indifferent electrolyte, $\Delta G_{CH_2} \approx -2.9$ kJ mole^{-1}, the aggregation being facilitated by the additional ions. With this type of data, it is possible to back-calculate the contribution to ΔG^0_{mic} of various polar head groups and, with these, to estimate the ΔG^0_{mic} and CMC values of other surfactants as the sum of group contributions. Extensive tabulations of these contributions are available.

8.5c Entropy Change During Micellization and the Hydrophobic Effect

We conclude this section with a brief discussion of the relatively large, *positive* values of ΔS^0_{mic}, which we have seen are primarily responsible for the spontaneous formation of micelles. At first glance it may be surprising that ΔS for Reaction (A) is positive; after all, the number of independent kinetic units decreases in this representation of the micellization process. Since such a decrease results in a negative ΔS value, it is apparent that Reaction (A) is incomplete as a description of micelle formation. What is not shown in Reaction (A) is the aqueous medium and what happens to the water as micelles form. The water must experience an increase in entropy to account for the observed positive values for ΔS^0_{mic}.

The key to understanding this entropy increase is the extensive hydrogen bonding that occurs in water. To a first approximation, the water molecule has a tetrahedral shape with the oxygen atom at the center, hydrogen atoms at two of the apex positions, and two lone pairs of electrons at the remaining apex positions. Hydrogen bonds form between the hydrogen atoms on one molecule and the oxygen lone pairs on another, effectively building a loose network of tetrahedra bound at the corners. Because of thermal fluctuations, various parts of this network continuously break and reform in liquid water, but at equilibrium a high average level of hydrogen bonding prevails. Many properties of water are due to this hydrogen bonding.

Next suppose a hydrocarbon moiety such as the tail of a surfactant is embedded in the water. Since water forms no hydrogen bonds with the alkyl group, the alkyl group merely occupies a hole in the liquid water structure. Our first thought might be that such cavities require the breaking of hydrogen bonds, and that their formation should involve a positive enthalpy contribution. As we have seen, a positive ΔS^0_{mic} is the dominant driving force behind surfactant clustering. We are therefore forced to conclude that water molecules at the surface of the cavity regenerate the hydrogen-bonded water network, as if the hydrocarbon chains were nucleation sites for network formation. As a result, water molecules become more ordered around the hydrocarbon with an attendant decrease in entropy. The standard state we have used means that the various Δ's measure the difference between surfactant in micelles and surfactant in water. Removing the surfactant from the water and placing it in a micellar environment allows the cavity to revert to the structure of pure liquid water. The highly organized cavity walls return to normal, hydrogen-bonded liquid with an increase in entropy. Incidentally, enhanced hydrogen bonding at the walls of the cavity largely compensates for the breaking of hydrogen bonds to form the cavity, so the enthalpy change is slight.

This explanation for the entropy-dominated association of surfactant molecules is called the *hydrophobic effect* or, less precisely, *hydrophobic bonding*. Note that relatively little is said of any direct "affinity" between the associating species. It is more accurate to say that they are expelled from the water and—as far as the water is concerned—the effect is primarily entropic. The same hydrophobic effect is responsible for the adsorption behavior of amphipathic molecules and plays an important role in stabilizing a variety of other structures formed by surfactants in aqueous solutions.

8.6 SOLUBILIZATION

Above the CMC, a number of solutes that would normally be insoluble or only slightly soluble in water dissolve extensively in surfactant solutions. The process is called *solubilization*, the substance dissolved is called the *solubilizate*, and—in this context—the surfactant is called the *solubilizer*. The result is a thermodynamically stable, isotropic solution in which the solubilizate is somehow taken up by micelles since the enhancement of solubility begins at the CMC. This observation, in fact, provides one method for determining the CMC of a surfactant; it

must be used cautiously, however, since the solubilizate may change the CMC from what would be observed in its absence. There is an upper limit to the amount of material that can be solubilized in a given surfactant solution; beyond this limit the excess solubilizate displays normal phase separation.

8.6a Location of the Solubilizate

There is considerable interest in establishing the location within a micelle of the solubilized component. As we have seen, the environment changes from polar water to nonpolar hydrocarbon as we move radially toward the center of a micelle. While the detailed structure of the various zones is disputed, there is no doubt that this gradient of polarity exists. Accordingly, any experimental property that is sensitive to the molecular environment can be used to monitor the whereabouts of the solubilizate in the micelle. Spectroscopic measurements are ideally suited for determining the microenvironment of solubilizate molecules. This is the same principle used in Section 8.3, in which the ultraviolet spectrum of solubilized benzene was used to explore the solvation of micelles. Here we take the hydration for granted and use similar methods to locate the solubilizate.

In addition to ultraviolet spectroscopy, resonance methods—such as electron spin resonance and nuclear magnetic resonance (NMR)—have been widely used to establish the site of solubilization. We briefly consider the use of proton NMR for such an investigation. Like the electron, the proton may have spin quantum numbers of $\pm 1/2$, resulting in two different nuclear quantum states. In a magnetic field these states differ in energy, the difference corresponding to the nuclear magnetic moment aligning with or against the magnetic field. If energy of the appropriate frequency is supplied, transitions from one state to the other are possible, and the radiation is absorbed. The absorbed frequency measures the separation of the two states by the familiar formula $\Delta E = h\nu$, where h is Planck's constant. The energy separation of the affected states depends on the strength of the applied magnetic field as modified by the local environment. Thus even protons in different parts of a molecule absorb at slightly different frequencies since the magnetic field each experiences is somewhat different from the applied field owing to the shielding effect of nearby electrons. Furthermore, neighboring protons cause a peak in an absorption spectrum to be split into a multiplet with characteristic relative intensities. This aids the identification of peaks in an NMR spectrum.

In addition to one part of a molecule influencing the NMR spectrum of another part, the medium in which the molecule is embedded also has an effect. Therefore the NMR spectrum of a solubilized molecule has the potential not only to reveal the location of a solubilized molecule in a micelle, but also to give information about its orientation.

Figure 8.6 shows some data measured for benzene solubilized in hexadecyl pyridinium chloride. The abscissa in the figure shows the extent of solubilization expressed as moles of solubilizate per mole of surfactant. The ordinate values show shifts in the resonance frequencies from water taken as an internal standard. Shifts of peaks arising from protons in the pyridinium ring, in benzene, and in methylene groups in the alkyl tail are shown versus the extent of solubilization.

The following points summarize these observations:

1. All peaks are shifted toward higher fields because of the diamagnetic effect of the solubilized benzene.
2. The slope of the shift versus extent of solubilization curve is much steeper for the protons on the benzene and pyridinium rings than for protons in methylenes in the tail. In fact, the first two increase in roughly parallel fashion.
3. Since the charged pyridinium ring must be at the surface, the benzene that parallels it in chemical shift must have a similar location.
4. The far less sensitive response of the chemical shift of the methylene protons to the extent of solubilization suggests that the core of the micelle is mainly composed of alkyl chains and is relatively unaffected by the benzene.

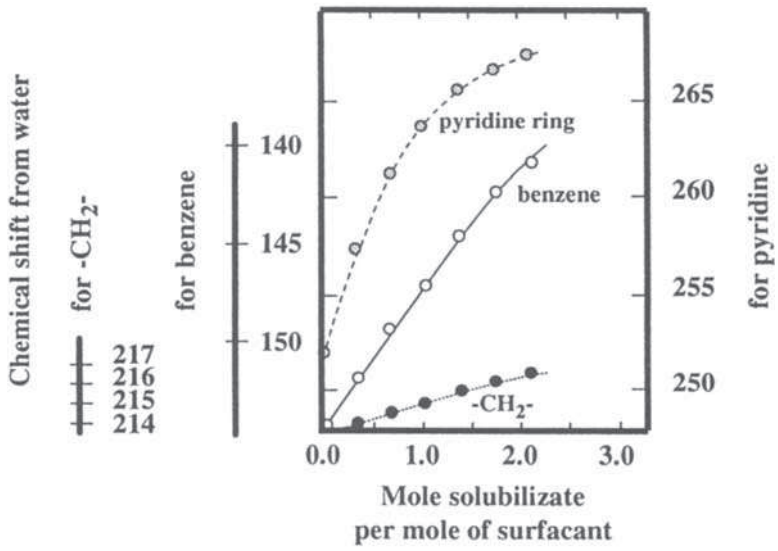


FIG. 8.6 Chemical shifts of protons in benzene, pyridium rings, and methylene groups versus the ratio of moles of solubilized benzene to moles of hexadecyl pyridinium chloride. (Redrawn, with permission, from J. C. Eriksson, *Acta Chem. Scand.*, **17**, 1478 (1963).)

In another related study (using hexadecyltrimethyl ammonium bromide micelles), isopropyl benzene was solubilized, and the chemical shifts of aromatic and alkyl protons were observed. The results suggest that the isopropyl benzene molecules are oriented such that the isopropyl groups are buried more deeply in the core of the micelle, while the benzene ring is in the more hydrated palisade layer. This plus the conclusion of Item 3 is consistent with the description presented in Section 8.3, which located the benzene in a relatively polar portion of the micelle.

8.6b Extent of Solubilization and Its Relation to Location of the Solubilizate

The extent of solubilization and the location of solubilizate in the micelle are related to one another. In the most hydrophobic inner core of the micelle, there simply is not as much room for the solubilizate compared to the hydrated palisade layer. In the case of polyoxyethylene nonionics, the core is surrounded by a mantle of aqueous hydrophilic chains, and solubilization may occur in both the core and the mantle. The relative amount of solubilization in these two regions of the nonionic micelle depends on the polarity of the solubilizate. Nonionics appear relatively more hydrophobic at higher temperatures, presumably owing to an equilibrium shift that favors dehydration of the ether oxygens. At some elevated temperature called the *cloud point*, solutions of these surfactants undergo phase separation. As the cloud point is approached, the solubilization of nonpolar solubilizates increases, probably because of an increase in the aggregation number of the micelles. For polar solubilizates, solubilization decreases owing to dehydration of the polyoxyethylene chains accompanied by their coiling more tightly. These observations demonstrate that nonpolar compounds are solubilized in the core of the micelle, while polar solubilizates are located in the mantle. Both of the temperature effects cited here are consistent with variations in the space available for the solubilized molecules in the micelle.

We saw in the last section how the removal of a surfactant molecule from water into a micellar environment has a negative $\Delta G°$ value. Qualitatively, it is not surprising that other

organic solutes—the solubilizates—should decrease their free energy by entering micelles. Next we briefly consider how this general notion is quantified.

8.6c Thermodynamics of Solubilization

Using the same logic as we applied to micellization in Section 8.3, we can write the following for the solubilizate (component 3):

$$\mu_3^{0,mic} - \mu_3^{0,W} = RT \ln \left(\frac{a_3^{mic}}{a_3^{W}} \right) = RT \ln \left(\frac{x_3^{mic} f_3^{mic}}{x_3^{W} f_3^{W}} \right) \tag{20}$$

Since solubility in water for many solubilizates is low, the aqueous phase may be treated as an ideal solution, that is, $f_3^W = 1$. As we have seen, however, the micellar phase is neither ideal nor simple, with f_3^{mic} varying from place to place within the micelle. Aside from noting that solubilization occurs spontaneously, which makes $\Delta\mu^0$ negative, we shall not pursue this approach to solubilization any further.

8.6d Micellar Phase Diagrams

Until now we have taken a rather narrow view of solubilization, having considered only the uptake of solubilizate starting with aqueous solutions of surfactant. This three-component system can clearly take on a much wider range of proportions than we have discussed so far. A ternary phase diagram describes the full range of possibilities. For surfactant systems these diagrams can be quite complex, even though we shall limit ourselves to relatively simple cases. We begin with a brief review of the method for reading a ternary phase diagram:

1. The apex points of an equilateral triangle are identified with the three pure components, say, A, B, and C.
2. Each of the sides represents one of the possible binary combinations, say AB, BC, and CA. The fraction of the distance along the side of the triangular diagram measures the fractional composition of the binary mixture.
3. The perpendicular distance from any point within the triangle to one of the bases is proportional to the amount in the mixture of the component opposite the base in question. Since the sum of the three perpendiculars from any point to the sides equals the height of the triangle, expressing these lengths as fractions of the height permits any point in the triangle to be described as being some fraction toward A, some fraction toward B, and some fraction toward C. These same fractions can be used to describe either the weight fraction or mole fraction compositions of a ternary mixture.
4. Since ternary phase diagrams are drawn at constant temperature and pressure, the phase rule allows one, two, or three phases to be present. In two-phase regions, tie lines connect points having the compositions of the equilibrium phases. Three-phase regions have triangular shapes; the coordinates of the corners of the triangular zone give the compositions of the equilibrium phases.

Readers desiring a more detailed review of ternary phase diagrams will find the topic discussed in most physical chemistry textbooks or in the concise but comprehensive monograph on ternary equilibrium diagrams by West (1982).

Figure 8.7 shows the ternary phase diagram for water, hexanoic acid, and sodium dodecyl sulfate at 25°C. Seven different areas are shown in the figure, which has been used to describe the solubilization of polar dirt by surfactant solutions in detergency applications. The following comments refer to these seven different regions and explain the labeling used in Figure 8.7:

1. *Two liquids*: This is a two-phase region in which two liquid solutions—each containing three components—are in equilibrium. The two solutions could exist as distinct layers (e.g., in a separatory funnel) or as an emulsion in which droplets of one phase are dispersed in the second.
2. *L_1 and L_2*: Each of these phases is a homogeneous, isotropic liquid solution containing three components. The sort of system we have focused on in this section is the L_1 region in which the acid is solubilized in aqueous micelles. In L_2 the reverse is true: The acid is the

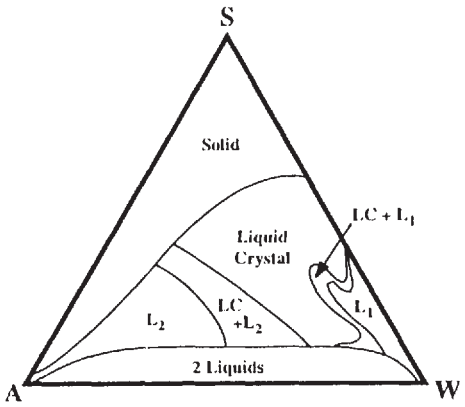


FIG. 8.7 Ternary phase diagram for water (W), hexanoic acid (A), sodium dodecyl sulfate (S) at 25°C. See text for a description of the various regions. (Redrawn, with permission, from A. S. C. Lawrence, *Chem. Ind.*, **44**, 1764 (1961).)

solvent, and water is solubilized in reverse micelles. We have more to say about the solubilization of water in reverse micelles in Section 8.8.

3. *Liquid crystal*: As the name implies, this is an ordered yet fluid phase in which water, surfactant, and solubilizate combine to form anisotropic, organized structures. These are called *lyotropic mesomorphic phases*, as opposed to *thermotropic mesomorphs*, which form when certain organic crystals are heated.

4. *L_1 or L_2 plus liquid crystal*: Each of these is a two-phase region in which the liquid crystalline phase exists in equilibrium with one of the isotropic liquid phases. Tie lines must be determined experimentally to give the precise compositions of the phases in equilibrium.

5. *Solid*: Eventually the system becomes saturated with surfactant and solid sodium dodecyl sulfate precipitates out.

The phase diagram shown in Figure 8.7 is simpler than many other ternary phase diagrams inasmuch as only one liquid crystal phase exists, and there are no three-phase triangles. By contrast, the water-decanol-sodium-caprylate phase diagram at 20°C shows 5 different liquid crystal phases and 10 triangular regions, in each of which 3 different sets of phases are in equilibrium.

8.6e Liquid-Crystalline Phases

Liquid crystals are a fascinating topic of study in their own right, but we limit our discussion to a brief description of the ordering in some of the possible structures. In all cases the amphipathic molecules are oriented in such a way as to minimize the contact between water and the alkyl chains. Whether the polar head points outward or not depends on which component dominates the continuous phase; the minor component is solubilized inside the micellar structures.

At relatively low concentrations of surfactant, the micelles are essentially the spherical structures we discussed above in this chapter. As the amount of surfactant and the extent of solubilization increase, these spheres become distorted into prolate or oblate ellipsoids and, eventually, into cylindrical rods or lamellar disks. Figure 8.8 schematically shows (a) spherical, (b) cylindrical, and (c) lamellar micelle structures. The structures shown in the three parts of the figure are called (a) the *viscous isotropic phase*, (b) the *middle phase*, and (c) the *neat phase*. Again, we emphasize that the orientation of the amphipathic molecules in these structures depends on the nature of the continuous and the solubilized components.

The "crystallinity" of liquid crystal phases refers to the large assortment of ways these micellar structures can be organized within a bulk phase. For example, spherical micelles of

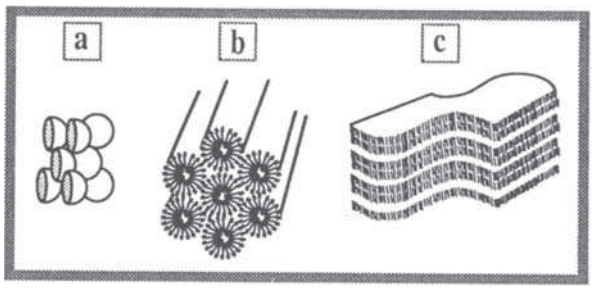


FIG. 8.8 Schematic representations of surfactant structures in (a) viscous isotropic, (b) middle, and (c) neat liquid crystal phases.

either type can form an ordered bulk phase by arranging themselves into either face-centered or body-centered cubic arrangements. Likewise, rodlike micelles of either type can be arranged into square or hexagonal packing. When all of the shape and packing possibilities are taken into account, it is no wonder that surfactant phase diagrams are complicated!

8.7 CATALYSIS BY MICELLES

For about four decades—between 1920 and 1960—the study of micelles was conducted mostly by researchers who would be classified as colloid scientists. Since about 1960 there has been an immense upswing of interest in micelles, with biochemists, organic chemists, inorganic chemists, and even physicists joining the ranks of those engaged in micellar research. The discovery that micellar media can affect the rates of chemical reactions has triggered this surge of interest in other branches of chemistry, and the notion that micelles might serve as models for enzymes has further stimulated research in this area. Studies have been conducted with a variety of objectives, including elucidation of reaction mechanisms, clarification of enzyme catalysis, investigations of micelles themselves, and applications to synthetic chemistry.

8.7a An Example: Enzyme Catalysis

Enzymes and micelles resemble each other with respect to both structure (e.g., globular proteins and spherical aggregates) and catalytic activity. Probably the most common form of enzyme catalysis follows the mechanism known in biochemistry as *Michaelis-Menton kinetics*. In this the rate of the reaction increases with increasing substrate concentration, eventually leveling off. According to this mechanism, enzyme E and substrate A first react reversibly to form a complex EA, which then dissociates to form product P and regenerate the enzyme:

$$E + A \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \mathrm{EA} \xrightarrow{k_2} \mathrm{P} + \mathrm{E}$$

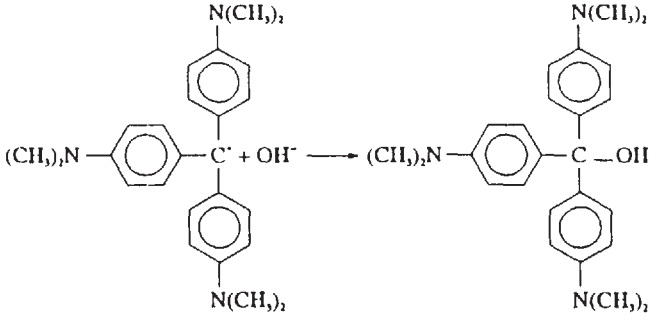
(E)

The various k's are the rate constants for the specific reactions shown. Standard kinetic analysis of this mechanism predicts that the rate of product formation is given by

$$\text{Rate} = \frac{k_1 k_2 [E]_0 [A]}{k_2 + k_{-1} + k_1 [A]} = \frac{k_2 [E]_0 [A]}{K_M + [A]} \tag{21}$$

in which $[E]_0$ is the total concentration of enzyme, and $K_M = (k_2 + k_{-1})/k_1$ is called the *Michaelis constant*. Note that as $[A]$ increases, K_M can be neglected in the denominator, allowing cancellation of $[A]$ and explaining the plateau in rate. The Michaelis constant inversely measures the affinity between enzyme and substrate: A small value of K_M means the enzyme binds the substrate tightly.

Figure 8.9 illustrates that micellar catalysts product similar effects. The chemical reaction under investigation in the figure is that between the crystal violet carbonium ion and the hydroxide ion to form the corresponding alcohol:



(F)

In Figure 8.9 pseudo-first-order rate constants at 30°C for the rate of this reaction in 0.003 M NaOH are plotted versus the concentration of various alkyl trimethyl ammonium bromides. Several things should be noted about these data:

1. The catalytic role of the surfactant begins more or less sharply at the CMC. The C_{10} and C_{12} surfactants also show catalytic activity above their respective CMCs at higher concentrations.
2. The enhancement of rate qualitatively follows Michaelis-Menton kinetics, with both the initial slope and the final plateau increasing with increasing length of the alkyl tails of the surfactant.
3. The initial (starting at the CMC) slope of these curves is inversely proportional to K_M and increases with increasing binding effectiveness between the micelle and the substrate (since K_M itself inversely measures affinity).

These results can be rationalized by picturing the crystal violet carbonium ion as solubilized and oriented in the micelle, followed by attack by the aqueous hydroxide ion. The catalytic effect of the micellar solution has a twofold origin: a concentration effect and an

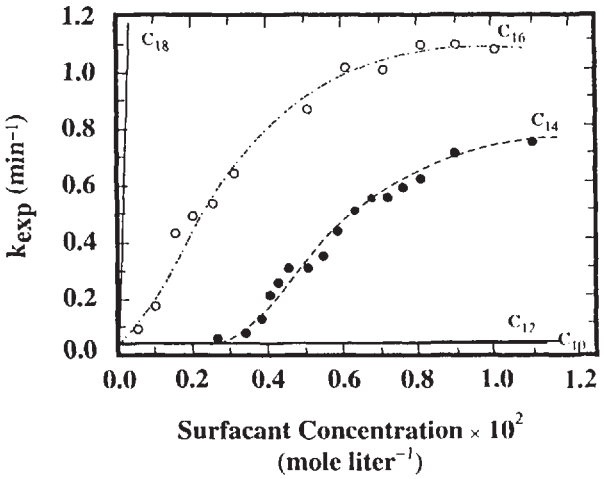


FIG. 8.9 Experimental rate constants for Reaction (F) versus concentration for alkyl trimethyl ammonium bromides with the indicated chain length. (Redrawn, with permission, from J. Albrizzio, J. Archila, T. Rodulfo, and E. H. Cordes, *J. Org. Chem.*, **37**, 871 (1972).)

effect on the transition state of the reaction. For the system shown in Figure 8.9 the micelles are positively charged, and the OH^- reactant is concentrated in the Stern layer of such micelles. In addition, the cationic crystal violet should be less stable in the cationic micelles than the zwitterionic transition state. Thus the transition state is electrostatically favored in the micelle compared to a nonsolubilized species. That electrostatic as well as hydrophobic effects are involved in the enhancement of the rate of Reaction (F) is evident from the fact that the relative rates with and without surfactant are 241/17 for cationic micelles and 1/17 for anionic micelles. It has also been observed that other anions inhibit the catalysis of Reaction (F), presumably by competing with OH^- for adsorption sites in the Stern layer.

8.7b Quantitative Analysis of Micellar Catalysis

Having looked at an example of micellar catalysis, let us next consider how such results are analyzed quantitatively. By analogy with Reaction (E), we visualize the micelle M and the substrate A entering a solubilization equilibrium characterized by an equilibrium constant K:

$$M + A \overset{K}{\rightleftharpoons} MA \xrightarrow{k_m} P \qquad MA \xrightarrow{k_0} P \qquad \text{(G)}$$

The solubilized substrate MA is the analog of the complex ES in Reaction (E), and the product P is formed in the micelle with the rate constant k_m. The product can also form from the substrate without involving the micelle; k_0 is the rate constant for the last process. The experimental (subscript *exp*) rate constant for Reaction (G) is then a weighted sum of the two constants k_m and k_0:

$$k_{exp} = f_0 k_0 + f_m k_m = f_0 k_0 + (1 - f_0) k_m \tag{22}$$

where f_0 and f_m are the fractions of substrate in the bulk solution and solubilized in the micelles, respectively. The second version of Equation (22) arises from the recognition that $f_0 + f_m = 1$. The equilibrium constant in Reaction (G) can be written as

$$K = \frac{[MA]}{[M][A]} = \frac{f_m[A]_0}{[M]f_0[A]_0} = \frac{f_m}{(1 - f_m)[M]} \tag{23}$$

where $[A]_0$ is the total substrate concentration. Solving Equation (23) for $f_0 = (1 + K[M])^{-1}$ and substituting into Equation (22) gives

$$k_{exp} = \frac{k_0 + Kk_m[M]}{1 + K[M]} \tag{24}$$

Subtracting k_0 from both sides of Equation (24) yields

$$k_{exp} - k_0 = \frac{k_0 + Kk_m[M]}{1 + K[M]} - k_0 = \frac{Kk_m[M] - Kk_0[M]}{1 + K[M]} \tag{25}$$

which, on inversion, gives

$$\frac{1}{k_{exp} - k_0} = \frac{1}{k_m - k_0} + \frac{1}{(k_m - k_0)K[M]} \tag{26}$$

Finally, the concentration of micelles can be eliminated from Equation (26) by noting that $[M]$ is given by the number of moles of surfactant in excess of the CMC value divided by the degree of aggregation of the micelle, or

$$[M] = (c - c_{CMC})/n \tag{27}$$

where c is the total concentration of surfactant. Substituting Equation (27) into Equation (26) gives

$$\frac{1}{k_{exp} - k_0} = \frac{1}{k_m - k_0} + \frac{n}{K(k_m - k_0)} \frac{1}{c - c_{CMC}} \tag{28}$$

Recall that k_{exp} and k_0 are rate constants with and without surfactant, respectively, for the reaction in question, and that c and c_{CMC} are surfactant concentrations in the reaction mixture and at the CMC, respectively. Therefore k_{exp}, k_0, c, and c_{CMC} are experimentally accessible. Equation (28) predicts that a plot of $(k_{exp} - k_0)^{-1}$ versus $(c - c_{CMC})^{-1}$ is a straight line with a slope and intercept that have the following significance:

$$\text{Intercept} = 1/(k_m - k_0) \tag{29}$$

$$\text{Slope} = n/[K(k_m - k_0)] \tag{30}$$

$$\text{Slope/intercept} = n/K \tag{31}$$

Since k_0 is known, Equation (29) allows k_m to be evaluated. Likewise, the equilibrium constant for the binding of the substrate to the micelle can be evaluated from Equation (31) if n is known from a separate experiment. This method of analysis of catalyzed reactions is called a *Lineweaver-Burke plot* after the corresponding technique in biochemistry. Example 8.5 illustrates the use of these relationships.

* * *

Example 8.5 *Estimation of Equilibrium and Rate Constants in Micellar Catalysis.* Using $k_0 = 0.050\ \text{min}^{-1}$ and $c_{CMC} = 1.0 \cdot 10^{-3}$ M, estimate k_m and K/n for the C_{16} data in Figure 8.9. Because of the scatter in the data, take points at regular intervals from the drawn curve in the figure as the basis for this estimate.

Solution: The following points are read from the C_{16} curve in Figure 8.9:

$c \cdot 10^3$ (M)	1.5	2.0	2.5	3.0	4.0	5.0	6.0	7.0	8.0	9.0
k_{exp} (min^{-1})	0.25	0.40	0.52	0.66	0.80	0.90	1.0	1.1	1.1	1.1

Calculate $(c - c_{CMC})^{-1}$ and $(k_{exp} - k_0)^{-1}$ using $c_{CMC} = 1.0 \cdot 10^{-3}$ M and $k_0 = 0.05\ \text{min}^{-1}$:

$(c - c_{CMC})$ 10^3 (M)	$(k_{exp} - k_0)$ (min^{-1})	$(c - c_{CMC})^{-1}$ (liter mole^{-1})	$(k_{exp} - k_0)^{-1}$ (min)
0.5	0.20	2000	5.00
1.0	0.35	1000	2.90
1.5	0.47	670	2.10
2.0	0.61	500	1.60
3.0	0.75	330	1.30
4.0	0.85	250	1.20
5.0	0.95	200	1.10
6.0	1.05	170	0.95
7.0	1.05	140	0.95
8.0	1.05	130	0.95

These data produce a reasonably linear Lineweaver-Burke plot of slope $2.2 \cdot 10^{-3}$ min mole liter^{-1} and intercept 0.62 min.

According to Equation (29), $k_m - k_0 = \text{intercept}^{-1} = 1/0.62 = 1.61$; therefore $k_m = 1.7\ \text{min}^{-1}$, a value 34 times larger than the estimated k_0 value.

According to Equation (31), $K/n = \text{intercept/slope} = 0.62/2.2 \cdot 10^{-3} = 280$ liter mole^{-1}. ■

* * *

The derivation of Equation (24) involves a number of assumptions, some of which are questionable in many cases. For example, in the above derivation, we assume the following:

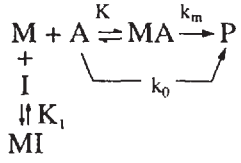
1. The association between substrate and micelle shown in Reaction (G) follows 1 : 1 stoichiometry; that is, only one substrate molecule is taken up per micelle. If this is not the case, n is not the actual aggregation number of the micelle, but rather the number of surfactant molecules per solubilizate molecule.

2. The substrate does not complex with monomeric surfactant, and rate enhancement begins sharply at the CMC. Reactant molecules, like other solubilizates, may change the CMC of a surfactant, and premicellar clusters could influence the rate.
3. There are no competitive processes that inhibit the reaction.

8.7c Electrolyte Inhibition in Micellar Catalysis

In discussing Reaction (F), we remarked that other anions are observed to compete with OH^- in the Stern layer. This sort of electrolyte inhibition is widely observed, and the dependence of the inhibition on both the size and charge of the ions generally corresponds to expectations. For example, in the base-catalyzed hydrolysis of carboxylic esters in the cationic micelles, anions inhibit the reaction in the order $NO_3^- > Br^- > Cl^- > F^-$. For acid-catalyzed ester hydrolysis in anionic micelles, the cation of added salt inhibits the reaction in the order $R_4N^+ > Cs^+ > Rb^+ > Na^+ > Li^+$.

Inhibition may be incorporated into the mechanism of micellar catalysis in the same way it is handled in enzyme kinetics. Representing the inhibitor by I, we can revise Reaction (G) as follows:



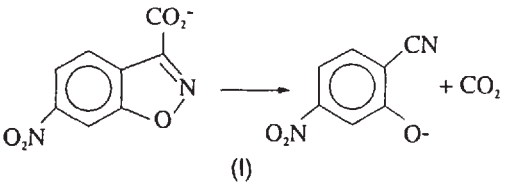
(H)

where K_I is the binding constant for the inhibitor. Assuming that solubilization of one inhibitor molecule into the micelle blocks the addition of substrate leads to

$$k_{exp} = \frac{k_0 + k_0K_I[I] + k_mK[M]}{1 + K_I[I] + K[M]} \tag{32}$$

which reduces to Equation (24) when $[I]$ or K_I equals zero. Equation (32) is also an oversimplification, but nevertheless it satisfactorily accounts for many observations.

Figure 8.10a is an example of some data in which the effect of added salt is more than a competitive ion-binding phenomenon. The reaction involved is the decarboxylation of 6-nitrobenzisoxazole-3-carboxylate, catalyzed by hexadecyl trimethyl ammonium bromide micelles:



(I)

In the presence of sodium tosylate (NaTos), both the rate constant k_{exp} and the viscosity of the solution show maxima at the same electrolyte concentration. The dramatic variation in η shown in Figure 8.10a suggests that the sodium tosylate alters the shape of the micelles, first producing rodlike structures that subsequently break up into more compact structures. The complicated phase diagrams of surfactants make this a plausible explanation. Effects such as this clearly complicate the picture not only of inhibition, but also of micellar catalysis in general.

Equilibrium constants for the binding between substrates and micelles—Reaction (G)—generally range from 10^3 to 10^6 for hydrophobic organic substrates. Furthermore, they are expected to increase as the hydrophobic character of the substrate increases. Figure 8.10b shows that this effect sometimes overshoots optimum solubilization. The figure shows, on a

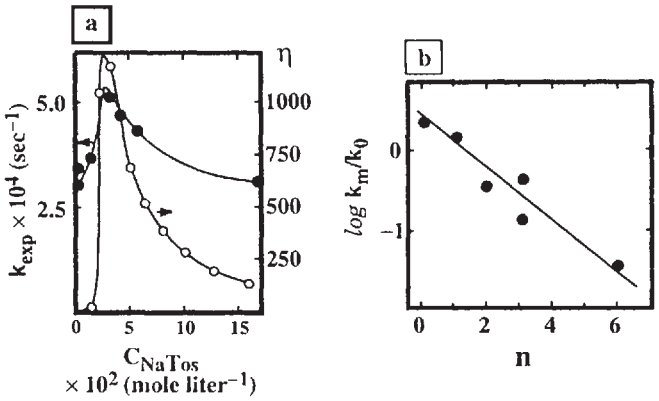
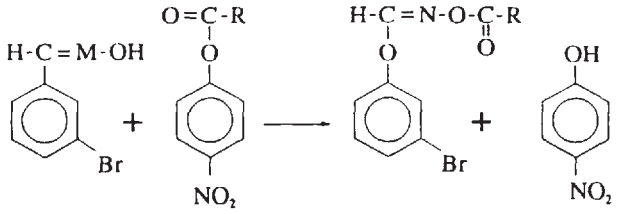


FIG. 8.10 Micellar catalysis: (a) the micelle catalyzed rate constant for Reaction (I) versus the concentration of sodium tosylate. The viscosity of the system versus sodium tosylate concentration is also shown (redrawn, with permission, from C. A. Bunton, In *Reaction Kinetics in Micelles* (E. H. Cordes, Ed.), Plenum, New York, 1973)); (b) ratio of k_m to k_0 for Reaction (J) versus the number of carbon atoms in the acyl group (redrawn, with permission, from A. K. Yatsimirski, K. Martinek, and I. V. Berezin, *Tetrahedron*, **27**, 2855 (1971)).

logarithmic scale, the ratio k_m/k_0 versus the number of carbon atoms in acyl groups for the following reaction:



(J)

The reaction is catalyzed by anionic micelles. In this case increasing the carbon content of the acyl group inhibits the reaction. Apparently, increasingly hydrophobic substrates are buried more and more deeply in the micelle such that they are inaccessible to the aqueous aldoxime anion.

8.7d Micellar and Enzyme Catalysis: Differences

Exploring the parallel between catalysis by enzymes and by micelles has been beneficial to both areas of research, but has also revealed two major differences between the two systems. First, the effectiveness of micelles in enhancing the rates of reactions is generally less than that of enzymes; for micelles, the effect rarely exceeds 100-fold. Second, micellar catalysts are less specific than enzymes. The lock-and-key model for enzyme activity implies specific binding sites in a protein molecule with a configuration maintained by covalent bonds. The corresponding features in micelles are zones of a particular polarity in rapidly changing equilibrium structures. Although the parallels between enzymes and micellar catalysts will undoubtedly continue to be profitably explored, it seems likely that future developments of micellar catalysts will probably emphasize synthetic chemistry.

8.8 REVERSE MICELLES

Examination of Figure 8.7—taken as a prototype of ternary phase diagrams involving surfactants—shows L_1 and L_2 phases occupying roughly symmetrical regions in the aqueous and organic corners of the diagram, although the sizes and shapes of these two areas are very different. The interpretation offered in Section 8.6 for these regions is that the minor component is solubilized in surfactant micelles dispersed in the major component as a solvent. Until now we have focused attention on aqueous micelles, emphasizing their relatively nonpolar cores—where nonpolar molecules are solubilized—and their polar surfaces. In the L_2 region of Figure 8.7, solubilized micelles are also present, but with the orientation of the surfactant molecules reversed. In nonaqueous media, amphipathic molecules cluster with their polar heads together in the micellar core and their tails in the organic continuous phase. Water is solubilized in the core of these structures, which are known as *reverse micelles*, terminology that emphasizes their difference from aqueous micelles.

Far more work has been done on aqueous than on nonaqueous micelles, partly because the number of amphipathic species that dissolve in nonpolar solvents is considerably fewer than in water-soluble surfactants. Aerosol OT, Triton X-100, and various Spans—the structures of which are given in Table 7.2—are examples of substances that have been widely used in the study of reverse micelles. Fundamental understanding of these structures is most likely to come from research involving pure, well-defined chemicals. Commercial products—petroleum sulfonates and the like—are widely studied from an applications point of view, but the interpretation of such experiments is often clouded by the uncertain effect of impurities. Even though we have approached this discussion from the perspective of water-containing systems, the behavior of anhydrous systems is obviously of interest. Through solvation of polar groups and hydrogen bonding, water can promote aggregation in nonpolar media; complete removal of water as an impurity in both the solvent and the surfactant can be very difficult. Since two-component systems are simpler than those with three components, let us consider some anhydrous solutions of surfactant in nonaqueous solvents.

8.8a Surfactant Aggregation in Nonaqueous Media

Surfactant aggregation in an anhydrous, nonpolar medium differs in several important respects from aggregation in water. The most apparent of these differences is that the hydrophobic effect plays no role in the formation of reverse micelles. The amphipathic species are relatively passive in aqueous micellization, being squeezed out of solution by the water. In contrast, surfactant molecules play an active role in the formation of reverse micelles, which are held together by specific interactions between head groups in the micellar core.

The differences in the solubility parameters (Chapter 3, Section 3.4b) of the hydrocarbon tail of the surfactant and the solvent have also been examined as contributing to reverse micelle formation. It is interesting to note that for the micellization of potassium benzene sulfonate in heptane, $\Delta H = -79.5$ kJ mole^{-1} and $\Delta S = -62.8$ J K^{-1} mole^{-1}. In contrast to aqueous systems, spontaneous micellization is largely due to a large *negative enthalpy change* with an unfavorable entropy change opposing micellization.

Another striking difference between aqueous and anhydrous, nonaqueous systems is the size of the aggregates that are first formed. As we have seen, n is about 50 or larger for aqueous micelles, while for many reverse micelles n is about 10 or smaller. A corollary of the small size of nonaqueous micelles and closely related to the matter of size is the blurring of the CMC and the breakdown of the phase model for micellization. Instead, the stepwise buildup of small clusters as suggested by Reaction (D) is probably a better way of describing micellization in anhydrous systems. When the clusters are extremely small, the whole picture of a polar core shielded from a nonaqueous medium by a mantle of tail groups breaks down.

It is sometimes argued that the "reverse micelle" terminology is an inappropriate comparison to aqueous micelles. Since water can be solubilized by these micelles, causing an increase in n, the reverse micelle model and vocabulary do seem useful for ternary systems.

Proton NMR has been used to measure both the onset of micellization and the tendency

toward solubilization in nonaqueous surfactant solutions. In an NMR spectrum the observed displacement of a resonance frequency—the chemical shift δ—depends on the environment of the molecule. Also, protons in the methylene groups of surfactants absorb at a different point in the spectrum than, say aromatic protons, so a surfactant can be readily monitored in, say, a benzene solution. If δ for a proton in a surfactant molecule is plotted against the concentration of the amphipathic species, a graph showing two linear regions of different slope results. In contrast to the properties of aqueous systems (Fig. 8.2), the transition between the two linear regions is gradual, but extrapolation to their point of intersection defines the CMC. If a solubilized molecule has an identifiable feature in an NMR spectrum, then solubilization can also be monitored by NMR. Example 8.6 explores this further.

* * *

Example 8.6 *A Nuclear Magnetic Resonance Study of Solubilization in a Reverse Micelle.* Let δ be the chemical shift produced by the proton of a solubilized molecule and differentiate between the experimental shift (subscript *exp*) and the δ's produced by molecules in bulk (subscript 0) or micellar (subscript *m*) environments. Arguing by analogy with the derivation of Equation (24) shows that

$$\delta_{exp} = (\delta_0 + \delta_m K[M])/(1 + K[M])$$

where K is the binding constant for the interaction of the solubilizate and the micelle. Suggest how experimental NMR data can be interpreted to give a quantitative value for K.

Solution: Incorporation of solubilizate B into micelle M can be represented as

$$B + M \overset{K}{\rightleftarrows} MB$$

for which K is the equilibrium constant. Define f_0 and f_m as the fraction of the solubilizate in the bulk solution and in the micelle, respectively. It follows by analogy with Equation (22) that $\delta_{exp} = f_0\delta_0 + f_m\delta_m = f_0\delta_0 + (1 - f_0)\delta_m$ since $f_m + f_0 = 1$.

By analogy with Equation (23),

$$K = f_m[B]_0/[M]f_0[B]_0 = (1 - f_0)/f_0[M]$$

with $[B]_0$ = total solubilizate concentration.

Finally,

$$\delta_{exp} = (\delta_0 + \delta_m K[M])/(1 + K[M])$$

with $[M] = (c - c_{CMC})/n$ by analogy with Equation (27).

This can be rearranged by analogy with Equation (28) to give

$$1/(\delta_{exp} - \delta_m) = 1/(\delta_m - \delta_0) + [n/K(\delta_m + \delta_0)][1/(c - c_{CMC})]$$

If δ is measured for different concentrations of surfactant and δ_0 and c_{CMC} are known, then this result can be plotted in a linear form and K/n evaluated from the intercept/slope ratio (compare Equation (31)). ■

* * *

8.8b Some Uses of Reverse Micelles

Some of the amphipathic species that have been used in the investigation of reverse micelles are capable of dissociation under suitable conditions. Metal carboxylates, alkyl aryl sulfonates, sulfosuccinates, and alkyl ammonium salts are examples of compounds with a high degree of ionic character. Coulomb's law describes the force between two charges q_1 and q_2 separated by a distance r in a medium of relative dielectric constant ε_r

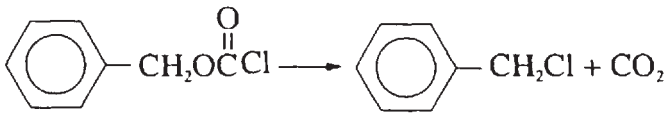
$$F_{Coul} \propto \frac{q_1 q_2}{\varepsilon_r r^2} \tag{33}$$

Electrolyte dissociation is expected to decrease as this force increases. This dissociation—as measured by pK for the dissociation process—should be inversely proportional to ε_r when a given electrolyte is studied in a series of solvents of various polarities. Since r in Equation (33) is the sum of anion and cation radii, the above statement is true only if the state of solvation

of the ions remains the same as the solvent varies. Alkali metal salts of dinonyl naphthalene sulfonic acid show decreasing dissociation as the solvent changes from ethanol ($\varepsilon_r = 24.3$) to acetone ($\varepsilon_r = 20.7$) to ethyl acetate ($\varepsilon_r = 6.02$). The quantitative magnitude of the difference between solvents is not the same for all cations, however, indicating that solvation of the cation also enters the picture. In solvents of very low polarity—like hydrocarbons ($\varepsilon_r = 2$–3)—dissociation is negligible.

The situation with respect to dissociation makes the small core of the reverse micelle a unique microenvironment, approaching the properties of ionic crystals, but at the same time readily accessible to solubilizates and reactants. This is why the interior of the reverse micelle is so effective in solubilizing water and explains why anhydrous systems are so difficult to obtain. Under nearly anhydrous conditions, as few as one solubilized water molecule per reverse micelle might be obtained. This degree of solubilization corresponds to 10 moles of solute per 0.018 kg water, or about 550 molal—an intriguing microenvironment for the investigation of catalytic effects! In addition, for amphipathic molecules with, say, weakly basic polar groups like sulfonates, cations with various strengths as Lewis acids, can be used as counterions.

One investigation of the catalytic activity of such a system used the reverse micelles formed in decane by sulfosuccinates with various cations. This medium was used to study the reaction



(K)

in which the benzylchloroformate reactant is basic and should therefore react strongly with acidic cations. The results were qualitatively similar to those shown in Figure 8.9 and could be analyzed similarly. The rate constants were evaluated at different temperatures, and activation energies were determined by a standard Arrhenius plot of ln k versus $1/T$. Table 8.4 shows the kinetic parameters so determined when the cations in the reverse micelle were Na^+, Al^{3+}, Ce^{3+}, and Zn^{2+}. Since the slope in this kind of plot is proportional to the activation energy and since the activation energy is very different for various cations, the lines in the Arrhenius plots can cross at accessible temperatures. This makes the relative effectiveness of the various cations in catalyzing Reaction (K) a matter of temperature.

TABLE 8.4 Activation Energies and Arrhenius Preexponential Factors for Reaction (K) Catalyzed by Reverse Micelles Containing the Indicated Cation

Cation	E_a (kJ mole^{-1})	ln A
Na^+	129	25.3
Al^{3+}	148	32.2
Ce^{3+}	84	11.7
Zn^{2+}	267	71.4

Source: F. M. Fowkes, D. Z. Becher, M. Marmo, C. Silebi, and C. C. Chao, In *Micellization, Solubilization and Microemulsions*, Vols. 1 and 2, (K. L. Mittal, Ed.), Plenum, New York, 1976.

Use of reverse micelles in synthetic chemistry to improve the rate and the yield of reactions seems likely to be a fruitful area of research in the future. In addition to catalysis, several other applications of reverse micelles can be cited. Just as nonpolar dirt is solubilized in aqueous micelles, so, too, polar dirt that would be unaffected by nonpolar solvents may be solubilized into reverse micelles. This plays an important role in the dry cleaning of clothing. Motor oils are also formulated to contain reverse micelles to solubilize oxidation products in the oil that might be corrosive to engine parts.

8.9 EMULSIONS AND MICROEMULSIONS

The term *microemulsion* was coined in 1958 to describe a fairly specific class of colloidal systems. Before we discuss these, a brief discussion of the broader term *emulsion* seems in order. If two immiscible liquids are shaken together, they will ordinarily separate rapidly into two distinct layers that can be divided in, say, a separatory funnel. Although any immiscible liquids might be considered, in this discussion we refer to the two as oil (abbreviation O) and water (abbreviation W). Next, instead of merely shaking the two liquids together, suppose we add a surfactant—often called an *emulsifying agent* in this context—and then vigorously mix the components in a blender or homogenizer of some sort. The milling together of the constituents causes one to be dispersed (the inner phase) in the other (the continuous phase): An emulsion is produced. There can be many variations in this procedure—in the nature of the components, their proportions, the milling process, the temperature, and so on; however, a few broad generalizations are possible.

1. The dispersed particles are spheres of great polydispersity. As noted in Appendix C (Section C.3b), this is generally the case in dispersions prepared by comminution such as this.
2. The average particle size is at the upper end of the colloidal size range (on the order of micrometers), and the particles are usually visible in a light microscope. We shall describe these as *coarse emulsions* when we want to emphasize their size range. Because the particles are relatively large and polydisperse, coarse emulsions look white when examined visually.
3. Emulsions are two-phase systems and—because of the free energy associated with the oil-water interface—are thermodynamically unstable with respect to separation into oil and water layers.
4. Oil may be the dispersed phase and water the continuous phase—designated an O/W emulsion—or water may be dispersed in oil (W/O). The form obtained depends on the specifics of the system, including the temperature. Compatibility with either oil or water on dilution is an easy way of establishing which phase is continuous.
5. The surfactant is adsorbed at the oil-water interface in the oriented fashion of monolayers. Judging from monolayer studies at the air-water interface, saturating the surface with surfactant lowers the surface tension γ by 25–50 mN m^{-1} (Fig. 7.6).
6. If the surfactant is ionic and imparts a charge to the interface, then the dispersed particle will be surrounded by an ion atmosphere. We see in Chapters 11 and 13 how an ion atmosphere surrounding a particle may slow down the rate at which such particles come together. This is one of the ways by which an emulsion may achieve some degree of kinetic stability.

Many of these concepts were introduced in Chapter 1 and seem fairly straightforward as abstract propositions. Things are not always so clear in concrete instances, however. The situation of microemulsions is a case in point.

8.9a Microemulsions

Historically, the term *microemulsion* was applied to systems prepared by emulsifying an oil in aqueous surfactant and then adding a fourth component, called a *cosurfactant*, generally an alcohol of intermediate chain length. Benzene, water, potassium oleate, and hexanol might be the components of a typical microemulsion formulation. What is observed experimentally is that the usual milky emulsion becomes transparent on addition of the alcohol. Light scattering and an assortment of other techniques reveal that the resulting system consists of either O/W

or W/O dispersions with particles having diameters in the 10- to 100-nm size range. Whether oil or water is continuous, the extent of uptake of the other component may be appreciable. In summary, the following differences between microemulsions and coarse emulsions should be noted:

1. Microemulsions contain particles at least an order of magnitude smaller than those in coarse emulsions.
2. Microemulsions are clear, and coarse emulsions are cloudy.
3. Microemulsions form spontaneously; coarse emulsions ordinarily require vigorous stirring.
4. Microemulsions are stable with respect to separation into their components; coarse emulsions may have a degree of kinetic stability, but ultimately separate.

The term *microemulsion* seems quite firmly established as the name for the sort of system described above. Still, there has been and continues to be a great deal of controversy as to the exact nature of these systems and the suitability of this vocabulary. The word *emulsion* implies the presence of two phases with an interfacial free energy associated with the phase boundary. Their small size makes the specific area large for microemulsions with a large free energy contribution from γ. Both the spontaneous formation of microemulsions and their stability with respect to separation are hard to reconcile with these considerations. It has even been suggested that the mixed film of surfactant and cosurfactant make the interfacial free energy negative. Alternatively, the increase in overall free energy with decreasing particle size may be offset by a favorable and hence negative $T\Delta S$ term in which ΔS describes the entropy of mixing the microemulsion particles with molecules of the dispersion medium. Since the number of microemulsion particles increases with decreasing particle size, the $T\Delta S$ term becomes more favorable with decreasing size. This idea is hard to apply quantitatively because of uncertainty as to the value—or meaning—of γ in these systems.

8.9b Microemulsions Viewed As "Swollen Micelles"

A totally different way of looking at microemulsions—and one that connects this topic with previous sections of the chapter—is to view them as complicated examples of micellar solubilization. From this perspective, there is no problem with spontaneous formation or stability with respect to separation. Furthermore, ordinary and reverse micelles provide the basis for both O/W and W/O microemulsions. From the micellar point of view, it is the phase diagram for the four-component system rather than γ that holds the key to understanding microemulsions.

The difference in perspective between the emulsion and micellar points of view is suggested by Figure 8.11. The small aggregate on the left represents a micelle with little or no solubilization. From left to right, the "particles" increase in size owing to increasing solubilization. The circle on the right represents an emulsion particle: an oil drop with a monolayer of surfactant on the surface. An actual continuum of states such as that suggested in Figure

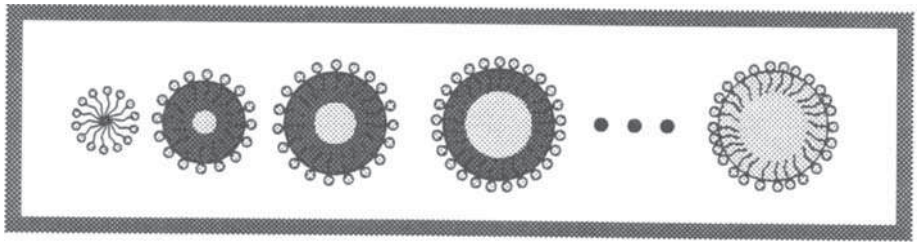


FIG. 8.11 Schematic progression from micelle (at left) to emulsion droplet (at right). Various degrees of solubilization, including microemulsions, lie between the two extremes.

8.11 is not physically attainable—most emulsions are prepared by comminution rather than condensation—but microemulsions do lie between the extremes. The conflicting schools of thought concerning microemulsions arise from the difference in perspective: One side looks from the emulsion point of view, and the other from the micellar point of view. *Swollen micelles* is another term for microemulsions; this terminology clearly reflects the different perspective of its originators.

8.9c Phase Diagrams for Microemulsions

We saw in Section 8.6 that phase diagrams are an effective way of representing the complex behavior of surfactant systems. Let us take a look at microemulsions in terms of phase diagrams. It turns out that nonionic surfactants form microemulsions at certain temperatures without requiring cosurfactants. Since only three components are present, these have somewhat simpler phase diagrams; this kind of system offers a convenient place to begin.

Figure 8.12 is a composite of both experimental and schematic, interpretive portions. The rectangular diagram shows the experimental behavior of the system water-cyclohexane-polyoxyethylene-(8.6)-nonyl phenol ether. All systems studied contained 5% surfactant with variable proportions of cyclohexane and water. Temperature was the experimental variable, and the nature of the phases present was recorded as the temperature was changed. The presence of different equilibrium phases is represented by different areas on the rectangular diagram, which is divided diagonally by a pair of lines that trace a ribbonlike pattern. This pair of lines cross twice, so the ribbon is divided into three areas. These, plus the regions

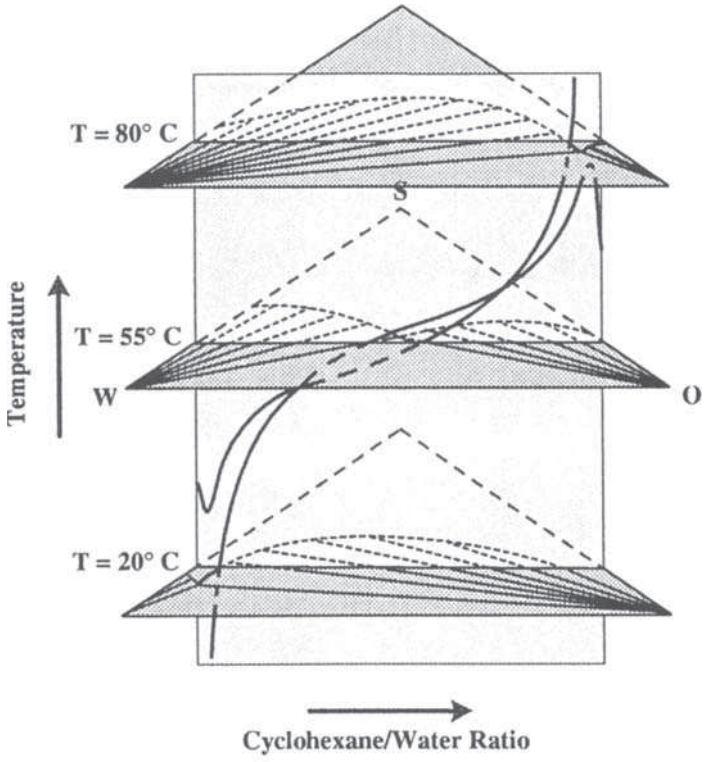


FIG. 8.12 Rectangular figure shows the phase diagram for water-cyclohexane (at 5% surfactant) versus temperature. Superimposed ternary phase diagrams offer an interpretation of the phases present. (Redrawn, with permission, from M. L. Robbins, In *Solution Chemistry of Surfactants*, Vols. 1 and 2 (K. L. Mittal, Ed.), Plenum, New York, 1979.)

above and below the ribbon, define five different phase situations encountered at various temperatures and compositions.

The idealized ternary phase diagrams that have been superimposed on the experimental plot in Figure 8.12 help us understand the five different regions of the rectangular diagram. The stacked triangles represent the oil-water-surfactant (abbreviation S) phase diagrams at different (constant) temperatures. Since the experimental data were collected at 5% surfactant, the rectangular plot slices through the triangles 5% of the distance toward S from the O-W base of the triangle. Example 8.7 offers practice in reading triangular phase diagrams and helps identify the phases present in the five different regions of the rectangular composition-temperature diagram.

* * *

Example 8.7 *Interpreting Phase Diagrams of Microemulsions.* At each of the three temperatures for which a ternary phase diagram is provided, describe the phases present when the diagram is crossed from the W-S side to the O-S side in the slice at 5% S. Describe the homogeneous phases in terms of the apparent solubilization.

Solution: Recall that on ternary diagrams the radial lines are tie lines representing two-phase regions. The triangular regions are three-phase regions, and the remaining area toward the S apex of the triangles is a homogeneous phase.

At 20°C, moving from left to right, we cross from a one-phase region (homogeneous micellar, O/W) into a two-phase region (oil plus homogeneous micellar, O/W).

At 55°C, moving from left to right, we cross from a two-phase region (water plus homogeneous micellar, O/W) into a three-phase region (oil plus water plus homogeneous micellar). Next we enter a different two-phase region (oil plus homogeneous micellar, W/O).

At 80°C, moving from left to right, we cross from a two-phase region (water plus homogeneous micellar, W/O) to a one-phase region (homogeneous micellar, W/O). ■

* * *

Since the micellar phase changes from water-continuous to oil-continuous with increasing temperature, it is an intriguing question how to describe the micelles in the three-phase region that exists at intermediate temperatures.

In this study it is the homogeneous micellar phases that comprise the microemulsions. For this 5% surfactant system it is only over a relatively narrow range of temperatures that it is possible to have fairly extensive solubilization of either oil or water in the micellar solutions.

It could be argued that the system described in Figure 8.12 is so different from those to which the name microemulsion was first applied as to make the figure irrelevant to the present discussion. However, one of the ways to represent a four-component phase diagram (p and T constant) is to use a trigonal prism in which one of the components—rather than T—varies along the rectangular faces. Figure 8.13a is a partial phase diagram for the system water-benzene-potassium oleate-pentanol. In this figure homogeneous micellar solutions are represented by unshaded areas in the individual triangles. These micellar systems are equivalent to the homogeneous areas in the ternary phase diagrams in Figure 8.12. At equilibrium, up to 50 wt% benzene can be incorporated into the system with only minor variations in the maximum water uptake. The resulting four-component microemulsion is the same as that produced by first emulsifying the oil and water with potassium oleate and then adding pentanol. This is true because we are considering equilibrium phase diagrams. As with all thermodynamic conclusions, the diagram tells nothing about the rate at which equilibrium is achieved.

Figure 8.13b shows the same data as Figure 8.13a, replotted with the variables interchanged. Note how the homogeneous area expands in size with increasing potassium oleate concentration up to 0.44 mole kg^{-1}. At still higher concentrations of potassium oleate the homogeneous area shrinks as other surfactant phases compete for the components.

8.10 SOME APPLICATIONS OF MICROEMULSIONS

Systems in which one liquid phase is finely dispersed in another under the stabilizing influence of one or more additional components find applications in countless areas. As consumers, we encounter many of these every day. Floor waxes, shaving lotions, beverage concentrates,

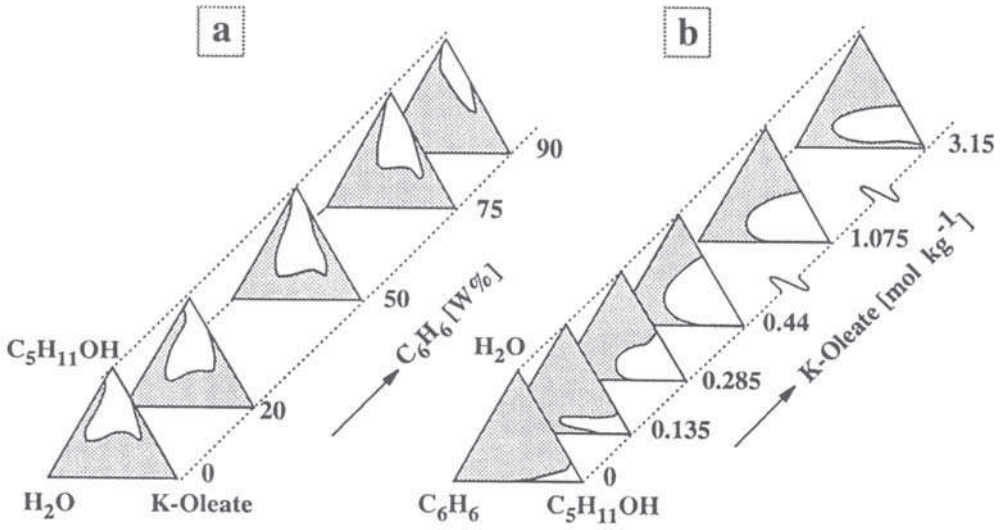


FIG. 8.13 Two representations of a portion of the phase diagram for the water-benzene-potassium oleate-pentanol system. The unshaded regions represent homogenous solutions. (a) Redrawn, with permission, from S. Friberg and I. Burasczeska, *Prog. Colloid Polym. Sci.*, **63**, 1 (1978). (b) Redrawn, with permission, from C. U. Herrmann, U. Wurz, and M. Kahlweit, In *Solution Chemistry of Surfactants*, Vols. 1 and 2 (K. L. Mittal, Ed.), Plenum, New York, 1979.

pesticide preparations, cold creams, and pharmaceutical products are a few of the more common examples. In recent years a great deal of research in this area has been directed toward the problem of *tertiary* oil recovery.

8.10a Tertiary Oil Recovery

First, the recovery of oil from natural reservoirs occurs in three stages. During the *primary* recovery stage the pressure of natural gases in the reservoir pushes the oil out. When the gas pressure is no longer adequate, water is pumped into the reservoir to force the oil out. This is called water flooding and represents the *second stage* of oil recovery. Primary and secondary oil recovery leave about 70% of the total oil in place, much of it trapped in the pore structure of the reservoir by capillary and viscous forces. It is estimated that in U.S. oil fields alone 300 billion barrels of oil are not recoverable by primary or secondary processes, and that of this 25 billion to 60 billion barrels are potentially recoverable through some tertiary process.

Numerous methods have been explored to recover at least some of this vast resource. Injection of oil-miscible fluids, gases under high pressure, and steam—either separately or in combination—have all been tried with various degrees of success. This is where microemulsions enter the picture. Under optimum conditions an aqueous surfactant solution—which may also contain cosurfactants, electrolytes, polymers, and so on—injected into an oil reservoir has the potential to solubilize the oil, effectively dispersing it as a microemulsion.

Any attempt to represent the trapped oil by a manageable model is bound to be an oversimplification. To see qualitatively, however, how capillary forces trap the oil and how surfactant solutions offer a potential for freeing it, imagine a cylindrical pore containing a slug of oil. Furthermore, assume the oil is in contact with water and that the interface is hemispherical. This assumption about the shape of the interface makes the water-oil-rock contact angle zero and is equivalent to neglecting θ ($\cos 0 = 1.0$). Although a primitive picture of oil in a rocky reservoir, this model can be described by a single size parameter r, the radius of the pore and the radius of curvature of the interface. The Laplace equation (Equation (6.29)) may then be used to describe the pressure across the oil-water interface, a pressure that must be exceeded to displace the oil. According to the Laplace equation, $\Delta p \propto \gamma/r$ and,

since r is small, Δp will be large unless r is offset by a small value of γ. Using r values that are sensible for geological structures, it has been estimated that $\Delta p \approx 500$ psi per foot if γ is on the order of 10 mN m^{-1}. This kind of pressure drop is unattainable under field conditions. Working backward and taking Δp as an attainable 1–2 psi per foot, the Laplace equation shows that γ must be less than about 0.1 mN m^{-1}, preferably closer to 10^{-3} mN m^{-1} for effective oil displacement.

Any surfactant adsorption will lower the oil-water interfacial tension, but these calculations show that effective oil recovery depends on virtually eliminating γ. That microemulsion formulations are pertinent to this may be seen by reexamining Figure 8.11. Whether we look at microemulsions from the emulsion or the micellar perspective, we conclude that the oil-water interfacial free energy must be very low in these systems. From the emulsion perspective, we are led to this conclusion from the spontaneous formation and stability of microemulsions. From a micellar point of view, a "pseudophase" is close to an embryo phase and, as such, has no meaningful γ value.

It is apparent that extrapolating laboratory studies on microemulsions to oil recovery is a formidable task. While laboratory research is conducted with pure solutes, distilled water, and at constant temperature, these are meaningless in the field, where the following applies:

1. The oil itself is a complex mixture containing surface-active components.
2. Commercial petroleum sulfonates, *a mixture of compounds*, are the most widely used surfactants.
3. Groundwater contains dissolved minerals, and, in practice, brine is used as the aqueous component.

In addition, viscosity considerations may be as important or more important than capillarity; fortunately, microemulsions also have relatively low viscosities!

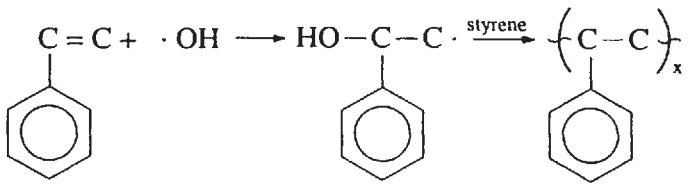
Despite the obvious difficulty of the tertiary oil recovery problem, this is a major area of surfactant research since the potential rewards for success are very great.

8.10b Polymer Synthesis

Another area of microemulsion application is in the synthesis of certain polymers. The process is called *emulsion polymerization*, a misnomer since micelles rather than emulsion drops are the site of the polymerization reaction. Because of the commercial importance of polymers, this process has been extensively researched and is quite well understood. We only consider some highlights of the process.

Emulsion polymerization is applicable only to monomers that are relatively insoluble in water, such as styrene. A coarse emulsion of monomer in aqueous surfactant is prepared with a water-soluble initiator, say, H_2O_2 in the solution. The surfactant concentration is above the CMC, so surfactant molecules are present as monomers, micelles, and emulsifiers at the oil-water interface. Even an insoluble liquid like styrene dissolves in water to some extent. Therefore the monomer is present in coarse emulsion drops, solubilized in micelles, and as dissolved molecules in water. A schematic illustration of the distribution of surfactant, monomer, and polymer in an emulsion polymerization process is shown in Figure 8.14.

The H_2O_2 molecules undergo thermal decomposition to form hydroxyl free radicals ·OH that initiate the polymerization. The overall reaction for the polymerization of styrene can be represented as



(L)

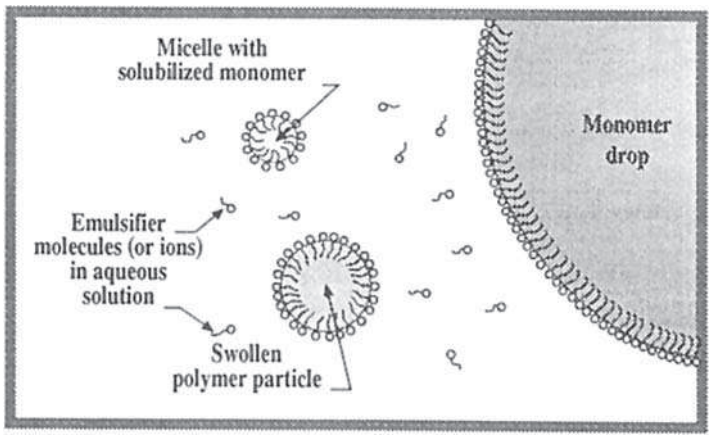


FIG. 8.14 Schematic representation of the distribution of surfactant and monomer in an emulsion polymerization. (Redrawn, with permission, from J. W. Vanderhoff, E. B. Bradford, H. L. Tarkowski, J. B. Shaffer, and R. M. Wiley, *Adv. Chem.*, **34**, 32 (1962).)

When x is large, the uniqueness of the end group(s) can be ignored; familiar polystyrene is the product.

In emulsion polymerization the first step in Reaction (L) takes place in water between dissolved monomer and initiator fragments. The resulting free radical is solubilized in micelles, in which it quickly reacts with solubilized monomer to form polymer. The low concentration of monomer in the aqueous phase prevents this from occurring to any appreciable extent in the water, although, by diffusion, there continues to be a flux of monomer from emulsion drops into micelles. Likewise, any polymerization that occurs in the coarse emulsion drops themselves is insignificant because of the much greater numerical abundance of micelles than the emulsified styrene droplets. The polymer chains grow in micelles until the process is terminated by reaction with another radical. Thus polymer growth is either propagating or terminating in micelles at any time; therefore half the micelles in a reaction mixture contain growing chains under stationary-state conditions. Both the rate of polymerization and the average molecular weight of the polymer depend on the surfactant concentration—via the concentration of micelles—in emulsion polymerization, while this has no effect on polymerizations conducted in nonmicellar solutions.

The aqueous polymer dispersion that results from emulsion polymerization is called a *latex*. In applications, the polymer may be separated, or the latex may be used directly as in paints and floor coatings.

As the conversion to polymer proceeds, the micelles become progressively more and more swollen by the polymer-monomer mixture. As with other microemulsions, it eventually becomes problematic as to whether the resulting dispersed particles should be called micelles or swollen polymer particles with adsorbed surfactant.

8.11 BIOLOGICAL MEMBRANES

It is neither feasible nor appropriate in a book like this to give a detailed presentation of biological membranes, which compartmentalize living matter and perform numerous cell functions as well. However, because of the impetus to the study of surfactants that the membrane-mimetic properties of surfactant structures have provided, it would be a mistake to exclude some mention of membranes in this chapter. We have already noted in connection with Figure 7.7 that a monolayer may collapse into a bilayer that leaves the surfactant in a tail-to-tail configuration. This is exactly the arrangement of molecules in the lipid portion of a cell

membrane. Protein molecules are arrayed in or on this lipid bilayer. Vignette 1.2 presented in Chapter 1 discusses this briefly. More details may be found in Bergethon and Simons (1990) and Goodsell (1993).

Figure 8.15 is a sketch of one possible relationship between the lipid bilayer and the membrane proteins. Molecules are free to move laterally in these membranes; hence the structure pictured in Figure 8.15 is called the *fluid mosaic model* of a cell membrane.

Cell membrane lipids are natural surfactants and display most of the properties of synthetic surfactants. The principal difference between these molecules and the surfactants that we discussed above in the chapter is that lipids contain two hydrocarbon tails per molecule. Table 8.5 shows the general structural formula of these cell membrane lipids and the names and formulas for some specific polar head substituents. The alkyl groups in these molecules are usually in the C_{16}–C_{24} size range and may be either saturated or unsaturated.

Much of our understanding of the chemical aspects of cell membranes has been derived from model systems based on surfactants, especially membrane lipids. In this section we are primarily concerned with the use of monolayers, bilayers, and especially black lipid membranes and vesicles as cell membrane models.

Fundamental membrane research has benefited greatly from the study of monolayers. One of the most important discoveries from this sort of research is the very existence of *two-dimensional phases and phase transitions*. Generally, studies of the sort that can be carried out with monolayers and bilayers cannot be directly extended to living cells, but some exceptional cases have shown that the extrapolation is valid. For example, it is known from monolayer studies that the presence of unsaturated hydrocarbon chains in lipid monolayers prevents some phase transitions from occurring as the temperature is lowered. Certain mutants of *Escherichia coli* are unable to synthesize fatty acids and hence can be manipulated through the compounds they are provided as nutrients. Abnormal levels of saturated hydrocarbon can

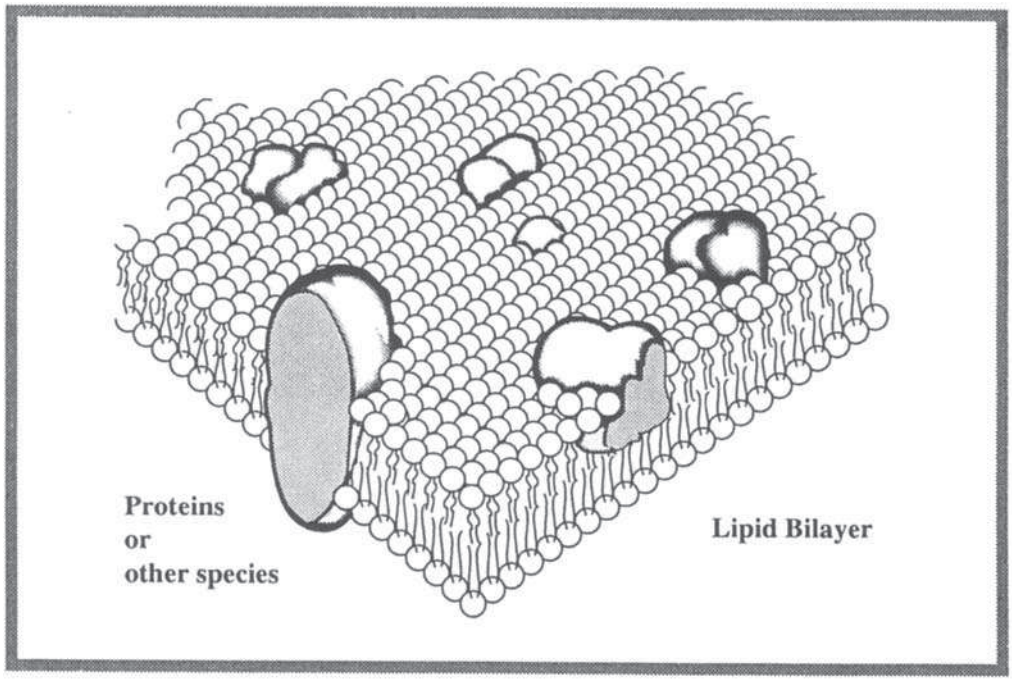
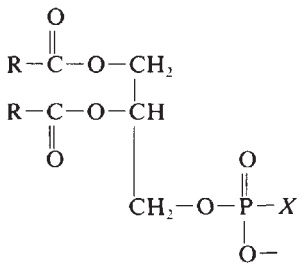


FIG. 8.15 Schematic representation of a biological membrane. The amphipathic phospholipid molecules form a bilayer with protein molecules embedded in it. (Redrawn, with permission, from S. J. Singer and G. L. Nicolson, *Science*, **175**, 720 (1972).)

TABLE 8.5 Names and Structures of Some Typical Phospholipid Surfactants



Name	X
Phosphatidic acid	$-OH$
Phosphatidylcholine	$-OCH_2CH_2\overset{+}{N}(CH_3)_3$
Phosphatidylethanolamine	$OCH_2CH_2NH_2$
Phosphatidylserine	$-OCH_2\underset{\displaystyle COOH}{\underset{\vert}{C}}HNH_2$
Phosphatidylthreonine	$-OCH-\underset{\displaystyle CH_3COOH}{\underset{\vert}{C}}H-NH_2$
Phosphatidylglycerol	$-OCH_2\underset{\displaystyle OH}{\underset{\vert}{C}}HCH_2OH$

thereby be introduced into the cell membranes of these organisms. Under appropriate conditions the same phase transitions occur in cell membranes as in model monolayers; still, the organisms survive. This is probably because patches of different two-dimensional phases segregate on the cell walls, with the unchanged regions carrying out normal biological functions.

Studies with monolayers and bilayers as models show that modifications of chain conformation can alter the effective thickness of the hydrocarbon portion of a bilayer; this, in turn, can have profound effects on its permeability. Suppose, for example, that a two-dimensional phase transition involved cooperatively inducing just one kink per molecule in otherwise fully extended chains. Such a kink shortens the extension of the chain by 0.127 nm, or thins a bilayer by 0.25 nm. We saw in Example 8.2 that this distance is about the length projected by two carbon-carbon bonds in the direction of the chain. Therefore, cooperative induction of such kinks has the same effect on the thickness of the hydrocarbon part of the bilayer as removing two carbon atoms from the tail groups.

Even closer to cell membranes than monolayers and bilayers are organized surfactant structures called *black lipid membranes* (BLMs). Their formation is very much like that of an ordinary soap bubble, except that different phases are involved. In a bubble, a thin film of water—stabilized by surfactants—separates two air masses. In BLMs an organic solution of lipid forms a thin film between two portions of aqueous solution. As the film drains and thins, it first shows interference colors but eventually appears black when it reaches bilayer thickness. The actual thickness of the BLM can be monitored optically as a function of experimental conditions. Since these films are relatively unstable, they are generally small in area and may be formed by simply brushing the lipid solution across a pinhole in a partition separating two portions of aqueous solution.

Because they are binary lipid films sandwiched in water, BLMs are excellent models for membranes in terms of electrical and permeability properties. Membrane potential, conductivity, and capacitance measurements have all been made on BLMs, and the effects on these quantities of electrolytes, proteins, and organic additives have all been studied. For example, 2,4-dinitrophenol, a known hydrogen ion carrier, has been observed to decrease the resistance significantly between electrodes on opposite sides of a BLM. Similarly, high applied voltages thin the BLM, with a simultaneous reversible increase in its permeability and conductance. The rate of permeation of water through BLMs is greater than that of nonpolar compounds; however, ions permeate BLMs more slowly than biological membranes.

The final surfactant structures we consider as models for biological membranes are *vesicles*. These are spherical or ellipsoidal particles formed by enclosing a volume of aqueous solution in a surfactant bilayer. When phospholipids are the surfactant, these are also known as *liposomes*, as we have already seen in Vignette I.3 in Chapter 1. Vesicles may be formed from synthetic surfactants as well. Depending on the conditions of preparation, vesicle diameters may range from 20 nm to 10 μm, and they may contain one or more enclosed compartments. A multicompartment vesicle has an onionlike structure with concentric bilayer surfaces enclosing smaller vesicles in larger aqueous compartments.

Phospholipid vesicles form spontaneously when distilled water is swirled with dried phospholipids. This method of preparation results in a highly polydisperse array of multicompartment vesicles of various shapes. Extrusion through polymeric membranes decreases both the size and polydispersity of the vesicles. Ultrasonic agitation is the most widely used method for converting the lipid dispersion into single-compartment vesicles of small size.

Phase transitions, electrical properties, and permeability have all been investigated for vesicle bilayers. Especially important are the molecular dynamics of transverse motion through the bilayer. Use of isotopic labels shows that lipid molecules can flip-flop from one surface to another in the bilayer. Techniques have been developed for incorporating protein and other molecules into liposomes, and it has been observed that proteins facilitate lipid flip-flop in the bilayer. Liposomes shrink and swell osmotically as the activity of water in the surrounding aqueous phase is changed by additives. Depending on the surface phase state, the addition of cholesterol to the liposome may decrease the water permeability of the bilayer in vesicles. Transport of ions through the vesicle walls is thought to occur either through channels or via carriers. Since these surfaces are orders of magnitude more permeable to H^+ and OH^- compared to other univalent ions, transmission through the vesicle wall by a hydrogen-bonded network of water molecules is proposed just as the high mobility of these ions in water is explained.

Only a few of the observed properties of vesicles have been enumerated here. The examples cited are sufficient, however, to illustrate how these "synthetic cells" are ideally suited as models for research into the structure and functioning of cell membranes.

REVIEW QUESTIONS

1. Explain why surfactant aggregates form in solution.
2. What is the *critical micelle concentration* (CMC), and how is it measured?
3. Do different types of measurements of CMC lead to the same value for CMC? Explain.
4. In aqueous systems, the enthalpy change due to micellization is usually positive, and micellization is driven by entropy change. Explain the reason for the positive entropy change.
5. What is meant by the *optimal head group area* of a surfactant? What is the *packing parameter* $\mathcal{P}$? Explain how packing considerations can be used to determine the possible shapes of the micellar aggregates.
6. Consider a spherical micelle with charged head groups. What changes in shape would you expect as you increase the concentration of added electrolytes?
7. Describe the details of the structure of a micelle.
8. What is a *palisade layer*? How does one determine if there is water penetration in a micelle and, if there is any, where the water molecules are?
9. Under what conditions can the formation of micelles be described in terms of *reaction-equilibrium formalism*? When is the *phase equilibrium model* appropriate?

10. What does *solubilization in a micelle* mean?
11. Draw a typical phase diagram for an oil-water-surfactant system and identify at least some of the phases that can be found.
12. Why is catalysis using micelles as "microreactors" advantageous sometimes? When is it advantageous?
13. What are *reverse micelles*? When would you expect them?
14. What are *microemulsions*? How do they differ from coarse emulsions?
15. What is a *swollen micelle*?
16. Give some examples of lipids that form biological membranes.
17. Why do biological membranes and bilayers usually consist of double-tailed surfactants?

REFERENCES

General References (with Annotations)

Ball, P., *Designing the Molecular World—Chemistry at the Frontier*, Princeton University Press, Princeton, NJ, 1994. (Undergraduate level. A *Scientific-American*-style tour of chemistry at the molecular level and its implications to science and technology. Chapter 7, entitled "A Soft and Sticky World: The Self-Organizing Magic of Colloid Chemistry," is an excellent and engaging overview of surfactant-and polymer-based colloidal systems of interest in technology, biology, and medicine.)

Clint, J. H., *Surfactant Aggregation*, Chapman and Hall, New York, 1992. (Graduate level. An advanced monograph on surfactant solutions and on self-assembly ["aggregation"] in surfactant solutions. A research reference at the graduate level.)

Evans, D. F., and Wennerström, H., *The Colloidal Domain: Where Physics, Chemistry, Biology, and Technology Meet*, VCH Publishers, New York, 1994. (Undergraduate and graduate levels. A textbook on colloids covering many of the topics discussed in the present book, but presented almost exclusively from the point of view of surfactant systems and self-assembly.)

Israelachvili, J. N., *Intermolecular and Surface Forces*, 2d ed., Academic Press, New York, 1991. (Undergraduate and graduate levels. This book is an excellent starting point for advanced concepts on modeling micellization and on molecular and surface forces that drive self-assembly. Graduate level, but portions of the material should be accessible to undergraduate students.)

Myers, D., *Surfaces, Interfaces, and Colloids: Principles and Applications*, VCH Publishers, New York, 1991. (Undergraduate level. A qualitative overview of micelles, microemulsions, and their applications.)

Rosen, M. J., *Surfactants and Interfacial Phenomena*, 2d ed., Wiley, New York, 1989. (Reference. A good source of reference for properties of surfactants and surfactant solutions. Brings a strong industrial perspective.)

Tanford, C., *The Hydrophobic Effect. The Formation of Micelles and Biological Membranes*, 2d ed., Wiley, New York, 1980. (Undergraduate level. A classic reference by a pioneer on the hydrophobic effect on the relevance of surfactants to biological membranes.)

West, D. R. F., *Ternary Equilibrium Diagrams*, 2d ed., Chapman and Hall, New York, 1982. (A "manual" of ternary phase diagrams. Concise.)

Other References

Bergethon, P. R., and Simons, E. R., *Biophysical Chemistry: Molecules to Membranes*, Springer-Verlag, New York, 1990.

Goodsell, S., *The Machinery of Life*, Springer-Verlag, New York, 1993.

Hartley, C. S., *Aqueous Solutions of Paraffin Chain Salts*, Hermann and Cie, Paris, 1936.

McBain, J. W., *Trans. Faraday Soc.*, **9**, 99 (1913).

Puvvada, S., and Blankschtein, D., *J. Phys. Chem.*, **93**, 7753 (1989).

Tatar, H. V., *J. Phys. Chem.*, **59**, 1195 (1955).

Wennerström, H., and Lindman, B., *Physics Reports*, **52**, 1 (1979).

PROBLEMS

1. Use the data for sodium dodecyl sulfate in Table 8.1 to test the equation developed in Example 8.1 and to evaluate α for these micelles. Criticize or defend the following proposition: The α

values in the table are more accurate than that evaluated here because they are based on individual light scattering experiments and reflect the variation of α with changing salt concentration; by contrast, the α value calculated here is an average value based on the assumption that m/n, and hence α, is independent of salt concentration.

2. Shinoda* examined the variation of the CMC of various potassium alkyl malonates, $RCH(COOK)_2$, with the addition of univalent salts. For $R = C_8$, C_{12}, C_{14}, and C_{16}, the log-log plots of CMC verus counterion concentration produce parallel straight lines of slope -1.12. Criticize or defend the following proposition: According to the analysis presented in Example 8.1, the slope of this type of plot equals $-(1 - \alpha)$, meaning that $\alpha = -0.12$; this negative function apparently means that the micelle binds an excess of counterions and has the opposite charge from that expected.

3. The spectra of substituted pyridinium iodides are characterized by charge transfer bands involving the interaction of pyridinium and iodide ions. Mukerjee and Ray† showed that this band is shifted about 90 nm toward the red for dodecyl pyridinium iodide, which forms micelles, compared to methyl pyridinium iodide, which does not. They measured λ_{max} for the micelles in mixed solvents of variable relative dielectric constant and obtained the following results:

λ_{max} (nm)	281	284	288	298
ε_r	43	38	33	34

Estimate the effective dielectric constant at the surface of the micelle from the fact that λ_{max} occurs at 286 nm for dodecyl pyridinium iodide micelles in water. In light of the value estimated in Section 8.3 for the dielectric constant in the vicinity of solubilized benzene, does it seem likely that the value of ε_r for bulk water applies in the Stern layer?

4. Ionescu et al.‡ measured the CMC of hexadecyl trimethyl ammonium bromide in water–dimethyl sulfoxide (DMSO) mixtures at 25 and 40°C:

Mole fraction of DMSO	CMC $\times 10^3$ (mole liter^{-1})	
	$T = 25$°C	$T = 40$°C
0.000	0.92	1.00
0.027	1.48	1.51
0.060	2.24	2.51
0.098	3.60	3.98
0.144	5.62	6.30
0.201	8.91	10.00
0.275	14.00	22.00
0.366	None	None

Use these data to evaluate ΔG^0_{mc} for hexadecyl trimethyl ammonium bromide and to estimate (remember, only two temperatures were measured) ΔH^0_{mc} and ΔS^0_{mc}. Discuss the values obtained in light of infrared and NMR experiments, which indicate formation of the stoichiometric compound $DMSO \cdot 2H_2O$ at $x_{DMSO} = 0.33$.

5. Both adsorption from solution and micellization occur as a result of the hydrophobic effect. To test the correspondence between these two effects. Rosen§ assembled ΔG^0 values for adsorption at the air-water interface and for micellization of a number of linear and branched surfactants. The following is a selection of these data:

---

*Shinoda, K., *J. Phys. Chem.*, **59**, 432 (1955).

†Mukerjee, P., and Ray, A., *J. Phys. Chem.*, **70**, 2144 (1966).

‡Ionescu, L. G., Tokuhiro, T., Czerniawski, B. J., and Smith, E. S., In *Micellization, Solubilization and Microemulsions*, Vols. 1 and 2 (K. Mittal, Ed.), Plenum, New York, 1979.

§Rosen, M. J., In *Micellization, Solubilization and Microemulsions*, Vols. 1 and 2 (K. Mittal, Ed.), Plenum, New York, 1979.

Compound	Temperature (°C)	ΔG^0 (kJ mole^{-1}) Adsorption	Micellization
n-C_9SO_4Na	25	−22.4	−17.3
n-$C_{14}SO_4Na$	25	−30.0	−25.2
p-n-$C_8\phi SO_3Na$	70	−27.7	−23.4
p-n-$C_{14}\phi SO_3Na$	70	−38.5	−34.1
n-$C_{10}SO_4Na$	50[a]	−27.8	−19.8
n-$C_{16}SO_4Na$	50[a]	−38.6	−30.4
p-n-$C_{12}\phi SO_3Na$	75	−35.6	−32.4
p-$C_6CHCH_2\phi SO_3Na$ (with C_4)	75	−34.8	−28.7
p-$(C-C-)_4\phi SO_3Na$ (with C)	75	−34.4	−27.5

[a]Measured at the hexane-water inferface.

Use these data to criticize or defend the following propositions: (a) The CMC is a good indicator of a surfactant's adsorption effectiveness since the ΔG^0 values for adsorption and micellization both show parallel changes with increasing chain length. (b) The dodecyl benzene sulfonates have some of the most favorable ΔG^0 values among the data shown; this shows that branching has no adverse effect on either adsorption or micellization.

6. Danbrow and Rhodes* used potentiometric titration data to determine the distribution of benzoic acid (*HB*) between water and nonionic micelles. The commercial surfactant used has the average formula $C_{16}(OC_2H_4)_{24}$. They obtained the following results in 4% surfactant solutions:

$[HB]_0$ (mmole liter^{-1})	4.055	8.804	13.63	18.18	24.30	28.21
HB_m (mmole)	0.1859	0.3497	0.5067	0.6660	0.7790	0.9460

These authors consider two postulates: (a) If the benzoic acid is solubilized in the micellar core, then dimerization should occur, in which case $HB_m/[HB]^2$ = const. (b) If the solubilization occurs at the micellar surface, then surface saturation analogous to Langmuir adsorption should occur. Test each postulate with the data provided and decide which gives the better fit.

7. Tokiwa and Aigami† used proton NMR to study the solubilization of benzyl alcohol, 2-phenyl ethanol, and 3-phenyl propanol in sodium dodecyl sulfate micelles. The upfield chemical shift of the aromatic protons in the micelle from their location in water was measured as a function of solubilization; it tends to increase with solubilization much like the results in Fig. 8.6. The following are some of the results obtained:

Compound	Chemical shift at 0.05 mole of solubilizate (100 g surfactant solution)$^{-1}$ (cps)	$(\text{alcohol})_m/(\text{alcohol})_w$	mole alcohol/ mole surfactant
ϕCH_2OH	3.0	2.28	4.48
ϕCH_2CH_2OH	8.5	4.35	4.08
$\phi CH_2CH_2CH_2OH$	17.0	11.80	2.81

Criticize or defend the following proposition: The more carbon atoms there are in the alcohols, the more hydrophobic these compounds become and the more enriched the micelles become relative to the aqueous phase; the magnitude of the chemical shift increases as the extent of solubiliation in the micelles increases owing to the diamagnetic effect of the phenyl groups.

*Danbrow, M., and Rhodes, C. T., *J. Chem. Soc.*, Supplement II and Indexes 6166 (1964).
†Tokiwa, F., and Aigami, K., *Kolloid Z. Z. Polym.*, **246**, 688 (1971).

8. Duynstee and Grunwald* present some experimental data for Reaction (F) in the presence of hexadecyl trimethyl ammonium bromide (CTABr, C = cetyl) and sodium dodecyl sulfate (NaLS, L = lauryl). Sodium hydroxide was the source of OH^- in all cases. A pseudo-first-order rate constant of $2.40 \times 10^{-2}\ s^{-1}$ is observed for k_{CTABr}. Use the following absorbance data to evaluate k_{NaLS} for this reaction:

t (min)	0	4	7	10	13	17	25	35	41	51
Absorbance	0.734	0.652	0.618	0.577	0.537	0.489	0.408	0.327	0.293	0.244
	60	68	85	124	· · ·					
	0.211	0.189	0.160	0.129	0.155					

Recall that for first-order kinetics a plot of ln (fraction unreacted) versus time has a slope $-k$. Also note that the reaction reaches an equilibrium characterized by an absorbance 0.115; the data must be corrected for this. For both the anionic and cationic micelles, qualitatively sketch, emphasizing the charge state, the micelle, the solubilized substrate, and the approaching OH^- reactant. Indicate how these pictures are consistent with the experimental rate constants.

9. In the same research described in Problem 8, the authors also examined the rate of Reaction (F) in pure water and in NaCl solution to guarantee that the ionic surfactants were not displaying an electrolyte activity effect as is ordinarily observed with ion combination reactions. For 10^{-5} M crystal violet and 0.01 M NaOH, they observed the following:

Solution	Pure water	0.01 M NaCl	0.01 M CTABr
$k \times 10^4\ (s^{-1})$	17.1	16.4	240

Use these data and the Debye-Hückel theory of electrolyte nonideality to criticize or defend the following proposition: Indifferent electrolytes always inhibit the rates of ion combination reactions because the activity coefficients are fractions. The data for CTABr show an enhancement of rate so this cannot be due to an activity effect. In these data, the k's for pure water and aqueous NaCl are essentially identical, so no activity effects operate in the absence of micelles either.

10. Bunton and Robinson† studied the effect of sodium dodecyl sulfate micelles on the rate of the reaction between OH^- and 2,4-dinitrochlorobenzene. These negative micelles have an inhibiting effect on the reaction, yet the kinetic data can be analyzed according to Equation (24). Use the following data and the authors' CMC value of 0.0064 M to estimate K/n, where K is the binding constant between the dodecyl sulfate micelles and the 2,4-nitrochlorobenzene:

$c_{NaC_{12}SO_4} \times 10^2$ (M)	0.0	1.41	1.85	2.80	3.90	5.60
$k \times 10^5$ (liter mole^{-1} s^{-1})	14.2	9.40	7.70	6.20	5.10	3.10

From the increased solubility of 2,4-dinitrochlorobenzene in sodium dodecyl sulfate solutions (without NaOH), the authors of this research estimate a K/n of 44. Criticize or defend the following proposition: The presence of OH^- could change the n value for sodium dodecyl sulfate micelles, so exact agreement between K/n values determined by the two methods is not necessarily expected.

11. The reaction of Problem 10 was studied at two different temperatures, and, from the temperature dependence of the rate constants, the authors† determined $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$, the enthalpy and entropy of activation, respectively. The following values of these parameters were obtained in pure water and in 0.01 M sodium dodecyl sulfate (NaLS) and 0.01 M hexadecyl trimethyl ammonium bromide (CTABr):

#Duynstee, E. F. J., and Grunwald, E., *J. Am. Chem. Soc.*, **81**, 4542 (1959).

†Bunton, C. A., and Robinson, L., *J. Am. Chem. Soc.*, **90**, 5972 (1968).

Solvent	$\Delta H‡$ (kcal mole^{-1})	$\Delta S‡$ (cal K^{-1} mole^{-1})
Water	21.3	−5.2
CTABr	16.4	−13.4
NaLS	21.3	−5.6

Use these values to criticize or defend the following proposition: The smaller endothermic value for $\Delta H‡$ in CTABr means the product molecules must be more readily expelled from these micelles, making the enthalpy contribution more favorable to the reaction in this case. The solubilized substrate has a higher entropy, so the decrease in entropy for the micellar reaction is larger. The last problem shows that the reaction occurs about twice as fast in water as in 0.01 M NaLS. The rate in water determines the kinetic parameters in the last case.

12. Chemistry students are certainly familiar with the regular tetrahedron (methane, sp^3 hybrids, etc.), but may not have considered this geometry as the basis for a quaternary phase diagram. McCarthy# discusses these as extensions of triangular phase diagrams: Each of the four faces of a regular tetrahedron is the ternary phase diagram that results as the concentration of component X in a system goes to zero. The opposite apex represents pure component X, and planes slicing through the tetrahedron parallel to any face have a constant percentage of X, the magnitude of which depends on their placement. Resketch one of the versions of Figure 8.13 using this kind of tetrahedral representation. This sould be done qualitatively and on a large enough scale to separate the various slices. Criticize or defend the following proposition: Both the tetrahedral and prismatic quaternary diagrams show essentially the same thing; in the tetrahedral diagram the advantage of having all four ternary diagrams contained therein is offset by the shrinking size of the slices as an apex is approached.

13. Proteins are polyamids formed from amino acids having the formula $H_2N-CHR-COOH$; therefore different R groups occur along the polymer backbone according to the amino acid sequence in the protein. Write structural formulas for the R groups in the following amino acids:

Alanine (Ala)*	Glycine (Gly)	Serine (Ser)*
Arginine (Arg)	Isoleucine (Ile)*	Threonine (Thr)
Asparagine (Asn)	Leucine (Leu)*	Tyrosine (Tyr)*
Aspartic acid (Asp)	Lysine (Lys)	Valine (Val)*
Glutamine (Gln)	Phenylalanine (Phe)*	

Kaiser and Kezdy† have suggested that helical structures for certain amino acid sequences in proteins may be stabilized by a hydrophobic interaction with cell membrane lipids. In the list above, R groups marked with an asterisk may be considered hydrophobic. Construct a model for residues 1–22 of human growth hormone releasing factor in the following way. Roll a piece of paper into a cylinder and sketch a helix on the surface. Evenly space the names of five consecutive amino acids along each turn of the helix. On a second cylinder, mark off portions of helix two turns long and carefully enter seven amino acids along each two-turn length. The models with 5 and 3.5 amino acid residues per turn are approximations of the π and α helical structures shown in protein crystals. Which of the two appears more effective in concentrating hydrophobic groups along one edge of the helix? The first 22 amino acids in this protein occur in the following order:

1. Tyr	9. Ser	16. Gln
2. Ala	10. Tyr	17. Leu
3. Asp	11. Arg	18. Ser
4. Ala	12. Lys	19. Ala
5. Ile	13. Val	20. Arg
6. Phe	14. Leu	21. Lys
7. Thr	15. Gly	22. Leu
8. Asn		

#McCarthy, P., *J. Chem. Educ.*, **60**, 922 (1983).

†Kaiser, E. T., and Kezdy, F. J., *Science*, **223**, 249 (1984).

14. On the basis of the model produced in Problem 13, criticize or defend the following proposition: A roughly 20-residue α-helix has a length approximately equal to the thickness of a membrane bilayer. A bundle of three, four, or more of these helices—with hydrophobic and hydrophilic sides—could orient themselves in such a way as to form a hydrophilic channel through the membrane (see Vignette I.2 and Figure 1.2). While the individual helices may be stable, their aggregation into this sort of channel involves bringing several membrane proteins together and is therefore entropically unfavorable.

9
Adsorption at Gas–Solid Interfaces

When I cut through your plane as I am now doing, I make your plane a section which you, very rightly, call a Circle. For even a sphere—which is my proper name in my own country—if he manifest himself at all to an inhabitant of Flatland—must needs manifest himself as a Circle.

From Abbott's *Flatland*

9.1 INTRODUCTION

9.1a Physisorption and Chemisorption

Adsorption at the solid-gas interface is traditionally subdivided into two broad classes: *chemisorption* and *physisorption* (i.e., physical adsorption).

- As the name implies, *chemisorption* comes very close to the formation of chemical bonds between the adsorbent (e.g., the solid) and the adsorbate (gas). In this case electron exchange between the adsorbent and the adsorbate occurs. Two consequences of this are that the associated heat effects are comparable to those that accompany ordinary chemical reactions and that the process is not always reversible. It is possible, for example, to adsorb (chemisorb) oxygen on carbon and desorb CO or CO_2.
- In *physical adsorption*, on the other hand, the energy effects are comparable to those that accompany physical changes such as liquefaction and are completely reversible for nonporous solids. In contrast to chemisorption, the adsorbent and the adsorbate in this case interact relatively weakly through van der Waals forces (see Chapter 10). Physical adsorption is the easier of the two types of adsorption and provides much background needed for an understanding of chemisorption.

A major portion of this chapter is concerned with physical adsorption, particularly from a global thermodynamic point of view. This is followed by a molecular-scale examination of crystalline surfaces and a brief discussion of chemisorption and its relevance to heterogeneous catalysis.

9.1b Focus of This Chapter

As should be evident from the discussions in Chapters 6 and 7, adsorption phenomena play a major role in colloid and surface chemistry. We also come across other examples in Chapters 11 and 13. Adsorption, especially at solid-gas interfaces, is very important in heterogeneous catalysis, as highlighted in Vignette IX. In this chapter, the focus is the introduction of quantitative measurement and the description of adsorption at solid-gas interfaces.

9.1b.1 Physisorption

There are several different ways in which the topics pertaining to physical adsorption can be subdivided. We are primarily concerned with nonporous solids, briefly discussing porous materials only in Section 9.7.

1. It is convenient to divide the extent of adsorption into three categories: submonolayer, monolayer, and multilayer. We discuss them in this order. The thermodynamics of adsorption may be developed around experimental isotherms or around calorimetric data. We begin with the definition of adsorption isotherms and how they are determined experimentally (Section 9.2).
2. Adsorption isotherms may be derived from a consideration of two-dimensional equations of state, from partition functions by statistical thermodynamics, or from kinetic arguments. Even though these methods are not fundamentally different, they differ in ease of visualization. We consider examples of each method in Sections 9.3 and 9.4.
3. Multilayer adsorption and the popular Brunauer-Emmett-Teller (BET) method of analysis are described in Section 9.5. This section also describes the determination of specific areas by gas adsorption. Low-temperature N_2 adsorption and the BET method of analysis are so widely used for this purpose that these topics will receive special attention.
4. A brief description of calorimetric analysis of adsorption follows in Section 9.6.
5. Adsorption in porous solids, an important topic in catalysis and other areas, is presented in Section 9.7, in which the adsorption hysteresis and capillary condensation are introduced.

9.1b.2 Chemisorption

The global thermodynamic approach used in the above sections is insensitive to details at the atomic level and can only yield a gross characterization of the surface. Properties such as the specific surface area and the presence or absence of pores can be determined using the above approach since only the average surface—not atomic details—is involved. The existence of a distribution of surface energy sites can also be inferred from adsorption data, but the method falls short when it comes to specifics about this distribution. Observations on an atomic scale are needed to learn more about the details of the surface structure. Such observations comprise the subject matter of the last two sections of the chapter.

1. A great assortment of high-technology (and high-cost!) methods exist for examining surfaces on a fine scale, and new methods are being developed continually, as we saw in Vignette I.8. The techniques available permit us to characterize solid surfaces in a variety of ways: pictorially, chemically, and crystallographically. Rather than attempt to catalog all of these and say a few things about each, it seems preferable to single out one example and develop it in a bit more detail. Accordingly, in Section 9.8 we discuss one such method—*low-energy electron diffraction* (LEED)—for studying surfaces.
2. There are numerous other techniques that we shall not discuss but that also tend to be known by their initials, as evident from Vignette I.8 on STM (scanning tunneling microscopy), SPM (scanning probe microscopy), and AFM (atomic force microscopy) and related discussions in Chapter 1. The literature on this subject contains what Adamson (1990) calls "a veritable alphabet soup of designations, many of which are contrived acronyms." Table 8.1 of Adamson (1990) (which is eight pages long) provides an excellent summary of the numerous techniques that are currently available for surface analysis. These include a wide variety of *spectroscopic* methods for analyzing the composition of a surface both qualitatively and quantitatively. An example is *Auger electron spectroscopy* (AES). What makes AES so suitable for chemical analysis of surfaces is the shallow escape depth of Auger electrons from solid surfaces. Only those electrons originating within two or three atomic layers of the surface are able to escape without suffering the sort of inelastic interactions that obscures the features of the spectrum. Several additional techniques, including the ones based on the scanning probe method highlighted in Vignette I.8, are described in Hubbard (1995).
3. Just as an assortment of experimental techniques exists, so too are there numerous

types of solid surfaces that might be examined. Not all methods can be used for all surfaces, and the techniques we consider are particularly well suited for the study of metal surfaces. Therefore we also limit our examples to metal surfaces (Section 9.9). Although this obviously excludes a number of interesting and important solid surfaces, it also includes some topics of great significance, including friction and corrosion in engineering materials, semiconductors for electronic devices, and catalytic surfaces for many chemical reactions.

VIGNETTE IX HETEROGENEOUS CATALYSIS AND CHEMICAL VAPOR DEPOSITION: Adsorption—Setting the Stage for Catalysis and Thin-Film Growth

Adsorption of gases at a solid surface can be a bother or a boon! It is a nuisance when one needs a very clean surface, devoid of any adsorbed species, for research or chemical analysis of a solid. On the other hand, it is a convenient tool in vacuum technology for pumping gases out of vacuum chambers (as in cryogenic pumping; see Section 9.8a.2). And, adsorption can be a vehicle for molecular engineering—as in the case of catalysis or in layer-by-layer fabrication of ordered films on surfaces! Perhaps more striking is the role of adsorption in the last context, and we highlight this in this vignette.

We have already noted in Vignette I.9 that selectivity of a catalyst is of extreme importance in practice. But, how does the surface of a solid serve as a catalyst? How does adsorption enter the picture? What issues are of interest in this context? We already touched on the answers to the first two questions in Vignette I.9. A further elaboration of these is presented schematically in Figure 9.1. The adsorption of a molecule on a surface contributes to the disassembly of the molecule and allows the molecule to react with the other reactants in the vicinity. Adsorption can either promote or *hinder* a reaction or may contribute to "side" reactions that generate undesirable products. This is the reason selectivity of a surface is of utmost importance. Depending on the strength of the adsorbate-adsorbent bond, an adsorbed molecule (or a disassembled part of it) may become free to wander along the surface and to participate in chemical reactions and in the nucleation and growth of islands of two-dimensional surface phases. Therefore, quantifying the energetics of adsorption, the amount of adsorbed material, the chemical nature of products (if any), and their relation to the atomic-level structure of the surface itself become issues of paramount importance not only in catalysis, but also in the technology of thin-film growth on surfaces. These issues serve as the motivation for the types of theoretical and experimental procedures we develop in this chapter and for the interest in tools such as scanning probe microscopy and spectroscopy highlighted in Vignette I.8.

As is evident from the above comments, adsorption is also the embryonic event that precedes nucleation and growth of oriented films—*epitaxial films*—on surfaces (if the conditions are right; see Figure 9.2). Chemisorption is again the stage-setter in this case (as is often the case in heterogeneous catalysis). Epitaxial growth (which is not restricted to gas-solid systems) has significant industrial importance and depends on, among other things, interactions among adsorbed molecules, adsorbent/adsorbed-molecule interactions, structure of the surface (also known as the substrate), temperature, and surface diffusion of adsorbed molecules.

Many of the issues relevant to the above examples of adsorption (and the attendant) processes are beyond the scope of the present chapter. The topics we cover—such as adsorption isotherms, relating bulk phase conditions to the extent of adsorption and surface phases, and low-energy electron diffraction (LEED) for studying surface structure—are the essential first steps in understanding the problems and the prospects engendered by adsorption phenomena.

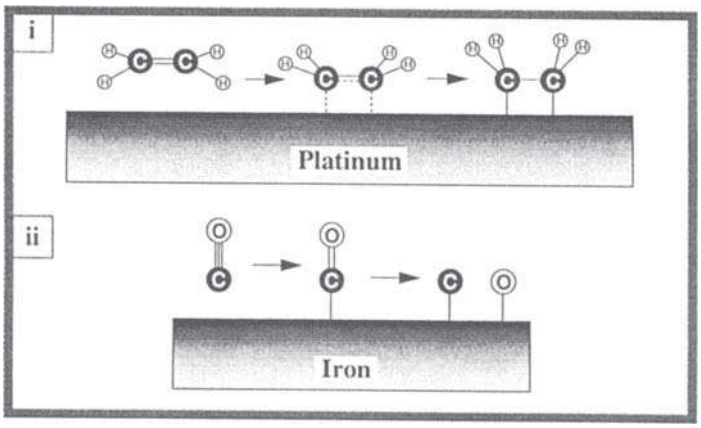


FIG. 9.1 A schematic illustration of the role of a substrate on adsorption and catalysis. Sketch (i), on chemisorption of ethylene on platinum, shows the breaking of C=C double bond so that the carbon atoms can form bonds to the surface. Sketch (ii) shows how iron substrate disassembles carbon monoxide into carbon and oxygen atoms. Breaking these bonds in gaseous phases requires a considerable amount of energy, whereas a catalytic surface allows these under moderate conditions. The broken fragments then may wander on the surface and react with other fragments or molecules in the gas. (Adapted and redrawn with permission from Ball 1994.)

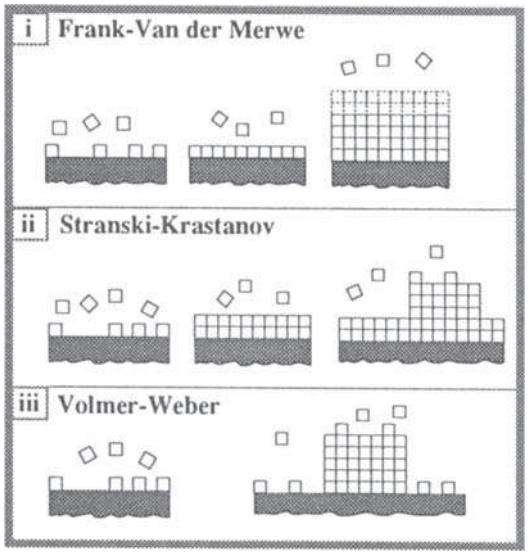


FIG. 9.2 A schematic representation of three topologically different epitaxial growth modes for crystallization on surfaces through adsorption. Sketch (i) illustrates layer-by-layer growth, known as *Frank-Van der Merwe growth*. Sketch (iii) illustrates just the opposite: Crystallites nucleate and grow on immediate contact between the substrate and the adsorbate. This is known as *Volmer-Weber growth*. Sketch (ii), known as *Stranski-Krastanov growth*, is an intermediate mode. (Adapted from A. Zangwill, *Physics at Surfaces*, Cambridge University Press, Cambridge, England, 1988.)

9.2 EXPERIMENTAL AND THEORETICAL TREATMENTS OF ADSORPTION: AN OVERVIEW

9.2a Adsorption Isotherms

Adsorption experiments are conducted at constant temperature, and an empirical or theoretical representation of the amount adsorbed as a function of the equilibrium gas pressure is called an *adsorption isotherm*. Adsorption isotherms are studied for a variety of reasons, some of which focus on the adsorbate while others are more concerned with the solid adsorbent. In Chapter 7 we saw that adsorbed molecules can be described as existing in an assortment of two-dimensional states. Although the discussion in that chapter was concerned with adsorption at liquid surfaces, there is no reason to doubt that similar two-dimensional states describe adsorption at solid surfaces also. Adsorption also provides some information about solid surfaces. The total area accessible to adsorption for a unit mass of solid—the specific area A_{sp}—is the most widely encountered result determined from adsorption studies. The energy of adsorbate-adsorbent interaction is also of considerable interest, as we see below.

9.2b Adsorption: Some Experimental Considerations

We saw in Chapter 6, Section 6.7, that solid surfaces are notoriously heterogeneous, particularly with respect to roughness and chemical composition. From the point of view of specific area determination, roughness is not too troublesome since the adsorbed gas molecules can generally cover the hills and valleys of the surface with ease. Pores with very small dimensions pose more of a problem. For now, we assume that such pores are absent; we take up the question of adsorption on porous solids in Section 9.7. Chemical heterogeneity affects the energetics of adsorption. For simplicity, we often assume that the surface is characterized by a single adsorption energy. Actually, a distribution of surface sites with differing adsorption energies may be present, and some indication of this may be extracted from adsorption data. As an approach to surface characterization, adsorption studies are indirect and give average rather than specific descriptions. We have already seen in Vignette I.8 that solid surfaces can be probed more directly for information on a molecular scale using scanning probe spectroscopy; we see below in this chapter (Section 9.8) how low-energy electron diffraction can be used to study the structure of surfaces on a molecular scale. For solids with high specific area, however, gas adsorption is the method of choice for quantifying this feature; we have more to say about it in Section 9.5.

9.2b.1 Experimental Determination of Adsorption Isotherms

Adsorption studies for the experimental determination of adsorption isotherms are conducted in a vacuum apparatus from which all gases can be removed prior to the addition of the adsorbate being studied. After the solid is introduced into the sample tube and the tube is attached to the vacuum line, it is generally pretreated by some sort of degassing procedure. This is a combination of heating and pumping to ensure the removal of physically adsorbed contaminants. Consideration must be given to the possibility of changing the solid when the temperature of degassing is selected. A heat treatment that is too vigorous may result in changes in any chemisorbed layer, which—from our point of view at least—amounts to a change in the adsorbent itself. If extensive enough, such changes may alter the surface area of a solid. Even less drastic changes are sufficient to alter the adsorption energy of a surface.

The range of pressures over which adsorption studies may be conducted is—in principle—from zero to p_0, the saturation pressure or the normal vapor pressure of the material at the temperature of the experiment. At the low-pressure end of this range adsorption will be slight, so the determination of the isotherm involves measuring small differences in pressure at low pressures. This is not easy to do experimentally, although relatively modern low-pressure techniques have greatly extended this region. As the pressure approaches p_0, adsorption often increases rapidly as if anticipating phase separation at the surface (resulting from multilayer adsorption in which most of the adsorbed molecules behave as if they were in the bulk liquid

state). As a matter of fact, if the solid is porous, the vapor may actually condense in the small pores at $p < p_0$ (see Section 9.7).

Figure 9.3 is a sketch of an apparatus that can be used to determine the equilibrium extent of gas adsorption as a function of pressure. We outline how such an experiment is conducted at ambient temperature, even though adsorption studies are frequently conducted at low temperatures, particularly when determination of A_{sp} is the objective of the experiment. A known mass of adsorbent is introduced into the sample tube and degassed as described above. Then the following set of pressure-volume readings are made, described here in terms of Figure 9.3.

1. The sample tube and gas burette are evacuated and then a nonadsorbing gas—frequently helium—is introduced into the gas burette. The burette is graduated with respect to volume and also serves as one leg of a manometer so that both the volume and pressure of gas in the burette can be measured. Ambient temperature is assumed to apply throughout.
2. The three-way stopcock is opened to connect the gas burette with the sample tube. The new pressure and volume are read. From this, the volume of the dead space—the volume beyond the three-way stopcock that is not occupied by sample—can be determined (see Example 9.1).
3. The nonadsorbed gas is pumped out and replaced by the adsorbate. Its volume, pressure, and temperature are measured, and from these the number of moles of gas introduced (initially) into the apparatus can be determined.
4. The three-way stopcock is opened to connect the burette and sample tube, and volume and pressure are measured again. Since the dead space is known, the (final) number of moles of gas can be calculated. The difference between the initial and final number of moles gives the number of moles adsorbed.
5. This amount of adsorbed gas is in equilibrium with bulk gas at the pressure read in Step 4.

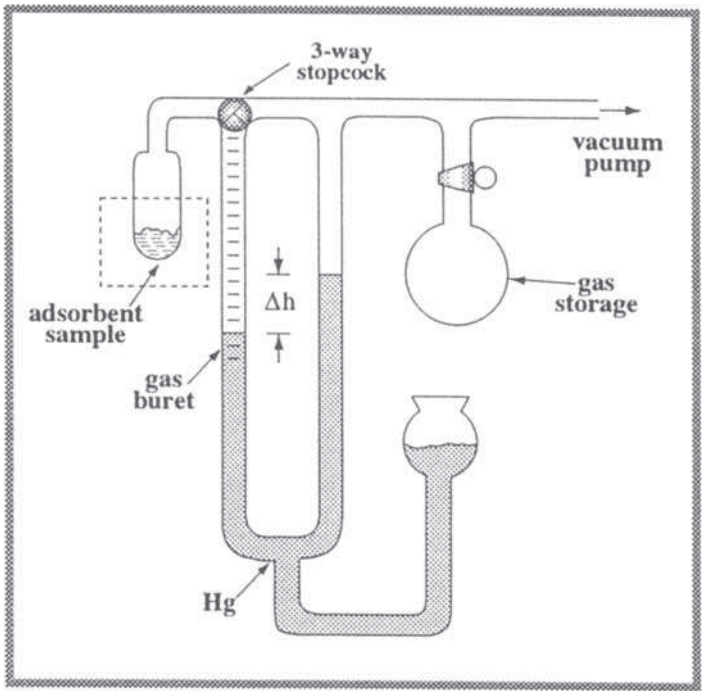


FIG. 9.3 Schematic illustration of a gas adsorption apparatus.

These steps describe the determination of a single point on an adsorption isotherm. By adjusting the mercury level, we can increase the pressure of the equilibrium gas with more adsorption occurring. Steps 3 and 4 are thus repeated until the full isotherm is mapped. Example 9.1 illustrates numerically how a point on the isotherm is established.

* * *

EXAMPLE 9.1 *Construction of Adsorption Isotherms*: The following pressure-volume (p-V) data were collected at a temperature of 22°C. The V's are volumes in the gas burette, and the numerical subscripts refer to the steps itemized above.

With helium, $p_1 = 21.71$ torr, $V_1 = 12.90$ cm^3, $p_2 = 16.50$ torr, $V_2 = 10.90$ cm^3. With adsorbate, $p_3 = 12.85$ torr, $V_3 = 13.70$ cm^3, $p_4 = 3.24$ torr, $V_4 = 5.00$ cm^3. What is the volume of the dead space? How many moles are adsorbed at the final equilibrium pressure, 3.24 torr?

Solution: Successive applications of the ideal gas law allow us to calculate the desired quantities.

The total volume to which the gas has access after the stopcock is opened is the sum of the burette volume and the dead space V_d. Therefore $V_d + V_2 = p_1V_1/p_2 = (21.71)(12.9)/(16.50) = 16.97$, or $V_d = 6.07$ cm^3. *It is convenient to use* $R = 62{,}360$ cm^3 torr K^{-1} mole^{-1} as the value of the gas constant in these calculations. The initial moles of adsorbate are given by $n_i = p_3V_3/RT = (12.85)(13.70)/(62360)(295) = 9.57 \cdot 10^{-6}$. After adsorption equilibrium is established, $n_f = (3.24)(5.00 + 6.07)/(62360)(295) = 1.95 \cdot 10^{-6}$. The difference $n_i - n_f = (9.57 - 1.95) \cdot 10^{-6} = 7.62 \cdot 10^{-6}$ mole gives the amount adsorbed at an equilibrium pressure of 3.24 torr. ■

* * *

If the adsorption isotherm is to be determined at some temperature other than room temperature—liquid nitrogen temperature, for example—the sample tube is placed in a suitable thermostat. This is indicated by the dotted line in Figure 9.3. In this case two sets of readings are made with the nonadsorbed gas, one at room temperature and one with the thermostat in place. In this way the partitioning of the dead space between the two temperature regions can be determined. Several additional considerations should be cited that are important in actual practice:

1. The gases used for adsorption must be of high purity.
2. The gas burette should itself be thermostated unless the laboratory has very good temperature control.
3. The nonideality for the adsorbate at low temperatures must be taken into account unless the volume of the low-temperature dead space is minimized.
4. Sufficient time must be allowed for equilibrium to be established. A way of checking this experimentally is to lower the pressure and observe that desorption follows the same curve as adsorption.

9.2b.2 Classification of Adsorption Isotherms

Experimental gas adsorption isotherms are traditionally classified into one of the five types shown in Figure 9.4.

- The Type I isotherm is reminiscent of Figure 7.16, the *Langmuir isotherm*. The plateau is interpreted as indicating monolayer coverage. We see that this type of behavior implies a sufficiently specific interaction between adsorbate and adsorbent to be more typical of chemisorption than physical adsorption.
- Type II adsorption, by contrast, is widely observed with physical adsorption and is interpreted to mean multilayer adsorption.
- We see in Section 9.5b (Example 9.5) that Type III behavior occurs when the heat of liquefaction is more than the heat of adsorption.
- Types IV and V are analogs of Types II and III, except for the leveling off that occurs at pressures below p_0. These cases are associated with porous solids in which the adsorbate condenses in the small pores at $p < p_0$.

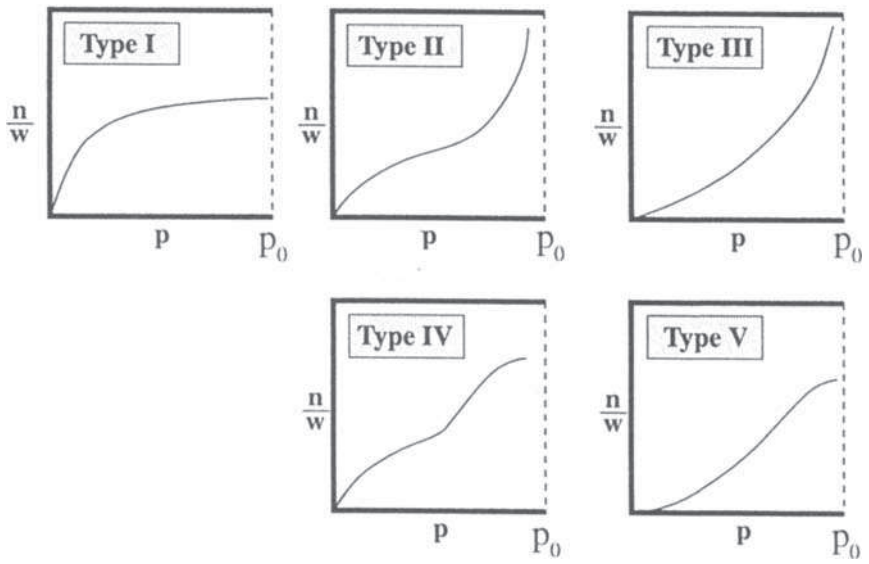


FIG. 9.4 Qualitative shapes of the five general types of gas adsorption isotherms. (See text for a discussion of their physical significance.)

If the specific area of the solid is the information sought, it is the amount of adsorption at monolayer coverage that must be measured. This is readily available in Type I adsorption, but requires considerable interpretation when multilayer adsorption takes place. Once determined, however, the number of molecules required to saturate a surface times the area occupied per molecule gives the surface area of a sample. This divided by the mass of the sample gives A_{sp}. In addition, once the adsorption at monolayer coverage is identified, all other extents of adsorption can be expressed as fractions or multiples of the monolayer.

9.2c Adsorption Isotherms: Theoretical Considerations

This makes a convenient point of contact with theory since models for adsorption inevitably subdivide the surface into an array of adsorption sites that gradually fill as the pressure increases. If θ is defined as the fraction of sites filled, then $\theta = 1$ corresponds to monolayer coverage, with $\theta < 1$ or $\theta > 1$ to submonolayer and multilayer coverages, respectively. Theoretical isotherms predict how θ varies with p in terms of some particular model for adsorption. It turns out that a set of experimental points can often be fitted by more than one theoretical isotherm, at least over part of the range of the data; that is, theoretical isotherms are not highly sensitive to the model on which they are based. A comparison between theory and experiment with respect to the temperature dependence of adsorption is somewhat more discriminating than the isotherms themselves.

Since it is relatively easy to fit experimental adsorption data to a theoretical equation, there is some controversy as to what constitutes a satisfactory description of adsorption. From a practical point of view, any theory that permits the amount of material adsorbed to be related to the specific surface area of the adsorbent and that correctly predicts how this adsorption varies with temperature may be regarded as a success. From a theoretical point of view, what is desired is to describe adsorption in terms of molecular properties, particularly in terms of an equation of state for the adsorbed material, where the latter is regarded as a two-dimensional state of matter.

As we see in the course of the chapter, these two approaches frequently clash. The adsorption isotherm of Brunauer, Emmett, and Teller (BET), which is discussed in Section 9.5, is an excellent example of this. The model on which the BET isotherm is based has been criticized by many theoreticians. At the same time, the isotherm itself has become virtually

the standard equation for determining specific areas from gas adsorption data. Ross (1971) summarizes the situation effectively by comparing it to a master chef who concocts a palatable dish out of an old shoe. For some, the end result is what matters: a palatable dish. For others, the starting material dominates their opinions: the old shoe. To present a relatively accurate picture of the current state of affairs in this area, it is necessary to present both of these viewpoints. We attempt not to take too one-sided a position, but to give some indication of each side.

9.3 THERMODYNAMICS OF ADSORPTION: PHENOMENOLOGICAL PERSPECTIVE

9.3a Relating Equations of State to Isotherms

To see how the equation of state of two-dimensional matter and the adsorption isotherm are related, we return to the Gibbs equation (Equation (7.46)):

$$-d\gamma = \Gamma_2 d\mu_2 \tag{1}$$

Since we are concerned here with adsorption from the gas phase, the chemical potential may be related to the pressure of the gas by

$$\mu_2 = \mu_2^0 + RT \ln (fp) \tag{2}$$

where f is the activity coefficient (see Equations (7.47) and (7.48)). In this discussion we assume that the gas behaves ideally, although in analyzing experimental results it may be necessary to include the correction required by the nonideality of the gas. Combining Equations (1) and (2) leads to

$$-d\gamma = RT\Gamma_2 \, d \ln p \tag{3}$$

In the present context it is convenient to use Equation (7.43) to eliminate Γ_2 from Equation (3) and express the surface excess as the number of moles adsorbed per unit area:

$$-d\gamma = \frac{n}{A} RT \, d \ln p \tag{4}$$

Next we substitute the product of sample weight times specific area for A in this equation to obtain

$$-d\gamma = \frac{RT}{A_{sp}} \frac{n}{w} d \ln p \tag{5}$$

For a given adsorbent and an isothermal experiment, T and A_{sp} are constants. The ratio n/w is the equilibrium amount adsorbed that will be a function of p. Therefore Equation (5) may be integrated as follows:

$$-\int d\gamma = \frac{RT}{A_{sp}} \int \frac{n}{w} d \ln p \tag{6}$$

The constant of integration may be evaluated by recognizing that n/w goes to zero as p approaches zero. Under these circumstances $\gamma \rightarrow \gamma_0$; therefore Equation (6) becomes

$$\gamma_0 - \gamma = \pi = \frac{RT}{A_{sp}} \int_0^p \frac{n}{w} d \ln p \tag{7}$$

If the experimental isotherm (n/w as a function of p) is known, then Equation (7) may be integrated either analytically or graphically to give the two-dimensional pressure as a function of coverage. This relationship therefore establishes the connection between the two-and three-dimensional pressures that characterize the surface and bulk phases. This is how adsorption data could be used to determine the film pressure in equilibrium with a drop of bulk liquid on a solid surface as discussed in Section 6.6b.

9.3b Ideal Behavior: The Henry Law Limit

As an illustration of the kind of information obtainable from Equation (7), suppose we consider the situation in which the equilibrium adsorption of a gas is described by the isotherm

$$n/w = mp \tag{8}$$

where m is a constant. Equations (7) and (8) may be combined to give

$$\pi = \frac{RT}{A_{sp}} \int_0^p m\,p\,\frac{dp}{p} = \frac{RT}{A_{sp}}\,m\,p = \frac{N_A\,k_B\,Tn}{A_{sp}\,w} \tag{9}$$

Since the quantity $(A_{sp}w/nN_A)$ equals σ, area per molecule, Equation (9) may be written

$$\pi\sigma = k_B T \tag{10}$$

the two-dimensional ideal gas law for the surface phase!

The adsorption isotherm—Equation (8)—associated with this surface equation of state is called the *Henry law limit*, in analogy with the equation that describes the vapor pressure of dilute solutions. The constant m, then, is the adsorption equivalent of the Henry law constant. When adsorption is described by the Henry law limit, the adsorbed state behaves like a two-dimensional ideal gas.

9.3c Deviations from Ideality

Equation (8) may also be written as

$$\theta = m'p \tag{11}$$

if the specific area of the adsorbent and the cross-sectional area of the adsorbate are known (Equation (7.71)). We see, therefore, that compliance with Henry's law implies that a log-log plot of θ versus p yields a straight line of unit slope. Figure 9.5 shows some experimental results for adsorption at 77.4K plotted in this way for pressures down to 10^{-10} torr. Note that even at these low pressures the Henry law limit is not yet reached (it would give a 45° line in Fig. 9.5), although argon on Pyrex appears to be approaching this limiting behavior. Note, further, that any errors introduced in evaluating θ would vertically shift the curves but would not change their slopes. We may conclude, therefore, that—just as with monolayers on aqueous substrates—the compliance of an adsorbed layer with the ideal gas law is a form of behavior that is extremely difficult to observe. The implication of this is that significant departures from two-dimensional ideality already set in at very low surface coverage.

9.3c.1 Adsorbent-Adsorbate Interactions

In general, there are two different types of interactions in which adsorbed molecules may participate. They are the interaction between the adsorbed molecules and the adsorbent and the interaction between the adsorbed molecules themselves. In the Henry law region, we have seen that the adsorbed layer behaves ideally. This is to be expected in view of the low surface concentration (θ very small) of adsorbate. The adsorbed molecules definitely do interact with the adsorbent, however, and at very low coverage the interaction energy might be very sensitive to surface heterogeneity. Any "hot spots" on the surface would adsorb first, less energetic patches next, then the normal surface sites. In Sections 9.4b and 9.6 we see how isotherms measured at several different temperatures may be interpreted to yield information on the energy of adsorption.

9.3c.2 Adsorbate-Adsorbate Interactions

The second type of interaction possible for adsorbed molecules is direct adsorbate-adsorbate interaction. Interactions of this sort are expected to lead to deviations from ideality in the two-dimensional phase just as they lead to deviations from ideal behavior for bulk gases. In this case surface equations of state, which are analogous to those applied to nonideal bulk gases, are suggested for the adsorbed molecules. The simplest of these allows for an excluded area correction (see Equation (7.23)):

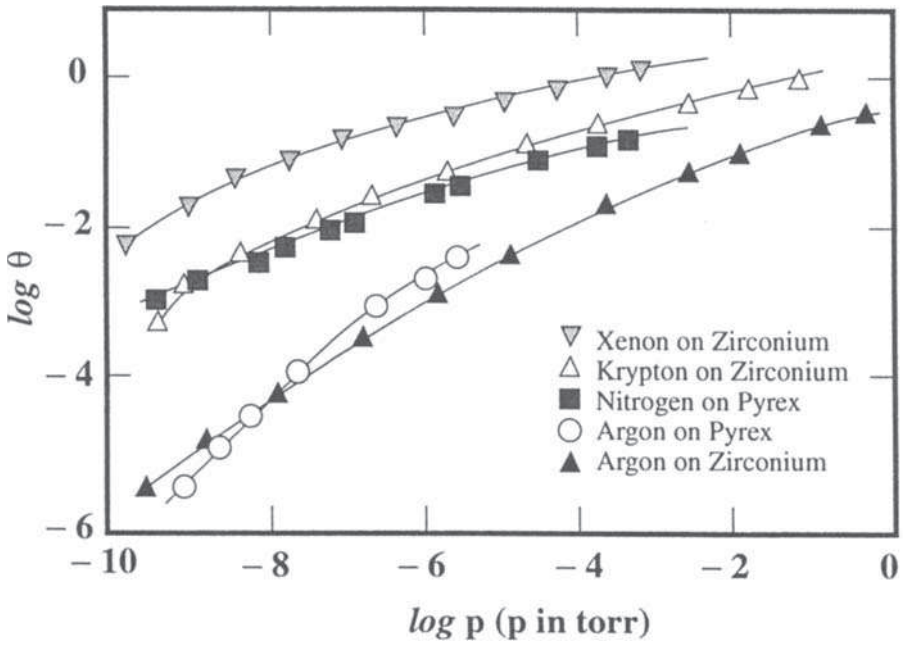


FIG. 9.5 A log-log plot of θ versus p for xenon, krypton, and argon on zirconium and nitrogen and argon on Pyrex. (Redrawn with permission from J. P. Hobson, "Physical Adsorption at Extremely Low Pressures." In *The Solid-Gas Interface*, Vol. 1 (E. A. Flood, Ed.), Marcel Dekker, New York, 1967.)

$$\pi(\sigma - \sigma^0) = k_B T \tag{12}$$

The use of the above equation of state to construct the corresponding adsorption isotherm is illustrated in Example 9.2.

* * *

EXAMPLE 9.2 *Evaluating Isotherms from Equations of State: Finite-Size Effects of Adsorbates.* Use Equation (12) along with Equation (7) to derive an adsorption isotherm that accounts for excluded-area adsorbate-adsorbate interactions.

Solution: According to Equation (7), $d\pi = -d\gamma$. *We now recall that*

$$\sigma = \frac{A_{sp} w}{n N_A}$$

Therefore Equation (5) may be written as

$$\frac{\sigma \, d\pi}{k_B T} = d \ln p$$

From Equation (12), $d\pi$ is given by

$$d\pi = -k_B T(\sigma - \sigma^0)^{-2} \, d\sigma$$

Combining the above two equations we get

$$-\frac{\sigma \, d\sigma}{(\sigma - \sigma^0)^2} = d \ln p$$

which on integration leads to

$$\ln\left(\frac{\theta}{1-\theta}\right) + \frac{\theta}{1-\theta} = \ln p + C$$

where

$$\sigma/\sigma^0 = 1/\theta$$

Next we evaluate the integration constant C. We know that $\theta \to 0$ as $p \to 0$. Therefore, at first glance, we are tempted to equate C to zero. It must be remembered, however, that Henry's law must apply as $p \to 0$. This condition is met if we let $C = \ln m'$, with m' defined by Equation (11):

$$\ln\left(\frac{\theta}{1-\theta}\right) + \frac{\theta}{1-\theta} = \ln p + \ln m'$$

This may be readily verified by examining the limit of the above equation as $\theta \to 0$. This equation may also be written

$$m'p = \left(\frac{\theta}{1-\theta}\right)\exp\left(\frac{\theta}{1-\theta}\right)$$

by taking the *antilog* of both sides of the equation. This is the equation for the adsorption isotherm we want. Two observations concerning this isotherm are worth noting here:

1. It is interesting to look at this form of the isotherm in the limit of small values of θ but still above Henry's limit. In this case the exponential term approaches unity, and the isotherm becomes

   $$m'p = (\theta/[1-\theta])$$

   or

   $$\theta = \frac{m'p}{1+m'p}$$

   which is identical in form to the Langmuir equation (see Equation (7.67)).
2. The isotherm also reveals that for $m'p \gg 1$, θ approaches unity as an upper limit. Thus at both the upper and lower limits, the isotherm gives the same results as the Langmuir equation. At intermediate values the two functions differ slightly, but it would probably be difficult to distinguish between them in fitting experimental data. ■

* * *

9.3c.2a Van der Waals Isotherm and Surface Phases. This approach can be extended to other forms of two-dimensional equations of state. For example, still greater nonideality in the two-dimensional equation of state might be represented by the *van der Waals* analog:

$$\left(\pi + \frac{a}{\sigma^2}\right)(\sigma - b) = k_B T \tag{13}$$

in which the van der Waals factor b and σ^0 from Equation (12) are identical in principle. Table 9.1 shows the isotherm resulting from this equation; the table also lists the result of Example 9.2 and one other example that leads to what is known as the *Harkins-Jura isotherm.* We return to the Harkins-Jura isotherm below in this section.

What makes the two-dimensional van der Waals equation of state especially interesting is the fact that there exists a temperature above which there is only one real root to Equation (13) and below which some values of π correspond to three values of θ. This situation is shown schematically in Figure 9.6a. With bulk gases and also insoluble monolayers on water, the three-root region is identified with a region of two-phase equilibrium. Is there any evidence for this type of phase equilibrium in the two-dimensional layer of adsorbed gas on a solid substrate? Figure 9.6b shows several data for the adsorption of krypton on specially treated graphite over a range of temperatures from about 77 to 91K. These plots are adsorption isotherms, not π versus σ diagrams, but nevertheless there is quite clear evidence of a two-phase region. A few additional assumptions lead to the prediction that the two-dimensional critical temperature should be one-half the value of the three-dimensional critical temperature

TABLE 9.1 Examples of Adsorption Isotherms Based on Some Two-Dimensional Equations of State

Two-dimensional equation of state	Isotherm	Comments
$\pi(\sigma - \sigma^0) = k_B T$	$m'p = \left(\frac{\theta}{1-\theta}\right)\exp\left(\frac{\theta}{1-\theta}\right)$	1. The equation of state accounts for only the excluded-area interaction. 2. Reduces to Equation (11) as $\theta \to 0$. 3. Reduces to the Langmuir equation for low θ's above the Henry limit.
$\left(\pi + \frac{a}{\sigma^2}\right)(\sigma - b) = k_B T$	$m'p = \frac{\theta}{1-\theta}\exp\left(\frac{\theta}{1-\theta} - \frac{2a\theta}{b k_B T}\right)$	1. The coefficients a and b are two-dimensional analogs of the three-dimensional van der Waals coefficients. 2. Reduces to the above case when $a = 0$.
$\pi = -C_1\sigma + C_2$	$\ln p = -\frac{C_1}{2N_A RT}A_{sp}^2\left(\frac{n}{w}\right)^{-2} + C_3$	1. Surface behaves like an incompressible bulk phase. 2. Known as the *Harkins-Jura Isotherm.*

according to this model. The two T_c values for Kr are 86K and 210K, and the corresponding values for argon are 65 and 151K.

There is another feature in Figure 9.6b that deserves additional comment. It is the existence of a second set of vertical segments in the isotherms at values of θ in the range of 0.65–0.78. This suggests a second phase equilibrium in the two-dimensional matter, perhaps something analogous to the transition between the *LE* and *LC* states of monolayers on water (see Chapter 7) or to a liquid-solid transition. Indeed, beyond the second two-phase region, the slope of the isotherm changes sharply in a way that corresponds to a much less compressible surface state.

Before exploring the consequences of this feature, it is first necessary to make certain that the result is not just an artifact arising from surface heterogeneities. We noted that the graphite substrate for which the data of Figure 9.6b were collected was "specially treated." A few details of this treatment should be mentioned. Graphite and anhydrous $FeCl_3$ are introduced into opposite ends of a Pyrex tube; the evacuated tube is then sealed and brought to about 300°C. A slight temperature gradient is maintained between opposite ends of the tube so the $FeCl_3$ distills to the graphite, where it reacts to form an interlamellar compound. Rapidly heating this compound to 900°C or higher results in the expulsion of the $FeCl_3$ and the attendant delamination of the graphite. The specific surface area of the final product—known as exfoliated graphite—is roughly two orders of magnitude greater than that of the initial graphite sample. The newly exposed graphite planes constitute a remarkably uniform solid adsorbent if all traces of $FeCl_3$ are removed. There is some evidence that the same vertical segments are observed in isotherms measured on the parent graphite as are observed in the exfoliated graphite. This argues that these features are not artifacts of the delamination process.

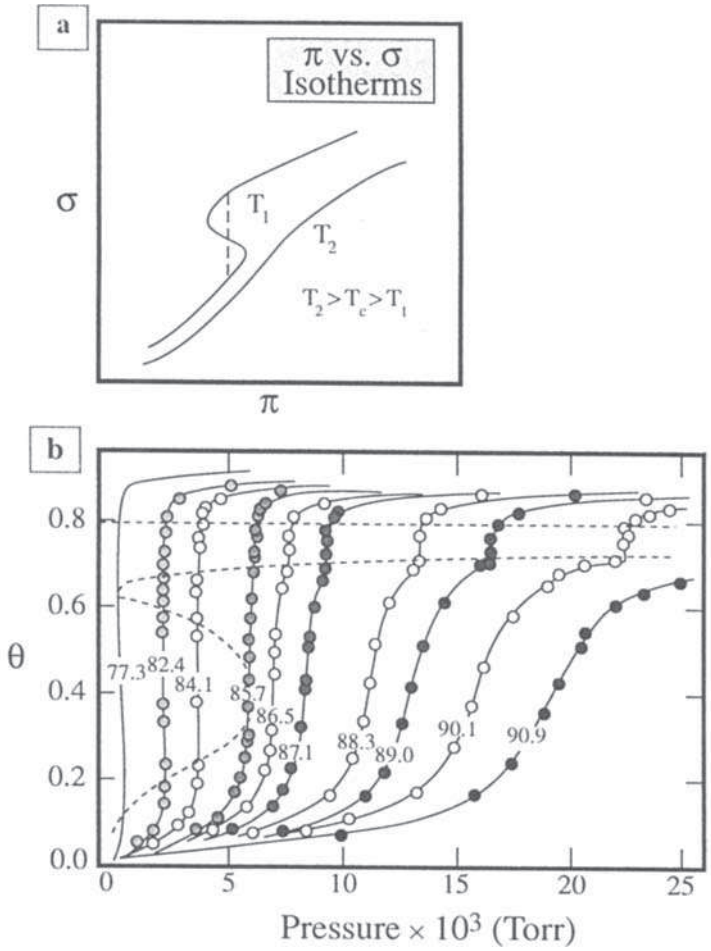


FIG. 9.6 Adsorption isotherms and surface phases. (a) schematic illustration of π versus σ isotherms in the vicinity of a two-dimensional critical temperature. (b) experimental data for the adsorption of krypton on exfoliated graphite showing similar features. (Data from A. Thomy and X. Duval, *J. Chem. Phys.*, **67**, 1101 (1970).)

9.3c.2b Harkins-Jura Isotherm. The Harkins-Jura equation (see Table 9.1) suggests that a plot of ln p versus $(n/w)^{-2}$ should be a straight line with a slope that is proportional to the square of the specific area. In experiments with solids of known area the linearity predicted by the Harkins-Jura equation has been observed. Furthermore, the proportionality constant relating the observed slopes to A_{sp}^2 is independent of the nature of the adsorbent to a first approximation. Thus solids for which A_{sp} is known may be used to "calibrate" this method for a particular adsorbed species; then the specific area of an unknown may be determined using the same adsorbate. The Harkins-Jura isotherm therefore introduces the possibility of determining specific areas by gas adsorption studies, although it is only one of many isotherms that permit the evaluation of A_{sp}.

Although it may give satisfactory values for A_{sp}, the Harkins-Jura equation leaves something to be desired at the molecular level. For example, the linear π versus σ equation of state—the starting point of the derivation of the Harkins-Jura isotherm—represents the relatively incompressible state of the surface phase (i.e., $\theta = 0.7$ in Fig. 9.6b). (This equation is obtained in analogy with the approximately linear π versus σ equation for insoluble monolayers discussed in Chapter 7.) However, in most instances of physical adsorption, no satura-

tion limit of adsorption appears. As $p \rightarrow p_0$ the amount of material adsorbed increases asymptotically. Multilayer adsorption is the only reasonable model for this observation. Except for special cases, the neighborhood of $\theta = 1$ is obscured by the onset of multilayer adsorption. In other words, there is a mismatch between the situation described by the Harkins-Jura model and that suggested by macroscopic observations. We noted above that many isotherms are insensitive to the assumptions of their derivation. In line with that observation is the fact that the Harkins-Jura equation does fit a fairly wide range of experimental data and gives reasonable values of specific surfaces despite these objections. This is one example of an "old shoemaster chef" situation; we see presently that it is only one of several such cases.

9.3d Summary of Observations

Several points might be noted in summarizing the results of Section 9.3:

1. In principle, it is possible to correlate an adsorption isotherm and a two-dimensional equation of state by working from either direction; that is, we may start with an experimental isotherm and develop the associated equation of state (as in going from Equation (8) to Equation (10)), or we may proceed from the equation of state to the isotherm (as illustrated in Example 9.2).
2. Relatively small increases in complexity for the equation of state result in considerably more complex equations for the adsorption isotherms. The gross features of the more complex isotherms are also given by simpler isotherms. This means that it is very difficult to choose among various isotherms in terms of the goodness of fit to experimental data. Therefore it is difficult to conclude from an experimental isotherm what the two-dimensional surface phases are like.
3. Two-dimensional equations of state are a useful source of isotherms, however, even though the test of the isotherm must be made in terms of some criterion other than an ability to describe adsorption. For example, the ability of an isotherm to predict the temperature dependence of adsorption or the specific area of an adsorbent is a more sensitive test of an isotherm than merely describing the way n/w increases with p.

Our approach until now has been to discuss adsorption isotherms on the basis of the equation of state of the corresponding two-dimensional matter. This procedure is easy to visualize and establishes a parallel with adsorption on liquid surfaces (Chapter 7); however, it is not the only way to proceed. In the following section we consider the use of statistical thermodynamics in the derivation of adsorption isotherms and examine some other approaches in subsequent sections.

9.4 THERMODYNAMICS OF ADSORPTION: A STATISTICAL PERSPECTIVE

The *partition function* is the central feature of statistical thermodynamics. From the partition function the various thermodynamic variables such as entropy, enthalpy, and free energy may be evaluated. It is also possible, in principle, to deduce the equation of state for a system from the partition function.

It should be apparent—since an adsorption isotherm can be derived from a two-dimensional equation of state—that an isotherm can also be derived from the partition function since the equation of state is implicitly contained in the partition function. The use of partition functions is very general, but it is also rather abstract, and the mathematical difficulties are often formidable (note the cautious "in principle" in the preceding paragraph). We shall not attempt any comprehensive discussion of the adsorption isotherms that have been derived by the methods of statistical thermodynamics; instead, we derive only the Langmuir equation for adsorption from the gas phase by this method. The interested reader will find other examples of this approach discussed by Broeckhoff and van Dongen (1970).

9.4a Preliminaries: Statistical Thermodynamics of Bulk Gases

A brief review of the statistical thermodynamics of ideal (bulk) gases will help us get started. In addition to reviewing some relevant physical chemistry, it will supply us with some expres-

sions that may be useful since the two-dimensional ideal gas law applies to adsorbed molecules as a limiting case.

The partition function Q is defined by the equation

$$Q = \sum_i g_i \exp\left(-\frac{\varepsilon_i}{k_B T}\right) \tag{14}$$

in which g_i and ε_i represent the *degeneracy* and the *energy*, respectively, of the ith state. An important property of a partition function is its factorability into contributions arising from translation and internal degrees of freedom:

$$Q = Q_{trans} Q_{int} \tag{15}$$

For N *indistinguishable* molecules that do not interact with each other, the total partition function Q_N is given by $(1/N!)$ times the Nth power of $(Q_{trans}Q_{int})$, where we take the Q's without the subscript N as the partition functions for individual molecules. The translational portion of the partition function is relatively easy to evaluate for ideal gases (see, for example, Atkins 1994). Substituting its value into Equation (15) gives

$$Q_N = \frac{1}{N!}(Q_{trans}Q_{int})^N = \left[V\left(\frac{2\pi m k_B T}{h^2}\right)^{3/2}\right]^N \left(\frac{e}{N} Q_{int}\right)^N \tag{16}$$

for N molecules of mass m in a volume V at temperature T. (The expression $(e/N)^N$ on the right-hand-side of Equation (16) comes from the use of Stirling's approximation for $\ln N!$ for large N, i.e., $\ln N! \approx N \ln N - N$, or $N! \approx N^N e^{-N} = (N/e)^N$.)

The easiest quantity to evaluate from this expression is the *Helmholtz free energy* F:

$$F = -k_B T \ln Q_N \tag{17}$$

A variety of other thermodynamic functions may be evaluated from this. For example, the chemical potential—the quantity equalized in equilibrium calculations—of species i in a multicomponent system is given by

$$\mu_i = \left(\frac{\partial F}{\partial N_i}\right)_T = -k_B T \frac{\partial \ln Q_N}{\partial N_i} \tag{18}$$

Also, to calculate the equation of state, we recall

$$p = -\left(\frac{\partial F}{\partial V}\right)_T \tag{19}$$

If we apply Equations (17) and (19) to Equation (16), for example, we get the ideal gas law:

$$p = k_B T\left(\frac{\partial \ln Q_N}{\partial V}\right)_T = \frac{N k_B T}{V} \tag{20}$$

9.4b Statistical Thermodynamics of Adsorption

The statistical thermodynamic approach to the derivation of an adsorption isotherm goes as follows. First, suitable partition functions describing the bulk and surface phases are devised. The bulk phase is usually assumed to be that of an ideal gas. From the surface phase, the equation of state of the two-dimensional matter may be determined if desired, although this quantity ceases to be essential. The relationships just given are used to evaluate the chemical potential of the adsorbate in both the bulk and the surface. Equating the surface and bulk chemical potentials provides the equilibrium isotherm.

We apply this method to the derivation of the Langmuir isotherm both to illustrate the method and to see the assumed nature of the surface energy states on which it is based.

9.4b.1 Localized, Monolayer Adsorption: The Langmuir Isotherm

The Langmuir isotherm is based on the assumption of localized adsorption. This means that an adsorbed molecule has such a high statistical preference for a certain surface site as to possess a negligible translational entropy in the adsorbed state. Localized adsorption is thus

seen to be very plausible for chemisorption, in which the adsorbed molecules and the adsorbent interact quite specifically. For nonspecific physical adsorption a nonlocalized or mobile layer seems to be a more plausible picture. We have already discussed the Henry law type of isotherm and the ideal gas equation of state that are associated with the simplest type of mobile adsorption.

The adsorption sites on the surface are assumed to be uniform and to bind the adsorbate with an energy ε per molecule or E per mole; that is, the potential energy of a molecule in the gaseous state is zero, and in the adsorbed state it is $-\varepsilon$. Note that this adsorption energy is a characteristic of the interaction between the adsorbed molecules and the adsorbent. As such, it is the same not only for all parts of the surface but also for all degrees of surface coverage. This is equivalent to saying that the adsorbed molecules do not interact with each other.

The surface is assumed to consist of M adsorption sites. Suppose we consider the case in which N of the sites are occupied; that is, N molecules are adsorbed. To write the partition function Q for the surface molecules, we must ask how these molecules differ from those in the gas phase (superscript g). Some of the internal degrees of freedom may be modified by the adsorption (Q^s_{int}), but the most notable difference will be in the translational degrees of freedom. From three equivalent translational degrees of freedom, the adsorbed molecule goes to two highly restrained translational degrees of freedom (remember the adsorption is localized) and one vibrational degree of freedom normal to the surface(s):

$$Q^g_{trans.\,3d} \rightarrow Q^s_{trans.\,2d}\, Q^s_{vib} \tag{21}$$

With these ideas in mind, we may assemble the partition function of the surface molecules as follows:

$$Q^s = g_N \left\{ Q^s_{trans.\,2d}\, Q^s_{vib} Q^s_{int} \exp\left[-\left(-\frac{\varepsilon}{k_B T} \right) \right] \right\}^N \tag{22}$$

In this expression, the degeneracy factor g_N represents the number of ways the N molecules may be placed on M sites, and is given by the combinatorial formula (see Equation (3.52) in Section 3.4a.2 or Equation (2.46) in Section 2.6a):

$$g_N = \frac{M!}{N!\,(M-N)!} \tag{23}$$

Combining Equations (22) and (23) gives the following expression for the partition function of the adsorbed molecules:

$$Q^s = \frac{M!}{N!\,(M-N)!} (Q^s_{trans.\,2d}\, Q^s_{vib}\, Q^s_{int})^N \exp\left(\frac{N\varepsilon}{k_B T} \right) \tag{24}$$

Application of Equation (17) to Equation (24) gives the Helmholtz free energy of the adsorbed molecules:

$$F^s = -k_B T \left(\ln M! - \ln N! - \ln(M-N)! + \frac{N\varepsilon}{k_B T} + N \ln (Q^s_{trans.\,2d}\, Q^s_{vib}\, Q^s_{int}) \right) \tag{25}$$

Since N and M are large, the factorials may be expanded by Sterling's approximation ($\ln x! \approx x \ln x - x$) to give

$$F^s = -k_B T \left(M \ln M - N \ln N - (M-N) \ln (M-N) + \frac{N\varepsilon}{k_B T} + N \ln (Q^s_{trans.\,2d}\, Q^s_{vib}\, Q^s_{int}) \right) \tag{26}$$

The chemical potential of the adsorbed molecules is given, according to Equation (18), by the derivative of F^s with respect to N:

$$\mu^s = k_B T \left[\ln \left(\frac{N}{M-N} \right) - \frac{\varepsilon}{k_B T} - \ln (Q^s_{trans.\,2d}\, Q^s_{vib}\, Q^s_{int}) \right] \tag{27}$$

The condition of equilibrium between the adsorbed molecules and molecules in the gas state requires that the chemical potential for the adsorbed species be the same in both the gas phase and the adsorbed state:

$$\mu^s = \mu^g \tag{28}$$

Applying Equations (17) and (18) to Equation (16) shows the chemical potential for an ideal gas to be

$$\mu^g = -k_B T \ln\left(\frac{k_B T}{p}\right)\left(\frac{2\pi m k_B T}{h^2}\right)^{3/2} Q_{int}^g \tag{29}$$

since $V = RT/p$.

Equating Equations (27) and (29) gives

$$\frac{N}{M-N} = \frac{p}{k_B T}\left(\frac{h^2}{2\pi m k_B T}\right)^{3/2} \frac{Q_{trans,2d}^s Q_{vib}^s Q_{int}^s}{Q_{int}^g} \exp\left(\frac{\varepsilon}{k_B T}\right) \tag{30}$$

We may group the following terms together to define a new quantity K:

$$K = \frac{1}{k_B T}\left(\frac{h^2}{2\pi m k_B T}\right)^{3/2} \frac{Q_{trans,2d}^s Q_{vib}^s Q_{int}^s}{Q_{int}^g} \exp\left(\frac{\varepsilon}{k_B T}\right) \tag{31}$$

in terms of which Equation (30) becomes

$$\frac{N}{M-N} = Kp \tag{32}$$

Now if we divide both the numerator and denominator of the left-hand side by M and recognize that $N/M = \theta$, we obtain

$$\frac{\theta}{1-\theta} = Kp \tag{33}$$

or

$$\theta = \frac{Kp}{1+Kp} \tag{34}$$

two forms of the *Langmuir adsorption isotherm* (see Equation (7.67)).

The quantity K defined by Equation (31) may easily be expanded somewhat further. First, we write the two-dimensional partition function by analogy with its three-dimensional counterpart, Equation (16). To do this, replace V by the area accessible to the adsorbed molecule σ and the exponent by 2/2 (= 1) in place of 3/2 since two rather than three degrees of freedom are involved. Therefore we obtain

$$Q_{trans,2d}^s = \sigma \frac{2\pi m k_B T}{h^2} \tag{35}$$

If the energy separating the vibrational quantum states is small relative to the thermal energy, the partition function for vibration is approximately given by

$$Q_{vib}^s = \frac{k_B T}{\varepsilon_{vib}} \tag{36}$$

For physical adsorption the approximation involved here is expected to be valid. Finally, the energy of vibration may be replaced by $h\nu$, where ν is the frequency with which the adsorbed molecules vibrate against the adsorbent:

$$Q_{vib}^s = k_B T/h\nu \tag{37}$$

Combining Equations (35) and (37) with Equation (31) gives

$$K = \frac{\sigma}{\nu}(2\pi m k_B T)^{-1/2} \frac{Q_{int}^s}{Q_{int}^g} \exp\left(\frac{\varepsilon}{k_B T}\right) \tag{38}$$

It should be noted that this quantity has the units length^2 force^{-1}, reciprocal pressure units, as required by Equation (34).

We saw from Equation (7.75) how to rearrange the Langmuir equation into a form that permits graphical evaluation of the parameters. For the adsorption of gases this becomes

$$p/(n/w) = mp + b \tag{39}$$

and predicts a straight line when $[p/(n/w)]$ is plotted versus p, with

$$\text{Slope} = m = N_A\sigma/A_{sp} \tag{40}$$

and

$$\text{Slope/Intercept} = m/b = K \tag{41}$$

Figure 9.7 shows some data for the adsorption of ethyl chloride on charcoal. Since these data were collected at different temperatures, the ratio p/p_0 is used as the independent variable in fitting the data to Equation (39). That is, to compare the adsorption at different temperatures, the pressure is expressed as a fraction of the equilibrium vapor pressure at that temperature. The data for each of the three temperatures in Figure 9.7 give quite good straight lines when plotted according to the linear form of the Langmuir equation (Equation (39)); the interpretation of this analysis is discussed in Example 9.3.

* * *

EXAMPLE 9.3 *Analysis of Adsorption Data Using the Langmuir Isotherm.* Slope and intercept values for the linearized plots of the data in Figure 9.7 are as follows:

Temperature (°C)	Slope (g g^{-1})	Intercept (g g^{-1})
−15.3	1.9	0.040
0.0	2.0	0.047
20.0	2.0	0.088

Note that since p/p_0 is used in the linearization, the abscissa is dimensionless and the slope and intercept have the same units. Figure 9.7 suggests that the data at −15.3 and 0.0°C converge to a common saturation value, but this is less clear for the data at 20°C. Do the linear plots clarify this situation? From the temperature dependence of the K's for these data, estimate the adsorption energy in this system.

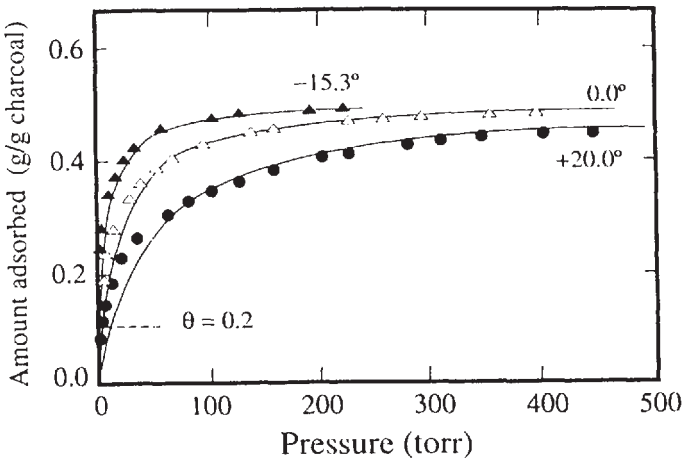


FIG. 9.7 Plot showing how the amount of ethyl chloride adsorbed on charcoal (in g g^{-1}) varies with pressure at −15.3, 0, and 20°C. (Data from F. Goldman, and M. Polanyi, *Z. Phys. Chem.*, **132**, 321 (1928).)

Solution: The slopes of these lines are identical (within experimental error), and the combination of Equations (40) and (7.72) shows that the reciprocal of the slope gives the adsorption at saturation. We conclude, therefore, that all three isotherms converge to the same saturation level of adsorption, namely, $(2.0)^{-1} = 0.50\ \text{g g}^{-1}$.

The ratio of the slope to intercept values gives K for the adsorption at the three temperatures: i.e., $K = 48$, 43, and 23 at -15.3, 0.0, and 20.0°C, respectively. Multiplying Equation (38) through by $T^{1/2}$ and taking logarithms gives $\ln (T^{1/2}K) = (\varepsilon/k_B)(1/T) + \text{const.} = (E/R)(1/T) + \text{const.}$ While the quality of the data and the number of points scarcely justify the interpretation, a value for E/R of about 1700 can be estimated, suggesting an apparent value for E of about 14 kJ mole^{-1}. The word *apparent* is included until the effect of using p/p_0 instead of p in the linearization of the Langmuir equation is clarified. ■

* * *

As noted above, the ability to predict correctly the temperature dependence of adsorption is a more stringent test of an isotherm than mere correlation of adsorption data. For this reason an independent measure of the energy of adsorption is clearly desirable. We return to this in Section 9.6.

Until now, we have focused our attention on those adsorption isotherms that show a saturation limit, an effect usually associated with monolayer coverage. We have seen two ways of arriving at equations that describe such adsorption: from the two-dimensional equation of state via the Gibbs equation or from the partition function via statistical thermodynamics. Before we turn our attention to multilayer adsorption, we introduce a third method for the derivation of isotherms, a *kinetic* approach, since this is the approach adopted in the derivation of the multilayer, BET adsorption isotherm discussed in Section 9.5. We introduce this approach using the Langmuir isotherm as this would be useful in appreciating the common features of (and the differences between) the Langmuir and BET isotherms.

9.4b.2 Langmuir Isotherm from Kinetic Theory

Suppose we consider a surface consisting of a total of M adsorption sites, M_1 of which are filled with adsorbed molecules to a depth of one molecule. Thus the number of bare sites M_0 is $(M - M_1)$. This situation is defined to be in equilibrium when the rate at which molecules attach to bare spots is the same as the rate at which they escape from monolayer regions. The rate of the adsorption process R_a is proportional to both the pressure and the number of available sites:

$$R_a = k_a p M_0 \tag{42}$$

The rate of desorption R_d is proportional to the number of sites occupied:

$$R_{d,1} = k_{d,1} M_1 \tag{43}$$

The subscript 1 for R_d and k_d in Equation (43) reminds us that these refer to layer 1 only (monolayer). At equilibrium the rates given by Equation (42) and (43) are equal:

$$k_a p M_0 = k_{d,1} M_1 \tag{44}$$

When adsorption is restricted to a monolayer, one has

$$M_0 + M_1 = M \tag{45}$$

a result that may be substituted into Equation (44) to give the Langmuir equation:

$$\frac{k_a}{k_{d,1}} p = \frac{M_1}{M - M_1} = \frac{\theta}{1 - \theta} \tag{46}$$

It is worthwhile to examine the expected form of the two rate constants k_a and $k_{d,1}$. The rate constant for desorption consists of a frequency factor and a Boltzmann factor. The former may be assumed to be proportional to a frequency ν, and the energy in the Boltzmann factor may be identified with the energy of interaction between the adsorbate and the adsorbent:

$$k_{d,1} \propto \nu \exp\left(-\frac{\varepsilon}{k_B T} \right) \tag{47}$$

From kinetic theory of gases (see Atkins 1994), the number of collisions per unit area per unit time Z between gas molecules and a wall equals

$$Z = (2\pi m k_B T)^{-1/2} p \tag{48}$$

If this is multiplied by a surface area σ, the result is the rate of surface collisions. Then by analogy with Equation (42), we may take k_a as

$$k_a \propto (2\pi m k_B T)^{-1/2} \sigma \tag{49}$$

The ratio $k_a/k_{d,1}$ may be set equal to the Langmuir coefficient K. Therefore, from Equations (49) and (47) we get

$$K = \frac{k_a}{k_{d,1}} \propto (2\pi\, m\, k_B T)^{-1/2} \frac{\sigma}{\nu} \exp\left(\frac{\varepsilon}{k_B T}\right) \tag{50}$$

a result that is identical to Equation (38) as far as translational factors are concerned.

This kinetic-theory-based view of the Langmuir result provides no new information, but it does draw attention to the common starting assumptions of the Langmuir derivation and the BET derivation (Section 9.5a). This kinetic derivation of the Langmuir equation is especially convenient for obtaining an isotherm for the adsorption of two gases. This is illustrated in Example 9.4.

* * *

EXAMPLE 9.4 *Kinetic-Theory-Based Description of Binary Adsorption.* Assume that two gases *A* and *B* individually follow the Langmuir isotherm in their adsorption on a particular solid. Use the logic that results in Equation (46) to derive an expression for the fraction of surface sites covered by one of the species when a mixture of the two gases is allowed to come to adsorption equilibrium with that solid.

Solution: The rate of desorption of *A* is given by Equation (43) as $R_{d,A} = k_{d,A} M_A$ and the rate of adsorption of *A* by Equation (42). Allowing for the fact that both *A* and *B* occupy surface sites, we have $R_{a,A} = k_{a,A} p_A M_0 = k_{a,A} p_A (M - M_A - M_B)$. In this expression p_A is the partial pressure of *A*. At equilibrium the rates of adsorption and desorption are equal; therefore $k_{d,A} M_A = k_{a,A} p_A (M - M_A - M_B)$. Defining $\theta_i = M_i/M$ and $K_i = k_{a,i}/k_{d,i}$, we can write this last result as $\theta_A = K_A p_A (1 - \theta_A - \theta_B)$. A similar expression can be written for θ_B: $\theta_B = K_B p_B (1 - \theta_A - \theta_B)$.

Next the ratio of θ_A/θ_B can be used to eliminate one of the θ's from the above expressions. Since $\theta_A/\theta_B = K_A p_A / K_B p_B$, $\theta_B = (K_B p_B / K_A p_A)\theta_A$. Using this to eliminate θ_B, we obtain $K_A p_A = \theta_A (1 + K_A p_A + K_B p_B)$, which rearranges to $\theta_A = K_A p_A / (1 + K_A p_A + K_B p_B)$.

A similar expression is obtained for θ_B: $\theta_B = K_B p_B / (1 + K_A p_A + K_B p_B)$.

The fraction of surface sites covered by one gas or the other is given by $\theta_A + \theta_B = (K_A p_A + K_B P_B)/(1 + K_A P_A + K_B P_{BB})$. Note that the expressions for θ_A, θ_B, and $\theta_A + \theta_B$ all reduce to simple Langmuir expressions when one of the p or K values is zero. ■

* * *

9.5 MULTILAYER ADSORPTION: THE BRUNAUER-EMMETT-TELLER EQUATION

9.5a Derivation of the Brunauer-Emmett-Teller Isotherm

As noted above, the range of pressures over which gas adsorption studies are conducted extends from zero to the normal vapor pressure of the adsorbed species p_0. An adsorbed layer on a small particle may readily be seen as a potential nucleation center for phase separation at p_0. Thus at the upper limit of the pressure range, adsorption and liquefaction appear to converge. At very low pressures it is plausible to restrict the adsorbed molecules to a monolayer. At the upper limit, however, the imminence of liquefaction suggests that the adsorbed molecules may be more than one layer thick. There is a good deal of evidence supporting the idea that multilayer adsorption is a very common form of physical adsorption on nonporous solids. In this section we are primarily concerned with an adsorption isotherm derived by Brunauer, Emmett, and Teller in 1938; the theory and final equation are invariably known by the initials of the authors: BET.

The BET isotherm has subsequently been derived by a variety of methods, but, as mentioned in the previous section, we follow the approach of the original derivation, namely, a

kinetic description of the equilibrium state. Like the Langmuir isotherm, the BET theory begins with the assumption of localized adsorption. There is no limitation as to the number of layers of molecules that may be adsorbed, however; hence there is no saturation of the surface with increasing pressure. In general, the derivation assumes that the rates of adsorption and desorption from each layer are equal at equilibrium and that adsorption or desorption can occur from a particular layer only if that layer is exposed, that is, provided no additional adsorbed layers are stacked on top of it.

If we allow the possibility that a surface site may be covered to a depth of more than one molecule, then the following modifications to the derivation presented in Section 9.4b.2 are required. First, we define M_i to be the number of sites covered to a depth of i molecules. Second, Equation (45) is modified to

$$M = M_0 + M_1 + M_2 + \ldots + M_n = \sum_{i=0}^{n} M_i \tag{51}$$

where the summation includes all thicknesses of coverage from zero to n, the maximum.

Now let us consider the composition of the second layer ($i = 2$). By analogy with Equation (42), the rate at which molecules adsorb to form layer 2 is proportional to p and M_1. For simplicity, the proportionality constant is assumed to be the same as that for adsorption on the bare surface. Therefore we write

$$R_a = k_a p M_1 \tag{52}$$

In a similar fashion, we assume the rate of desorption from the second layer to be proportional to M_2:

$$R_{d,2} = k_{d,2} M_2 \tag{53}$$

The constant $k_{d,2}$ is assumed to be of the same form as that given by Equation (47) for the first layer, with one important modification. While desorption from the first layer involves detaching a molecule from the adsorbent as a substrate, desorption from the second layer involves detachment from another adsorbed molecule of the same kind. The adsorption energy ε was used in the Boltzmann factor in the first case; the energy of vaporization ε_v is a more appropriate value to use for the analogous quantity in the second case. We shall assume the frequency factor to be unchanged and write

$$k_{d,2} = \nu \exp\left(- \frac{\varepsilon_v}{k_B T} \right) \tag{54}$$

At equilibrium, the rate of adsorption and the rate of desorption from the second layer are also equal; therefore

$$k_a M_1 p = k_{d,2} M_2 \tag{55}$$

As a matter of fact, the same expression also applies to the third, fourth, . . . , ith levels. As a first approximation, the "activation energy" for desorption is the same for all layers after the first. This leads to the generalization

$$k_a M_{i-1} p = k_{d,i} M_i \tag{56}$$

for $2 \leq i < n$.

Equation (56) enables us to relate M_i to M_{i-1} just as Equation (44) relates M_1 to M_0. Moreover, through repeated application of Equation (56) we may relate M_i to M_1 or M_0 as follows:

$$M_i = \left(\frac{k_a}{k_{d,i}} p \right)^{i-1} M_{i-(i-1)} = \left(\frac{k_a}{k_{d,i}} p \right)^{i-1} M_1 = \left(\frac{k_a}{k_{d,i}} \right)^{i-1} \frac{k_a}{k_{d,1}} p^i M_0 \tag{57}$$

Since k_a is assumed to be the same for each layer and $k_{d,i}$ differs from k_d only in the value of the energy in the Boltzmann factor, this may be written

$$M_i = \frac{k_a^i p^i M_0}{\nu^i [\exp(-\varepsilon_v/k_B T)]^{i-1} \exp(-\varepsilon/k_B T)} \tag{58}$$

Multiplying the numerator and denominator of the right-hand side by $\exp(-\varepsilon_v/k_B T)$ we get

$$M_i = \frac{k_a^i p^i M_0}{[\nu \exp(-\varepsilon_v/k_B T)]^i} \frac{\exp(-\varepsilon_v/k_B T)}{\exp(-\varepsilon/k_B T)} \tag{59}$$

which may be written

$$M_i = x^i c M_0 \tag{60}$$

with c and x defined as follows:

$$x = \frac{k_a p}{\nu \exp(-\varepsilon_v/k_B T)} = \frac{k_a}{k_{d,i\geq 2}} p \tag{61}$$

and

$$c = \exp\left(\frac{\varepsilon - \varepsilon_v}{k_B T}\right) \tag{62}$$

Further, Equation (60) may now be substituted into Equation (51) to give

$$M = M_0 + M_1 + M_2 + \ldots + M_n = M_0 + \sum_{i=1}^{n} x^i c M_0 \tag{63}$$

Under some circumstances there may be a reason to restrict adsorption to a finite number of layers, that is, assign some specific value to n. In general, however, $n \to \infty$ as $p \to p_0$ is usually taken as the upper limit for this summation.

At this point we may define two other quantities in terms of the variables involved in Equation (59): the total volume of gas adsorbed and the volume adsorbed at monolayer coverage, V and V_m, respectively. The total volume is obviously the sum of the volume held in each type of site V_i, which is proportional to iM_i:

$$V = \sum_{i=1}^{n} V_i \propto \sum_{i=1}^{n} i M_i \tag{64}$$

The volume adsorbed at monolayer coverage is simply proportional to the total number of sites irrespective of the depth to which they are covered:

$$V_m \propto M = M_0 + \sum_{i=1}^{n} M_i \tag{65}$$

that is, V_m equals the volume of gas that would be adsorbed if a monolayer were formed. However, note that in writing Equation (65) it is not assumed that the monolayer is completely filled before other layers are formed. On the contrary, the picture allows for the coexistence of all types of patches, with adsorbed molecules stacked to various depths on each. There is no implication that a filled monolayer is required for multilayer formation.

Taking the ratio of Equation (64) to Equation (65) eliminates the unspecified proportionality constant and gives

$$\frac{V}{V_m} = \frac{\Sigma_i i M_i}{M_0 + \Sigma_i M_i} = \frac{c \Sigma_i i x^i}{1 + c \Sigma_i x^i} \tag{66}$$

The ratio V/V_m may be identified with θ, which may have values greater than unity in the case of multilayer adsorption. All that remains to be done to complete the BET derivation is to evaluate the summations in Equation (66).

To assist in the evaluation of the summations, suppose we consider the quantity $[x(1 - x)^{-1}]$. The factor in parentheses may be expanded as a power series (see Appendix A) to give

$$x(1 - x)^{-1} = x(1 + x + x^2 + \ldots) = \sum_{i=1}^{n} x^i \tag{67}$$

Next we consider the quantity $\{x(d/dx)(\Sigma_i x^i)\}$. Carrying out the indicated differentiation gives

$$x \frac{d}{dx} \sum_i x^i = x \sum_i i x^{i-1} = \sum_i i x^i \tag{68}$$

Equation (67) may now be used as a substitution for $\Sigma_i x^i$ in Equation (68) to yield

$$\sum_i i x^i = x \frac{d}{dx}[x(1 - x)^{-1}] = \frac{x}{(1 - x)^2} \tag{69}$$

Substituting Equations (67) and (69) into Equation (66) gives

$$\frac{V}{V_m} = \frac{c x (1 - x)^{-2}}{1 + c x (1 - x)^{-1}} \tag{70}$$

This last result may be simplified further to become

$$\frac{V}{V_m} = \frac{c x}{(1 - x)[1 + (c - 1)x]} \tag{71}$$

Equation (71) is the result generally defined as the Brunauer-Emmett-Teller (BET) equation.

The condition that $V \to \infty$ is seen to correspond to $x = 1$ by Equation (71). Recalling the definition of x given by Equation (61) and that $V \to \infty$ as $p \to p_0$ permits us to write

$$1 = \frac{k_a}{k_{d,i \geq 2}} p_0 \tag{72}$$

Therefore Equation (61) becomes

$$x = \frac{p}{p_0} \tag{73}$$

Thus the independent variable in the BET theory is the pressure relative to the saturation pressure. Therefore the BET equation describes the volume of gas adsorbed at different values of p/p_0 in terms of two parameters V_m and c. Furthermore, the model supplies a physical interpretation to these two parameters.

9.5b Testing the Brunauer-Emmett-Teller Theory

Basically there are three criteria against which the success of the BET theory may be evaluated: its ability to "fit" adsorption data, correct prediction of the temperature dependence of adsorption, and correct evaluation of specific area. We discuss these three issues in this section.

9.5b.1 Evaluation of Brunauer-Emmett-Teller Constants

The easiest way to evaluate the BET constants is to rearrange Equation (71) into the following linear form:

$$\frac{1}{V} \frac{x}{1 - x} = \frac{c - 1}{c V_m} x + \frac{1}{c V_m} \tag{74}$$

This form suggests that a plot of $(1/V)[x/(1 - x)]$ against x should yield a straight line, with

$$\text{Slope} = m = \frac{c - 1}{c V_m} \tag{75}$$

and

$$\text{Intercept} = b = 1/cV_m \tag{76}$$

Equations (75) and (76) may be solved to supply values of V_m and c from the experimental results:

$$V_m = \frac{1}{m + b} \tag{77}$$

and

$$c = \frac{m}{b} + 1 \tag{78}$$

In the following few paragraphs we first examine some of the general features of gas adsorption as predicted by the BET theory. Next we consider how well the theory actually fits experimental data. The use of experimental V_m values in the evaluation of A_v, is taken up in the following section.

9.5b.2 Predictions of Brunauer-Emmett-Teller Theory: Success in Fitting Experiments

Figure 9.8 enables us to observe some of the general features of the BET isotherm. The most apparent aspect concerning all the curves in this figure is their rapid increase as $p/p_0 \rightarrow 1$. It is also apparent, however, that the shape of the curve is sensitive to the value of c, especially for low values of c. Note particularly that the BET equation encompasses both Type II and Type III isotherms (Fig. 9.4). For values of c equal to 2 or less, the curves show no inflection point, whereas the inflection becomes increasingly pronounced as c increases above 2. In view of the wide diversity of curve shapes that are consistent with the BET equation and the relative insensitivity of adsorption data to the model underlying a particular equation, we might expect that the BET equation will fit experimental data rather successfully.

Figure 9.9a is a plot of some actual experimental data showing the volume of N_2–

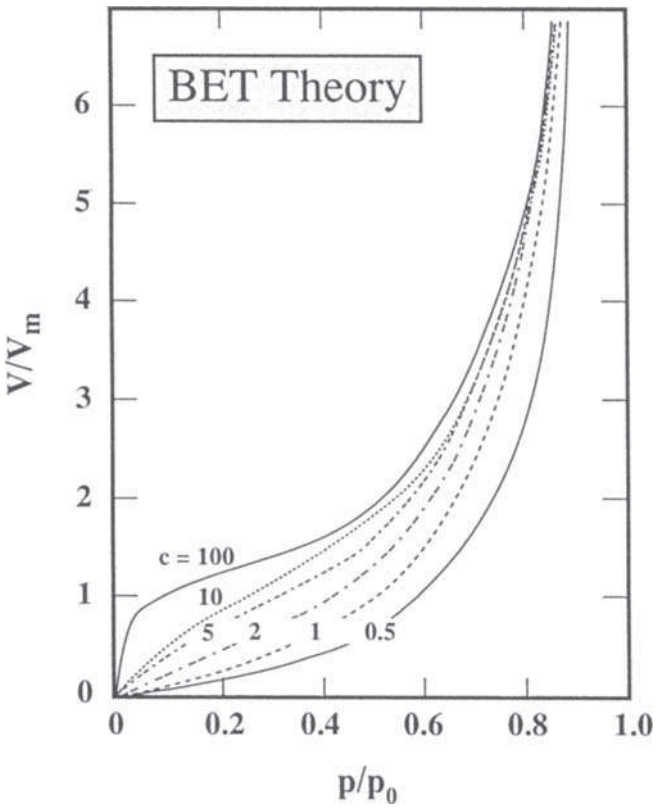


FIG. 9.8 Plots of V/V_m versus p/p_0 for several values of the parameter c, calculated according to the BET theory by Equation (71).

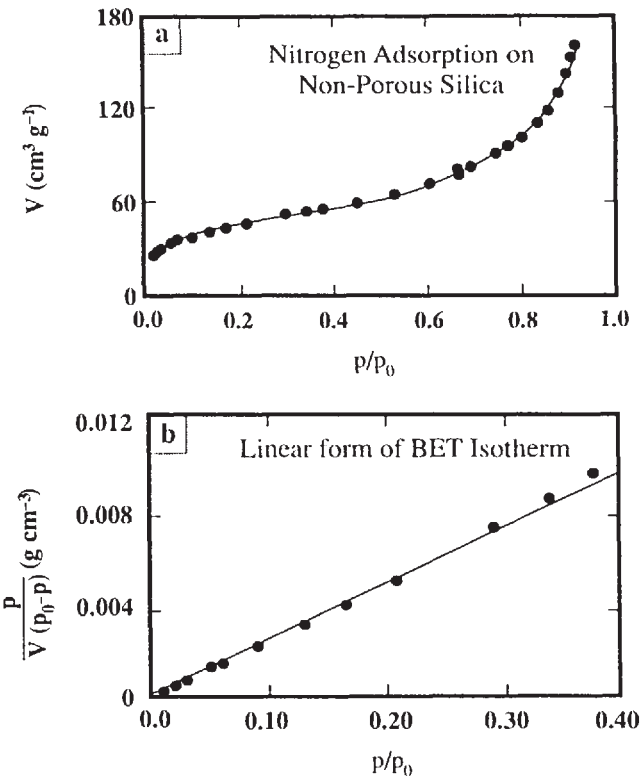


FIG. 9.9 Nitrogen adsorption on nonporous silica at 77K. (a) volume per gram (in cm^3 per gram at STP) versus p/p_0; and (b) according to the linear form of the Brunauer-Emmett-Teller equation (Equation (74)). (Data from D. H. Everett, G. D. Parfitt, K. S. W. Sing, and R. Wilson, *J. Appl. Chem. Biotechnol.*, **24**, 199 (1974).)

expressed as cubic centimeters at STP per gram – adsorbed by a sample of nonporous silica at 77K. Figure 9.9b shows these same results plotted according to the linear form of the BET equation, given by Equation (74). The BET equation is seen to fit the adsorption data in the range $0.05 < p/p_0 < 0.30$. From the values of the slope (0.0257 g cm^{-3} at STP) and intercept ($2.85 \cdot 10^{-4}$ g cm^{-3} at STP), Equations (77) and (78) may be used to evaluate V_m and c for this system:

$$V_m = \frac{1}{(257 + 2.85) \cdot 10^{-4}} = 38.5 \text{ cm}^3 \text{ g}^{-1} \text{ at STP} \tag{79}$$

$$c = \frac{257 \cdot 10^{-4}}{2.85 \cdot 10^{-4}} + 1 = 91.2 \tag{80}$$

The range of pressures over which the linear form of the BET equation fits the experimental data in Figure 9.9 is fairly typical for a variety of gases and adsorbents. At relative pressures below the range of fit, the BET equation underestimates the actual adsorption, whereas above $p/p_0 = 0.35$, the equation overestimates it. This range of p/p_0 values encompasses the region in which $V = V_m$ for $2 < c < 500$. This is an important range for applying the BET equation to surface area determination.

9.5b.3 Brunauer-Emmett-Teller Theory and Temperature Dependence of Adsorption

The parameter c describes the temperature dependence of adsorption. In the derivation already presented, certain simplifying assumptions were made as to the constancy of k_a and ν_i in each

of the layers. Various modifications of this assumption might be made, but they would involve minor temperature effects at best. To a first approximation, then, one should be able to evaluate $(\varepsilon - \varepsilon_v)$ from a knowledge of experimental c values, or vice versa. Proceeding in the first manner leads to values of $(\varepsilon - \varepsilon_v)$ that are too low, perhaps half their expected value. When we consider the region of fit on which the evaluation of c is based, it is not difficult to see why values of c are too low. The data are not fitted to the earliest stages of adsorption that are associated with larger values of ε; hence c is underestimated. The definition of c allows us to conclude some unfinished business pertaining to Example 9.3. This is examined in Example 9.5.

* * *

EXAMPLE 9.5 *Calculating the Adsorption Energy from the Brunauer-Emmett-Teller Isotherm.* The BET analysis uses p/p_0 rather than p as a variable just as we used this pressure ratio to compare Langmuir adsorption at different temperatures in Example 9.3. What corrections, if any, are needed in the "apparent adsorption energy" of about 14 kJ mole^{-1} as calculated in Example 9.3?

Solution: The answer to this is obtained by comparing Equations (31) and (62). When p/p_0 is used instead of p, we obtain c, not K, from the linearized Langmuir equation via Equation (41). The exponential energy term evaluated from the analysis of ln K versus $1/T$ is seen by Equation (62) to be the difference between the adsorption energy and the energy of vaporization of the adsorbate. For ethyl chloride $E_v = 23$ kJ mole^{-1}; therefore $E = 23 + 14 = 37$ kJ mole^{-1}; this is the actual adsorption energy for the ethyl chloride-charcoal system discussed in Example 9.3.

Note also that small values of c—which account for Type III adsorption isotherms—result when $\varepsilon_v > \varepsilon$ in Equation (62). This was anticipated in our remarks about Figure 9.4. ■

* * *

9.5b.4 Brunauer-Emmett-Teller Theory and Specific Surface Area

The final criterion for judging the success of a theoretical isotherm is its ability to measure specific surface areas. Since the BET equation has become a standard in this regard, we devote this subsection to the discussion of this topic.

With monolayer adsorption, we saw how the saturation limit could be related to the specific surface area of the adsorbent. The BET equation permits us to extract from multilayer adsorption data (by means of Equation (77)) the volume of adsorbed gas that would saturate the surface if the adsorption were limited to a monolayer. Therefore V_m may be interpreted in the same manner that the limiting value of the ordinate is handled in the case of monolayer adsorption. Since it is traditional to express both V and V_m in cubic centimeters at STP per gram, we write (see Equation (7.72))

$$V_m = \left(\frac{n}{w}\right)_{sat} (22{,}414\ \text{cm}^3\ \text{mole}^{-1}) = \frac{A_{sp}\ (22{,}414)}{N_A\ \sigma^0} \tag{81}$$

Note that it is assumed that V_m has been expressed on a "per gram" basis in writing Equation (81). If this is not the case, the area of the actual sample rather than its specific area is given by Equation (81). If the area occupied per molecule on the surface is known, the specific surface may be evaluated from Equation (81):

$$A_{sp} = \frac{V_m N_A \sigma^0}{22{,}414} \tag{82}$$

This last quantity is something that may be determined by independent methods; therefore the BET theory may be tested at this point.

It is clear that the same value of A_{sp} must be obtained for a particular adsorbent regardless of the nature of the adsorbate used. Studies of this sort lead to quite consistent values of A_{sp}, provided the adsorbed species all have access to the same surface. On a porous surface, for example, large adsorbate molecules may not be able to enter small cavities that are accessible to smaller molecules.

In the preceding discussion, it has been assumed that a value of σ^0 is known unambiguously. This quantity is obviously the "yardstick" by which moles of adsorbed gas are converted

to areas. Any error in this quantity will invalidate the determination of A_{sp}. What is generally done is to assume the adsorbed material has the same density on the surface that it has in the bulk liquid at the same temperature and to assume the molecules are close packed on the surface. In view of the earlier discussion of surface phases, this is seen to be a somewhat risky procedure. The safest way to proceed would be to evaluate σ^0 for a particular adsorbate from independent measurements of V_m and A_{sp}.

In view of the difficulty in translating measured gas adsorption into absolute specific surface areas, it is not surprising that self-consistency is often normative in this matter. In this sense, at least, an area of 16.2 Å^2 for nitrogen has become something of a standard. Values for other common adsorbed species may be found in the works of Adamson (1990), Broeckhoff and van Dongen (1970), and Kantro et al. (1967). It is probably not surprising that polar molecules such as water display values of σ^0 that are sensitive to the nature of the substrate. Using nitrogen adsorption to evaluate A_{sp} and then using the latter to evaluate σ^0 for water has led to values of 12.5, 10.4, and 11.4 Å^2 for amorphous silica, calcium hydroxide, and calcium silicate hydrate, respectively.

In spite of a variety of objections to the BET theory, V_m values from N_2 adsorption studies have become a very common means for determining specific surface areas. As a matter of fact, an IUPAC commission was organized in 1969 to study N_2 adsorption with the objective of preparing reference standards for surface area determinations. In this project four silicas and four carbon blacks of different particle size were investigated independently in 13 different laboratories. Nitrogen adsorption data were analyzed by the BET method over the best fit region, and values of V_m were converted to A_{sp} using 16.2 Å^2 as the value of σ^0. Table 9.2 shows the values of A_{sp} and c obtained in the 10 laboratories that studied the silica of Figure 9.9. The average value of the specific surface from these determinations is 163.4 ($\pm 6\%$) $m^2 g^{-1}$. This value agrees with our analysis of Figure 9.9 obtained by combining Equations (79) and (82):

$$A_{sp} = \frac{(38.5)\ (6.02 \cdot 10^{23})\ (16.2 \cdot 10^{-20})}{22{,}400} = 168\ \text{m}^2\,\text{g}^{-1} \tag{83}$$

Samples of this material as well as others investigated are now available for calibration purposes as gas adsorption standards. Details may be found in the reference cited for the data of Table 9.2.

TABLE 9.2 Values of A_{sp} and c as Determined in 10 Different Laboratories by the Brunauer-Emmett-Teller Method for the Same Silica Sample Shown in Figure 9.9

Laboratory	A_{sp} ($m^2 g^{-1}$)	c
A	166.4	92
B	162.8	101
C	174.0	100
D	148.5	166
E	173.5	70
F	166.0	98
G	167.9	91
H	143.6	113
I	169.5	62
J	161.7	122
Average	163.4	102
Standard deviation	10.0 (6%)	29 (28%)

Source: D. H. Everett, G. D. Parfitt, K. S. W. Sing, and R. Wilson, *J. Appl. Chem. Biotechnol.*, **24**, 199 (1974).

At first glance, it may seem surprising that the BET method is as successful as it is in the evaluation of A_{sp}. After all, the values of c are not in particularly good agreement with expected values, nor, conversely, are $(\varepsilon - \varepsilon_v)$ values calculated from experimental c's in good agreement with expectations. Probably a significant part of the discrepancy between theory and experiment with respect to c arises from the heterogeneity of the surface. The BET equation—like the Langmuir equation to which it reduces in certain limits—assumes that a single adsorption energy applies to all surface sites. As we saw in Chapter 6, Section 6.7, a distribution of surface energy states is not uncommon with solids. In this regard, both the Langmuir and BET models are unrealistic. However, sites with high adsorption energies are apt to be covered by adsorbate molecules first. If these represent a relatively small fraction of the surface sites, this effect gets masked fairly quickly as the first layer fills. In addition, lateral interactions between adsorbed molecules start at zero and increase as the surface begins to fill. Even these interactions tend to level off after a certain degree of coverage is achieved. Therefore, even though a single adsorption energy is unrealistic, there are compensating effects that tend to minimize the variation.

Although we have often mentioned the adsorption energy in this chapter, we have not yet discussed any procedure by which this can be measured quantitatively except as some sort of an average quantity. Since most solid surfaces are heterogeneous, it is desirable to be able to examine adsorption energy in a more discriminating way, for example, as a function of coverage or pretreatment. In the following section we see how this can be done.

9.6 ENERGETICS OF ADSORPTION

As we have seen, an adsorption isotherm is one way of describing the thermodynamics of gas adsorption. However, it is by no means the only way. Calorimetric measurements can be made for the process of adsorption, and thermodynamic parameters may be evaluated from the results. To discuss all of these in detail would require another chapter. Rather than develop all the theoretical and experimental aspects of this subject, therefore, it seems preferable to continue focusing on adsorption isotherms, extracting as much thermodynamic insight from this topic as possible. Within this context, results from adsorption calorimetry may be cited for comparison without a full development of this latter topic.

9.6a Definitions of Heat of Adsorption

The approach we follow is essentially that used to derive the Clapeyron equation (Atkins 1994). Suppose we consider an infinitesimal temperature change for a system in which adsorbed gas and unadsorbed gas are in equilibrium. The criterion for equilibrium is that the free energy of both the adsorbed (subscript s) and unadsorbed (subscript g) gas change in the same way:

$$dG_s = dG_g \tag{84}$$

The following equations may be written for these two quantities if the temperature change is assumed to cause no change in the amount of adsorbed material:

$$dG_g = -S_g dT + V_g dp \tag{85}$$

and

$$dG_s = -S_s dT + V_s dp \tag{86}$$

Substituting Equations (85) and (86) into Equation (84) and rearranging gives

$$\left(\frac{\partial p}{\partial T}\right)_{n_s} = \frac{S_g - S_s}{V_g - V_s} \tag{87}$$

In writing this last result, it has been explicitly noted that the number of moles of adsorbed gas n_s is constant. If the process under consideration is carried out reversibly, $S_g - S_s$ may be replaced by q_{st}/T, where q_{st} is known as the *isosteric* (the same coverage) *heat of adsorption*:

$$S_g - S_s = q_{st}/T \tag{88}$$

Combining Equations (87) and (88) leads to the result

$$\left(\frac{\partial p}{\partial T}\right)_{n_s} = \frac{q_{st}}{T(V_g - V_s)} \tag{89}$$

Equation (89) may be integrated if the following assumptions are made: (a) $V_g \gg V_s$, so that V_s may be neglected, (b) the gas behaves ideally so that the substitution $V_g = RT/p$ may be used, and (c) q_{st} is independent of T. With these assumptions, Equation (89) integrates to

$$\ln\left(\frac{p_1}{p_2}\right) = -\frac{q_{st}}{R}\left(\frac{1}{T_1} - \frac{1}{T_2}\right) \tag{90}$$

Equation (90) shows that the isosteric heat of adsorption is evaluated by comparing the equilibrium pressure at different temperatures for samples showing the same amount of surface coverage. The data of Figure 9.7 may be used as an example to see how this relationship is applied.

For an arbitrarily chosen extent of adsorption, a horizontal line such as the dashed line in Figure 9.7 may be drawn that cuts the various isotherms at different pressures. The pressure coordinates of these intersections can be read off the plot. According to Equation (90), a graph of ln p versus $1/T$ should be linear with a slope of $(-q_{st}/R)$. From Figure 9.7, for example, when the adsorption is 0.10 g ethyl chloride (g charcoal)$^{-1}$ (which corresponds to $\theta = 0.2$), the equilibrium pressures are 0.20, 0.63, and 2.40 torr at -15.3, 0, and 20°C, respectively. When plotted in the manner just described, these data yield a line of slope -5330 K. Multiplication by R gives $q_{st} = 44.3$ kJ mole^{-1} as the isosteric heat of adsorption for this system at $\theta = 0.2$. Table 9.3 lists values of q_{st} for different θ values as calculated from the data in Figure 9.7.

9.6b Comparison of Adsorption Energies Measured Using Different Methods

Table 9.3 is based on the same data that were analyzed according to the Langmuir equation in Section 9.4b.1. Examples 9.3 and 9.5 show that these data are consistent with an adsorption energy of about 37 kJ mole^{-1} according to the Langmuir interpretation.

TABLE 9.3 Values of the Isosteric Heat of Adsorption at Different Values of θ for the Data Shown in Figure 9.7 as Evaluated by Equation (90)

θ	q_{st} (kJ mole^{-1})
0.06	56.9
0.08	47.3
0.10	46.4
0.20	44.3
0.30	41.4
0.40	40.2
0.50	40.2
0.60	41.0
0.70	38.1
0.80	37.2

In presenting this result initially, the remark was made that an independent determination of the adsorption energy would be desirable. Although the present reinterpretation of the same data is not exactly an "independent" determination, it does extract an energy quantity from the experimental results that is free of any assumed model for the mode of adsorption. Accordingly, it is informative to compare the two interpretations. The Langmuir model assumes that a single energy applies to all adsorption sites. Therefore any data analyzed according to this model cannot yield more than one energy. The isosteric heat of adsorption, on the other hand, is evaluated at different degrees of surface coverage. The data in Table 9.3 show that this quantity definitely varies with coverage, tending to level off as $\theta \rightarrow 1$. This is consistent with the picture of the more active "hot spots" being covered first. The average energy that the Langmuir analysis yields is approximately the same as the energy toward which q_{st} converges.

There are several additional thermal quantities besides q_{st} and E that may be generically called "heats of adsorption." One of these is the *integral heat of adsorption* Q_n, which is related to the isosteric heat of adsorption as follows:

$$Q_n = \int_0^n q_{st}\, dn \tag{91}$$

It applies to the process in which n moles of adsorbate are transferred from the bulk gas to the surface, starting from a bare surface. The integral heat of adsorption is determined from data such as those contained in Table 9.3 by graphical integration.

In addition, there are a number of different calorimetric methods to determine heats of adsorption. For example, we may distinguish between isothermal and adiabatic heats depending on the type of calorimeter involved. Of course, thermodynamic relationships exist among these various quantities. We shall not pursue these topics, but one should be aware of the differences and seek precise definitions if the need arises.

The data shown in Figure 9.10 indicate both the kind of data that may be obtained by direct calorimetric study of gas adsorption and some evidence of the effect of preheating on the properties of surfaces. The figure shows the calorimetric heat of adsorption of argon on carbon black. The broken line indicates the behavior of the untreated black, and the solid line is the "same" adsorbent after heating at 2000°C in an inert atmosphere, a process known as *graphitization*. The horizontal line indicates the heat of vaporization of argon.

There are several interesting aspects of this figure that are quite generally observed.

1. The untreated carbon black shows the effect of surface heterogeneity, an effect that becomes smeared out as the coverage increases. The surface of the untreated black contains a certain amount of oxygen in a variety of functional groups (e.g., ether, carbonyl, hydroxyl, and carboxyl). Graphitization results in both the reduction of these oxygen-containing groups and the sharpening of both the basal and prismatic crystallographic planes. Electron micrographs of carbon black before and after graphitization are shown in Figure 1.9.
2. After graphitization, the most notable feature is the sharp discontinuity at monolayer coverage. Beyond the monolayer the heat of adsorption is close to the heat of vaporization (as required by the BET theory) but does show some influence of the surface as well. At coverage below $V/V_m = 1$ the heat of adsorption increases with increasing coverage, probably due to lateral interactions between the adsorbed molecules.

Figure 9.10 is an extreme example of the effect of heating on the properties of an adsorbent. Degassing prior to measuring an isotherm is done under far less severe conditions; nevertheless, these conditions should always be reported when adsorption studies are conducted because of the possibility of surface modification on heat treatment.

Another calorimetric technique for measuring the heat of adsorption consists of comparing the heat of immersion (see Chapter 6, Section 6.6c) of bare solid with that of a solid preequilibrated with vapor to some level of coverage. Table 9.4 summarizes some results of this sort. The experiment consisted of measuring the heats of immersion of anatase (TiO_2) in benzene with the indicated amount of water vapor preadsorbed on the solid. Small quantities of adsorbed water increase the heat of immersion more than threefold so that it approaches the value for water itself. Most laboratory samples will be contaminated with adsorbed water.

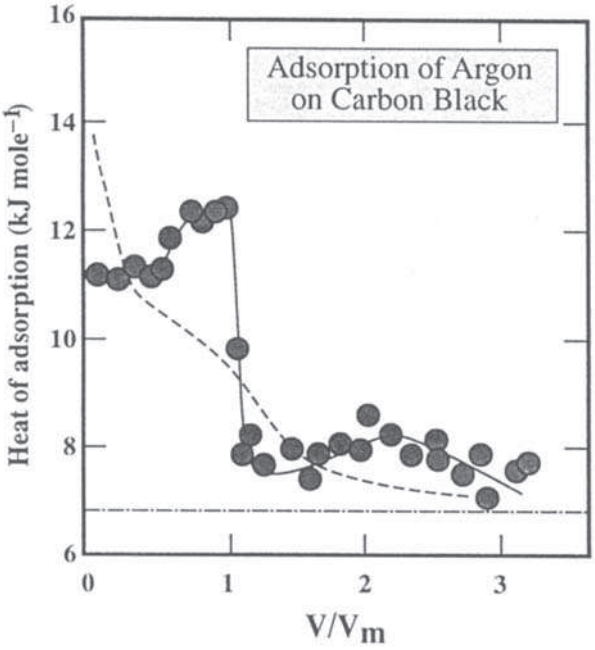


FIG. 9.10 Calorimetric heats of adsorption as a function of coverage for argon on carbon bla at 78K. The dashed line represents untreated black; the solid line is after graphitization at 2000° The horizontal line is the heat of vaporization of argon. (Redrawn with permission from R. Beebe and D. M. Young, *J. Phys. Chem.*, **58**, 93 (1954).)

Unless the material has been carefully pretreated, the actual nature of the surface may be qu different from what would be expected nominally.

9.7 ADSORPTION IN POROUS SOLIDS

High-specific-area solids of the type studied by gas adsorption consist of small particles f which the radius of an equivalent sphere is given by Equation (1.2). This figure may be a fai reasonable measure of a characteristic linear dimension even of irregularly shaped particl

TABLE 9.4 Effect of Traces of Adsorbed Water on the Heat of Immersion of TiO_2 in Benzene

Amount of water adsorbed (mmol kg^{-1})	Heat of Immersion (mJ m^{-2})
0.0	150
2.0	250
4.0	320
10.0	450
17.0	506
Pure H_2O	520

Source: G. E. Boyd, and W. D. Harkins, *J. Am. Chem. Soc.*, **64**, 1195 (1942).

provided the gas adsorption is restricted to the exterior surface of the particles. Any cracks or pores in the solid particles will expose additional adsorbing surfaces and increase the total specific area of the material. Because of this complication, we have specified nonporous solids at several places in this chapter. The particles of high-specific-area solids are small at the beginning, so it follows immediately that any pores in these solids will necessarily have very small dimensions.

In Chapter 6, Section 6.9, we discussed the pressure required to force a liquid—most commonly mercury—into the pores of a solid. Our emphasis in that section was on the pores *between* powder particles when the particles are tightly pressed into a plug. In this section we are not concerned with these interstitial pores since powders are not densely packed in adsorption studies; instead, our interest is in the pores *within* the particles themselves. The difference is a matter of emphasis, however, and both liquid intrusion and gas adsorption complement one another for the study of porous solids.

9.7a Adsorption Hysteresis and Capillary Condensation

We restrict our attention to only one aspect of the adsorption behavior of porous solids: the *hysteresis* they display in their adsorption isotherms. A schematic illustration of the phenomenon is shown in Figure 9.11. Although the region enclosed by the hysteresis loop may have a variety of shapes, in all cases there are two quantities of adsorbed material for each equilibrium pressure in the hysteresis range. That branch of the loop that corresponds to adsorption (increasing pressure) inevitably displays less adsorption at any given pressure than the desorption branch (decreasing pressure). In many cases hysteresis loops such as this are reproducible, although they are not reversible in the thermodynamic sense. Since irreversible processes are involved, the substitution of q_{st}/T for ΔS in the derivation of Equation (90) is not valid. One can go through the motions of evaluating $\partial p/\partial T$ for a system that displays hysteresis, but the "apparent q_{st} values" so obtained (one for each branch) are not easily related to calorimetric adsorption energies.

Adsorption hysteresis is often associated with porous solids, so we must examine porosity for an understanding of the origin of this effect. As a first approximation, we may imagine a pore to be a cylindrical capillary of radius r. As just noted, r will be very small. The surface of any liquid condensed in this capillary will be described by a radius of curvature related to r. According to the Laplace equation (Equation (6.29)), the pressure difference across a curved interface increases as the radius of curvature decreases. This means that vapor will condense

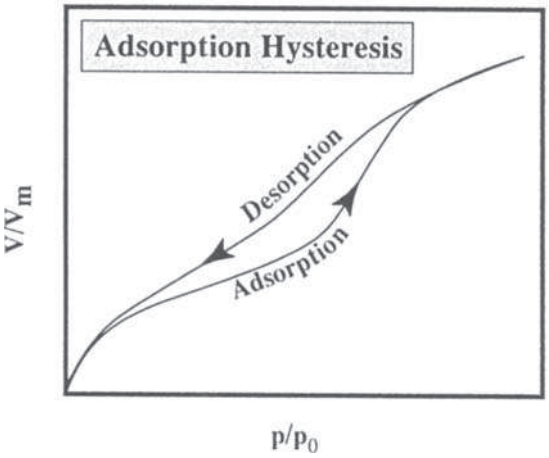


FIG. 9.11 A schematic illustration of hysteresis between the adsorption and desorption branches of an experimental isotherm.

in small capillaries at pressures less than the normal vapor pressure p_0, which is defined for flat surfaces. The condensation of vapors in small capillaries is an equilibrium phenomenon, however, so capillary condensation in itself does not account for hysteresis. It does point out the fact that a liquid-vapor surface is also involved in the adsorption on porous solids for $p < p_0$. In fact, the leveling off of the adsorption in Type IV and Type V isotherms before p_0 is reached is the result of liquid condensation in small pores at these pressures.

For simplicity, let us assume that the liquid condensed in a pore has a surface that is part of a sphere of radius R_s, with $R_s > r$. For a spherical surface we may use the Kelvin equation (Equation (6.40)) to calculate p/p_0

$$N_A k_B T \ln\left(\frac{p}{p_0}\right) = -\frac{2M\gamma}{\rho R_s} \tag{92}$$

A minus sign has been introduced in the Kelvin equation because the radius is measured outside the liquid in this application, whereas it was inside the liquid in the derivation of Chapter 6. In hysteresis, adsorption occurs at relative pressures that are higher than those for desorption. According to Equation (92), it is as if adsorption-condensation took place in larger pores than desorption-evaporation. Since the pore dimensions are presumably constant, we must seek some mechanism consistent with this observation to explain hysteresis.

9.7b Some Models of Capillary Condensation

Figure 9.12 contains sketches for several different models of pores that will be useful in our discussion of capillary condensation. Figure 9.12a is the simplest, attributing the entire effect just described to variations in pore radius with the depth of the pore. That is, when liquid first begins to condense in the pore, the larger radius R_a determines the pressure at which the adsorption-condensation occurs. Once the pore has been filled and the desorption-evaporation branch is being studied, the smaller radius R_d determines the equilibrium pressure. Although bottlenecked pores of this sort may exist in some cases, this model seems far too specialized to account for the widespread occurrence of hysteresis.

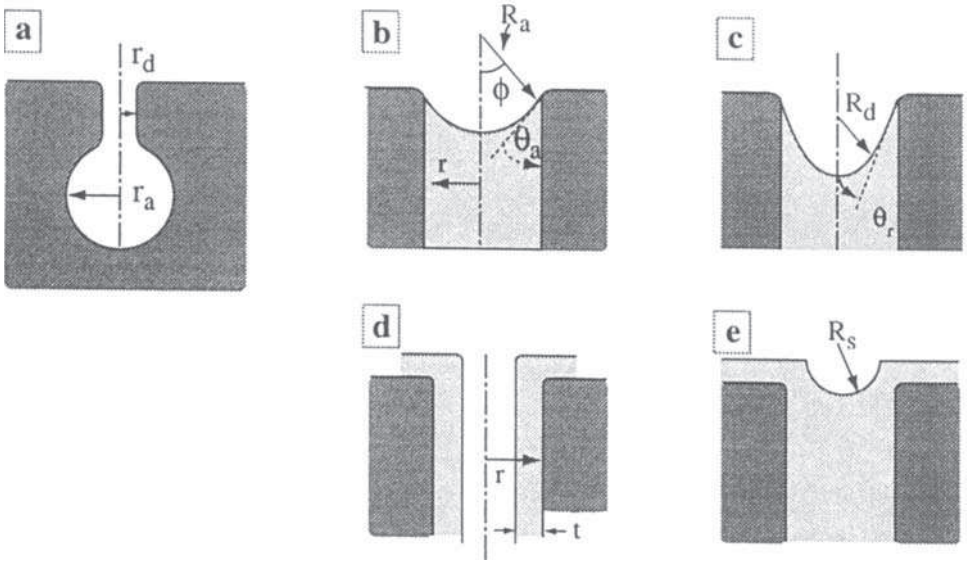


FIG. 9.12 Five models for capillary condensation. The radius of the pore equals r, the radius of curvature of the spherical meniscus is R_s, and t is the thickness of the adsorbed layer. The subscripts a and d refer to adsorption and desorption.

Figures 9.12b and 12c represent another model, based on contact angle hysteresis. These sketches represent the surface of liquid in a capillary during the adsorption and desorption stages of hysteresis, respectively. In Figure 9.12b the capillary is filling; in Figure 9.12c it is emptying. Accordingly, the advancing and receding values of the contact angle apply to adsorption and desorption, respectively. The radius R_s of the spherical surface and the radius r of the capillary are related through the contact angle θ and its complementary angle ϕ (see Fig. 9.12):

$$r = R_s \sin \phi = R_s \cos \theta \tag{93}$$

Substituting Equation (93) into Equation (92) gives

$$N_A k_B T \ln \left(\frac{p}{p_0} \right) = -\frac{2 M \gamma \cos \theta}{\rho r} \tag{94}$$

For a pore of constant radius the equilibrium pressure decreases as cos θ increases or as θ decreases. It will be recalled from Chapter 6 that advancing contact angles are larger than receding ones. Therefore this model is consistent with the observation that desorption-evaporation occurs at lower relative pressures than adsorption-condensation. The only objection to this explanation of adsorption hysteresis is that it makes no reference whatsoever to adsorption!

Figures 9.12d and 12e illustrate another model for adsorption hysteresis that considers multilayer adsorption explicitly. During adsorption the capillary is viewed as a cylinder of radius $(r - t)$, with t the thickness of the adsorbed layer at that pressure. This is represented by Figure 9.12d. For such a surface the Kelvin equation becomes

$$N_A k_B T \ln \left(\frac{p}{p_0} \right) = -\frac{M \gamma}{\rho (r - t)} \tag{95}$$

since one of the radii of curvature is infinite. As multilayer adsorption proceeds, however, the thickness of the adsorbed layer equals r. Once the pore is thus filled, its surface becomes a meniscus that may be treated by Equation (94) as a portion of a sphere, as shown in Figure 9.12e. In the event that $\theta = 0$, Equation (92) applies with $R_s = r$. Comparison of adsorption and desorption in this situation is particularly easy if $t \ll r$ in Equation (95). In that case, the right-hand side of Equation (95) is proportional to $1/r$, and comparison with Equation (92) reveals that

$$\left(\frac{p}{p_0} \right)_d^{1/2} = \left(\frac{p}{p_0} \right)_a \tag{96}$$

Since both of these ratios are less than unity, $p_a > p_d$. This model is qualitatively consistent with the observed hysteresis but is difficult to apply quantitatively because it neglects differences between the adsorbed material on the surface and that in the bulk liquid.

Gas adsorption data may be analyzed for the distribution of pore sizes. What is generally done is to interpret one branch of the isotherm and use an appropriate equation to calculate the effective pore radius at a given pressure. The amount of material adsorbed or desorbed for each increment or decrement in pressure measures the volume of pores with that effective radius.

9.8 ADSORPTION ON CRYSTAL SURFACES

Until now, our treatment in this chapter of the solid-gas interface has been very one-sided, focusing almost entirely on the adsorbed layer(s). To be sure, we have extracted some information about the solid, e.g., the specific surface area and the presence or absence of pores. In addition, we have extracted values for the energy of interaction between the solid adsorbent and the adsorbate molecules but, other than this, the influence of the solid has been ignored. Likewise, the solids we have considered have been high-surface-area powders, presenting

either a multitude of crystal faces or amorphous surfaces to the adsorbing gas molecules. Because of this approach, we have not asked whether the order of a crystalline surface carries over to monolayer or submonolayer adsorption. As a matter of fact, the notion that there may be order in a layer of adsorbed molecules actually contradicts the models on which the isotherm presented above were based.

In this section we discuss the adsorption on crystal surfaces. First, we begin with low-energy electron diffraction (LEED)—an experimental method for examining crystal surfaces—and introduce some basic crystallographic concepts needed to interpret the experimental measurements. Then we look at the implication of the adsorbate structure to adsorption and the structure of adsorbed layers using LEED measurements.

9.8a Low-Energy Electron Diffraction

Diffractometry provides an excellent tool for examining structure so we turn now to *low-energy electron diffraction* to study the order at a specific face of a single crystal, with and without adsorbed molecules. For the remainder of the chapter, we focus attention on the faces of the metal crystals. There are several reasons for this choice:

1. Many metals crystallize in the relatively simple cubic (either primitive, face-centered, or body-centered) structures.
2. Metal surfaces have been studied extensively by LEED in the context of a number of practical applications such as corrosion, friction, and semiconductor devices.
3. Metal surfaces are important as catalysts, and this discussion of LEED enables us to introduce some ideas about chemisorption and catalysis.

9.8a.1 Basic Principle

In LEED a beam of low-energy electrons rather than x-rays is used to form the diffraction pattern, but otherwise many of the concepts, relationships, and vocabulary are based on x-ray diffraction (Van Hove and Tong 1979). Accordingly, our discussion of LEED includes a review of pertinent aspects of this topic. Since diffractometry can get quite involved, we tailor our review to those subjects most helpful in getting us started and leave more advanced concepts for further study.

The de Broglie concept of wave-particle duality enables us to calculate the wavelength of an electron. According to the *de Broglie equation*,

$$\lambda = \frac{h}{p} = \frac{h}{[2\,\mathrm{m}\,(KE)]^{1/2}} \tag{97}$$

in which h is Planck's constant, and the momentum of the electron, $p = [2m(KE)]^{1/2}$, can be calculated in terms of its mass and the accelerating voltage that accounts for its kinetic energy. A straightforward application of this formula enables us to calculate that a voltage of 150 V gives electrons a wavelength of 0.1 nm, the order of magnitude of interatomic spacings in crystals:

$$\lambda = 6.63 \cdot 10^{-34}\ \mathrm{J\ s}/[2(9.11 \cdot 10^{-31}\ \mathrm{kg})(1.6 \cdot 10^{-19}\ \mathrm{C})(150\ \mathrm{V})]^{1/2} = 10^{-10}\ \mathrm{m} = 0.10\ \mathrm{nm}$$

Thus the electrons used in LEED are usually in the range 10 to 300 eV, lower in energy by one to two orders of magnitude than those used in scanning electron microscopy (SEM; Chapter 1) and Auger electron spectroscopy (AES; see Section 9.1b.2). Davisson (Nobel Prize, 1937), working with a series of collaborators, is generally credited with the first application of electrons to diffraction studies.

What makes LEED of particular interest in the study of surfaces is the fact that low-energy electrons such as these are only able to penetrate and reemerge by diffraction from a few atomic thicknesses of the target material. Even at that, the diffracted beam acquires a background of secondary electrons picked up through inelastic interactions with the solid. Figure 9.13 is a schematic of a LEED apparatus. Note that several grids are interposed between the specimen and the viewing screen. One of these is adjusted in potential so that only

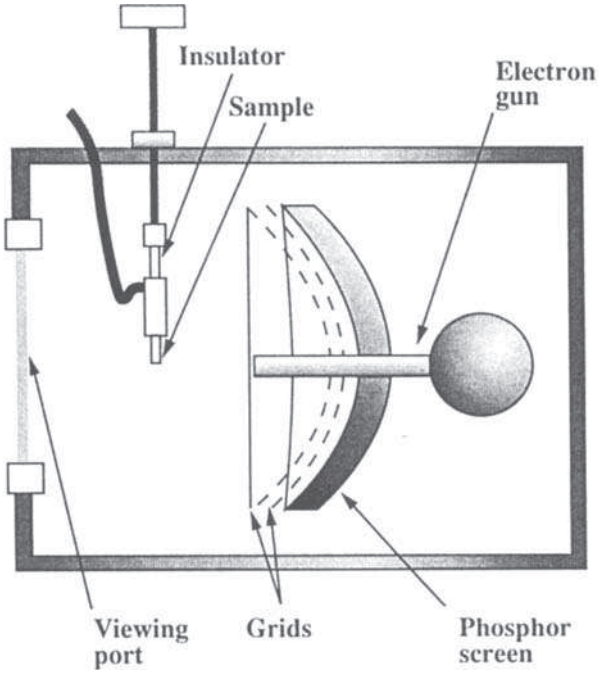


FIG. 9.13 Schematic illustration of a low-energy electron diffraction (LEED) apparatus. (Redrawn with permission from Atkins 1994.)

electrons of the original energy—those elastically diffracted—are allowed to pass; secondary electrons of lower energy are blocked. The diffracted beam is accelerated by a final, positively charged grid to produce an image on the fluorescent screen.

We saw in connection with the discussion of Figure 9.5 that measurable gas adsorption occurs even at gas pressures as low as 10^{-10} torr. As a matter of fact, the two-dimensional density of the adsorbed molecules is not low enough to conform to the two-dimensional ideal gas law even when the pressure is on the order of 10^{-10} torr. A question of considerable practical importance, then, is how low the pressure must be for an initially clean surface to remain that way for a reasonable period of time. The above reference to adsorption cites equilibrium data that are not useful for answering questions of rate.

Instead, we must turn to the kinetic molecular theory of gases for an estimate of the frequency with which molecules collide with a solid surface. We shall not be misled, however, if we anticipate that this pressure is low. Example 9.6 is a numerical examination of gas collisions with walls.

* * *

EXAMPLE 9.6 *Rate of Atomic Collisions as a Function of Pressure.* Assuming 10^{19} atoms per square meter as a reasonable estimate of the density of atoms at a solid surface, estimate the time that elapses between collisions of gas molecules at 10^{-6} torr and 25°C with surface atoms. Use the kinetic molecular theory result that relates collision frequency to gas pressure through the relationship $Z = 1/4\ \bar{v}N/V$, for which the mean velocity of the molecules $\bar{v} = (8RT/\pi M)^{1/2}$ and N/V is the number density of molecules in the gas phase and equals pN_A/RT. Repeat the calculation at 10^{-8} and 10^{-10} torr.

Solution: If 10^{19} collisions occurred per square meter per second, each surface atom would be hit an average of once each second. Therefore we must divide 10^{19} by Z at the required pressures to obtain the elapsed time. Assuming a molecular weight of 28 g mole^{-1} for the gas, we obtain

$$(8RT/\pi M)^{1/2} = [8(8.314\ \text{J K}^{-1}\ \text{mole}^{-1})(298) / \pi(0.028\ \text{kg mole}^{-1})]^{1/2}$$
$$= 475\ \text{m s}^{-1}$$

Converting the gas pressure to SI units gives, for $p = 10^{-6}$ torr,

$$p = [10^{-6}\ \text{torr} \cdot 1.013 \cdot 10^{5}\ \text{N m}^{-2}] / 760\ \text{torr}$$
$$= 1.33 \cdot 10^{-4}\ \text{N m}^{-2}$$

and

$$N/V = \frac{(1.33 \cdot 10^{-4}\ \text{Nm}^{-2})(6.02 \cdot 10^{23}\ \text{molecules mole}^{-1})}{(8.314\ \text{J K}^{-1}\ \text{mole}^{-1})(298\ \text{K})}$$
$$= 3.24 \cdot 10^{16}\ \text{molecules m}^{-3}$$

Therefore

$$Z = (1/4)\bar{v}N = (475)(3.24 \cdot 10^{16})/4 = 3.85 \cdot 10^{18}\ \text{collisions m}^{-2}\ \text{s}^{-1}$$

and

$$\text{Elapsed time} = 10^{19}/3.85 \cdot 10^{18} = 2.60\ \text{s}$$

Since N/V and Z are directly proportional to p and the elapsed time is inversely proportional to p, the values of these quantities at other pressures can be written by inspection. At $p = 10^{-8}$ torr, $N/V = 3.24 \cdot 10^{14}$ molecules m^{-2}, $Z = 3.85 \cdot 10^{16}$ m^{-2} s^{-1}, and time = 260 s = 4.3 min; at $p = 10^{-10}$ torr, time $= 2.6 \cdot 10^{4}$ s = 7.2 hr. ■

* * *

9.8a.2 Ultrahigh Vacuum Requirement

The calculation in Example 9.6 shows that until gas pressure is lowered below, say, 10^{-8} torr, the rate of surface bombardment is too great for the surface to remain clean long enough for an experiment to be carried out. Pressures on the order of 10^{-6} torr are achieved in ordinary vacuum lines with diffusion and mechanical pumping. Any pressure below this is considered ultrahigh vacuum (UHV); today devices are commercially available that routinely achieve pressures in the range 10^{-9} to 10^{-12} torr.

An assortment of different pumping arrangements is employed to achieve UHV, including ion pumps, cryopumps, and getter pumps. With each of these the pumped gas is stored within the system rather than being removed, as is the case with diffusion pumps. This means that caution must be exercised so that the stored gas is not re-released into the system during the course of an experiment. Ion pumping is achieved by ionizing with electrons the gas to be removed and then collecting these ions at a metallic cathode, where they remain adsorbed. Getter pumps selectively adsorb or dissolve the gas without using ionization to collect the molecules, and cryopumps use low temperatures to promote adsorption. These three modes of pumping are thus the highly specific reverse of outgassing procedures: Adsorption at a specific site (the pump) is promoted to lower the pressure.

It is not enough merely to reach a low pressure; it must also be possible to measure it. Ionization gauges are almost always used for this purpose. In these the residual gas is ionized, collected at an electrode, and the resulting current measured. The current varies linearly with the gas pressure down to about 10^{-11} torr. If the ions are separated by mass—making the gauge a mass spectrometer—then the partial pressures of various gases in the vacuum chamber can be determined.

Next let us consider the preparation of the sample itself. Above we referred to the desirability of keeping a surface clean; nothing was said about the problem of obtaining such a surface in the first place.

9.8a.3 Sample Preparation

Two preferred methods for preparing clean surfaces consist of generating the surfaces under high vacuum at the beginning. Using remote control manipulators to crush a sample under vacuum exposes fresh surface to an environment in which adsorption equilibrium is very slow. This technique produces a heterogeneous array of crystal faces, however. Far more suitable for the specific and localized examination that diffraction methods offer is crystal cleavage

under vacuum. Again by use of remote manipulators, a sample may be cleaved along a predetermined crystal plane to expose a fresh surface of known character. Surface repair through annealing may be necessary after this relatively violent procedure. The need for annealing also means that UHV chambers must possess the capability for heating samples in a controlled and measurable way. Sample preparation by cleavage plus annealing gives the best defined surfaces for subsequent examination.

Prior to a discussion of the use of LEED for structural measurements, a review of a few concepts from crystallography is needed for interpreting data from LEED.

9.8b Some Basic Concepts from Crystallography

A crystal is an orderly array of atoms or molecules but, rather than focusing attention on these material units, it is helpful to consider some geometrical constructs that characterize its structure. It is possible to describe the geometry of a crystal in terms of what is called a *unit cell*: a parallelepiped of some characteristic shape that generates the crystal structure when a three-dimensional array of these cells is considered. We then speak of the lattice defined by the intersections of the unit cells on translation through space. Since we are interested in crystal surfaces, we need to consider only the two-dimensional faces of these solids. In two dimensions the equivalent of a unit cell is called a *unit mesh*, and a net is the two-dimensional equivalent of a lattice. Only four different two-dimensional unit meshes are possible.

Figure 9.14 shows a set of points in one plane of a cubic structure. Much of our discussion in this section is based on cubic unit cells for simplicity, but the results are quite general. It also turns out that many metals crystallize with cubic unit cells. We may regard this as representing the surface net of a cubic lattice. Several sets of parallel lines connecting various points in the net have been drawn in the figure. These rows of atoms—like lines etched on a diffraction grating—are the origin of the LEED diffraction.

The lines in Figure 9.14 are, of course, the edges of crystallographic planes that slice through the crystal and are identified by the so-called Miller indices just as the lines themselves are. Since we are working in two dimensions, only two indices are needed to characterize them. The sets of numbers labeling the lines in Figure 9.14 are called the *Miller indices*—represented as hk—for that line. The easiest way to remember the significance of these indices is to note that h and k count the number of lines (of type hk) crossed in moving from one net point to the next in the x and y directions, respectively, for cubic cells. Thus the 11 line in Figure 9.14 cuts across the surface in such a way that one of these lines is crossed in moving from one net point to the next in the x direction and one is crossed in moving from point to point in the y direction.

Before we leave Figure 9.14, there are two additional points to be made concerning the spacing and atom density of the various lines. The first thing to note for the simple array of points in Figure 9.14 is that the actual spacing of the net points is exactly the same in all directions, the edge length of the unit cell (a cube) δ. Each of the various sets of lines that can be drawn through this net has its own characteristic spacing, say, d_{hk} for the hk line. What makes this significant is that in LEED studies electrons are reflected off the various crystallographic planes and constructive interference occurs when the spacing between the diffracting surfaces d_{hk} and the angle of incidence satisfy the Bragg equation (see Equation (1.21)),

$$n\lambda = 2d_{hk} \sin \theta \tag{98}$$

where λ is the wavelength of the electrons and n ($= 1, 2, 3, \ldots$) is the order of the diffraction. The Bragg equation is discussed in connection with microscopy in Chapter 1, Section 1.6. The point here is this: First-order ($n = 1$) diffraction results in reinforcement at different angles for different reflecting planes, even when the actual net is characterized by a single distance δ. Our interest is in the relationship between δ and the spacing d_{hk} for planes with those indices. It is possible to show that there are only three area-filling, two-dimensional shapes: square, rectangular, and oblique parallelogram. Table 9.5 shows how the inter-row spacing d_{hk} is related to the Miller indices and the side lengths a and b for these different shapes of two-dimensional unit cells.

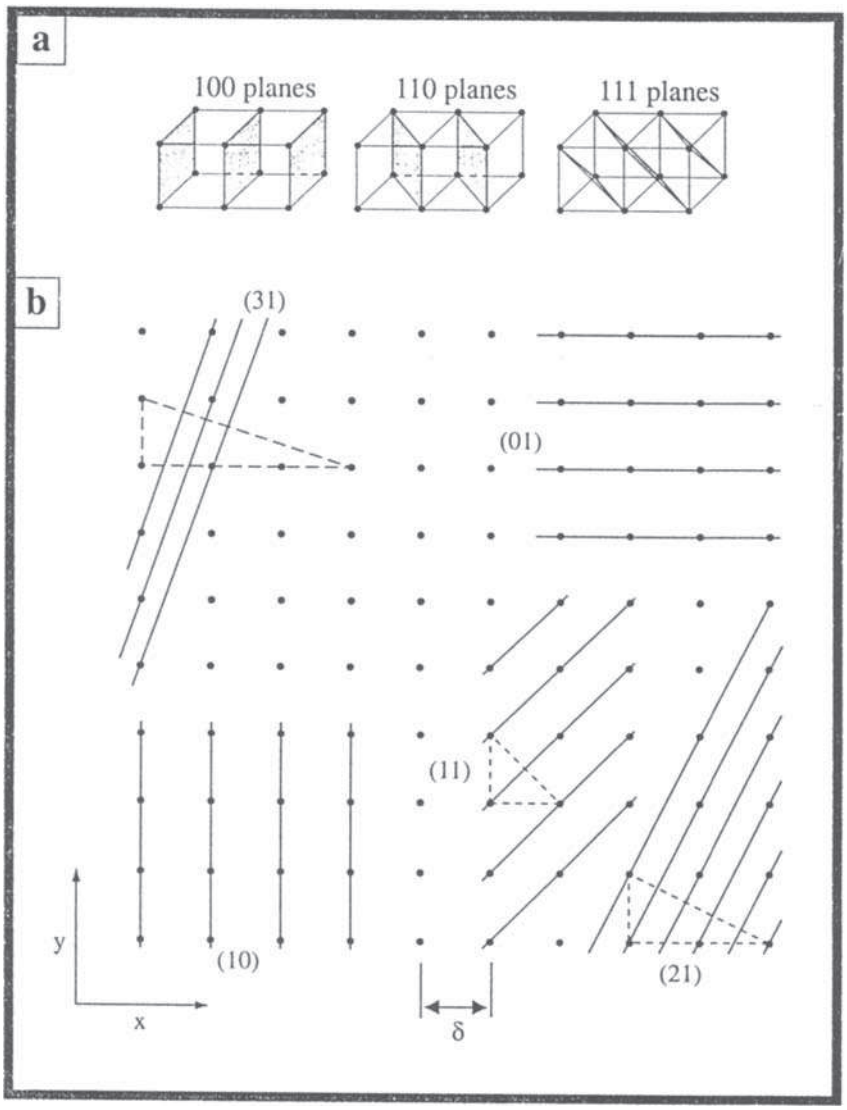


FIG. 9.14 Some crystallographic planes through a simple cubic lattice in which the atomic spacing is δ. The numbers in parentheses are the Miller indices of (a) planes and (b) lines defined by the edges of planes.

A second result that is evident from an inspection of Figure 9.14 is the fact that planes with lower Miller indices have higher atomic densities. Those planes characterized by high Miller indices strike out across the lattice at acute angles and cross several planes of atoms before actually intercepting an atom. This makes the average distance between atoms greater in planes that have high Miller indices. The surface tension is less in planes of low Miller index than in planes of higher index, in which the interatomic spacing is larger. If it were possible to deform the former surface into the latter, work would have to be done against the attractive forces between atoms to increase this separation. We may think of this work as an increment in surface free energy in the planes of high index that is not present when the Miller indices are lower. Surfaces of lower γ and hence lower Miller index are thermodynamically more stable. Surface studies conducted on the specific face of a crystal generally involve the 100, 110, or

TABLE 9.5 Relationship Between the Spacing d_{hk} and the Side Lengths δ and δ' in the Two-Dimensional Unit Cells of Square, Rectangular, and Oblique Parallelogram Nets

Geometry	Relationship
Square ($\delta = \delta'$)	$d_{hk}^{-2} = (h^2 + k^2)/\delta^2$
Rectangle ($\delta \neq \delta'$)	$d_{hk}^{-2} = (h/\delta)^2 + (k/\delta')^2$
Oblique, general ($\delta \neq \delta'$; $\alpha \neq 90°$)[a]	$d_{hk}^{-2} = (h^2/\delta^2 \sin^2\alpha) + (k^2/\delta'^2 \sin^2\alpha) - (2hk \cos\alpha/\delta\delta' \sin^2\alpha)$
Oblique, hexagonal ($\delta = \delta'$; $\alpha = 120°$)[a]	$d_{hk}^{-2} = 4(h^2 + hk + k^2)/3\delta^2$

[a]The angle between sides is α in the oblique case and equals 120° for the hexagonal element.

111 plane for this reason. (Remember, three indices are necessary to describe the planar face of a crystal.)

9.8c Interpretation of Low-Energy Electron Diffraction Patterns

Generally, LEED experiments are conducted on specified faces of single crystals. When this is done, the diffraction pattern produced consists of a series of spots with a location, shape, and intensity that can be interpreted in terms of the surface structure. We focus attention on what can be learned from the location and shape of the spots since the study of intensity is beyond the scope of this book. It is generally assumed that the surface examined by LEED is an extension of an already-known bulk crystal structure. The correctness of this assumption can be tested, and results are often expressed in terms of modifications of the three-dimensional structure at the surface. Before we turn to the LEED patterns below, we must first figure out how they are read.

Figure 9.15a follows the path of the diffracted electron beam from the edge of the planes—separated by the distance d_{hk}—to the photographic surface where the LEED pattern is observed. The diffracting planes are shown in greatly exaggerated size in this version, and the angle of incidence is purposely drawn to be exceptionally small. Small angles like these are not typical of LEED, but the argument that follows will help index the LEED pattern.

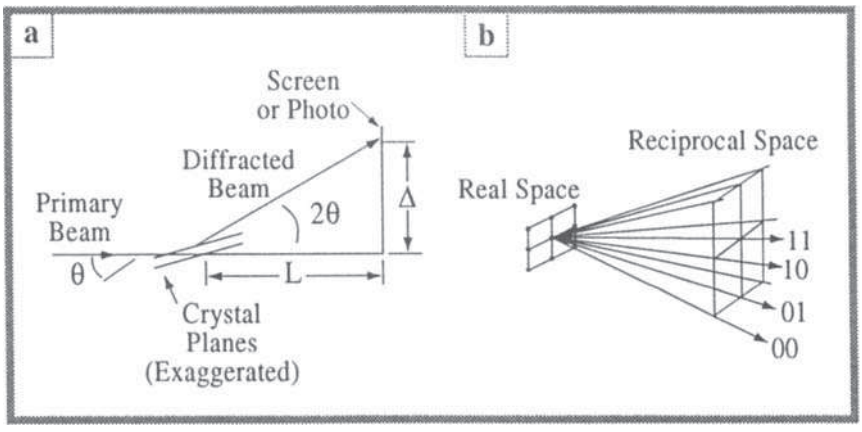


FIG. 9.15 Path of a diffracted electron beam in low-energy electron diffraction (LEED), and indexing of points in reciprocal space. (a) interaction of a diffracted beam with a photographic plate for small angles of incidence; and (b) illustration of the indexing of points in reciprocal space relative to the primary beam, labeled 00.

In Figure 9.15a the diffracted beam produces a spot on the photographic plate a distance Δ from the point at which the primary beam strikes. From trigonometry, it is evident that $\tan 2\theta = \Delta/L$. For the small angles considered here $\tan x \approx x$; therefore $2\theta \approx \Delta/L$, or $\theta \approx \Delta/2L$. For small angles, $\sin x \approx x$ also; therefore $\sin \theta \approx \Delta/2L$ for the situation shown in the figure. Combining this small-angle approximation with the Bragg condition for first-order diffraction,

$$2 d_{hk} \sin \theta = \lambda \tag{99}$$

we obtain

$$\lambda = d_{hk} \frac{\Delta}{L} \tag{100}$$

Since d_{hk} and L are constant, Equation (100) predicts that Δ is inversely proportional to d_{hk}, with Δ the distance of the diffraction spot from the spot produced by the primary. Figure 9.14 shows that d_{hk} is largest for planes of low Miller indices; since Δ varies inversely with d_{hk}, it follows that spots nearest the primary spot are due to low Miller index planes. Likewise, more distant spots are due to planes of higher index. There is a reciprocal relationship between the location of the spot on the photographic plate and the separation of the planes responsible for the spot.

Since every point in an LEED photograph is associated with a set of rows across a net, it is standard to identify the points by the same index numbers as their originating lines. Figure 9.15b shows how the points in a photograph are identified by a pair of indices hk. The primary beam is labeled 00, and the distance Δ in the approximation described above is measured outward from 00.

Now that the theory and vocabulary of LEED are established, let us look at some simple LEED patterns and see what they tell us about solid surfaces.

9.8d Low-Energy Electron Diffraction Applied to Metal Surfaces

Figure 9.16 is a set of photographs of LEED patterns on clean metal surfaces and surfaces with adsorbed gases. Several variables are in effect to produce the LEED patterns shown:

1. Parts (a) and (d) of Fig. 9.16 are the 100 surface of tungsten and the 111 surface of platinum, respectively. The symmetry of these patterns characterizes the cubic and hexagonal packing of the crystal faces.
2. LEED patterns could be used to distinguish among different crystal faces.
3. Parts (a) and (d) of Fig. 9.16 are clean surfaces, while parts (b) and (c) of the figure are surfaces with adsorbed species present.
4. LEED patterns could be used to distinguish between clean surfaces and those with adsorption, at least in some cases.
5. Part (b) of Fig. 9.16 is the tungsten 100 face with adsorbed oxygen, while part (c) of the figure is the same surface with adsorbed hydrogen.
6. LEED patterns could be used to distinguish among adsorbed species, at least in some cases.

In most instances there are much simpler ways to determine the orientation of a crystal and the presence or absence of adsorbate molecules and their nature than to use LEED. The unique power of LEED is its ability to measure order at a surface. Therefore we may state the following:

1. LEED enables us to determine the structure of the solid surface and to compare this with the bulk structure.

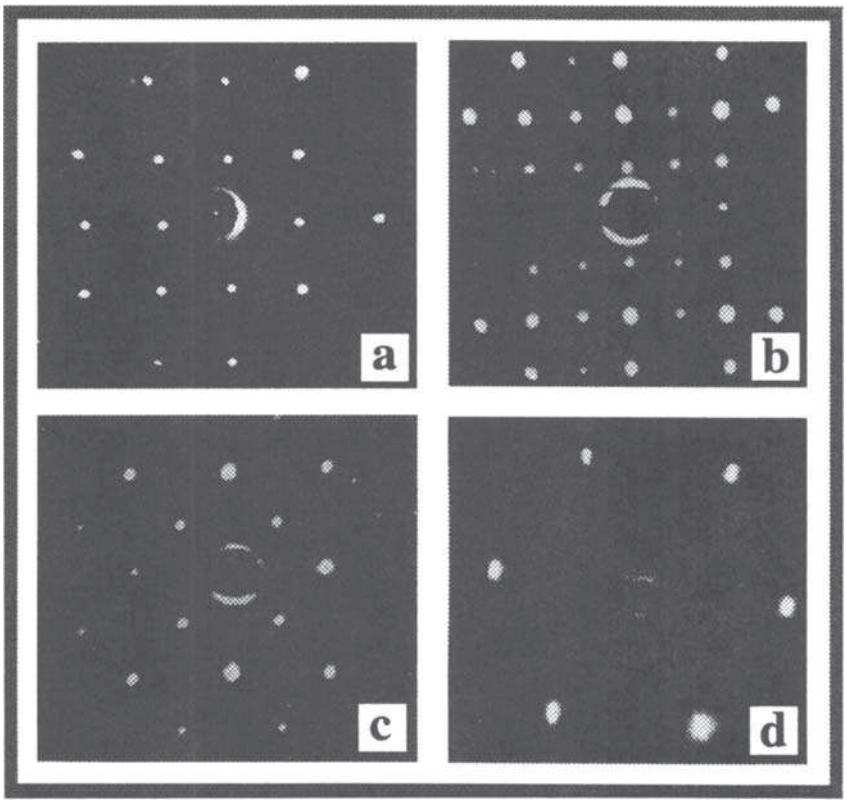


FIG. 9.16 Low-energy electron diffraction (LEED) patterns for clean metal surfaces and surfaces with adsorption: (a) clean W(100) surface; (b) W(100) with adsorbed oxygen; (c) W(100) with adsorbed hydrogen; and (d) clean Pt(111). (Figs. 19.16a–19.16c reprinted with permission from P. J. Estrup, In *Modern Diffraction and Imaging Techniques in Materials Science* (S. Amelincks, R. Gevers, G. Remaut, and J. Van Landuyt, Eds.), North Holland, Amsterdam, Netherlands, 1970; Fig. 9.16d reprinted with permission of Somorjai 1981.)

2. LEED allows us to observe and measure the order that exists in some adsorbed monolayers.

Item 2 in this last list is discussed somewhat further in Example 9.7.

* * *

EXAMPLE 9.7 *Comparison Between Bulk and Surface Structures Using Low-Energy Electron Diffraction Patterns.* If we accept that an LEED pattern has the same symmetry as the net of surface atoms responsible for its formation, what additional information is needed to complete the comparison between bulk and surface structures for the tungsten surface shown in Figure 9.16a?

Solution: We assume the crystal structure of the bulk metal, including the dimensions of the unit cell, are known. Comparison, then, depends on determining this quantity for the surface.

1. Taking photographic magnification into account, measure the distance of the diffraction spots from the origin, the spot produced by the primary beam.
2. These distances are multiples of the crystallographic spacing d_{hk}. The spacing may be calculated by evaluating a magnification factor from the experimental geometry in the manner suggested by Figure 9.15a.

3. The d_{hk} values so calculated are expressed in units of λ, so the wavelength of the electrons must be known to obtain absolute values for these spacings.
4. To convert d_{hk} values to the dimensions of the two-dimensional unit cell by a formula from Table 9.5, the spots in the LEED pattern must be indexed.
5. Since the pattern is a simple square in Figure 9.16a and since the spots nearest the origin have the lowest indices, the pattern can be indexed just as rectilinear graph paper might be marked off. Negative indices are possible; these are written above the number affected. Figure 9.17a shows how the points in Figure 9.16a are indexed.

From the index numbers and the experimental parameters, the dimensions of the two-dimensional unit cell can be determined. ■

* * *

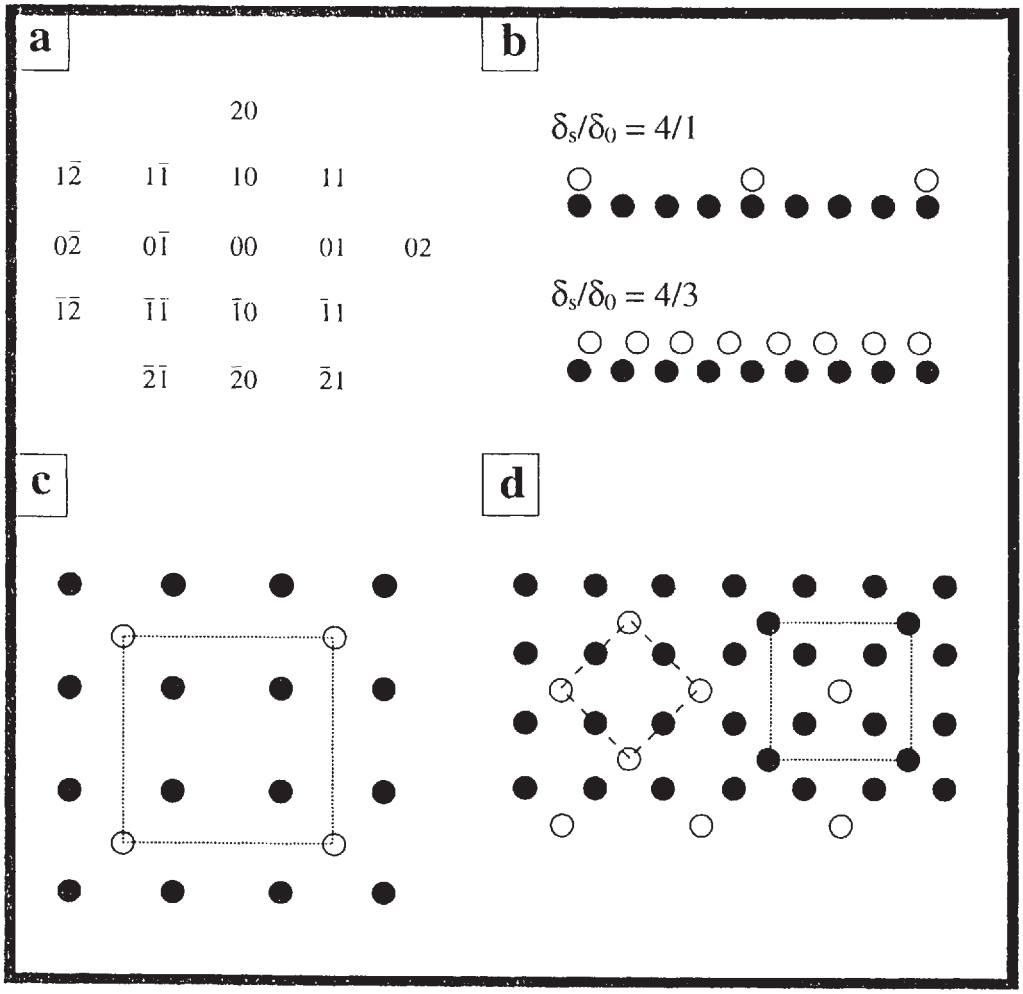


FIG. 9.17 Illustrations of LEED spots and coherent structures formed by adsorbed species. (a) indexing of the LEED spots of W(100) pattern in Figure 9.16a as described in Example 9.7; (b) side view of coherent structures formed by adsorbed (open circles) species on metal surface; (c) unit mesh for W(100) with adsorbed oxygen (open circles); (d) unit mesh for W(100) with adsorbed hydrogen (open circles).

For the 100 face of tungsten, written W(100), it has been found that the characteristic dimension of the unit cell in the two-dimensional lattice is 6% less at the surface than in the bulk crystal, although the packing geometry remains the same in both. The change in separation without a change in symmetry for surface atoms is called *relaxation* and is widely observed, particularly in crystal faces with relatively low packing efficiency. For example, the contraction of the Al(110) and Mo(100) surfaces are 5–15% and 11–12%, respectively. By contrast, the high atomic density 111 surfaces of silver and platinum show no contraction, and the Fe(111) surface contracts by 1.5%. This phenomenon may be explained as an attempt by surface atoms to compensate for their lower coordination number by pulling closer to their neighbors for more effective bonding. In some instances, it is observed that clean surfaces have totally different structures from what would be expected on the basis of bulk structures; this is called *surface reconstruction*.

Although UHV is required for LEED measurement, there is considerable interest in applying this technique to surfaces that carry adsorbed species. In view of our discussion of adsorption equilibrium above in the chapter, there is a difficulty here since adsorbed molecules imply an equilibrium gas phase. One way around this problem is to study surfaces at a sufficiently low coverage that the equilibrium gas pressure is compatible with the LEED technique. When higher pressures are desired, the surface is first equilibrated, then the excess gas is pumped out, and the surfaces before and after adsorption are compared through LEED. Chemisorption is better suited for study by this method than physical adsorption because the adsorbed layer remains intact when the equilibrium gas is removed.

Figures 9.16b and c are examples of this, and we discuss this figure in detail. First, let us consider the circumstances in which we might expect LEED to be sensitive to adsorbed species. From the perspective of approaching gas molecules, the surface of a single crystal presents an ordered array of adsorption sites (see, for example, Fig. 1.22 in Vignette I.8). We might anticipate that this order is carried over into the adsorbed layer, at least at low coverage, in which crowding and lateral interactions are not complications. Note also that this implies highly specific adsorption; this is another reason why LEED is especially valuable in the study of chemisorption. Since LEED responds to surface order, randomly adsorbed molecules such as those held by weak physical adsorption might go undetected. It was suggested at the beginning of this section that the LEED pattern of a clean surface might be taken as evidence that no adsorption had occurred. Actually, it may only indicate the absence of order in an adsorbed layer. A supplementary chemical analysis is a more definitive proof of surface cleanliness.

Figure 9.17b schematically represents a cross-sectional view of the surface of a solid and represents the topmost layer of atoms by shaded circles. The open circles represent molecules in an ordered pattern on the solid substrate. Since the adsorbed molecules are ordered, their structure on the surface is characterized by what is called a *supernet*. Suppose we define δ_0 as the characteristic spacing of the substrate and δ_s the equivalent quantity for the supernet. Then the two arrangements in Figure 9.17b are described by the ratios $\delta_s/\delta_0 = 4/1$ and $\delta_s/\delta_0 = 4/3$. Building on the notion of reciprocal distances as developed in the discussion of Figure 9.15, it follows that the adsorbed layer with $\delta_s/\delta_0 = 4/1$ should produce spots with a separation that is 1/4 that of the substrate. Likewise, for the case when $\delta_s/\delta_0 = 4/3$, a pattern of spots with a separation that is 3/4 that of the substrate is predicted. Thus, if the substrate produces spots at, say, 0 and 1, extra spots would be expected at 1/4, 2/4, and 3/4 for the $\delta_s/\delta_0 = 4$ case, and at 3/4, 6/4, and 9/4 when $\delta_s/\delta_0 = 4/3$. The cases illustrated here are called *coincident structures* since the two patterns coincide periodically. When there is no correlation between two structures, they are said to be *incoherent*.

With this background we are now in a position to make sense of the LEED pattern shown in Figure 9.16b. The substrate is the same W(100) surface shown in Figure 9.16a, but in Figure 9.16b the surface carries adsorbed oxygen. In the LEED pattern additional spots appear midway between the spots produced by the clean tungsten. The extra spot at the 1/2 position means that $\delta_s/\delta_0 = 2$ for the oxygen in this experiment; that is, the dimension of the oxygen unit mesh is twice that of the unit mesh of the substrate. Figure 9.17c is a schematic illustration of a possible supernet that is consistent with this description. It should be pointed out that

spot location alone does not tell us where in the unit mesh of the supernet the adsorbed oxygens are located, only that they display a periodic structure with two times the repeat distance of the solid. The relative positions of the adsorbate and substrate atoms within their respective nets may be calculated from intensity data.

This example of a supernet makes it evident that an assortment of these supernets is possible with different values of δ_s/δ_0. This suggests that a system of designation is needed to distinguish among the various possibilities. One such system uses the notation $p(n \times m)$ in which the letter p (for primitive) indicates that the adsorbed net has the same primitive unit mesh as the substrate. In this system of notation n and m are integers, not necessarily the same, that express the expansion of the supermesh dimensions relative to the substrate along the mesh axes. Since LEED patterns of this type portray specific crystal faces under definite adsorption conditions, they are labeled to identify both the solid and adsorbate structures. Thus Figure 9.16b is a photograph of the W(100)$-p(2 \times 2)-$oxygen LEED pattern. This system of labeling is known as the *Wood notation*.

Figure 9.16c is another example of an LEED pattern of the W(100) surface carrying an adsorbate, this time hydrogen. By their greater brightness, the spots originating from the substrate can be identified, and it is seen that the extra spots caused by the adsorbate once again lie midway between the tungsten spots, but this time with a different orientation from the array produced by the substrate. The midpoint positioning of these extra spots suggests that this adsorbate also shows a 2×2 supernet, but one that is rotated by 45° compared to the substrate net. A schematic illustration of the two nets is shown in Figure 9.17d, in which the substrate atoms are represented by filled circles and the adsorbed species by open circles. The dashed square in the figure shows the unit mesh of the supernet and emphasizes its 45° orientation relative to the substrate. The unit mesh dimensions in the dashed square are the same as those of the substrate, so the distances seem to be wrong even though the symmetry is correct. The solid square in the figure shows that there is a second way of looking at the identical supernet: as a 2×2 enlargement of the substrate net but possessing a centered rather than primitive packing. This is a preferable way of describing the supernet since it accounts for both the symmetry and the separation of the unit mesh. These two descriptions do not involve different supernets, but merely different ways of looking at the same net.

Since the version represented by the solid square in Figure 9.17d best characterizes the adsorption, it is described as a $c(2 \times 2)$ net, the c reminding us that this is a *centered* structure. The full description of the LEED pattern in Figure 9.16c in the Wood notation is therefore written W(100)$-c(2 \times 2)-$hydrogen.

The solid squares in Figures 9.17c and 9.17d represent the unit mesh of the adsorbed layers of oxygen and hydrogen, respectively, on the W(100) surface. Both are seen to have the square symmetry of the underlying tungsten surface, but with oxygen showing a primitive net and hydrogen a centered structure. An additional detail about the surface with the adsorbed hydrogen is that the $c(2 \times 2)$ structure shown in Figures 9.16c and 9.17d persists until the surface coverage is half a monolayer. This seems too neat for an accidental circumstance. On the other hand, if we assume the H_2 is dissociated and the adsorbed species are hydrogen atoms, the same amount of adsorbed material is enough to form a monolayer. The chemisorbed hydrogen atoms form a surface layer characterized by a square, centered unit mesh. In Wood notation the surface layer is W(100)$-c(2 \times 2)-$H, indicating the fact that atomic hydrogen rather than H_2 is the adsorbate. Oxygen is also adsorbed in the monatomic state on the W(100) surface.

In the adsorption studies we have discussed, the expansion of the unit mesh is the same in both directions, but this need not be the case. Examples in which the expansion along different axes of the mesh varies are $p(4 \times 2)-$O for the adsorption of O_2 on Mo(111), $p(3 \times 15)-$O for O_2 on Pt(111), $c(4 \times 2)-$S for H_2S on Au(100), and $c(9 \times 5)-$CO for CO on W(110). Somorjai (1981, 1994) has assembled extensive tables of this sort of information. Note that many but not all adsorbates are dissociated on the metal surfaces. Finally, it is not necessary for the supernet and the substrate to show the same angles between sides of their respective meshes. The Wood notation does not apply in these cases, but an alternative notation exists

that we shall not pursue. Example 9.8 deals with a surface at which the adsorbate and the substrate display different nets.

* * *

EXAMPLE 9.8 *An Example of the Lattice Structures of the Adsorbate and Adsorbed Layers* The unit mesh of the Pt(111) surface is a parallelogram with $\delta = \delta' = 0.277$ nm and having angles of 60° and 120°. This surface adsorbs *n*-butane with a unit mesh that is a parallelogram having $\delta = 0.480$ nm and $\delta' = 0.733$ nm and angles of 71° and 109°. Show that these two nets come into periodic register if the short side of the supermesh coincides with the diagonal of the substrate mesh.

Solution: Use the law of cosines to determine the length of the two diagonals of the substrate unit mesh:

$$L^2 = \delta^2 + \delta'^2 - 2\delta\delta' \cos 60° = 2(0.277)^2 - 2(0.277)^2(0.500)$$

$$L = 0.277 \text{ nm}$$

$$L' = [2(0.277)^2 - 2(0.277)^2(-0.500)]^{1/2} = 0.480 \text{ nm}$$

Since the short side of the supernet mesh has $\delta = 0.480$ nm, this mesh registers with the substrate along the diagonal of the latter. Since the long diagonal of the substrate and the short side of the super mesh register, this common line makes a convenient reference line to test the register in the perpendicular direction. Half the length of the short diagonal, 0.277/2 = 0.139 nm, measures the perpendicular distance from this base to the corner atom. Also, the long side of the supernet mesh makes an angle of 71° with the base. Therefore the long side of the supernet mesh projects a length (0.733 sin 71) = 0.693 nm onto the short diagonal of the substrate. The ratio 0.693/0.139 = 5.00 shows that the edge of the supermesh and the diagonal of the substrate net will coincide again at the fifth row of atoms perpendicular to the diagonal that serves as the base for this calculation. Even though the two parallelograms have different side lengths, angles, and orientations, they do become coincident at periodic intervals. ■

* * *

Next, we see some additional examples of LEED applied to the study of metallic catalysts.

9.9 METAL SURFACES AND HETEROGENEOUS CATALYSIS

The topic of heterogeneous catalysis is the point at which the study of surfaces and the study of catalysts meet, as we have illustrated in Vignette IX (see also Somorjai 1994). It has been recognized for a long time that heterogeneous catalysts owe their activity to the properties of their surfaces. Until fairly recently, however, catalyst particle size and hence specific area were the major surface parameters that could be varied and monitored in the study of the surface effects. Much of the interest in characterizing high-surface-area solids through physical adsorption that we discussed at the beginning of this chapter originates from this application. With the advent of LEED and various types of surface spectroscopy, the role of surface science in the study of these catalysts is greatly expanded.

A wide variety of solid surfaces is used as catalysts in an even wider assortment of industrial processes (see, for example, Richardson 1989 and Somorjai 1994); we limit our discussion to metal catalysts. While these represent only a fraction of all catalytic systems, they do include a number of industrially important examples. Table 9.6 lists some metals used as commercial catalysts and indicates briefly the types of reactions for which they are employed. In this section we emphasize the effect on catalytic activity of the chemical and crystallographic properties of metal surfaces.

It should be emphasized here that there is a vast difference between the microenvironment of the catalyst surface as examined by the type of analytical techniques mentioned in Section 9.1 and the overall surface that influences commercial processes. Until the modern techniques became available, however, catalyst preparation was mostly a matter of trial and error; we have now entered an era in which science has a chance to catch up with technology. It seems fairly safe to predict that a greatly increased understanding of heterogeneous catalysis will emerge as modern surface chemistry matures.

TABLE 9.6 Examples of Some Metal Catalysts and the Reactions They Catalyze

Metal	Reaction
Cobalt	Fischer-Tropsch synthesis of hydrocarbons from CO and H_2
Iron	Haber synthesis of ammonia from N_2 and H_2
Platinum	Hydrogenation of vegetable oils
Platinum-palladium	Oxidation of hydrocarbons and CO and reduction of NO_x in automobile exhaust
Platinum-rhenium and platinum-tin	Reforming alkanes to aromatic hydrocarbon
Platinum-rhodium	Oxidation of NH_3 to HNO_3

9.9a Catalysis at Surfaces

In our discussion of micellar catalysts in Chapter 8, we noted that effective catalysts have two features: the ability to accelerate the rate of a reaction and the ability to do so selectively. Chemistry students are familiar with the general notion that catalysts modify the mechanistic path of a reaction in such a way as to lower the activation energy and make the conversion of reactants to products more probable. One of the easiest places to see this is in reactions of diatomic gas molecules. In the gas phase the mechanism for the reaction of hydrogen and oxygen to form water involves the following steps, among others:

$$H_2 + O_2 \rightarrow HO_2^{\bullet} + H^{\bullet} \qquad \text{(A)}$$

$$H_2 + HO_2^{\bullet} \rightarrow HO^{\bullet} + H_2O \qquad \text{(B)}$$

$$H_2 + HO^{\bullet} \rightarrow H^{\bullet} + H_2O \qquad \text{(C)}$$

How much simpler things would be if these were monatomic gases and there was no need for all the juggling between intermediate species to dispose of unused molecular fragments! We saw in our discussion of LEED that molecules of this sort are chemisorbed at metal surfaces in the dissociated state. The combination of chemisorbed hydrogen and oxygen atoms to form water clearly follows a different mechanism than in the gas phase. The fact that the reaction occurs rapidly in the presence of platinum and not at all when the reactants are mixed without the metal shows that the activation energy has been lowered tremendously by this modification.

A similar reaction of great commercial importance is the synthesis of ammonia from the diatomic elements. The catalysts that are used commercially in this reaction are mixtures of iron and iron oxide with the oxides of potassium and aluminum. Indicating the adsorbed species by the subscript (a), the mechanism for the surface-catalyzed synthesis of ammonia is thought to involve the steps

$$N_{2(g)} \Leftrightarrow N_{2(a)} \qquad \text{(D)}$$

$$H_{2(g)} \Leftrightarrow H_{2(a)} \qquad \text{(E)}$$

$$N_{2(a)} \Leftrightarrow 2N_{(a)} \qquad \text{(F)}$$

$$H_{2(a)} \Leftrightarrow 2H_{(a)}$$

(G)

$$N_{(a)} + H_{(a)} \Leftrightarrow NH_{(a)} \Leftrightarrow NH_{2(a)} \Leftrightarrow NH_{3(a)}$$

(H)

$$NH_{3(a)} \rightarrow NH_{3(g)}$$

(I)

The reaction between chemisorbed atoms dispenses with the problem of disrupting the triply bonded N_2 molecule in the gas phase, but Reaction (F) may also be more complicated than what is shown here.

The above are equilibrium reactions, and their successful exploitation requires that they be carried out under conditions in which the equilibrium favors the product. Specifically, this requires that the adsorbed species in Reactions (D)-(I) not be held so tightly on the catalyst surfaces as to inhibit the reaction. On the other hand, strong interaction between adsorbate and catalyst is important to break the bonds in the reactant species. Optimization involves finding a compromise between scission and residence time on the surface. Although we are especially interested in metal surfaces, those constituents known as promoters in catalyst mixtures are also important. It is known, for example, that the potassium in the catalyst used for the ammonia synthesis shifts Equilibrium (F) to the right and also increases the rate of Reaction (D) by lowering its activation energy from 12.5 kJ mole^{-1} to about zero.

In addition to affecting reaction energetics, other atoms at metal surfaces also influence the selectivity of the catalyst. Platinum catalysts are excellent examples of this since the platinum surface is capable of catalyzing a number of hydrocarbon reactions, including hydrogenation, dehydrogenation, isomerization, ring opening, and dehydrocyclization. Which of these processes is most favored by a particular catalyst is sensitive to the presence of foreign atoms at the platinum surface. Auger electron spectroscopy and other surface spectroscopies are clearly important techniques for the study of these effects. It is known, for example, that a partial monolayer of oxygen on platinum enhances reactions involving scission of the carbon-carbon bond and inhibits dehydrogenation reactions. Conversely, gold on platinum blocks C—C scission without affecting dehydrogenation or isomerization. Sulfur, which often poisons catalysts, increases the selectivity of Pt-Re catalysts.

At low pressures hydrocarbon reactions on Pt are controlled by the clean metal surface. At higher pressures a carbonaceous layer controls the selectivity of the catalyst. Since the second condition describes how catalysts are generally used, the carbonaceous layer becomes an intrinsic part of the catalyst. At low temperatures hydrocarbons are physically adsorbed on Pt surfaces; at high temperatures a graphitic coating poisons the catalyst. Under intermediate conditions of temperature, various hydrocarbon fragments are present on the surface that determine its reactivity. It is noteworthy that the hydrogen atoms in these surface fragments are readily available for reaction, as evidenced by isotope exchange studies with deuterium. Carbonaceous matter, oxygen, and other atoms are thought to affect catalyst activity in several ways. For example, the electronic structure of the metal may be altered, surface restructuring may occur, and specific surface sites may be blocked.

9.9b Crystallographic Structure of Catalysts and Catalytic Behavior

The crystallographic character of surfaces has been shown to be of great importance in determining their catalytic behavior. Particularly interesting are those surfaces designed with well-characterized "roughness." To see how this is accomplished, consider slicing through a face-centered cubic platinum crystal at a small angle relative to, say, the 111 plane. This would result in a surface of high Miller index, low atomic density, and high surface free energy. Such a slice can be stabilized, however, by forming terraces of 111 planes separated by steps of, say, 100 planes. If the angle of the cut is small relative to the 111 plane, the steps are only one atom

high; the width of the terrace depends on the angle of the cut, with narrower terraces resulting from steeper cuts. Application of crystallographic concepts shows that a slice through a face-centered cubic crystal along a plane with Miller indices 7,5,5 is equivalent to a series of 111 terraces five atoms wide separated by 100 steps one atom high.

The slice through a bulk crystal can differ from both the 111 plane and the 100 plane by small angles. This produces a kink in the face of the step. By an extension of the analysis that leads to step characterization, these kinks can also be characterized. For example, a plane with Miller indices 10,8,7 has 111 terraces seven atoms wide, 110 steps one atom high, and kinks of 100 orientation every two atoms. Because of the greater thermodynamic stability of the planes of low Miller index, these surfaces of ordered roughness are stable and can be prepared and studied. Since it is sensitive to periodicity over a domain about 20 nm in diameter, LEED "sees" the pattern associated with terraces of various widths and may be used to characterize these surfaces. Satisfactory LEED patterns do not require absolute uniformity of terrace width but may be obtained with experimental surfaces that display a distribution of widths.

The idea that catalyst surfaces possess a distribution of sites of different energies has been around since the 1920s, but it has not been possible until fairly recently to show that adsorption sites on terraces, steps, and kinks differ in energy. For example, hydrogen shows stronger bonding to steps and kinks on platinum than on the 111 terraces. In addition, the activation energy for H_2 dissociation is about zero on the step face and about 8.4 kJ mole^{-1} on the terrace plane. In addition, carbon monoxide is adsorbed with dissociation on the kinks of Pt, but in the molecular form on the steps and terraces.

The behavior of catalyst surfaces with respect to selectivity and poisoning by foreign atoms is now known to depend on the location of these atoms with respect to the terrace, step, and kink structure of the surface. By using the well-characterized faces of single crystals as catalytic surfaces, the effects of these variables have been examined for some reactions. Figure 9.18 is an example of the results of such a study. The hydrogenation of cyclohexene to cyclohexane is the reaction involved, and the figure shows how the ordered roughness and oxygen content of the platinum surface affect this reaction at 150°C. In Figure 9.18 the ordinate gives the number of molecules converted to product per second per Pt atom—the *turnover number* of the catalyst—and the abscissa gives the oxygen content of the surface as measured by Auger spectroscopy.

The three different lines in the figure correspond to platinum surfaces of different Miller

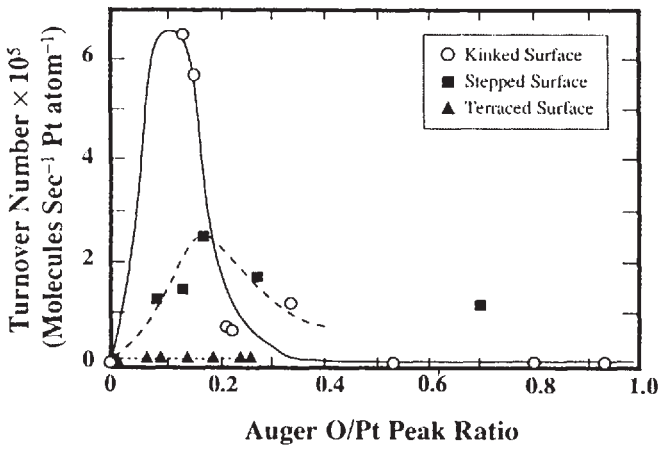


FIG. 9.18 Number of molecules of cyclohexene converted to cyclohexane per second per Pt atom versus oxygen content of platinum surface. The oxygen content is expressed as the ratio of O to Pt Auger peak heights. The data shown are for kinked (circles), stepped (squares), and terraced (triangles) platinum surfaces. (Redrawn with permission from Somorjai 1981.)

indices, corresponding to different degrees of ordered roughness. The data points marked by circles were measured on an 11,9,8 surface that consisted of terraces, steps, and kinks; the squares on a 7,5,5 surface that had terraces and steps; and the triangles on 111, which corresponds to terrace sites only.

The experiments described in Figure 9.18 were conducted under low-pressure conditions so the following generalizations apply:

1. In the absence of oxygen all three surfaces show undetectable catalytic activity.
2. The Pt(111) surface continues to show undetectable activity as the oxygen content increases, although the activity passes through a maximum for both the stepped and kinked surfaces with increasing oxygen content.
3. Compared at their respective maxima, the kinked surface is more active than the stepped surface.
4. The oxygen content at the maximum is slightly higher for the stepped surface than for the kinked surface.

These observations suggest that maximum catalytic activity for this reaction occurs at a kinked surface with about one-third of the surface covered by oxygen. For a competitive reaction, the dehydrogenation of cyclohexene to benzene, the kinked surface is also the most active. In this case, however, the optimum oxygen content is at a coverage of about one-half compared to one-third for the hydrogenation reaction on which Figure 9.18 is based. Differences such as these show that the selectivity of various catalysts is traceable to the chemical and physical nature of the surface. Even though the reasons for this behavior are not yet understood, these types of data hold out the promise that catalysts may someday be custom designed for maximum effectiveness. Polymer and pharmaceutical chemistry offer many examples of molecules that are designed with a specific function in mind. The methods and results of this chapter, augmented by some of the methods described in Chapter 1, suggest that this may become possible for catalysts in the not too distant future. Many of the modern developments in the area of heterogeneous catalysis may be found in Somorjai (1994).

REVIEW QUESTIONS

1. What is the difference between *chemisorption* and *physisorption*?
2. What is an *adsorption isotherm*?
3. Describe an experimental procedure for determining adsorption isotherms.
4. Describe the various types of isotherms observed in experiments. Describe the significance of the different forms of the isotherm.
5. How is an equation of state related to an adsorption isotherm? What is the basic thermodynamic principle that governs the equilibrium between the surface phase and bulk phase?
6. Discuss the implication of the *van der Waals equation of state* for a two-dimensional phase to the corresponding adsorption isotherm and to the analysis of adsorption data.
7. How is statistical thermodynamics used for deriving adsorption isotherms? What are the similarities and differences between this procedure and the one based on phenomenological thermodynamics? How is the kinetic theory of gases used for deriving adsorption isotherms?
8. What are the assumptions implicit in the *Langmuir adsorption isotherm*?
9. What is meant by the *Henry law limit* in the case of an adsorption isotherm?
10. What is the *Harkins-Jura isotherm*? What assumption does it make concerning the nature of the surface phase? Is the assumption consistent with the experimental observations?
11. What is the *BET isotherm*? How does it differ from the Langmuir isotherm?
12. How is the BET isotherm used to determine the surface area of powders?
13. What is meant by *calorimetric* analysis of adsorption?
14. What is *isosteric heat of adsorption*? How is it related to the pressure-versus-temperature relationship at constant surface coverage?
15. What is *adsorption hysteresis*?
16. What is *capillary condensation*? Describe some models of capillary condensation?
17. Describe the basic principles of operation of *low-energy electron diffraction (LEED)*.
18. Why are very low pressures needed for LEED experiments?

19. Describe how LEED patterns are interpreted.
20. What is meant by *heterogeneous catalysis*?
21. What is meant by *selectivity* of a catalyst?
22. Describe how adsorption assists in catalysis.
23. List a few examples of metal catalysts and what reactions they catalyze.

REFERENCES

General References (with Annotations)

Adamson, A. W., *Physical Chemistry of Surfaces*, 5th ed., Wiley, New York, 1990. (Graduate level. A good source of extended discussions on the topics considered in this chapter. This volume includes more advanced material as well.)

Atkins, P., *Physical Chemistry*, 5th ed., W. H. Freeman, New York, 1994. (Undergraduate level. An excellent, introductory-level book on physical chemistry for basic concepts not described in detail in this chapter.)

Ball, P., *Designing the Molecular World: Chemistry at the Frontier*, Princeton University Press, Princeton, NJ, 1994. (Undergraduate level. A highly readable, *Scientific-American*-style, popular introduction to the many facets of molecular design using chemistry. Chapter 2 discusses chemical reactions and heterogeneous catalysis.)

Richardson, J. T., *Principles of Catalyst Development*, Plenum Press, New York, 1989. (Undergraduate level. An introduction to catalysis and catalyst development from an engineering perspective. The emphasis throughout is on the practical aspect of the subject, rather than on theory.)

Somorjai, G. A., *Introduction to Surface Chemistry and Catalysis*, Wiley, New York, 1994. (Undergraduate level. This in-depth treatment of surface chemistry and catalysis brings the experience and perspectives of a pioneer in the field to the general audience. The book is meant to be an introductory-level description of modern developments in the area for students at the junior level. However, it is also an excellent source of the current literature and contains numerous, extensive tables of data on kinetic parameters, surface structure of catalysts, and so on. Chapter 3, "Thermodynamics of Surfaces," and Chapter 7, "Catalysis by Surfaces," cover information relevant to the present chapter. Chapter 8 discusses applications in tribology and lubrication (not discussed in this chapter).)

Other References

Broeckhoff, J. C. P., and van Dongen, R. H., Mobility and Adsorption on Homogeneous Surfaces. In *Physical and Chemical Aspects of Adsorbents and Catalysts* (B. G. Linsen, Ed.), Academic Press, New York, 1970.

Hubbard, A. T. (Ed.), *The Handbook of Surface Imaging and Visualization*, CRC Press, Boca Raton, FL, 1995.

Kantro, D. L., Brunauer, S., and Copeland, L. E., BET Surface Areas: Methods and Interpretations. In *The Solid-Gas Interface*, Vol. 1 (E. A. Flood, Ed.), Marcel Dekker, New York, 1967.

Lowell, S., *Powder Surface Area and Porosity*, 3d ed., Chapman and Hall, New York, 1991.

Prutton, M., *Surface Physics*, 2d ed., Oxford University Press, Oxford, England, 1983.

Ross, S., Monolayer Adsorption on Crystalline Surfaces. In *Progress in Surface and Membrane Science*, Vol. 4 (J. F. Danelli, M. D. Rosenberg, and D. A. Cadenhead, Eds.), Academic Press, New York, 1971.

Somorjai, G. A., *Chemistry in Two Dimensions: Surfaces*, Cornell University Press, Ithaca, NY, 1981.

Van Hove, M. A., and Tong, S. Y., *Surface Crystallography by LEED*, Springer-Verlag, Berlin, 1979.

PROBLEMS

1. Consider the following linear equation of state appropriate for a relatively incompressible surface state (e.g., above $\theta \approx 0.70$ in Fig. 9.6b)

$$\pi = -C_1\sigma + C_2$$

where C_1 and C_2 are known constants. Derive the corresponding adsorption isotherm, known as the *Harkins-Jura isotherm* (see Table 9.1),

$$\ln p = -\frac{C_1}{2N_A RT} A_{sp}^2 \left(\frac{n}{w}\right)^{-2} + C_3$$

(where C_3 is a constant of integration) for this linear equation of state.

2. An isotherm that is not too difficult to derive by the methods of statistical mechanics assumes an adsorbed layer that obeys the two-dimensional analog of the van der Waals equation. The result of such a derivation is the equation (see Table 9.1)

$$\ln p = \ln\left(\frac{\theta}{1-\theta}\right) + \frac{\theta}{1-\theta} - \frac{2a}{bk_B T}\theta + \ln K$$

where a, b, and K are constants, the first two being the two-dimensional van der Waals constants. Like its three-dimensional counterpart, this equation predicts that at the critical point for the two-dimensional matter

$$\left(\frac{\partial \ln p}{\partial \theta}\right)_{T_c} = 0 \quad \text{and} \quad \left(\frac{\partial^2 \ln p}{\partial \theta^2}\right)_{T_c} = 0$$

From the second of these derivatives, evaluate θ_c as predicted by this model. Use this value of θ_c and the first of these derivatives to evaluate the relationship between T_c and the two-dimensional a and b constants. How does this result compare with the three-dimensional case? The van der Waals constant b is four times the volume of a hard-*sphere* molecule. What is the relationship between the two dimensional b value and the area of a hard-*disk* molecule?

3. Use the linear form of the Langmuir equation to evaluate $(n/w)_{sat}$ and K for the adsorption of pentane on carbon black from the higher pressure values in the following data. Use the ratio p/p_0 rather than p only to normalize pressures relative to the equilibrium vapor pressure of pentane at different temperatures.* (All pressures are in torr.)

T(°C)	−63.7		0		5.24		20.5	
p_0	3.48		187.5		235.6		445.1	
	p	g C_5 g^{-1}	p	g C_5 g^{-1}	p	g C_5 g^{-1}	p	g C_5 g^{-1}
	0.024	0.2827	0.0533	0.1062	3.8	0.2299	0.284	0.1061
	0.028	0.2904	0.1405	0.1322	7.2	0.2535	0.675	0.1317
	0.067	0.3162	0.203	0.1427	19.6	0.2939	0.95	0.1421
	0.103	0.3276	0.89	0.1908	36.7	0.3147	3.62	0.1884
	0.233	0.3459	2.5	0.2305	53.6	0.3231	9.7	0.2262
	0.288	0.3507	5.0	0.2506	88.5	0.3313	15.9	0.2469
	0.671	0.3581	14.6	0.2964	199.0	0.3418	62.9	0.3002
	1.460	0.3647	29.2	0.3184			90.4	0.3107
			45.4	0.3272			155.2	0.3204
			80.7	0.3353			328.7	0.3321
			161.0	0.3433			428.7	0.3372

*Polanyi, M., and Goldmann, F., *Z. Phys. Chem.*, **132**, 321 (1928).

4. Examine the temperature variation of K values from the preceding problem by means of the procedure given in Example 9.3. (To decrease computational effort, various members of the class may be assigned different temperatures to analyze in Problem 3. The K values may then be pooled for this problem.)

5. The accompanying data* give the volume of N_2 at STP adsorbed on colloidal silica at the temperature of liquid nitrogen as a function of the ratio p/p_0. Plot these results according to the linear form of the BET equation. Evaluate c, V_m, and A_{sp} from these results, using 16.2 Å^2 as the value for σ^0.

V at STP (cm^3 g^{-1})	p/p_0	V at STP (cm^3 g^{-1})	p/p_0
44	0.008	117	0.558
52	0.025	122	0.592
57	0.034	130	0.633
61	0.067	148	0.692
64	0.075	165	0.733
65	0.083	194	0.775
70	0.142	204	0.792
77	0.183	248	0.825
78	0.208	296	0.850
85	0.275		
90	0.333		
96	0.375		
100	0.425		
109	0.505		

6. The following data give the volume at STP of nitrogen and argon adsorbed on the same nonporous silica at −196°C:†

	V at STP (cm^3 g^{-1})	
p/p_0	Nitrogen	Argon
0.05	34	23
0.10	38	29
0.15	43	32
0.20	46	38
0.25	48	41
0.30	51	43
0.35	54	45
0.40	58	50
0.45	58	54
0.50	61	55
0.60	68	62
0.70	77	69
0.80	89	79
0.90	118	93

Using 16.2 Å^2 as the N_2 cross section, calculate A_{sp} for the silica by the BET method. What value of σ^0 is required to give the same BET area for the argon data?

*Everett, D. H., Parfitt, G. D., Sing, K. S. W., and Wilson, R., *J. Appl. Chem. Biotechnol.*, **24**, 199 (1974).

†Payne, D. A., Sing, K. S. W., and Turk, D. H., *J. Colloid Interface Sci.*, **43**, 287 (1973).

7. Use the accompanying data* to criticize or defend the following proposition: Self-consistent A_{sp} values for nonporous solids are obtained at 77 and 90K by using values of σ^0 for N_2 equaling 16.2 and 17.0 Å^2, respectively. These are consistent with the density of liquid N_2 at these two temperatures. For the same self-consistency in A_{sp} using O_2 as the adsorbate, σ^0 values of 14.3 and 15.4 Å^2 must be used at these two temperatures. This suggests that O_2 is somewhat more loosely packed on the surface than in the liquid state at 90K compared to 77K.

	N_2		O_2	
T (K)	ρ (g cm^{-3})	V_{sp} (cm^3 g^{-1})	ρ (g cm^{-3})	V_{sp} (cm^3 g^{-1})
77	0.808	1.238	1.204	0.831
90	0.751	1.332	1.140	0.877

8. Use the data from Problem 3 to estimate the isosteric heat of adsorption of pentane on carbon black at $\theta \simeq 0.3$, 0.6, and 0.9. Under what conditions would greater variation of q_{st} be expected? What prevents these conditions from being examined in this problem? How does q_{st} compare with the energy of adsorption for this system as determined in Problem 4? How does the difference between q_{st} and E_{ads} compare with ΔH_v for pentane?

9. Colloidal carbon was formed by the slow pyroloysis of polyvinylidene chloride. Surface oxidation occurs during subsequent storage in air. By heating at elevated temperatures, the following gases are evolved† (heats of immersion were also measured after the different thermal pretreatments):

	Cumulative amount desorbed (mmol g^{-1})			Heat of immersion
T (°C)	CO	CO_2	H_2	(cal g^{-1})
100	0.03	0.02	–	11.0
200	0.05	0.14	–	9.8
300	0.15	0.33	–	9.2
400	0.35	0.44	–	8.7
500	0.63	0.51	–	8.2
600	0.95	0.55	–	7.7
700	1.300	0.55	–	7.0
800	1.580	0.55	–	6.2
900	1.800	0.55	0.03	6.0
1000	1.823	0.556	0.269	5.9

Using 7.9 and 9.1 Å^2 as the areas of oxide desorbing as CO and CO_2, respectively, estimate the area (in m^2 g^{-1}) occupied by each of these oxide types. If the specific area of the carbon is 1100 m^2 g^{-1}, what percentage of the surface is covered with oxide? Discuss the variation of the heat of immersion with the removal of the surface oxides.

*Hanna, K. M., Odler, I., Brunauer, S., Hagymassy, T., and Bodor, E. E., *J. Colloid Interface Sci.*, **45**, 27 (1973); Also see Osipow, L. I., in Ross, S., Monolayer Adsorption on Crystalline Surfaces, in *Progress in Surface and Membrane Science,* (Danell, T. E., Rosenberg, M.D. and Cadenhead, D. A., Eds.), Academic Press, New York, 1971.

†*Barton, S. S., Evans, M. J. B., and Harrison, B. H., *J. Colloid Interface Sci.*, **45**, 542 (1973).

10. Samples of rutile TiO_2 were outgassed at elevated temperatures; in some instances this was followed by subsequent exposure to O_2 at 150°C. The following results were obtained by the BET analysis of N_2 adsorption after various preliminary treatments:*

Pretreatment	A_{sp} ($m^2\ g^{-1}$)	c
150°C + O_2	11.0	200
210°C + O_2	11.7	100
200°C, no O_2	11.9	520
240°C, no O_2	11.3	370
470°C + O_2	11.7	450
560°C + O_2	11.9	1210

Criticize or defend the following proposition: After the initial removal of physically adsorbed water, partial dehydroxylation of the surface occurs with increasing temperature. In the absence of O_2, Ti^{3+} cations are present that interact with N_2 in pretty much the same way as the isolated hydroxyl groups on the surface. At still higher temperatures surface diffusion permits the hydroxyl groups to migrate into patches, which show stronger interactions with N_2. The specific surface is not essentially changed throughout treatment.

11. Low-energy electron diffraction diagrams reveal that O_2 adsorbs on a Pd(110) surface through a series of ordered structures that interchange reversibly with increasing surface coverage obtained by varying the pressure and temperature:

$$\text{Clean Pd(110)} \rightleftharpoons (1 \times 3) \rightleftharpoons (1 \times 2) \rightleftharpoons c(2 \times 4) \rightleftharpoons c(2 \times 6)$$

The equilibrium pressure-temperature coordinates of the transitions between one LEED pattern and another were measured by Ertl and Rau† and were found to obey the two-dimensional Clausius-Clapeyron equation. When $\log_{10} p$ is plotted versus T^{-1}, straight lines of slope -1.75×10^4, -1.68×10^4, -1.35×10^4, and -1.05×10^4 K, respectively, are obtained for the four transitions above. Use these data to evaluate ΔH for each of the phase transitions of the adsorbed oxygen layer. Criticize or defend the following proposition: Since the only difference between these structures is the extent of oxygen coverage, the values of ΔH may be viewed as isosteric heats of adsorption for the surface at their respective degrees of coverage.

12. The object of this problem is to demonstrate that only certain shapes are possible for the unit mesh of a surface net. The following steps outline the proof: (a) Draw two parallel lines; one is the x axis, the other is at y = const. (b) Draw a transverse cutting across the two lines and making an angle θ with the x axis. (c) Cutting the x axis at another location, draw another transverse, this time making an angle $-\theta$ with the axis. The two transverse lines should be mirror reflections of each other. (d) Imagine a surface atom at each of the four intersections of these lines. If the unit mesh is to be a regular polygon, then three of the lines must be equal to each other and equal to the interatomic spacing d. Make any necessary adjustments in the drawing so that the three shorter lengths—say, the base and the two transverses—are equal. (e) The longer side of the drawing has a length $d + 2d \cos \theta$ since each of the transverse lines project a length $d \cos \theta$ in the x direction. For this array of atoms to come into periodic register, the ratio $(d + 2d \cos \theta)/d$ must be an integer. What values of n and θ satisfy this condition, and what shapes do the corresponding meshes possess?

13. Firment and Somorjai‡ showed that the C_4–C_8 n-alkanes adsorb on the Pt(111) surface in ordered monolayers for which the unit mesh is a parallelogram with the following dimensions:

---

*Parfitt, G. D., Urwin, D., and Wiseman, T. J., *J. Colloid Interface Sci.*, **36**, 217 (1971).

†Ertl, G., and Rau, P., *Surface Sci.*, **15**, 443 (1969).

‡Firment, L. E., and Somorjai, G. A., *J. Chem. Phys.*, **66**, 2901 (1977).

Hydrocarbon	δ (Å)	δ' (Å)	Degree
n-Butane	4.80	7.33	71.0
n-Pentane	4.80	16.79	74.0
n-Hexane	4.80	9.99	75.6
n-Heptane	4.80	22.29	78.0
n-Octane	4.80	12.69	79.5
Pt(111)	2.77	2.77	60.0

Look up the formula for the area of a parallelogram and calculate the area per mole, assuming one molecule per unit mesh. Compare this area with the Pt(111) unit mesh for which the dimensions are also given. Do the data make any more sense if it is assumed that some of these alkanes form surface structures with two molecules per unit mesh? Prepare a plot of the area per molecule versus the number of carbon atoms in the chain. Criticize or defend the following proposition: The amount of free area per unit mesh in these packings is equivalent to one Pt mesh.

14. The authors of the research cited in Problem 13 point out that the width of the alkane molecule in the bulk crystals is 4.79, 4.78, and 4.70 Å for *n*-octane, *n*-heptane, and *n*-hexane, respectively. This corresponds almost exactly to one of the dimensions of the unit mesh of these molecules on the Pt(111) surface. The data tabulated in Problem 13 show that the longer dimensions of these cells tend to increase with the chain length. To establish whether this is a quantitative correlation, plot δ' versus the number of methylenes N in the alkane and determine the slope and intercept of the resulting graph. This gives an equation of the type $\delta' = mN_{CH_2} + b$, where m is the length contribution per methylene and $b/2$ is the length contribution per terminal methyl group. Compare these lengths with the dimensions of a fully extended hydrocarbon chain as presented in Example 8.2. What does this reveal about the configuration of these molecules on the Pt(111) surface?