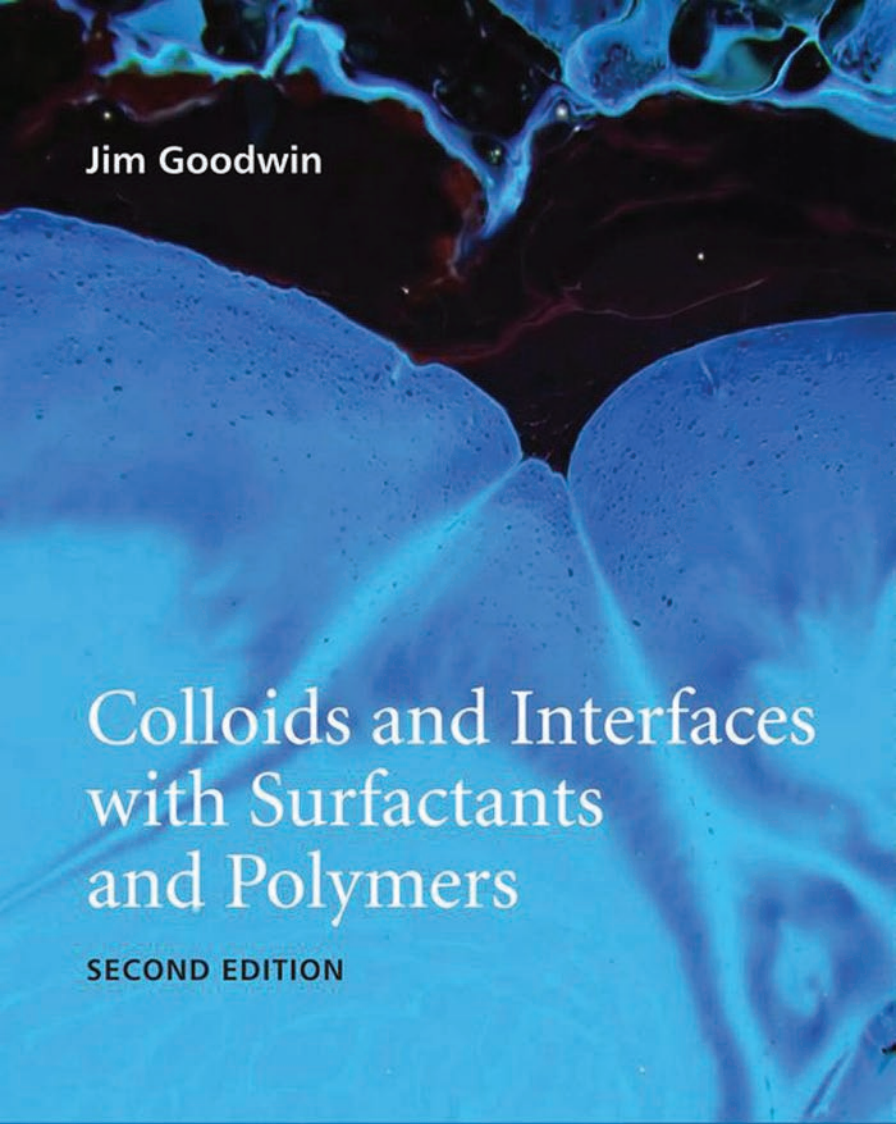


Jim Goodwin

Colloids and Interfaces with Surfactants and Polymers

SECOND EDITION

WILEY

Colloids and Interfaces with Surfactants and Polymers

Second Edition

Colloids and Interfaces with Surfactants and Polymers

Second Edition

Jim Goodwin
Life Technologies Corporation

WILEY
A John Wiley and Sons, Ltd, Publication

This edition first published 2009

Registered office
John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom

Library of Congress Cataloging-in-Publication Data
Goodwin, James W. (James William)
Colloids and interfaces with surfactants and polymers / Jim Goodwin. – 2nd ed.
p. cm.
Includes bibliographical references and index.
ISBN 978-0-470-51880-9 (cloth)
1. Colloids. 2. Surface active agents. 3. Surface chemistry. 4. Polymers. I. Title.
QD549.G57 2009
541′.345—dc22
2009015523

A catalogue record for this book is available from the British Library.

ISBN 978-0-470-51880-9 (HB)
978-0-470-51881-6(PB)

Typeset in 10.5/13pt Sabon by Aptara Inc., New Delhi, India.
Printed in Great Britain by TJ International Ltd, Padstow, Cornwall, UK

Contents

Preface to the Second Edition

The ubiquitous nature of colloidal systems means that we tend to accept their behaviour without recognizing the unique nature of the colloid state in that it is the fourth state of matter. The heterophase character coupled with size of the dispersed phase or phases results in an extraordinary range of systems that we rely on every day, whether naturally occurring or manufactured, with the size of the component phases lying between the atomic and the macroscopic. Few academic courses include more than a cursory study of the subject, with the result that many scientists working with colloids are unaware that there are many others working on similar problems on apparently quite unrelated materials or applications. So Food Scientists may have many interests in common with Material Scientists, Physicists working with soft matter or complex fluids, Paint Technologists or synthetic Inorganic Chemists.

Colloid Science as a subject reached a significant level of maturity during the second half of the twentieth century, with a good understanding of the forces acting between the components in many systems made up of, not only dilute phases, but also of condensed phases. The activity in recent years has redoubled as applications for nanoparticles, self-assembled systems, complex fluids, and soft solids have permeated the conventional disciplines. The need for a general basic text is greater than ever and this volume is intended to fulfil that need. This second edition is an expansion of the first edition and includes material on biopolymers, nanoparticles, fluorescence and confocal microscopy, in addition to the analysis of sedimentation behaviour. The mathematics used is of a straightforward nature so that many of the quantitative descriptions of the properties can be derived. Where more complex mathematical treatments are required, references to the original texts are included.

The volume should be accessible to undergraduates and postgraduates working in both the physical and life sciences.

Jim Goodwin
Oregon, 2009

Preface to the First Edition

We take for granted the behaviour of colloidal systems from a very early age and as we are formally taught the various scientific disciplines, the nature of colloids is rarely mentioned. This is surprising as it impacts on so many products that we use everyday. The processing and delivery of the correct properties are often dependent on the material being in the colloidal state and yet few courses cover the subject in depth, if at all. As a result, many scientific workers have to acquire their knowledge piecemeal while working on other problems. An introductory text is what is most used in this situation. Specialist texts are often left on the shelf as we search for attempts to solve a particular problem. This present text covers a little more ground than some of the older introductory volumes that are still available but I have attempted to help the reader in the more complicated regions by providing a strategy for any calculations. Some derivations are outside the scope of an introduction, but, where they are straightforward, they are included in order to help readers gain as much insight as possible. None of the mathematics is at a high level.

The references are mainly to specialist volumes in the area. This is not to discourage the reading of the original work but it is because the general volumes are more readily available in many libraries and often help to put the work in context. There will inevitably be some areas that are neglected, as with any introductory volume. This is in part due to what areas are perceived to be currently most generally needed and the composition of this volume has been governed by the questions which are currently asked when discussing problems in industry.

I would like to acknowledge with gratitude the help and encouragement of all of my previous colleagues in the Department of Physical Chemistry at the University of Bristol, UK, especially Professors Ron

Ottewill and Brian Vincent, the past and current holders of the Leverhulme Chair in that Department and also to Dr Paul Reynolds, the manager of the Bristol Colloid Centre.

Jim Goodwin
Portland, Oregon, 2003

1

The Nature of Colloids

1.1 INTRODUCTION

Colloids are all about size. They consist of at least two phases and the dimensions of the dispersed phase have traditionally been considered to be in the sub-microscopic region but greater than the atomic size range, that is, within the range 1 nm to 1 μm. The term 'colloid' was coined for 'glue-like' materials which appeared to consist of only one phase when viewed under the microscopes of the day. Of course, now we are able to see much smaller particles as microscopy has advanced. However, the size range is still taken to be the same, although 10 μm would be a more appropriate upper limit as the unique behaviour of colloidal particles can still be observed with particle dimensions greater than 1 μm.

The particle size is similar to the range of the forces that exist between the particles and the timescale of the diffusive motion of the particles is similar to that at which we are aware of changes. These two factors, as we shall see later in this volume, are the key to understanding why so many colloidal systems have interesting behaviour and textures. Typically, the range of the interparticle forces is 0.1–0.5 μm whether they are forces of attraction between the particles or forces of repulsion. When we look at a colloidal sol in the microscope, we observe the particles to move around with a random motion. This is known as *Brownian motion* as it was recorded by the botanist Brown while studying a suspension of pollen grains in the microscope. The cause of this motion is, in turn, the motion of the molecules making up the suspending fluid. All the

Colloids and Interfaces with Surfactants and Polymers, Second Edition James Goodwin

atoms or molecules are in random or thermal motion and at any given instant the local concentration of a small volume element of the fluid will be either higher or lower than the global average concentration. The thermal motion of the colloidal particles will tend to be in the direction of the lower molecular densities. As these fluctuate in a random manner, so does the directional motion of the colloidal particles and the velocity is governed by the hydrodynamic drag. We know that diffusion tends to be away from high concentrations to low concentrations, so that if we have a high concentration of particles then there will be a directional drift away from this region. Now, for a sphere, the *Stokes drag factor*, s_v, is a function of the radius of the sphere, a, and the viscosity of the fluid, η, as follows:

$$s_v = 6\pi\eta a \tag{1.1}$$

The motion is random as we have already noted and the net velocity, v, is the average distance moved, $\bar{x}$, in the time interval t, namely:

$$v = \bar{x}/t \tag{1.2}$$

The work that has been done in moving a particle is simply the hydrodynamic force, $f_v = vs_v$, multiplied by the average displacement $\bar{x}$. The thermal energy available for this motion is $k_B T$, where k_B is the Boltzmann constant and T is the absolute temperature. Hence we can write

$$k_B T = \bar{x} f_v \tag{1.3}$$

Substituting for v and f_v and rearranging:

$$D = \frac{\bar{x}^2}{t} = \frac{k_B T}{6\pi\eta a} \tag{1.4}$$

Equation 1.4 is the Stokes–Einstein equation for the diffusion coefficient, D, and has units of $m^2\,s^{-1}$. We can define a characteristic timescale for this diffusive motion if we calculate the time it takes for a particle to diffuse a distance equal to the particle radius. This is done by a straightforward substitution of a for $\bar{x}$ in Equation 1.4 as follows:

$$t = \frac{6\pi\eta a^3}{k_B T} \tag{1.5}$$

Table 1.1 Types of colloidal dispersions

Phase	Gas (bubbles)	Liquid (droplets)	Solid (particles)
Gas	Molecular solution	Liquid aerosol (mist)	Solid aerosol (smoke)
Liquid	Foam (shampoo)	Emulsion (mayonnaise)	Sol (ink)
Solid	Solid foam (packaging)	Solid emulsion (butter)	Solid sol (stained glass)

This is known as the Einstein–Smoluchowski equation. For an isolated particle in water at 20 °C with a diameter of 1 μm, it takes about 0.5 s to diffuse one radius. When the colloidal dispersion becomes concentrated, the interactions with the neighbouring particles (hydrodynamic, electrostatic if the particles are charged or simply van der Waals forces) will slow the movement down. The timescale of our perception is approximately 1 ms to 1 ks so we should expect to observe interesting temporal behaviour with colloidal systems. We will re-visit this point later in this volume.

When we consider the number of possible phase combinations of our heterophase systems, we find that there should be eight different possibilities. This is illustrated in Table 1.1, where either phase could be a gas, solid or a liquid. Two gas phases will mix on a molecular level and do not form a colloidal system. Each of the other combinations results in systems with which we are familiar.

Gas bubbles and liquid droplets are spherical due to the surface tension forces unless the phase volume is very high. Solid particles may be spherical but are often non-spherical. The shape is a function of the history of the formation. Opals are an example of a solid sol with spherical silica particles in an aqueous silicate matrix. The silica particles are amorphous silica, and the distribution of sizes of the particles is narrow and the particles form a face-centred cubic array. It is diffraction of light by this highly regular structure which gives the characteristic colours seen in the polished gemstones. Colloidal dispersions in which the standard deviation on the mean size is less than 10% of the mean are usually considered to be 'monodisperse'. If the particle size distribution is broader than this, the dispersion is considered to be 'polydisperse'. Although this cut-off appears arbitrary, monodisperse systems have the ability to form colloidal crystals whereas polydisperse systems do not. Bimodal systems can also form crystalline structures if the size ratio is suitable. When the particles are formed by a crystallization process, other shapes are found. Silver chloride can be produced as a colloidal dispersion in water as monodisperse cubes. Haematite can form as ellipsoidal plates. Clays are

naturally occurring aluminosilicates that usually form plates. Kaolinite particles ('china clay') are hexagonal plates with an axial ratio of ~10:1. Montmorillonite particles can have much greater axial ratios and under the right conditions can be dispersed as crystals of one or two unit layers thick. Attapulgite has a lath shape and longer rod-like structures can be seen with chrysotile asbestos. These shaped particles show colloidal behaviour when the size is within the colloid range. For spheres or cubes we have three-dimensional colloidal size, with rods that is reduced to two dimensions whereas for plates only one dimension needs to be in the appropriate size range. This last case may seem strange but soap films are a good example of a system with two dimensions well within the macroscopic size range but with the third in the colloid range and being governed by colloidal forces.

This last example of a colloidal system brings into focus systems other than particles that have common ground with particulate colloids. Surface-active molecules or surfactants, such as soaps, detergents and lipids, can self-assemble to form multi-molecular aggregates of colloidal size and show the effects of colloidal forces in addition to their individual phase behaviour.

1.2 COLLOIDS IN ACTION

It will serve as a useful illustration to take some examples of colloidal systems and discuss why the colloidal state is used, what the important aspects are and what characterization is desirable. Although each colloidal material appears to be very different from others, there are frequently generic aspects which we can identify and so we can learn from solutions developed for quite disparate systems.

1.2.1 Decorative Paint

The function of this type of coating is twofold. First, it is intended to protect the surface from damage from environmental conditions. Second, it is intended to cover marks and produce an attractive colour. By choosing a colloidal system, we are able to manufacture and apply this very simply. A polymer film provides the surface protection. Synthesizing the polymer as colloidal particles dispersed in water can efficiently produce this. This material is known as a *latex* and is manufactured by the emulsion polymerization of vinyl monomers. The latter are dispersed as

an emulsion using surface-active materials (*surfactants*), which adsorb at the surface of the droplets and prevent them from coalescing. Once the polymerization reaction is initiated, the size and stability of the subsequent particles are also controlled by the surfactants. The advantages of using this colloidal synthetic route are excellent heat and mass transfer and simple handling of the product, which can easily be pumped out of the reactor and into storage tanks. Here we have to understand how the surfactants adsorb on different organic phases and operate at different temperatures.

The covering power of the film is provided by a white pigment and the colour by tinting with coloured pigments. Light scattered from the white pigment particles (usually titanium dioxide) hides the underlying surface. The particles must be fine enough to give a smooth film but not too fine or insufficient light will be scattered – 200 nm is about the optimum size. To manufacture this we must understand the control of crystal growth and the subsequent drying process to ensure easy redispersion of the dry powder down to the sub-micron level. The surface of the titanium dioxide is usually covered by a layer of alumina or silica to reduce catalytic breakdown of the polymer film when exposed to sunlight. The dispersion of dry powders in liquids requires surfactants and energy. Here, we have to understand how particles scatter light, the separation of colloidal particles and the 'wetting out' of dry powders followed by their subsequent redispersion. Hence this means how surfactants control the wetting of surfaces and how shear forces break up aggregates. The coloured pigments may be organic and therefore require different surfactant systems and so we may put together a system with three different surfactant materials and there will be ample opportunity for exchange at the various interfaces.

The final aspect of our paint is the application. At this point, the sedimentation of the pigment must be controlled and the viscosity has to be such that the wet film thickness is sufficient to give good hiding power. In addition, the brushmarks have to level out as much as possible and the polymer particles in the dry film must coalesce to give a coherent film. Soluble polymers are added to adjust the viscosity and to control sedimentation. This is partly due to the increase in the viscosity of the medium as a result of the entanglements of the long polymer molecules, but a major effect is for the polymers to induce a weak flocculation of the particles in a process known as *depletion flocculation*. Now, we must also understand how polymer molecules behave in solution, how they interact with particle surfaces and effect the particle–particle interaction forces.

The generic problems that we find when studying this coating areas are:

(a) control of particle size (of both inorganic and polymeric organic particles);
(b) surfactant behaviour in solution and adsorption;
(c) drying and the redispersion of powders;
(d) solution properties of polymers;
(e) particle interaction forces and the effect of surfactants and polymers on these;
(f) sedimentation in concentrated systems;
(g) flow properties of concentrated systems.

1.2.2 Paper

Paper is another material of colloidal origin, which we use without a second thought. It may be in the form of newsprint, a cardboard box, a glossy magazine or the high-quality material that our degree certificates are printed on. It is formed from cellulose, a naturally occurring sugar-based polymer most frequently obtained from trees. When wood is pulped for the manufacture of paper, the cellulose is separated into fibres with sizes stretching down into the colloidal domain. The fibres are filtered to give a mat and dried in a high-speed continuous process. The fibres are negatively charged and this plays a role in the tendency of fibres to aggregate, with the latter being an important feature in the formation of a dense filter mat in which the particles are aligned to give maximum strength in the direction of the moving sheet. The understanding of both particle aggregation and filtration is paramount for successful production in high-speed modern equipment.

Pigments such as titanium dioxide are added to give a white sheet. As the fibres are hollow, some of the pigment particles end up inside the fibres. Removal of this can become a problem when the paper is subsequently recycled. Ink from printing on the exterior of the paper is less of a problem but does require the removal by detergent action of surfactant materials. The attachment and detachment of particles from surfaces require an understanding of the interparticle forces and how we can manipulate them whether by chemical environment or surfactant type.

Glossy paper requires additional colloidal treatment. Well-dispersed kaolinite platelets are coated on to the surface and give a filler aligned parallel to the paper surface. Kaolinite consists of particles with negatively charged faces and positively charged edges, so they tend to stick

together very firmly to give a strong open particle network. This aggregation is controlled either by inorganic ions such as phosphates or organic polyelectrolytes [poly(acrylic acid), for example] and again the ability to manipulate interparticle forces is important. A binder is used with the clay surface to give a sealed, smooth and glossy final surface. A colloidal dispersion of polymer particles makes a suitable material. Emulsion polymerization is the normal route for this type of material. The application of the coating mix requires an understanding of the flow of concentrated dispersions.

Some of the generic problems that we may identify here are as follows:

(a) control of particle-particle forces;
(b) separation of colloidal systems;
(c) interaction of surfactants with surfaces and detergent action in the removal of particulates;
(d) hetero-aggregation and its control;
(e) particle size control.

1.2.3 Electronic Inks

Modern hybrid circuits are built up from sequential printing of fine circuits and layers of insulating material. The circuits are printed using inks with metallic colloidal particles dispersed in organic media. For example, gold or palladium has first to be produced as fine particles, separated and dried. Sufficient knowledge to allow the control of particle size and the subsequent separation of the colloidal particles is paramount here.

To make it into an ink suitable for printing, the system is dispersed in organic solvents with the aid of a surfactant to prevent the particles from sticking together. The mechanism of the stabilization must be understood. The viscosity of the concentrated dispersion has to be suitable for both flow during the screen-printing and the production of the correct film thickness. After drying, the circuits are completed by sintering the particles to give optimum conductivity. This process has parallel problems to film formation with polymer particles in other coatings, and also in the firing of ceramic materials, whether these are derived from clays or other oxides which are employed in high-grade ceramics used, for example, as chip bases in the electronics industry. The generic colloidal problems that we can immediately identify are as follows:

(a) particle size control;
(b) separation and drying of particles;
(c) wetting of dry powders;

(d) adsorption of surfactants;
(e) stabilization of particles in a dispersion;
(f) control of flow properties;
(g) wetting of surfaces;
(h) sintering of fine particles.

1.2.4 Household Cleaners

A large amount of surfactant is sold for domestic cleaning purposes, whether for clothes, skin or other surfaces. Each of these will have a different detailed formulation, of course, and as an example we will choose a cleaner for a surface such as a sink. The first requirement is that there is a high surfactant concentration. This is needed to solubilize grease and resuspend particulate material. Hence an understanding of detergent action is essential. Abrasive particles are required to break up the films that are responsible for staining the surface being cleaned, but these particles should not be of such a size that they produce deep scratches or produce a 'gritty' feel. Particles of a micron or two will be satisfactory. The creamy feel is also achieved by the formation of long branching 'worm-like' assemblies of the surfactant molecules, which requires a sufficient understanding of surfactant phase behaviour to optimize this requirement.

The size and density of the abrasive particles are such that sedimentation will occur in a short period, and to prevent this the system can be gelled by the addition of a soluble polymer. This has the side benefit of enhancing the texture or feel of the material. The solution behaviour of polymers and the control of the flow properties have to be understood in order to optimize the formulation. The generic problems here can be identified as follows:

(a) phase behaviour of surfactants in solution;
(b) detergent action;
(c) control of particle size;
(d) solution behaviour of polymers;
(e) control of flow properties.

1.2.5 Butter

Milk is a colloidal dispersion of fat droplets which are stabilized by the protein casein. The protein prevents the coalescence of the fat drops

by a combination of electrostatic repulsion and a steric barrier as the protein layers make contact. On standing, the fat drops rise to the top in a process known as *creaming*, which is analogous to sedimentation. So far, colloid stability and creaming (*sedimentation*) can be identified as areas of importance.

In the churning process, a phase inversion is produced and a water-in-oil emulsion is formed from an oil-in-water system. The saturated animal fats have a molecular weight such that they crystallize at temperatures close to body temperature. This is the reason why butter is difficult to spread at low temperatures. Many spreads are produced by blending in lower molecular weight vegetable oils with a lower melting point. The generic colloidal aspects are as follows:

(a) interaction forces between particles;
(b) coalescence of emulsion droplets;
(c) phase inversion of emulsions;
(d) flow behaviour of concentrated dispersions.

There are many other materials that are colloidal at some stage of their use, but the colloidal problems can still be reduced to just a few generic problems. It is important to recognize this in spite of the complexity of a particular system. At first sight, it is often difficult to understand how the apparently abstract physics and chemistry presented in most courses and texts can apply to a 'practical system'. The application of the general principles, however, are usually sufficient to enable the problems to be both defined and tackled in a systematic manner. All of these points will be addressed in the following chapters.

1.3 CONCENTRATED COLLOIDAL DISPERSIONS

Traditionally, our ideas of colloidal interactions have stemmed from the study of the behaviour of dilute systems of colloidal particles and the theoretical work based on two isolated particles interacting. This is nearly always in quite a different concentration region from the systems in which we employ colloids. However, in recent years this situation has changed and we now have a great body of work on concentrated dispersions. Of course, most of the academic work has been on model systems, but general principles apply to the more complicated systems that are in everyday use.

As a starting point, it is important to describe what we mean by a dilute dispersion. This is not based on just the value of the percentage by weight of the disperse phase or even the volume fraction. It is based on the mean separation of the particles compared with the range of the interaction forces between the particles. In the dilute state, the particles are well separated so that the particle interactions are negligible at the mean separation. The consequence of this is that the particles diffuse in a random fashion due to the Brownian motion, with a diffusion constant that can be described by Equation 1.4. The distribution of the particles in space can be considered as uniform, that is, randomly distributed, and the spatial correlations are very weak. Now, this is strictly true only for dispersions of particles which approximate to hard spheres. If there are forces of either attraction or repulsion acting between particles there will be some deviation from a random distribution as particles collide. This point can be important but we do not need to consider it in detail at this stage; we only need to be aware of the possibility. In a fluid continuous phase, the motion of particles can be described by the hydrodynamics appropriate to an isolated particle. This is true for diffusion, sedimentation or viscous flow. The behaviour of the dispersion can be thought of as analogous to that of a gas except that the motion is Brownian and not ballistic, that is, any two particles will experience many changes of direction before colliding. This means that the concept of *mean free path* is difficult to apply.

If we now steadily replace the continuous phase by more particles, as the concentration increases our colloid becomes a condensed phase and we have a more complicated behaviour. This is a familiar concept to the physical scientist, who will immediately recognize this behaviour as similar to that which occurs when a molecular gas is compressed until it forms a liquid and finally a solid. Many of the thermodynamic and statistical mechanical ideas translate well from molecular liquids to colloids in the condensed state. However, some caution is required as forces can be quite different. A liquid medium, for example, can result in hydrodynamic forces with a range of a few particle diameters. A very attractive feature though is that the colloidal forces can be readily manipulated by changes in the chemical environment of our colloidal particles. This, in turn, can dramatically alter the behaviour and thus it provides the means of manipulating our material to suit our needs more closely.

Now, in this condensed phase there will always be strong interactions between the particles. This is the case whether the interactions are repulsive or attractive. Such a situation gives rise to strong spatial correlations

and we have a shell of nearest neighbours. The number of particles in this shell is the coordination number and this reflects both the magnitude and type of force in addition to the concentration or particle number density. For example, if the particles are of very similar size and the forces are repulsive, colloidal crystals can be formed with very long-range order. The spatial arrangement is face-centred cubic and, if the lattice spacing is of the order of the wavelength of light, strong diffraction will be seen. Opal is a naturally occurring colloid where this effect is utilized as a gemstone. When the particles are in a liquid medium, 'exciting behaviour' can be seen. Three modes of diffusive motion can be identified. The particles are all moving due to the thermal or Brownian motion but are generally constrained to be within their individual coordination shell. This motion is fairly rapid and is known as the *short-time self-diffusive* motion. The motion is still random and, if we were to take a series of 'snapshots' of a particular volume, we would we see that the number density of particles in that region would fluctuate about the global mean for the dispersion. The diffusion of these regions is the *collective diffusion* and the constant is slower than for short-time self-diffusion. All liquids behave in this way and it is this local density fluctuations in the continuous phase that produces the Brownian motion of the particles. Occasionally, the fluctuations will allow sufficient separation in a coordination shell for a particle to move through and change its neighbours. This is known as *long-time self-diffusion*.

The flow properties reflect this interesting behaviour. To illustrate the point, let us consider a simple system of uniform particles with strong repulsive forces at a high concentration. The particles are highly spatially correlated in a face-centred cubic structure. If we deform the structure into a new shape, the arrangement of particles is distorted. We have had to do work on the structure and the energy is stored by the movement of the particles to a higher energy configuration. An elastic response is observed. Over time, the particles can attain a new low energy configuration in the new shape by the long-time self-diffusion mechanism. The system now will remain in the new shape without applying the external force, that is, the structure has relaxed and the elastically stored energy has dissipated (as heat). The time for this process is known as the *stress relaxation time* and the material is behaving as a *viscoelastic* material. In other words, we are saying that the material is now exhibiting a 'memory' and it takes several relaxation times before the original shape is 'forgotten'. The information that is 'remembered' or 'forgotten' is in the detailed interaction between individual particles or molecules. When this timescale falls within that of our normal perception we are aware of

the textural changes and many concentrated colloids are manipulated to take advantage of this.

The transition from a dilute to a condensed phase can be very sharp and is a function of the range of the forces, as noted above. We may now move back to consider a system of hard spheres – a system, incidentally, which can only really be attained in a computer simulation but which we can get fairly close to under very limited conditions. In a computer simulation, it is possible to take a fixed volume and increase the fraction of that volume which is occupied by particles, all in random Brownian motion of course. The volume fraction of the 'dispersion' is simply the product of the number of particles per unit volume, N_p and the particle volume, v_p:

$$\varphi = N_p v_p \tag{1.6}$$

The simulations show that a liquid–solid transition occurs at $\varphi_t \approx 0.5$. Below the transition, we have a viscoelastic liquid and above it a viscoelastic solid. How does this relate to systems with colloidal particles stabilized by long-range electrostatic repulsion or with extensive polymer layers that prevent the particles from coming together? We can introduce the concept of an *effective volume fraction*, which is calculated from the particle volume which has been increased by a volume from which neighbouring particles are excluded due to repulsion. For example, we can easily visualize the case for a dispersion of spherical particles, each of which has an attached polymer layer which physically prevents approach of another particle. Figure 1.1 illustrates this schematically.

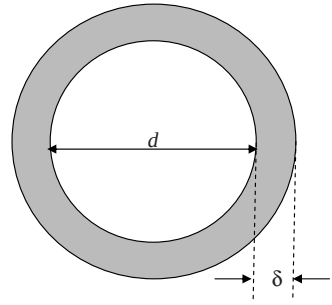


Figure 1.1 Schematic of a particle with an adsorbed polymer layer which increases the effective volume fraction of the system

The thickness of the polymer layer is denoted by δ, which gives the effective hard sphere diameter as $(d + 2\delta)$. The effective hard sphere volume fraction is now

$$\varphi_{HS} = N_p \frac{\pi (d + 2\delta)^3}{6} \tag{1.7}$$

and the liquid–solid transition would fall to a lower value of the volume fraction calculated from the core particles. Thus

$$\varphi_{HS} \approx 0.5$$

so

$$\varphi_t \approx 0.5/(\varphi_{HS}/\varphi)$$

and then

$$\varphi_t \approx \frac{0.5}{\left(1 + \frac{2\delta}{d}\right)^3} \tag{1.8}$$

When the stability is due to long-range electrostatic repulsion between particles, we may also define an effective hard sphere diameter. The simplest approach in this case is to recognize that the principle of the equipartition of energy applies to colloidal particles so that a particle moves with kinetic energy $k_B T/2$ along each of the x, y and z coordinates. Thus, an *average* value of the energy of a Brownian collision would now be $k_B T$. We may then take the distance δ as the distance at which the repulsive energy reaches this value and again define an effective hard sphere diameter as $(d + 2\delta)$. This now enables us to try to estimate the concentration of the liquid–solid transition. Figure 1.2 illustrates the result for a particle with a radius of 100 nm. We will return to this in more detail in a later chapter, but we should note at this point that because the electrostatic interactions are relatively 'soft' the material will form a *soft solid*. That is, the application of an external force can cause large deformations which can be permanent. This is a natural consequence of the range of the interparticle interactions compared with the particle size. The further we move to the right in Figure 1.2, the harder the solid becomes.

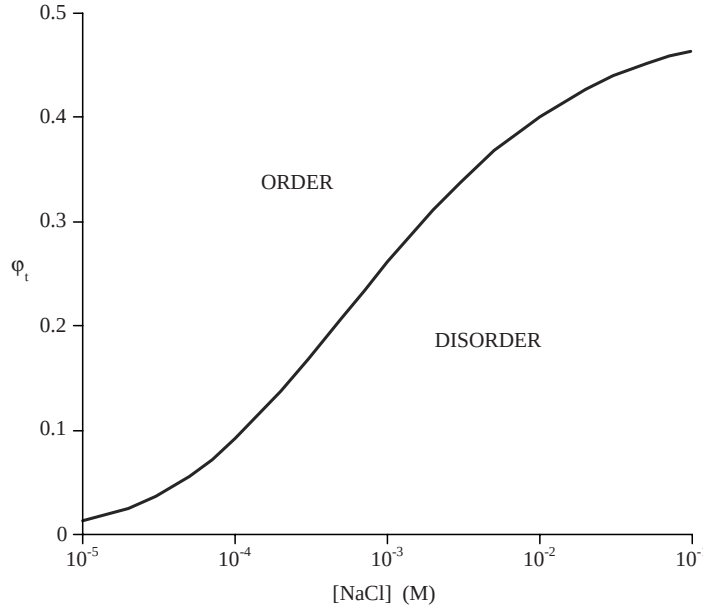


Figure 1.2 Order–disorder regions calculated for a 100 nm particle

1.4 INTERFACES

As soon as we consider a fine dispersion of one phase in another, the issue of the interface between the two phases becomes of major importance. As an illustration of the points that arise, consider the atomization of water into fine droplets in air. The area per unit mass is known as the *specific surface area* (SSA). The disperse phase is in the form of spherical particles because there are surface tension forces that we will discuss in a moment. The calculation of the SSA is based on the area of a sphere of diameter d (πd^2) divided by its mass [$(\pi d^3/6)\rho_{H_2O}$], where ρ_{H_2O} is the density of water. This gives

$$\text{SSA} = \frac{6}{d\rho_{H_2O}} \tag{1.9}$$

Thus, for 1 l of water (i.e. about 1 kg) before atomization, SSA $\approx$ 0.05 m^2. After spraying to give droplets of 1 μm, SSA $\approx 6 \times 103\ m^2\ kg^{-1}$ and we are now dealing with an interfacial area larger than the area of a football field! It is easy to see why the effectiveness of a catalyst is maximized when in a finely divided form and also why the oxidation of finely

divided materials such as metals or flour can be a dangerous problem due to the exothermic reaction becoming uncontrollable. If the droplet size were reduced to the order of 10 nm, the SSA would be $\sim 10^6\ m^2\ kg^{-1}$. It is interesting now to consider the fraction of the molecules that would be at the interface as the size of the drop is made smaller. The approximate number is shown in Figure 1.3 and it is a significant fraction for drops in the colloidal size range – particularly when the droplets would be in the nanoparticle size range, that is, up to a few tens of nanometres in diameter. This looks just like a simple exercise in geometry so far, but the implications are important. To illustrate this, let us think about the amount of work we would have to do to take our 1 kg of water down to droplets of 300 nm in diameter where ~0.1% of the water molecules are at the surface. Remember that the intermolecular forces in water are dominated by hydrogen bonding – giving the tetrahedral structure – and at 4 °C when the density is $1000\ kg\ m^{-3}$ this would be nearly complete. Thus, if we make the crude assumption that each surface molecule is one hydrogen bond short and that the energy of a hydrogen bond is $\sim 40\ kJ\ mol^{-1}$, then we may estimate how much work we would have to do to disperse the water into a fog. (Note that there is a factor of 2 as each hydrogen bond broken would result in two fresh surface areas.) The result is also illustrated in Figure 1.3. Of course, if we had broken

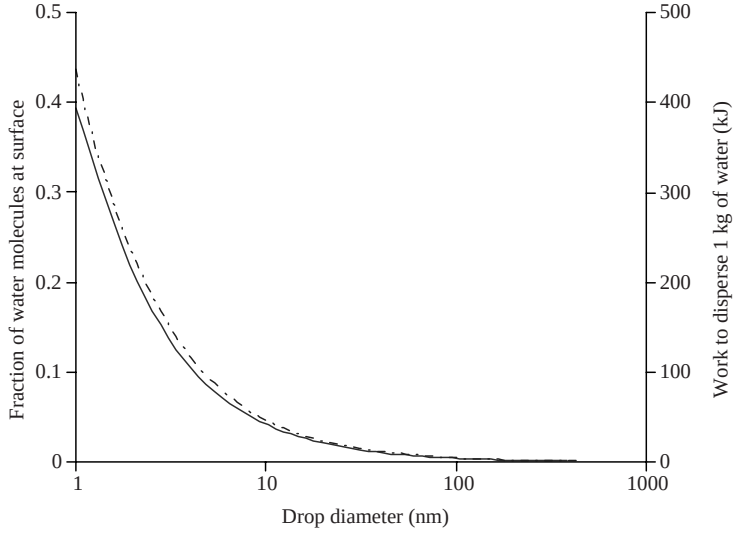


Figure 1.3 The fraction of water molecules in a drop that are located at its surface. Solid line, fraction of water molecules at the surface; dashed line, work to disperse 1 kg of water

all the hydrogen bonds, we would have boiled the water (this would take $\sim 2.5 \times 10^3$ kJ), but a lot of work is required to reduce bulk water to drops in the sub-micron level.

The above illustrates that we have to do work to create a new surface and that the origin is the work done against the intermolecular forces. This is a key concept when we consider surfaces or interfaces. Here, the term 'surface' is taken to refer to a surface of a liquid or solid in contact with a gas or vapour, and the term 'interface' is used to describe the region between two condensed phases, whether two liquids, two solids or a liquid and a solid. In the bulk of a condensed phase, the intermolecular forces act between the atoms or molecules in essentially a symmetrical fashion. At the surface or interface, there is an imbalance as the local chemical environment changes. If we think of the intermolecular forces as molecular springs, the imbalance in attractive force results in a surface tension, γ_l. This acts to minimize the surface area. Now, when the surface area of the liquid is increased by an amount ∂A against this surface 'spring' tension, the amount of work is given by

$$\partial W = \gamma_l \partial A \tag{1.10}$$

This is the case only for a pure material. If there are dissolved species present, we must consider the presence of those species at the surface or interface, as we shall see when we explore surfactants. The units of the surface tension are $J\,m^{-2}$ (i.e. energy per unit area) and, as energy is force multiplied by the distance moved, the dimensions are also written as $N\,m^{-1}$, which is the spring constant. Water, for example, has a value of $\gamma_l = 72\,mN\,m^{-1}$. If we integrate Equation 1.10 up to an area of $1\,m^2$, we have the energy required to create a fresh surface of unit area, and we see that if the area is the SSA of droplets 300 nm in diameter, we require 1.4 kJ. This value compares favourably with the simplified estimate illustrated in Figure 1.3.

In water, the hydrogen bonding interaction is the strongest intermolecular force but it is not the only contribution. The usual van der Waals forces also play a role and contribute about 25% of the surface energy of water. These are the forces that cause an interaction between all atoms and molecules, even the inert gases. They are the *London dispersion forces*, which are due to the coupling of the fluctuations of the charge density of the electron clouds of one atom with its neighbours. This will be discussed in some detail in Chapter 3, with aspects of surface energy being discussed in Chapter 6. An important feature of the recognition that an appreciable amount of work is required to generate new surfaces

is that the process is endothermic and that the dispersed state is not the lowest energy condition. In other words, there is a natural tendency for droplets to coalesce and for particles to aggregate. To maintain the material in the colloidal state, we must set up the correct conditions.

We have just begun to explore the molecular implications of an interface or surface. The structure of the liquid surface in equilibrium with its vapour cannot be as well defined as that of a crystalline solid and the concept of a well-defined plane is a convenience rather than a reality as there is really an interfacial *region*. When a surface is expanded or contracted, diffusional motion to or from the bulk accompanies the changes and the intensive properties of the interface remain unchanged. With a solid surface, the situation can be more complex and crystal structure, for example, can result in anisotropy. The surface free energy described above appears to be straightforward. However, equating the surface free energy just with the surface tension can only hold for a pure liquid. Whenever another species is present, the distribution becomes important as this controls the details of the intermolecular forces in the interfacial region. If the concentration of solute species is lower in the surface region than in the bulk phase, the species is termed *lyophilic* as it 'prefers' the bulk phase to the surface phase. The solute species is *negatively adsorbed* at the surface or interface. Indeed, the stronger interaction between the lyophilic solute species and the solvent can even lead to a small increase in the surface tension. If the molecules tend to accumulate at the interface, they are termed *lyophobic*. This tendency for the solute species to accumulate at the interface implies that the intermolecular interactions are most favourable if there is a separation of the solvent and solute into the region of the surface. This is particularly marked with *amphiphilic* (also termed *amphipathic*) molecules. These are a class of molecules known as *surfactants* or *surface-active agents*. In this case, there are two distinct moieties making up the molecule: part of the molecule is lyophilic and the another part is lyophobic. In water, a polar group such as the salt of a carboxylic acid group would be a lyophilic moiety. In water, this is also described as being *hydrophilic*. A linear paraffin chain or an aromatic hydrocarbon would be a typical lyophobic or *hydrophobic* moiety. The increase in concentration at the interface is known as the *surface excess*.

The surface tension of water is lowered as solute molecules accumulate in the surface region. Water is an associated liquid and the solute molecules do not display the relatively strong hydrogen bonding forces that occur between adjacent water molecules. Hence, even if the London dispersion forces are stronger, the surface tension is lowered. A

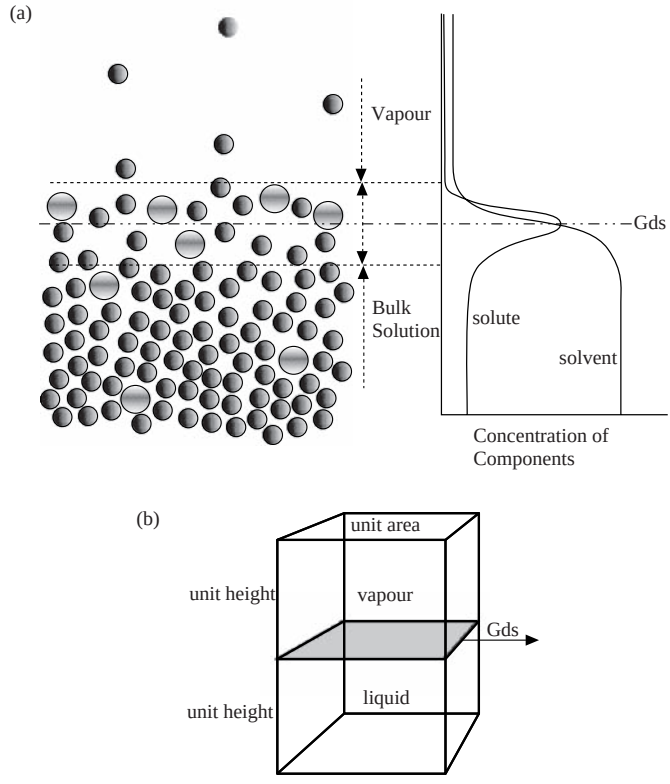


Figure 1.4 Representations of a simple model for the liquid–vapour interface. Gds indicates the Gibbs dividing surface (see text for details)

diagrammatic picture of the surface of a solution is shown in Figure 1.4. Of course, this picture is not restricted to the surface of an aqueous solution.

There are some important ideas illustrated in this figure. The interface between the liquid phase and the vapour phase is not a plane when we work at the molecular level. Rather, it is a region a few molecules in thickness – say five or six – where the molecular density or concentration profile changes from that of the liquid to that of the vapour. Hence we can think of there being a *surface phase*. When there are two molecular species present, we can expect the concentrations to vary with the nature of the solute species, as indicated in the previous paragraph. In the figure, we have large solute molecules which are lyophobic and so there is a surface excess concentration. This illustrated with the peak in the concentration profile (Figure 1.4a) and, as shown, the large molecules have a much lower vapour pressure than the solvent molecules, but this,

of course, is not a prerequisite. When we know the local concentration, in principle we can estimate the surface tension. Direct measurement of the concentration profiles is not something that has been achieved with precision so far, but it is possible to estimate the surface excess from measurement of surface tension. To do this we need to use just a little thermodynamics, as clearly laid out in the book by Everett [1].

First, we are going to choose a volume for our system at equilibrium which contains saturated vapour, v, the solution phase, ℓ, and the surface phase, s. Our problem is to define the volume of this surface phase. What we are going to do is to model it as though it were just a planar surface with all the material in the surface phase located in that plane. This plane is known as the *Gibbs Dividing Surface* – the Gds line in Figure 1.4(a) – and for simplicity we will consider a volume with unit area Gds as in Figure 1.4(b). As this is a model, we may choose the location of the Gds to be the most convenient, i.e. to make the calculations as simple as possible. The appropriate concentration terms are defined as follows:

Γ_{1s} is the number of moles of solvent species per unit area at the Gds;
Γ_{2s} is the number of moles of solute species per unit area at the Gds;
$c_{1\ell}$ is the concentration of solvent in the liquid phase;
c_{1v} is the concentration of solute in the vapour phase;
$c_{2\ell}$ is the concentration of solute in the liquid phase;
c_{2v} is the concentration of solute in the vapour phase;
c_1 and c_2 are the total concentrations of solvent and solute in the system, respectively.

Thus, we have $c_1 = c_{1\ell} + c_{1v} + \Gamma_{1s}$ and $c_2 = c_{2\ell} + c_{2v} + \Gamma_{2s}$, that is,

$$\Gamma_{1s} - c_1 = -(c_{1\ell} + c_{1v}) \quad \text{and} \quad \Gamma_{2s} - c_2 = -(c_{2\ell} + c_{2v})$$

$$\frac{\Gamma_{1s} - c_1}{c_{1\ell} + c_{1v}} = \frac{\Gamma_{2s} - c_2}{c_{2\ell} + c_{2v}}$$

$$\Gamma_{2s} = c_2 + (\Gamma_{1s} - c_1)\left(\frac{c_{2\ell} + c_{2v}}{c_{1\ell} + c_{1v}}\right)$$

which gives

$$\Gamma_{2s} - \Gamma_{1s}\left(\frac{c_{2\ell} + c_{2v}}{c_{1\ell} + c_{1v}}\right) = c_2 - c_1\left(\frac{c_{2\ell} + c_{2v}}{c_{1\ell} + c_{1v}}\right) \tag{1.11}$$

In principle, the latter term is experimentally accessible, but we can simplify Equation 1.11 if we choose the location of our Gibbs dividing surface carefully. We will define our dividing surface so that the excess number of solvent molecules on the vapour side is exactly matched by the deficit on the liquid side. This gives the value of Γ_{1s} as 0 and then we call the surface excess of the solute, Γ_{2s}, *the relative adsorption of solute* at the surface.

The Helmholtz free energy of the system is just the sum of the free energy of each phase:

$$F = F_v + F_\ell + F_s \tag{1.12}$$

The surface term is of importance for our colloidal systems where the surface area is large. For the bulk phases, we have the usual equation for the change in free energy with the amount n of species i:

$$dF_v = -S_v dT_v - p_v dV_v + \sum_i \mu_{vi} dn_{vi} \tag{1.13}$$

and an analogous equation for the surface:

$$dF_s = -S_s dT_s + \gamma_s dA_s + \sum_i \mu_{si} dn_{si} \tag{1.14}$$

Here, the pressure term has become the surface tension term and the sign has to change as it is a tension instead of a pressure. The phase volume is replaced by the area of the surface. The temperature is constant so when we integrate Equation 1.14 we have the Gibbs–Duhem equation for the surface:

$$F_s = \gamma_s A_s + \sum_i \mu_{si} n_{si} \tag{1.15}$$

Differentiating Equation 1.15 generally gives

$$dF_s = \gamma_s dA_s + A_s d\gamma_s + \sum_i \mu_{si} dn_{si} + \sum_i n_{si} d\mu_{si} \tag{1.16}$$

We can now equate Equations 1.16 and 1.14, recalling that the 'SdT' term is zero as we are working at constant temperature, to give the

following:

$$A_s d\gamma_s + \sum_i n_{si} d\mu_{si} = 0 \tag{1.17}$$

Dividing through by A_s gives us the relative adsorption of the components as follows:

$$d\gamma_s = -\sum_i \Gamma_{si} d\mu_{si} \tag{1.18}$$

With a system with just two components, we can choose the Gds to give $\Gamma_{s1} = 0$ and so remove the solvent from the equations. In addition, it is convenient to use the chemical potential of the solute in the liquid phase (at equilibrium the chemical potential of each species, μ_i, is the same in each phase) and we have the *Gibbs adsorption isotherm*:

$$d\gamma_s = -\Gamma_{s2} d\mu_{\ell 2} \tag{1.19}$$

The chemical potential is related to how much of the solute we have in the liquid phase, that is, the activity of the solute:

$$d\mu_{\ell 2} = RT d\ln a_{\ell 2} \tag{1.20}$$

This now gives us a convenient means of estimating the relative adsorption of the solute at the surface by measuring the slope of the curve of the surface tension as a function of the natural logarithm of the activity:

$$\Gamma_{s2} = -\frac{1}{RT}\left(\frac{d\gamma_s}{d\ln a_{\ell 2}}\right) \tag{1.21}$$

This equation is frequently used to estimate the amount of strongly adsorbed material such as surfactants at the liquid surface. It will only be approximate if the molar concentration is used as even though the solution concentrations are usually low there are problems such as they are far from being ideal solutions with an activity coefficient of unity. When there are several components present the algebra is only slightly more complicated, and general expressions are clearly given in the book by Everett [1].

1.5 SURFACTANTS

Surfactants are molecules which have a chemical structure that makes it particularly favourable for them to reside at interfaces, hence they are termed *surface-active agents* or simply *surfactants*. These molecules are a frequent component of colloidal systems whether synthesized or naturally occurring, so it is of great importance to know how much resides at the interfaces in our systems. It was shown above that the rate of change of surface tension with the logarithm of the activity gives an estimate of the amount of the solute adsorbed at the interface. Now, we should use Equation 1.21 to make all the algebraic manipulation worthwhile and to get a feel for what the equation can tell us. The example that we will take is the experimental data plotted in Figure 1.5 for a simple cationic surfactant in water.

The surfactant is hexadecyltrimethylammonium bromide (C_{16}TAB). It consists of a straight 16-carbon aliphatic chain with the quaternary ammonium group as the terminal group at one end. The ionic terminal group carries a positive charge and is strongly solvated so that the long

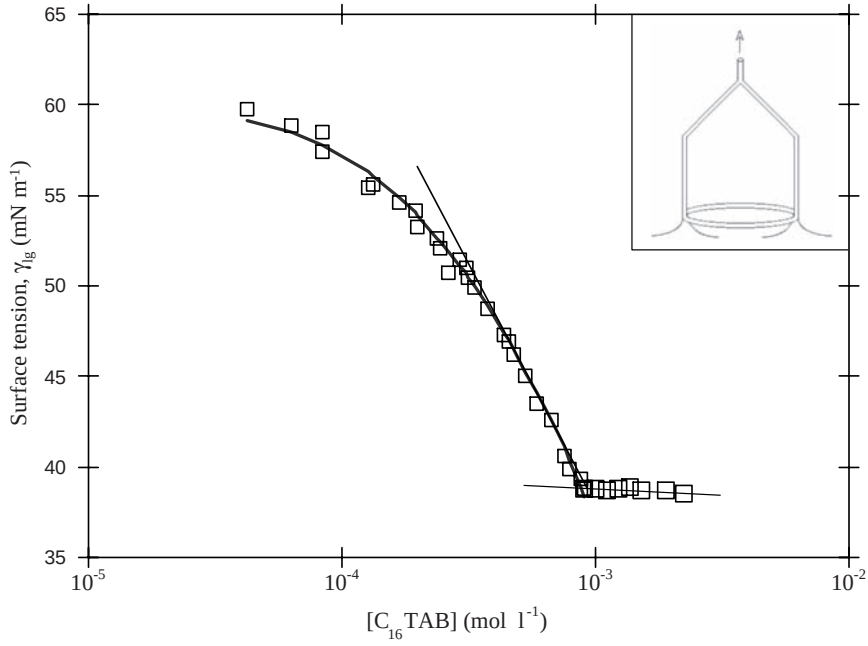


Figure 1.5 Surface tension as a function of concentration for hexadecyltrimethylammonium bromide

aliphatic chain is carried into solution in water. The solution behaviour of such surfactant molecules will be discussed in more detail in Chapter 2, but it is a good example for our current use. An aliphatic chain of 16 carbon atoms is not very soluble in water and the result is that there is strong adsorption at the water–vapour interface. The experimental curve of surface tension versus concentration is shown in Figure 1.5. The surface tension shows a monotonic decrease up to a concentration of 9×10^{-4} mol l^{-1}. Beyond this, the curve is almost parallel to the x-axis. This point at which this abrupt change in slope occurs is known as the *critical micelle concentration* (cmc). We will come to this shortly, but let us concentrate on the first section of the curve. First, we must recognize that we are using molar concentrations and *not* the activities. Although the concentrations are low, the activity coefficient will be slightly less than 1, so our results will only be approximate, but still of use. The limiting slope of the curve prior to the cmc is 1.18×10^{-2}, which yields a value from Equation 1.12 for $\Gamma_{s2} = 4.6 \times 10^{-6}$ mol m^{-2}. At 35 °C we have the area occupied by a C_{16}TAB molecule as 0.36 nm^2; this is about twice that found for an undissociated fatty acid which gives a close packed layer at 0.19 nm^2. The first thing to note is that the trimethylammonium head group is larger than a carboxylic acid group, but is it twice as big? – well, perhaps not. Hence the second feature that we should consider is that the group is positively charged. Like charges repel, and this acts to reduce the packing density.

Let us consider the charge in more detail. We have a surface for which we estimate from the surface tension measurements that there would be one positive charge (i.e. 1.6×10^{-19} C) for every 0.36 nm^2 of surface. This gives a measure of the *surface charge density*, σ_s, of $\sim$45 μC cm^{-2}. Experiments with solids, such as silver iodide, or oxides, such as titanium dioxide, yield surface charge densities in the range 1–15 μC cm^{-2}, so this clearly would be a very highly charged surface. Of course, the head groups are just one half of the ion pair, while the bulky bromide ion is the counter-ion to the surface charge and will be strongly attracted to the positively charged surface. The binding of the counter-ions reduces the repulsion between head groups. The charge on the surface attracts the counter-ions but, as the concentration of the latter is high, diffusion acts in the opposite direction, tending to dilute the concentration at the surface. The model for the surface now consists of the hexadecyltrimethylammonium ions located in the surface with the hydrocarbon tails extended into the vapour and the head groups in a densely organized layer which is highly charged. The charge is balanced by many counter-ions which are closely bound to the surface and the

remaining counter-ions in a more diffuse layer where the remaining electrostatic attraction is balanced by diffusion. This concept of a charged surface with a layer of counter-ions, some of which may be strongly bound and the remainder are in a diffuse array, is a key concept which helps us to understand the behaviour of charged particles in a dispersion. It is known as the *electrical double layer* and will be discussed more fully in subsequent chapters.

This is an appropriate point to discuss the measurement of the tension of the surface. The data in Figure 1.5 were obtained by measuring the force exerted when attempting to pull a platinum ring out of the surface. The equipment is a DuNoüy tensiometer and this is just one technique for determining the surface tension of a liquid. Chapter 6 give details for several other methods. The inset in Figure 1.5 illustrates the geometry of the measuring element. As a force is exerted on the ring supported perpendicular to the surface, the surface resists the displacement of the ring. In principle, the force at which the ring will detach is given by the surface tension in Newtons per metre multiplied by twice the circumference of the ring (in metres) (remember the surface makes contact with both sides of the platinum wire of the ring). A computer-controlled microbalance does the job of measuring the maximum force very well. However, the points that we need to keep in mind here arise from the usual condition in thermodynamic calculations that at some point we have required the system to be at equilibrium. Thermostatting is, of course, a prerequisite. The first problem that we must deal with is that the vapour phase should be saturated. Hence our system should be enclosed and sufficient time taken for the vapour phase to come to equilibrium. This is particularly important if the vapour pressure of the solute is significant compared with the solvent. This is not a problem with large molecules such as C_{16}TAB, however. The second problem of equilibrium is important in this case as at low concentrations of surfactant a significant time passes before the molecules in solution diffuse to the surface and equilibrium becomes established. Each point on the curve shown in Figure 1.5 usually follows a dilution of the solution and mixing. At concentrations close to the cmc there are many surfactant molecules close to the surface and equilibrium is quickly attained. However, at the other end of the curve several minutes are needed for consistent measurements to be achieved, repeat readings are necessary to confirm the values and the time taken to produce the full curve can stretch into hours!

The slope of the surface tension–log(concentration) curve increases steadily as the surfactant concentration is increased. This tells us that the relative adsorption of the C_{16}TAB is increasing as more is added to

the water. However, at the cmc there is an abrupt change in slope and what then occurs is that the surface tension changes very little with more concentrated solutions. What we find here is that above the cmc, where the surface is closely packed, there are small aggregates of surfactant molecules in solution. In other words, surfactant in excess of that required to give a concentration equal to the cmc has self-assembled into macro-ions. These aggregates are called micelles with a surface charge, the nature of the charge reflecting that of the hydrophilic head group of the surfactant molecule. Typically, the aggregation number of surfactant molecules in a micelle is around 50–100 when the solution concentration is close to the cmc and the diameter of each micelle is a few nanometres. The core of the micelle can be pictured as rather like a small oil droplet with the surfactant head groups located at the surface. The latter moieties are strongly hydrated and the first two or three carbon atoms of the tail close to the head group are close enough to be influenced by the head group hydration. In fact, on the nanometre scale the concept of a clear distinction between the outer edge of the hydrocarbon core and the aqueous phase breaks down. This ability of surface-active species to self-assemble into various structures is extremely important in a wide range of applications, from cell membranes to washing clothes.

It is also possible to use the variation in surface tension with surfactant concentration to monitor the adsorption of the surfactant on the surfaces of particles in suspension. At equilibrium concentrations up to the cmc, the procedure can be similar to a titration where a surfactant solution of known concentration is added and the surface tension monitored without separating the solids from the liquid. However, beyond the cmc, the phases must be separated, for example by centrifugation, and an aliquot of the supernatant removed and diluted carefully to below the cmc prior to the measurement. The data in Figure 1.6 show the adsorption isotherm of C_{16}TAB on a sample of china clay. For comparison, data obtained from radiochemical assay are also given. The faces of the clay particles were negatively charged and the edges positively charged at the pH of the experiment, so the adsorption occurs on the particle faces. The isotherm shape is typical of that of a high-affinity isotherm. Initially, the attachment is by the head groups of the surfactant molecules leading to a monolayer, which results in a hydrophobic surface, and further adsorption occurs to give a bilayer. This coverage occurs at an equilibrium concentration of the surfactant in the solution which is approximately half the value of the cmc. At much higher concentrations, there is evidence of yet further adsorption. The clay surfaces are not simple, however, as they possess 'steps' and the adsorption

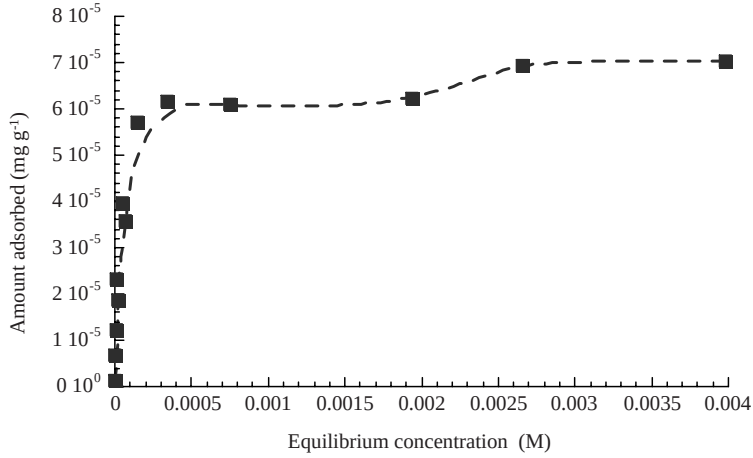


Figure 1.6 Adsorption isotherm of hexadecyltrimethylammonium bromide on sodium kaolinite at 35 °C. Data for adsorbed amounts below 4×10^{-5} mg g^{-1} were obtained by independent radiochemical assay

close to the step edges may require higher equilibrium concentrations. However, prior to the adsorption of the surfactant, the clay particles are aggregated edge-to-face in a 'house of cards' structure. As soon as the adsorption plateau is reached, the particles are completely dispersed and the surfactant titration technique is well suited to provide this type of adsorption data rapidly. At the plateau, the area occupied by each molecule (calculating the face area from the specific surface area measured by gas adsorption and reducing this by the fraction corresponding to the edge area) is $\sim$0.5 nm^2 in each layer (note that this is fairly close to that found at the air–water interface at the same equilibrium concentration).

One of the main uses of surfactants is to provide stability to dispersions of colloidal particles and the above titration technique provides a quick method to determine how much surfactant is required. However, the molecules are only physisorbed and *not* chemisorbed, so care has to be taken when additions to the system are made. If the system is diluted with solvent, then surfactant will desorb until a new equilibrium is attained. To prevent this, dilution should be carried out with a solvent phase containing the equilibrium concentration of surfactant required to maintain the value where the adsorption plateau occurs. In addition to the provision of colloidal stability, surfactants are also used to aid the wetting and hence the dispersion of powders in liquids, in addition to aiding the break-up of oil droplets in emulsification processes, as we shall see in later chapters.

1.6 SOLUTION POLYMERS

Macromolecules or polymers, like surfactants, are often a key component in colloidal systems and so it is important to introduce them here in this early part of the text. The robustness of the stability against aggregation of many colloids of biological origin is due to the presence of proteinaceous macromolecules on their surfaces. As an example, we have to look no further than the stabilization of the fatty acid droplets in milk which are stabilized by casein. We often add polymers which will adsorb on particles for this purpose. However, Nature has provided a very effective mechanism for keeping particles apart by three components. Only part of the macromolecule adsorbs, that is, is attached. This leaves the rest, which is solvated to expand away from the interface and prevent other particles from close approach. The proteins are also charged, and the charges repel other particles too, adding to the effectiveness of the stabilizing layer. That charges are located at the outer edge of the adsorbed layer brings an additional advantage in that the origin of the repulsion now occurs at a distance from the core particle surface, which is where we consider the origin of the attractive forces to be.

Synthetic polymers are also used as stabilizers. Homopolymers are not much use as stabilizers, because if they are readily soluble in the continuous phase, they will not form strong, effective attachments to the surface. Hence we emulate the smaller molecules like surfactants and synthesize the polymers with some lyophobic blocks along the chain. Frequently, these polymers are of relatively low molecular weight, typically in the range 5×10^3–50×10^3.

Polymers of higher molecular weights are also in common use. These are employed to alter the flow or sedimentation behaviour of colloidal systems. For this reason, they are termed 'thickeners' or 'rheology modifiers'. A polymer in solution increases the viscosity of that solution and high molecular weight material is particularly effective at this so that only a small amount is required. When molecular weights $>10^6$ are utilized, however, problems in rheological behaviour become apparent. For example droplets do not break away from the bulk cleanly: we see a 'stringy' behaviour, which is due to a marked resistance to stretching. That is, the extensional viscosity is high and applications such as spraying become difficult. One solution to this problem is to use a lower molecular weight polymer and make it behave like much higher molecular weight polymers under quiescent conditions but like lower molecular weight polymers on application. This is done by having a small molar percentage of lyophobic polymer on the backbone of the polymer, which

results in a weak assembly of these regions so all the polymers are associated with each other. This has similarities to the self-assembly of surfactant molecules and is becoming increasingly widely utilized.

It is interesting that when a soluble polymer is added as a rheology modifier to a colloidal dispersion, *a synergistic* effect is often observed. That is, the relative increase in viscosity of the dispersion is markedly greater than the relative increase found for the polymer solution on its own. What occurs here is that a solution polymer, which does not adsorb to the disperse phase, produces a weak, reversible aggregation of the disperse phase and this increased interaction is observed as a further change in the rheological behaviour.

Polymers with charged groups are known as *polyelectrolytes* and these can be added either as stabilizing agents for particulate dispersions or to cause aggregation. For example, poly(acrylic acid) produces good dispersion of china clay by adsorbing on the edges carrying a positive charge. Positively charged polyacrylamide can be used to remove negatively charged particulates by a bridging mechanism which holds particles together and makes them easy so separate. The polymer concentration required to do this is extremely low. Too high a level could give complete coverage of the surfaces by the cationic polymer and provide (unwanted) stability of the system.

1.7 THE WORLD OF NANOPARTICLES

The smaller end of the colloidal size range has been important both in Nature and in manufactured products. However, it has only been recognized in the last decade or so that the small size can give rise to special properties that we may exploit. An important feature of nanoparticles is that they are small when compared with the range of the forces that govern the physical behaviour of colloidal dispersions. The term 'nanoparticle' is currently widely used to describe particles which have at least one dimension >100 nm. Figure 1.3 shows that the fraction of the water molecules making up a spherical drop that occupy a position at the surface of the drop increases rapidly as the drop diameter becomes >100 nm. This is a key factor in many of the applications as the molecules at the surface in a higher energy state and we can see the relationship between surface energy, surface tension and wetting, the macroscopic demonstration of which is discussed in Chapter 6. In water droplets, the intermolecular forces are much weaker than in a solid particle consisting of an inorganic crystal. The high surface energy in combination with the

high specific surface area accounts for the high catalytic activity of some metallic nanoparticles. A more exciting result of the marked difference in energy of an atom in the surface region compared with one in the bulk is that some metastable polymorphs, which will necessarily have a lower surface energy, can become the more stable form of the crystal if the size is reduced to the nanoscale so that a large fraction of the atoms have the lower energy. This phenomenon can be observed with alumina, titania, zirconia and the oxides of iron [2, 3]. For example, of the hydrated iron oxides, haematite is the most stable crystal structure until the particle size is reduced to a level where the specific surface area is $>75\,m^2\,g^{-1}$. Goethite then becomes the more stable structure until the diameter is $>4.5\,m^2\,g^{-1}$, when akageneite is then more stable. The crystal structure controls the magnetic properties. The amount of information that can be stored on magnetic media, for example, is a function of particle size and so the interplay between structure and size is clearly important. The binding of atoms or molecules to the surface of crystals is a function of the surface energy and so the crystal structure will affect this. For example, a smaller extent of binding of lead to a surface would be disadvantageous if we were attempting to sequester contaminants, but an advantage if we were looking for a longer lived catalyst.

1.8 PREPARATION OF NANOPARTICLES

A huge range of materials are now available commercially as nanoparticles, albeit many in small quantities at present. The discussion will therefore be limited to some general issues as the individual chemical issues can easily occupy a whole text and there is currently a huge research and development effort in this area.

1.8.1 From the Vapour Phase

Many metals, oxides and even carbon are prepared by condensation from the vapour phase to give nanoparticle dispersions in powder form. The use of the established plasma-gas process gives uniform nanosized particles. The problem here is that the dry powders that result have the nanoparticles in large agglomerates and we have to try to disperse these materials in a variety of media for the final application. The usual first step is to try to produce a liquid dispersion at as high a solids loading as possible. The object is to disperse the agglomerates and break them down

to the single particles, but retain the original surface character. Significant damage to the surface of the crystals can be the precursor to crystal structure change and a loss of activity. Nano-titania is a good example. It has a very high level of catalytic activity in the nanocrystal form, but if the deagglomeration process has been of high enough energy to damage the crystal surface, the material becomes amorphous, reagglomerates and loses its activity.

The process with the lowest energy inputs is the use of a high-speed mixer, which may be sufficient to produce a good dispersion. Higher energy inputs may be required and an ultrasonic probe may be required. In small volumes, heating effects may limit the time or energy input. With larger volumes, a flow-through cell of small internal volume may be used to process large volumes of material. This ensures that all volumes of the material have similar exposures to the energy input and both temperature and exposure time can be controlled by the flow rate. In many cases, however, a higher energy process is required and we may have to resort to bead milling.

Ball milling has been a used for many years to disperse pigments in paint media. In its simplest form, the mill consists of a rotating cylinder which contains the paint mix and 1 cm diameter ceramic balls. The tumbling action of the balls produces collisions with enough energy to fracture aggregates. The relative motion between the surfaces of adjacent balls creates a high-shear region which aids the dispersion. However, it is the impact energy of the collisions that does most of the work. In a conventional ball mill, the surfaces of the pigment particles can be badly damaged. This usually results in some reagglomeration so that well-dispersed nanosized particles are not produced. Smaller beads are more effective and can be run to produce a gentler and less damaging process.

The modern generation of bead mills for ultrafine grinding uses beads in the size range 50–100 μm. A flow-through design is used in which the energy input is via a central rotor with multiple triangular vanes. The outlet pipe has a fine screen so that the grinding medium is retained. In order to prevent build-up of the medium on the screen and concomitant back-pressure build-up, the outlet is located in a slotted cylindrical cavity in the end of the rotor. This configuration acts as a centrifugal pump which returns the grinding medium to the agitated region. The critical principle here is that the microbeads are very large in number and so make multiple low-energy impacts with the aggregates that are being disrupted. In addition, at the bead loadings used, the grinding medium acts as a dynamic filter that holds the aggregates back in the grinding

zone and prevents them from reaching the outlet screen. These mild milling conditions enable the aggregates of nanoparticles to be dispersed successfully with no surface damage if the energy input is carefully controlled. Beads may be made from yttria-stabilized zirconia or quartz and are effective in the preparation of nanoparticle dispersions of materials such as diamond for polishing applications, titania for catalysts and nickel for capacitors.

1.8.2 Templated Growth

Carbon nanotubes have interesting electronic properties and have great potential in areas such as microelectronics and wearable electronics. Single-walled carbon nanotubes have a high current-carrying capacity along the tube, that is, metallic conductance, but they can also be produced as semiconductors. The challenge is to produce bundles of aligned nanotubes of similar dimensions. A quartz substrate can be used as a template for growing single-walled carbon nanotubes by chemical vapour deposition in a direction perpendicular to the crystal surface [4]. Kocabas *et al.* [5] used the step edge of mis-cut quartz crystals to permit orientated nanotube growth horizontally along the surface. When iron catalyst is printed in certain regions, the growth in these regions gives a random orientation, with the result that these regions are highly conducting zones within the semiconducting area of the aligned nanotubes, and offer potential contacts for a carbon nanotube-designed transistor.

1.8.3 Formation from Solution

The formation of nanoparticles from solution is a widely used technique and can be applied to both the formation of nanosized crystals and the precipitation of polymers. Although the chemistries will depend on the type of particle being produced, there are some common physical process characteristics that are required to produce nanoparticles with a narrow size distribution. It should be noted that in order to utilize the properties of the nanoparticles fully, monodispersity, that is, a narrow size distribution, is desirable. The first process requirement is to achieve a rapid nucleation process. This is then followed by a slower growth process. If further nucleation can be avoided during this phase of the formation process, the nuclei will grow simultaneously and a monodisperse system will be formed. Finally, stability of the particles against aggregation is

required and this can be achieved by the adsorption of surfactants or by the build-up of charged species at the growing interface. At this point. it will be instructive to describe briefly some nanoparticle syntheses.

1.8.3.1 Gold nanoparticles

Colloidal gold is one of the earliest nanoparticle dispersions formed in a controlled way in the laboratory, originally by Michael Faraday. A sol of monodisperse 40 nm gold particles can be formed by heating a 0.1% solution of chloroauric acid ($HAuCl_4 \cdot 3H_2O$) to 100 °C followed by the addition of 1% sodium citrate solution to give a final citrate concentration of 0.25%. The citrate reduces the gold chloride and a stable sol is produced which has negatively charged particles with a diameter of ~40 nm [6]. The gold sol has a strong ruby colour. Surface functionality can be added by reaction with a variety of thiols.

Gold nanoparticles are being used in the medical diagnostics industry for point-of-care devices based on lateral flow platforms. These assays use gold particles with physically adsorbed antibodies that are dried down on a membrane. When hydrated with body fluids (blood, urine or saliva), the particles move chromatographically through the membrane. If antigens from the condition of interest are present in the fluid, they attach to the antibodies on the particles and are carried forward with these to a read-out zone which has some antibodies attached to the membrane matrix. These capture the gold particles, which now have an antigen-covered surface and a red–brown band develops if the condition is present. If not, the gold particles pass through and can be captured in a narrow zone with antigens linked to the matrix. This second region provides a negative control, indicating that no antigens were present in the body fluid and that the capture antibodies were not saturated by a very high level of free antigen. Each manufacturer has their own variant on this general format, but in all cases lateral flow test units are cheap, give results in a few minutes and can be used in the field, clinics or hospitals.

1.8.3.2 Silver halides

When ~20 mM solutions of silver nitrate and sodium halide are rapidly mixed at room temperature, a silver halide sol is formed. The particle–solution interface functions like a reversible electrode so that the charge is a function of the ion type that is in excess in solution (silver or halide) and its concentration. For example, monodisperse 50 nm

cubes of silver iodide can be formed with a high enough negative surface charge if the iodide concentration in solution has p$I \approx 4$. If the silver nitrate is in excess, stable, positively charged silver iodide particles are produced if pAg ≈ 4. The particles are single crystals and have a cubic shape with the occasional tetrahedron formed from twinned crystals.

The largest use of silver bromide nanoparticles is in the photographic industry as the particles decompose when exposed to electromagnetic radiation to give silver. The particles are dispersed in a gelatin gel which is used to coat either polymer film or glass plates. Sodium thiosulfate is used to dissolve unreacted silver bromide leaving the well-known negative image.

1.8.3.3 Iron oxide

Magnetite is often produced in nanocrystal form. The crystal contains iron in both of its oxidation states so that we may write the formula as $FeO{\cdot}Fe_2O_3$. To produce this material, a mixture of ferric and ferrous chlorides is prepared in water in a 3:2 mole ratio and ammonium hydroxide is added to precipitate the magnetite [7]. At this stage, the nanocrystals are heavily agglomerated. For an application such as ferrofluids, the particles need to be dispersed on an organic medium such as tetradecane or hexadecane. A strategy to achieve this is to add oleic acid (hexadecenoic acid) to the mixture and slowly heat with stirring to evaporate the water and ammonia. The dry powder can then be dispersed in the organic phase by ultrasonication, giving hydrophobic, stable nanocrystals with a diameter of $\sim$10 nm.

Small single crystals of magnetite are super paramagnetic, that is, they respond strongly to a magnetic field but show zero residual magnetism when that field is removed. Organic dispersions of magnetite nanoparticles have the potential to be used as ferrofluids. These are dispersions with a moderately high solids loading so that they flow with a viscosity which is only slightly higher than that of the continuous phase but, when subjected to a strong magnetic field, each particle becomes a small magnet and all the magnetic dipoles align with the field so that the particles stick together in a linear structure perpendicular to the flow direction, forming a solid or, at least, a strongly pseudoplastic material. As soon as the field is removed, the sol changes from a solid to a fluid again and flow can resume. The response time is very short, but the greatest problem to their use in units such as high-speed actuators is the mass of the electromagnets required for the high fields.

Magnetite nanoparticles have a strong potential for use in biological applications; however, a limitation is their sensitivity to the chemical environment. Core–shell particles consisting of an iron core with a gold shell can be used successfully as they are inert, biocompatible and can be surface functionalized. Example applications are magnetic resonance imaging (MRI) enhancement, cell sorting and targeted drug delivery [8–10]. Iron–platinum core–shell nanoparticles have a strong potential use in microelectronics as very high-density storage materials when self-assembled into arrays. This is due to the high degree of magnetocrystalline anisotropy of the particles.

1.8.3.4 Semiconductor nanocrystals

Nanocrystals made of semiconductors are known generically as quantum dots and show fluorescence with a high quantum yield. The fluorescence emission maximum is a function of the crystal size. The most common nanocrystals have a cadmium–selenide core with an outer shell of zinc sulfide. High temperatures are used to produce rapid nucleation of the crystals followed by a lower temperature growth stage. Originally pyrophoric organometallic materials were used for the core synthesis [11], but cadmium salts are now used in conjunction with an organic solvent that is stable at high temperature [12]. The colloid stability of the crystals is provided by the adsorption of trioctylphosphine oxide on the surface. After the crystals have been grown to the desired size, as determined from the fluorescence spectrum, the zinc sulfide shell is synthesized by the decomposition of a zinc organometallic material and a sulfur compound. Both spherical and anisometric nanocrystals can be produced, with the former being the easiest to produce with uniform shells. The photostability and quantum yields are a function of the shell quality and the narrow emission spectrum is governed by the degree of monodispersity.

Even in a crystal as small as $\sim$5 nm in diameter, there is of the order of 10^6 atoms and hence a large number of electrons in the valence orbitals so that there is a band with a continuum of energy levels. Electrons in crystalline solids behave as waves as they move through the solid. As with all semiconductors, the excited electrons 'jump' into a conductance band, which also consists of a continuum of energy levels. The width of these bands is, of course, a function of the number of electrons as there can only be a maximum occupancy of two at any energy level. The minimum energy difference between the highest level of the valence band and the lowest level of the conduction band, the 'band gap', is a function

of the particular types of atoms making up the semiconductor but also the size of the particle. We can model the electrons in the quantum dot as a 'particle in a box' problem so that the kinetic energy, E, is [13, 14]

$$E \propto \frac{\hbar \pi^2}{m^* L^2} \tag{1.22}$$

in which $\hbar$ is Planck's constant/2π, L is the length of the box, which in this case is the diameter of the quantum dot or the minor axis if the dot is asymmetric, and m^* is the effective mass of the electron in the crystalline solid (~8% of the free electron mass). With our box dimension in the size range 5–8 nm, our band gap energy is in the visible part of the electromagnetic spectrum. Note that it is higher for a 5 nm than a 10 nm crystal, so that when an electron drops from the bottom of the conduction band to the top of valence band we see emission in the blue and into the red with an 8 nm nanocrystal. The surface states of the crystal can act as surface traps so that we can lose electrons from the 'box'. To prevent this, a shell of a larger band gap semiconductor material is grown on the cadmium–selenide core crystal. Any electrons escaping from the core are funnelled down through the zinc sulfide conduction band and then move from the lowest energy in that band back into the cadmium–selenide conduction band. The cadmium–selenide band gap is the minimum downward jump. Figure 1.7 illustrates the band structure of a core–shell quantum dot. The higher the quality of the nanoengineering of the nanocrystal core–shell, the brighter and more photostable are the quantum dots.

Currently, the most widespread use of quantum dots is in the life sciences area, where they are being used in cellular imaging and as molecular tags for research and diagnostic purposes. This use demands that the dots are dispersible as single particles in aqueous buffers. To achieve this, the dots are coated with an amphiphilic polymer which provides colloidal stability to the particle in an aqueous environment and provides chemical functionality for covalently linking biomolecules such as antibodies and nucleic acid oligomers. The brightness of imaged cells is equal to or greater than that which can be achieved by organic dyes; however, the dots open up different and new possibilities because they can all be excited by a single wavelength source at the blue/UV end of the spectrum. (Organic dyes have excitation and emission wavelengths fairly close to each other.) This immediately introduces the possibility of multiplexing and producing images of different labelled parts of a cell without using composite images obtained with different filter sets.

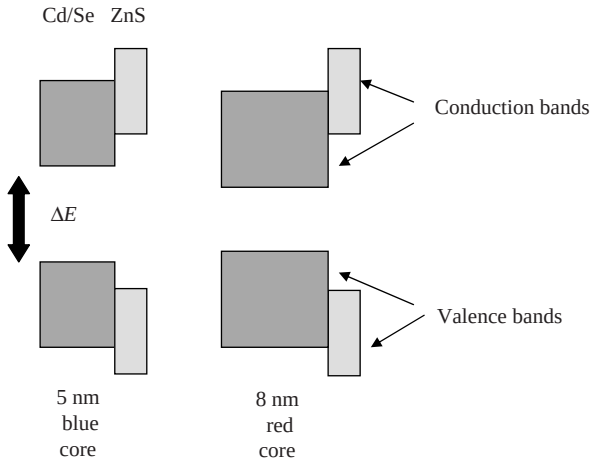


Figure 1.7 A representation of the electronic band gaps of zinc sulfide-coated cadmium–selenide nanocrystals showing how the band gap reduces with the size of the nanocrystal, which indicates that the emission moves towards the red. The larger band gap of the zinc sulfide shell serves to direct excitons back into the core

In well-engineered dots, such as Q-Dots®, the photostability is remarkably good and archiving of fluorescently stained samples without loss of definition is now easily attained.

In the non-life sciences areas, the dots have promise in electro-optical devices and displays. The possibility now exists of photovoltaic cells being manufactured by painting nanocrystals on to a metallized polyester film to make large solar panels. In the future, conducting polymers could replace the metallized film so that solar cells could be spray painted directly on to the roofs and walls of buildings. The greatest barrier to these applications is currently the ability to make large quantities of material with the core–shell quality that is available in the life sciences areas where the volume demand is much smaller.

1.8.3.5 Polymer nanoparticles

Polymer nanoparticles are synthesized from vinyl monomers using the technique of 'emulsion polymerization'. An extensive recent review is available in the book edited by van Herk [15]. Common vinyl monomers used to produce polymer nanoparticles are styrene and acrylate esters. Although many other vinyl monomers are polymerized using emulsion polymerization, the final particle size is larger than the nanosize range

of $<10^2$ nm. The technique is straightforward and consists of using an emulsion of a monomer with a low water solubility and a source of free radicals which initiates polymerization in the aqueous phase. Persulfate ions thermally decompose to two sulfate free radicals at a convenient rate at temperatures between 50 and 100 °C. The ionic free radical attacks the π-bond of the vinyl group attaching the initiator fragment to one of the carbon atoms, leaving the other carbon atom as the radical. The chain rapidly propagates as other dissolved monomer molecules come into contact with the radical. Nucleation occurs as the rapidly growing chains become insoluble and precipitate – this occurs after the addition of about five monomer units for styrene but is a function of the solubility of the monomer. The precipitation of a very large number of live polymer chains occurs at a very short time into the reaction. The nuclei adsorb monomer and will aggregate with each other as there is insufficient charge to make them colloidally stable. However, there is also a surfactant present at a concentration in excess of the cmc; sodium dodecyl sulfate is a common example, and this adsorbs on the growing nuclei and suppresses further aggregation so that most of the reaction is dominated by particle growth both by the adsorption of monomer into colloidally stable particles and by adsorption of newly precipitated nuclei. Again, the principle of a short nucleation phase followed by a long growth phase leads to the highest degree of monodispersity. About 20 nm is the smallest particle size attainable by this technique and, with the high nucleation rates required for this, some monodispersity has to be sacrificed and distribution becomes narrower as the size is increased towards the upper end of the nanosize region.

An alternative method to polymer nanoparticles is similar to the comminution process used for some solid systems described earlier in this section. The monomer phase containing an oil-soluble initiator is mixed with an aqueous phase containing a high level of surfactant. The surfactant decreases the interfacial tension markedly, but normal homogenizers will give a droplet size of around 1 μm; however, ultrasonics can deliver an emulsion with a narrow size distribution with a mean droplet diameter <100 nm. The curvature of the smaller droplets is more marked than that of the larger droplets, so the Laplace pressure (see Chapter 6) is higher, which drives the monomer phase through the aqueous phase and into the larger droplet fraction. This is 'Ostwald ripening' and will drive a broadening of the size distribution which, in most cases, is undesirable. The solution to this problem is to add a water-insoluble component to the monomer phase. It might seem that a good candidate would be some preformed polymer, but this is not the case. The reason is that as a droplet

shrinks, the osmotic pressure increases at a rate which is dependent on the mole fraction of solute for simple solutions but with the volume fraction in the case of polymer solutions (see Chapter 2). A mole fraction of $>1\%$ is required to provide a sufficiently high osmotic pressure and if a high molecular weight polymer were to be added at a volume fraction $>1\%$, the result would be a marked increase in the viscosity of the oil phases and concomitant changes in the emulsification dynamics. An extensive review of the preparation of polymer nanoparticles by the use of miniemulsion processes has been provided by Landfester [16].

Microemulsions are a thermodynamically stable nanoparticle heterophase system and can be used in polymerizing systems [17]. Microemulsion droplets are not static entities and as polymerization occurs drop sizes move to the upper end of the nanoparticle size range. A major advantage of the technique is that the molecular weights obtained are higher than by other heterophase techniques rather than in the size range produced.

1.8.3.6 Swelling clay particles

These types of clay minerals are formed from a 2:1 layer lattice structure and montmorillonite is typical of this class. The clay structure consists of an alumina layer located between to silica layers. The silicon and aluminium atoms are linked via oxygen atoms. Isomorphous substitution within the layers results in a net negative charge which is balanced by cations located between the trilayer sheets. When the cations are monovalent, the trilayer structure can expand as water is drawn in between the sheets by the osmotic pressure difference between the outer medium and that inside the clay particles. The individual sheets are 1.1 nm in thickness; however, the length and breadth are of the order of microns, hence such particles are not usually described as nanoparticles but should included here as some of their properties, for example specific surface area, are characteristic of nanoscale materials.

Clearly in the class of nanoparticles is the synthetic hectorite clay laponite. This clay can be crystallized from a solution of salts containing the correct ion concentrations. At moderate temperatures, crystals are produced which disperse in stacks of three or four sheets. Small-angle X-ray measurements show that the stack dimensions are of thickness ~ 3.5 nm and diameter ~ 40 nm [18]. Both lithium and fluoride atoms are present in the clay sheets and the stack structure is thick enough that the edge shows a positive charge (controlled by the fluoride level)

and the face has a negative charge (controlled by the lithium level). Laponite is used commercially as a transparent rheology modifier and in specialist paper coatings and is showing some promise as a medium for the synthesis of silver nanoparticles by photoreduction [19].

1.9 NANOCOMPOSITES

A huge range of objects that are in use today are manufactured from or are coated with polymers. It is also common practice to modify the polymer properties by the addition of a filler material. For example, carbon black is added to car tyres to make the elastomer more abrasion resistant, clays to solid polymers to increase the hardness and titanium dioxide to paints to enhance the optical properties. In many cases the particulates are in, or are bordering on, the nanosize range. The drive now is to push the size of the disperse phase down in order to enhance the properties further.

The key feature of the fillers is not simply the loading but what is occurring at the filler–bulk interface. Good adhesion is required otherwise weak points are added, from which stress cracks can propagate. If there is good adhesion, the inclusions can be effective in stopping cracks propagating and, clearly, as the particle size is reduced for a given disperse phase loading, the number density increases and each particle shares in shutting off crack propagation. It is even possible to toughen polymers by inducing small-sized inclusions from phase separation of an incompatible polymer. At high loadings of nanoparticles, the polymer is dominated by the properties of the polymer in the interfacial region as this is now where most of the polymer resides. Hence the way in which the polymer is arranged at the interface can, for example, align chains and increase the net van der Waals forces and enhance mechanical properties such as tensile strength. We will now consider a few examples of nanocomposites which make use of the special properties that are the result of at least one component being in the nanoscale size range.

1.9.1 Polymer–Clay Nanocomposites

Sodium montmorillonite is the most common clay to be used in this role. In order for these composites to be successful, the clay needs to be very well dispersed so that most of the high specific surface area is in contact with polymer. The isomorphous replacement of some of the

trivalent aluminium ions by divalent magnesium ions in the octahedral layer results in a net negative charge balanced by cations sandwiched between the tetrahedral silicate layers of adjacent sheets. When these cations are exchanged for cationic surfactant molecules, an oleophilic clay is formed that can be finely dispersed in a polymer matrix with the oleophilic chains acting as steric stabilizing moieties. Single-, di- or tri-chain cationics can be used with the head consisting of a quaternary ammonium group. If high-temperature applications are contemplated, such as in some automobile components, an imidazole head group will be more stable as it is not subject to catalytic degradation at the clay surface at temperatures in excess of 200 °C [20]. The book edited by Pinnavaia and Beal provides a wide range of examples of applications for a variety of polymer matrices [21].

1.9.2 Quantum Dot–Polymer Nanocomposites

We discussed in Section 1.8.3.5 how polymer nanoparticles can be prepared, but putting a secondary disperse phase inside such particles at a high loading and with a high degree of dispersion can be challenging. Semiconductor nanocrystals or quantum dots as initially synthesized are hydrophobic and sterically stabilized with strongly bound surfactants on their surfaces. At this stage, they are easily dispersed in a wide range of organic solvents. One strategy for producing a composite is to use cross-linked polymer nanoparticles in a swelling solvent and mix with the organic quantum dot dispersion. Bradley *et al.* showed that when the very heavily swollen particles had become highly expanded microgels, it was possible to obtain a large number of quantum dots dispersed throughout the gels [22]. Deswelling such gels and transferring the particles back into an aqueous phase, although possible, is not very easy, especially if the polymer particles are small. An alternative is to disperse the particles in a monomer first and then to use the miniemulsion route [16] to prepare nanoparticles containing quantum dots by using ultrasonics, an oil-soluble initiator and the correct surfactant. It is worth noting that other nanoparticle materials such as carbon and titania can be prepared as polymer nanocomposites by this route.

Quantum dots inside a polymer matrix are extremely well protected from oxidants that can cause quenching, so their optical behaviour is extremely robust. An interesting application of these materials was described by Fleischhaker and Zentel [23], where nanocomposites of quantum dots in the cores of core–shell polystyrene–poly(methyl

methacrylate) nanospheres were then taken to a high total solids content. Monodisperse charged stabilized nanospheres maximize their interparticle separation due to the long-range repulsion and, when concentrated, can form cubic colloidal crystals. Although body-centred cubic structures can be formed, face-centred cubic is the most common form. After this stage was reached, further drying produced a regular lattice structure with quantum dot-rich regions precisely spaced on a three-dimensional lattice. The self-assembled structure produced photonic crystals with photonic stop bands which modified the photoluminescence of the quantum dots, so leading the way to novel optoelectronic devices.

1.9.3 Polymer–Cement Nanocomposites

In the initial state, Portland cement is ground into the microparticle size range with a specific surface area of $\sim 0.5\,m^2\,g^{-1}$. When setting occurs, calcium and aluminosilicates start to crystallize. Water-soluble polymers such as poly(ethylene glycol) can be mixed in at moderate to high concentrations and, if the system is subjected to ultrasonics, very fine structures can be formed such that the resulting set composite has hardness and strength greater than those of steel, although it is brittle. It is possible to form thin, light structures such as gutter or springs from these materials. It is surprising that these fine structure composites have been ignored industrially whereas much coarser composites using polymer particles have wide use.

1.9.4 Polymer–Iron Nanocomposites

Polymer microspheres with embedded magnetic iron oxide are now a well-established system for the separation of biological materials such as cells and proteins and are available commercially in a format that lends itself to automation. Ugelstad *et al.* [24] pioneered the methodology of building the nanocomposite *in situ* by using porous polystyrene particles with nitro functionality in the pores, which converted iron salts to a magnetic oxide in nanocrystalline form. The resulting paramagnetic particles are readily functionalized with antibodies and can then be used to attach specifically to particular cells that can subsequently be separated from a matrix [25]. The iron content of these particles can be as high as 35% and the small domain size of the iron oxide ensures excellent paramagnetic behaviour. A high iron loading and large particles

(i.e. diameter >500 nm) make separation very effective even in flowing systems. Ramírez and Landfester used a miniemulsion process to encapsulate magnetite in polymer particles with diameters in the nanosize range [26]. Particles as small as 40 nm containing 30% magnetite were obtained.

1.9.5 Polymerized Pickering Emulsions

A 'Pickering emulsion' is an emulsion in which the droplets are stabilized by particles residing at the liquid–liquid interface. When the drops consist of a vinyl monomer, they may be polymerized by the inclusion of a thermally activated oil-soluble initiator. Silica nanoparticles have been used as the stabilizer for polystyrene particles [27]. To aid the stabilizing behaviour, 4-vinylpyridine was added as a minor comonomer. The nitrogen group can protonate to form a pyridinium cation, which adsorbs strongly to the negatively charged silica surface resulting in stable 'armoured' polymer nanocomposite particles with diameters of ~150 nm. Titania nanoparticles have been used to stabilize polystyrene particles [28]. In this case, the polymer controls the crystal form and the subsequent catalytic properties of the titania nanoparticles.

1.10 JANUS PARTICLES

The term 'Janus particle' is used to describe particles which have opposing faces with different physico-chemical properties such as hydrophobicity/hydrophilicity or positive/negative charge. The silica nanoparticles producing the Pickering emulsion described in Section 1.9.5 is a good example as the half of a particle in the oil is rendered hydrophobic by the interaction with vinylpyridine whereas the other half of the particle in the water is hydrophilic. Naturally occurring particles, such as kaolinite, which has negatively charge basal surfaces and positively charged edges so that an open aggregate structure resembling a 'house of cards' is formed, begin to indicate some of the types of colloidal behaviour that can be anticipated. However, these examples of split surface behaviour only serve to encourage us to look more closely into how we might engineer biphasic particles of different forms. We already know that amphiphilic molecules show important behaviour at interfaces and so we are immediately led to ask 'how do biphasic nanoparticles behave at interfaces?'. Large particles can be surface modified so that only part

of the surface has been affected, but we shall not discuss these here as the nanoscale particles appear to have more promise in applications.

Currently, the most studied form of Janus nanospheres is synthesized from terpolymers with a well-defined triblock structure. The centre block is a cross-linkable moiety such as polybutadiene or polyisoprene and the end blocks are made up of linear polymers with different degrees of hydrophilicity. These polymers can self-assemble into micellar units; the cores are then cross-linked to give insoluble cores and the nanoparticles are dispersed with ultrasonics [29]. The nanoparticles have a 'corona' of polymer with different character in each hemisphere. Casting the terpolymer as a film on to a flat substrate allows the polymers to self-assemble into layers. After cross-linking of the butadiene layers, they were swollen with solvent and then sonicated to give discs with the two basal surfaces having different degrees of hydrophilicity [30]. Cylinders have also been prepared from the same terpolymer so that the Janus character is of two hemi-cylinders connected by a cross-linked core [31].

The amphiphilic nature of Janus particle makes them good candidates for study at oil–water interfaces [32] and air–water interfaces [33]. The adsorption energy of a particle at an interface is made up of the sum of the contributions from all the moieties in each phase and so the displacement from the interface is much more difficult than it would be for a single amphiphilic molecule. As a result, we can expect good stabilization of the interface as shown by high emulsion or foam stability, and the particles would not be good emulsifiers as their mobility to an interface is slower than that of small amphiphilic molecules.

1.11 SUMMARY

This introductory chapter has defined what we mean by colloidal systems and has illustrated how widely different systems can fit into this form of matter. The related systems of surface-active molecules and macromolecules have also been introduced and it has been shown how they are intimate adjuncts to colloidal dispersions. A few common systems have been described, which, although they appear to be widely disparate, have some basic or generic aspects. These will be a focus of this book and will show why the subject has a marked interdisciplinary flavour.

Nanoparticles have been strongly featured in this introductory chapter. This is a relatively new and fast-moving area of colloid science and the fascination in this area lies in the combination of the range of the colloid forces being the same as or larger than the particle size and the

fact that a very large fraction of the atoms making up these small particles are located at an interface. The latter changes the interatomic forces compared with atoms in bulk material and we can see new physical properties for the materials in this very finely divided form.

REFERENCES

1. Everett, D. H. (1988) *Basic Principles of Colloid Science*, Royal Society of Chemistry, London.
2. Navrotsky, A. (2003) *Geochem. Trans.*, **4**, 34.
3. Mazeina, L. and Navrotsky, A. (2005) *Clays Clay Miner.*, **53**, 113.
4. Li, W. Z., Xie, S. S., Qian, L. X., Chang, B. H., Zhou, B. S., Zhou, W. Y., Zhao, R. A. and Wang, G. (1996) *Science*, **274**, 1701.
5. Kocabas, C., Shim, M. and Rogers, J. A. (2006) *J. Am. Chem. Soc.*, **128**, 4540.
6. Handley, D. A. (1989) Methods for the synthesis of colloidal gold, in *Colloidal Gold*, Vol. **1** (ed. M. A. Hayat), Academic Press, New York, pp. 13–32.
7. Reimers, G. W. and Kalafalla, S. E. (1974) US Patent 3 843 540.
8. Cho, S. J., Jarret, B. R., Louie, A. Y. and Kauzlarich, S. M. (2006) *Nanotechnology*, **17**, 640.
9. Niemeyer, C. M. (2001) *Angew. Chem. Int. Ed.*, **40**, 4128.
10. Mornet, S., Vasseur, S., Grasset, F. and Duget, E. (2004) *J. Mater. Chem.*, **14**, 2161.
11. Murray, C. B., Morris, D. J. and Bawendi, M. G. (1993) *J. Am. Chem. Soc.*, **115**, 8706.
12. Peng, Z. A. and Peng, X. G. (2001) *J. Am. Chem. Soc.*, **123**, 183.
13. Kouwenhoven, L. P. and McEuen, P. L. (1999) Single electron transport through a quantum dot, in *Nanotechnology* (ed. G. Timp), Springer, New York, pp. 471–536.
14. Brus, L. (1999) Chemical approaches to semiconductor nanocrystals and materials, in *Nanotechnology* (ed. G. Timp), Springer, New York, pp. 257–283.
15. van Herk, A. M. (ed.) (2005) *Chemistry and Technology of Emulsion Polymerisation*, Wiley-Blackwell, Oxford.
16. Landfester, K. (2006) *Annu. Rev. Mater. Res.*, **36**, 231.
17. Candau, F. (1992) Polymerization in microemulsions, in *Polymerization in Organized Media* (ed. C. M. Paleos), Gordon & Breach Science Publishers, Philadelphia, pp. 215–282.
18. Saunders, J. M., Goodwin, J. W., Richardson, R. M. and Vincent, B. (1999) *J. Phys. Chem. B*, **103**, 9211.
19. Yang, Y. and Huang, H. (2007) *Key Eng. Mater.*, **334**, 813.
20. Xie, W., Gao, Z., Pan, W. P., Hunter, D., Singh, A. and Vaia, R. (2001) *Chem. Mater.* **13**, 2979.
21. Pinnavaia, T. J. and Beal, G. W. (eds) (2000) *Polymer–Clay Nanocomposites*, John Wiley and Sons, Ltd, Chichester.
22. Bradley, M., Bruno, N. and Vincent, B. (2005) *Langmuir*, **21**, 2750.
23. Fleischhaker, F. and Zentel, R. (2005) *Chem. Mater.*, **17**, 1346.
24. Ugelstad, J., Berge, A., Ellingson, T., Schmid, R., Nilsen, T.-N., Mørk, P. C., Stenstad, P., Hornes, E. and Olsvik, Ø. (1992) *Prog. Polym. Sci.*, **17**, 87.

25. Ugelstad, J., Stenstad, P., Kilaas, L., Prestvik, W. S., Herje, R., Berge, A. and Hornes, E. (1993) *Blood Purif.* **11**, 349.
26. Ramírez, L. P. and Landfester, K. (2003) *Macromol. Chem. Phys.*, **204**, 22.
27. Tiarks, F., Landfester, K. and Antonietti, M. (2001) *Langmuir*, **17**, 5775.
28. Zhang, M., Gao, G., Zhao, D. C., Li, Z. Y. and Lui, F. Q. (2005) *J. Phys Chem. B*, **109**, 9411.
29. Gohy, J.-F. (2005) Block copolymer micelles, in *Block Copolymers II* (ed. V. Abetz), Springer, New York, pp. 65–136.
30. Walther, A., André, X., Dreschler, M., Abetz, V. and Müller, A. H. E. (2007) *J. Am. Chem. Soc.* **129**, 6187.
31. Lui, Y., Abetz, V. and Müller, A. H. E. (2003) *Macromolecules*, **36**, 7894.
32. Binks, B. P. and Fletcher, P. D. I. (2001) *Langmuir*, **17**, 4708.
33. Xui, H., Rainer, E., Abetz, V., Müller, A. H. E. and Goedel, W. A. (2001) *Langmuir*, **17**, 6787.

2

Macromolecules and Surfactants

2.1 INTRODUCTION

Macromolecules and surfactants are added to nearly all formulations that include colloidal particles. Macromolecules may be added to control the rheological behaviour and also to control the stability of the dispersion. Surfactants can be used as wetting agents to disperse dry powders, as emulsifiers for oil-containing formulations and to provide stability against aggregation, in addition to being added as cleaning agents. Indeed, these molecules are so important to colloidal systems that we need to discuss their solution properties prior to the discussion of colloidal particles.

2.2 MACROMOLECULAR DEFINITIONS

The term macromolecules is used here to include synthetic polymers such as poly(ethylene oxide), naturally occurring macromolecules such as proteins like gelatin or polysaccharides like ethylhydroxycellulose or oligomers such as cyclodextrin. In each case, the monomer or building block of the macromolecule is a small molecule. With synthetic polymers, the chains are often hydrocarbons with side groups which give the correct properties. For example, poly(acrylic acid) is water soluble because of the polar carboxylic acid group on every second carbon atom along the chain, whereas polystyrene is soluble in aromatic hydrocarbons and

Colloids and Interfaces with Surfactants and Polymers, Second Edition James Goodwin

not water as there is a benzene ring attached to every second carbon atom. Sugar rings – glucose is a common example – are the monomeric units of the polysaccharides, whereas amino acids are the building blocks of proteins. We use the term 'oligomer' to indicate that there is only a small number of monomeric units that are linked – maybe 10 or 20, for example. Some chains are simple linear molecules whereas others may be branched. This branching is taken to an extreme with *dendrimers*, where each chain branch branches again and again to give a large, approximately spherical unit. Chains can be cross-linked and, if this is achieved by covalent bonds, *elastomers* are produced in which the rigidity is a function of the density of cross-links. Uses of these materials range from synthetic rubber to ion-exchange resins. In the latter case, the cross-link density is so high that the polymer matrix has sufficient rigidity to maintain an open-pore system even under high applied pressures. Also, covalently cross-linked material can be prepared as small, swollen gel particles and these are termed *microgels*. If the chains are lightly cross-linked by placing a small amount of insoluble species on the chains, we have a weak self-assembly and we have synthesized an 'associative thickener' of the type that is becoming common in decorative paints. With a higher level of self-association, we can produce highly swellable gels such as the 'super absorbers' which are used in disposable diapers, for example.

If there are N segments bonded together in one chain (i.e. the degree of polymerization is N) and the molar mass of each segment (monomer) is M_m, the molecular weight of that chain is given by the following:

$$M = NM_m \tag{2.1}$$

During the polymerization process, a distribution of chain lengths is always produced [1]. Usually, the distribution is broad, but some ionic-initiated polymerizations can be controlled to give a narrower distribution than, for example, a free radical-initiated polymerization of vinyl monomer. Hence we need to define the various kinds of average molecular weight:

		Measured by :
Number average	$M_n = \sum_i \frac{M_i n_i}{n_i}$	Osmotic pressure
Viscosity average	M_v	Viscometry
Weight average	$M_w = \sum_i \frac{M_i w_i}{w_i} = \sum_i \frac{M_i^2 n_i}{M_i n_i}$	Light scattering
Z-average	$M_z = \sum_i \frac{M_i^3 n_i}{M_i^2 n_i}$	Ultracentrifugation

(2.2)

These averages increase in the order $M_n < M_v < M_w < M_z$, so it is important to define which method has been used to determine the molecular weight. Although the detailed distribution is often not known in detail, the width of the distribution is often characterized by the polydispersity, P, which is defined in terms of two of the commonly measured averages:

$$P = \frac{M_w}{M_n} \tag{2.3}$$

Polymer molecular weight standards, used for calibrating equipment, for example, would have a value of $P < 1.1$, but polymers for bulk usage usually have a polydispersity with a value of 3 or more. It should be noted though that even a value of $P = 1.1$ represents a fairly broad distribution and will, of course also depend on the details of the 'skew' in that distribution. In the practical usage of polymers, for example as thickeners, this wide distribution can be useful as the changes in viscosity with rate of shearing of the system is slower if the distribution is broad. Any property which is dependent on the diffusive motion of the components will be affected similarly.

2.3 CONFORMATION IN DILUTE SOLUTIONS

The books by Flory [2, 3] present the classical description of the solution properties of polymers in dilute solution, and other important texts include those by Yamakawa [4], deGennes [5] and Doi and Edwards [6]. The starting point for the description of the conformation of a large polymer molecule in solution is to use the statistics of a three-dimensional random walk. At this stage, the problem is simpler than a description of random motion, such as diffusion, because the step sizes are equal as each step has a dimension equal to that of the monomer unit in the chain. By considering each bond as a vector and summing the squares, the mean square distance between the starting point of the chain can be calculated, so that

$$\langle r_c^2 \rangle = Nl^2 \tag{2.4}$$

where l is the segment length. (Note that by squaring the vectors, we eliminate the problem of the directional sign convention that is inherent in the Cartesian coordinate system, otherwise the vectors indicating a negative direction would cancel those indicating a positive direction.)

Equation 2.4 describes a *freely jointed chain* and no account has been taken of finite bond angles or the excluded volume interactions of both

Table 2.1 Values of the characteristic ratios for various polymers

Polymer	c_∞
Freely jointed chain	1
Tetrahedral bond angle	2
Poly(ethylene oxide)	4.1
Polydimethylsiloxane	5.2
Poly(12-hydroxystearic acid)	6.1
Polystyrene	9.5

neighbouring segments and distant segments along the chain that interact as the 'walk' takes them back to cross the chain. For any real polymer chain, there are fixed bond angles and rotation around bonds is markedly reduced if bulky side groups are present, so the 'walk' is much more spatially extended. In other words, a real chain is much stiffer than a freely jointed chain and the conformation is expanded, with the mean square end-to-end distance can be expressed as follows:

$$\langle r_c^2 \rangle = c_\infty N l^2 \tag{2.5}$$

were c_∞ is the 'characteristic ratio' and some typical values are given in Table 2.1.

The solvent chosen to dissolve the polymer is also important. In a *good* solvent, a chain segment is surrounded by the maximum number of solvent molecules that can be packed around it. In a *poor* solvent, there is an increased probability of there being other chain segments around any particular segment. (Thus, as the quality of the solvent decreases, the polymer chains become insoluble.) The conformation of a polymer chain in solution is a dynamic, slightly elliptical, spheroidal coil which can be characterized by the root mean square end-to-end chain dimension, as follows:

$$\begin{aligned} \langle r_c^2 \rangle^{0.5} &= (c_\infty N)^{0.5} l \quad \text{Ideal solution} \\ \langle r_c^2 \rangle^{0.5} &= (c_\infty N)^{0.6} l \quad \text{Good solvent} \end{aligned} \tag{2.6}$$

We may also characterize the coil dimension in terms of it radius of gyration, R_g. This is the average distance of the polymer segments from the centre of mass of the coil:

$$R_g^2 = \frac{\sum_i m_i r_i^2}{\sum_i m_i} = \frac{\langle r_c^2 \rangle}{6} \tag{2.7}$$

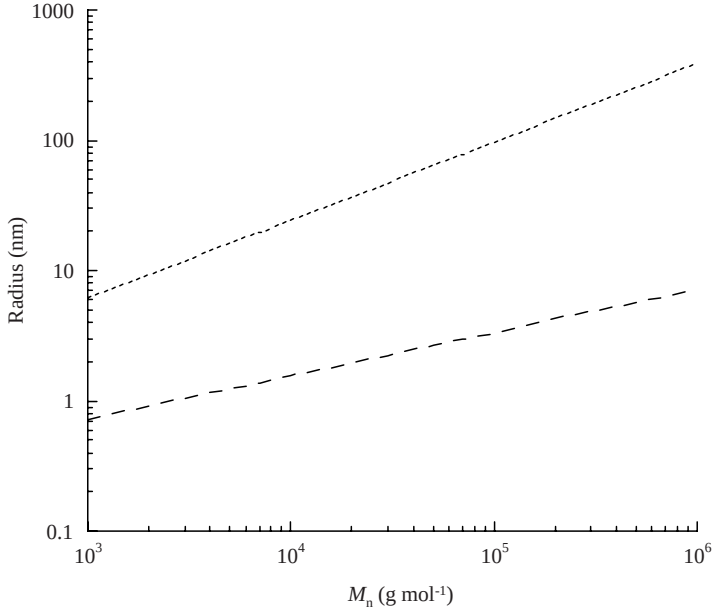


Figure 2.1 The radius of a polystyrene molecule as a function of the (number-average) molecular weight. Dotted line, radius of gyration in a good solvent; dashed line, radius of the equivalent amorphous sphere

The radius of gyration can be obtained experimentally from light scattering measurements of dilute polymer solutions. It is interesting to compare the dimension of a polymer molecule dissolved in a good solvent with what we would expect from the bulk density of the same polymer (Figure 2.1). We can see from this figure that the dimensions of the molecule in solution are very much greater than they would be in the dry state and so the concentration of polymer within the coil in solution is very low indeed. It is important to keep in mind that the 'connectivity' along the chain demands this very open structure.

2.3.1 The Gaussian Chain

At first sight, it may appear that the model of a freely-jointed chain would not be a good picture of a real polymer with rigid bond angles, even though an expansion factor has been included. However, if groups of several bonds are considered, the cooperative effect is to add flexibility. Hence the artifice is to consider the chain segments as a larger unit

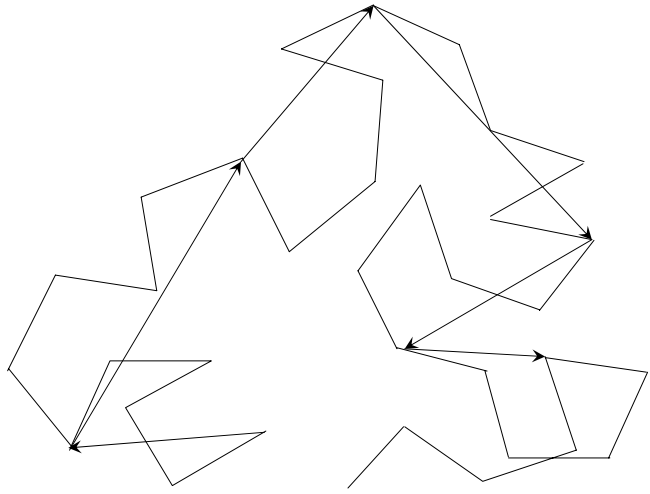


Figure 2.2 Schematic of a three-dimensional random walk with fixed bond angles and fixed step lengths. The arrows illustrate the increased flexibility introduced by designating sections containing five bonds as one 'chain segment' at the cost of variable step lengths

which could contain, for example, five bonds. This allows the flexibility to be reintroduced. The complication is that the bond length is now a variable so our random walk no longer has the constraint of equal step lengths. This is now closer to the diffusion problem and the result is that there is a Gaussian distribution of step lengths. Figure 2.2 is a schematic illustration of part of a chain. The number of segments in the example is $N/5$ with the mean step length from the Gaussian distribution as l' and $l'/\left(l \times 5^{0.6}\right)$ is taken care of in the value of c_∞ that we use.

2.4 THE FLORY–HUGGINS THEORY OF POLYMER SOLUTIONS

This theory [2, 3, 7] describes the thermodynamics of polymer solutions in sufficient detail for the purposes of this text. The starting point is a simple lattice model. A cubic lattice is normally used as this is easy to visualize and there is no real gain in understanding from using other lattice structures such as a tetrahedral one. The key feature of this lattice is that the solvent molecules are assumed to be the same size as the segments of the polymer chain. The entropy of mixing is estimated from

the number of possible configurations on the lattice and the enthalpy from the interactions between the various components.

2.4.1 The Entropy of Mixing

The entropy of mixing is calculated from the number of ways in which a polymer molecule can occupy the sites on a filled lattice. Sites not occupied by chain segments must be occupied by solvent molecules. Hence there are n_s lattice sites occupied by solvent molecules with n_p sites occupied by polymer chains (illustrated in Figure 2.3). The result is as follows:

$$\Delta S_{mix} = -k_B \left(n_s \ln \varphi_s + n_p \ln \varphi_p \right) \tag{2.8}$$

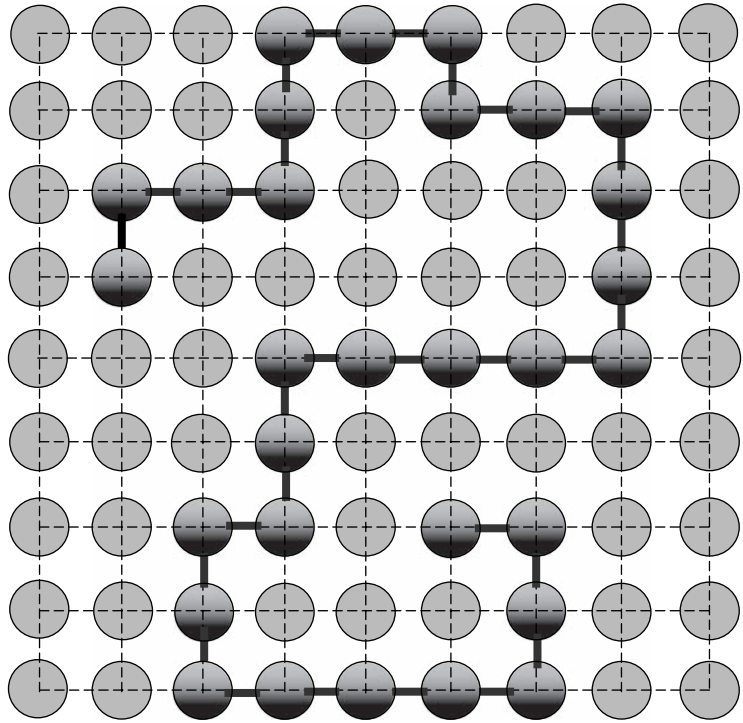


Figure 2.3 Illustration of a filled cubic lattice with solvent molecules surrounding a polymer chain. Note that the lattice must be fully occupied and the solvent molecular size is equal to the chain segment size

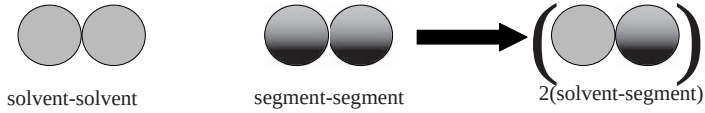


Figure 2.4 Interactions occurring on mixing polymer segments with solvent

The volume fractions, φ, are given by the following:

$$\begin{aligned} \varphi_s &= \frac{n_s}{n_s + Nn_p}; \quad \varphi_p = \frac{Nn_p}{n_s + Nn_p}; \\ \frac{\varphi_s}{\varphi_p} &= \frac{n_s}{Nn_p} \quad \text{and so} \quad n_s\varphi_p = Nn_p\varphi_s \end{aligned} \tag{2.9}$$

This is analogous to the problem of mixing two liquids except that volume fractions have been used instead of mole fractions. The mole fraction of a high molecular weight polymer in solution would always be very small, even when the volume fraction is large, due to that high molecular weight. [If the segments of the chain were not connected, we would just be mixing two miscible liquids and $\Delta S_{mix} = -k_B(n_1 \ln x_2 + n_2 \ln x_1)$, where x is the mole fraction of either the solvent, 1, or the solute, 2.]

2.4.2 The Enthalpy of Mixing

Flory [2, 3, 7] calculated this by considering the local energy changes as we mix polymer segments with solvent, as shown in Figure 2.4. Therefore, the change in the internal energy is given by

$$\Delta U_{sp} = U_{sp} - \frac{U_{ss} - U_{pp}}{2} \tag{2.10}$$

and the enthalpy of mixing by

$$\begin{aligned} \Delta H_{mix} &= \Delta U_{mix} + P\Delta V \\ \Delta H_{mix} &= N_c \Delta U_{sp} + \sim 0 \end{aligned} \tag{2.11}$$

The PV term can be approximated to zero, and the number of contacts per unit volume, N_c, is as follows:

$$N_c = \varphi_s n_p N z \tag{2.12}$$

that is, the number of contacts is equal to the probability of a site being occupied by a solvent molecule, φ_s, multiplied by the number of polymer segments, $n_p N$, and a coordination number for the lattice contacts, z. Therefore, the enthalpy of mixing becomes (with Equation 2.9)

$$\Delta H_{mix} = \varphi_p n_s \chi k_B T \tag{2.13}$$

where we have defined χ, which is known as the *Flory–Huggins interaction parameter*, as the internal energy change per segment on mixing relative to thermal energy as

$$\chi = \frac{\Delta U_{sp} z}{k_B T} \tag{2.14}$$

We may now write the free energy of mixing as follows:

$$\begin{aligned} \Delta G_{mix} &= \Delta H_{mix} - T\Delta S_{mix} \\ \Delta G_{mix} &= k_B T \left(n_s \ln \varphi_s + n_p \ln \varphi_p + n_s \varphi_p \chi \right) \\ \Delta G_{mix} &= RT \left(N_s \ln \varphi_s + N_p \ln \varphi_p + N_s \varphi_p \chi \right) \end{aligned} \tag{2.15}$$

where N_s and so on are in molar quantities. Formally, Equation 2.10 should have used a free energy component, so χ really contains an entropy term; however, as it is determined experimentally, we do not generate practical problems by this approximation. For a polymer in a good solvent, the value of χ is found to be between 0.5 and 0.1. Now that we have the free energy of mixing as a function of the solution concentration, we may calculate the osmotic pressure of the dilute polymer solution [2]:

$$\Pi = \frac{RT}{\bar{v}_s} \left[\frac{\varphi_p}{N} + \left(\frac{1}{2} - \chi \right) {\varphi_p}^2 + \cdots \right] \tag{2.16}$$

where $\bar{v}_s$ is the molar volume of a solvent molecule. The significance of the χ parameter can be immediately appreciated from Equation 2.16. When we have the condition that $\chi = 0.5$, the polymer–polymer interaction term vanishes from the osmotic pressure for the dilute solution and the osmotic pressure relationship is similar to the van't Hoff expression. Using the relationships from the lattice model, Equation 2.16 can be

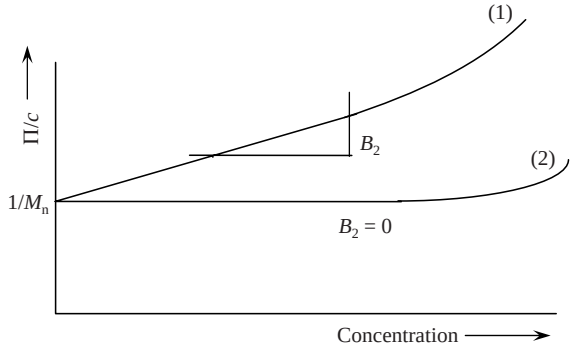


Figure 2.5 Schematic curves for the reduced osmotic pressure as a function of concentration for the conditions (1) $\chi < 0.5$ and (2) $\chi = 0.5$

recast in the more familiar form

$$\frac{\Pi}{c_p} = RT\left[\frac{1}{M_n} + \left(\frac{1}{2} - \chi\right)\left(\frac{\bar{v}_p N}{M_n{}^2}\right) c_p + \cdots\right] \tag{2.17}$$

where c_p is the polymer concentration in mass per unit volume and $\bar{v}_p$ the molar volume of the polymer. The coefficient of the polymer concentration on the right-hand side of Equation 2.17 is the osmotic second virial coefficient, B_2, and is the slope of the curve of the reduced osmotic pressure as a function of concentration, as illustrated in Figure 2.5.

As the solvent properties are changed, by changing the temperature, pressure or composition, for example, the value of χ changes and the 'quality' of the solvent can be defined as follows:

1. $\chi < 0.5$ – we have a 'good' solvent for the polymer.
2. $\chi \approx 0.5$ – the solvent is termed a θ-solvent.
3. $\chi > 0.5$ – the solvent is a 'poor' solvent and as the value increases much above 0.5, the polymer solubility reduces even though it may be swollen by the solvent.

For example, polystyrene is soluble in cyclohexane. The θ-temperature is 38.5 °C, so at 45° cyclohexane is a good solvent for polystyrene. At the θ-temperature, the conformation of the polymer molecule is minimally disturbed by solvent–chain segment interactions and is as close to a random coil as obtainable by that molecule.

2.5 POLYMER SOLUTION PHASE BEHAVIOUR

When we consider the solubility of a solute in a solvent, our normal experience is for the solubility to increase as the temperature is increased and, conversely, if we cool a solution, at some temperature we will observe the solute phase coming out of solution. This is the usual pattern with polymers in a good solvent. When the system is cooled below the θ-temperature, the solvent becomes progressively poorer and two phases will be observed with the polymer-rich phase being polymer swollen with solvent. The phase boundary is known as the *upper consolute solution temperature* (UCST). Above this temperature, a single phase is formed. In many aqueous systems and occasionally in some polar organic solutions of polymer, another phase boundary – at the *lower consolute solution temperature* (LCST) – can be found where phase separation can occur as the solution is heated. Water-soluble polymers contain polar groups such as hydroxyl, carboxylic acid or ether groups which can take part in the hydrogen (H)-bonding structure of water. As the temperature increases, the H-bonding is reduced, the polymer ceases to be in a good solvent and phase separation can occur. A general solubility diagram is sketched in Figure 2.6. This type of behaviour is also observed with non-ionic surfactants in aqueous solution, with the LCST being termed the *cloud point*.

However, because of the connectivity of the groups making up a high polymer, there are further subtleties in the phase behaviour of polymer solutions. Consider a polymer in a θ-solvent. In dilute solution, the

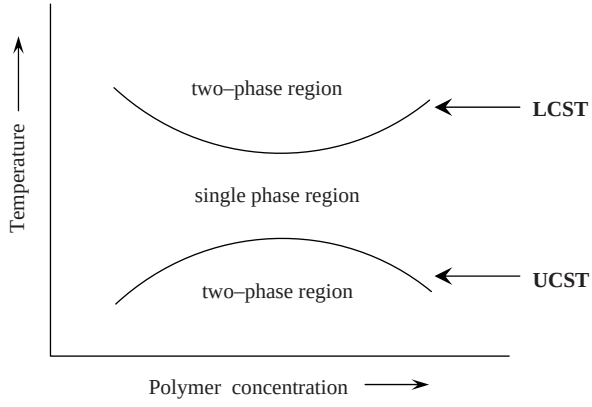


Figure 2.6 Illustration of a general solubility diagram. LCST, lower consolute solution temperature; UCST, upper consolute solution temperature

polymer is as close to a random coil configuration as it can attain. As the concentration is increased, the polymer molecules interpenetrate extensively as the interactions between polymer segments, polymer solvent and solvent molecules have similar energy in a 'θ-condition'. When the polymer is in a 'better than θ-solvent', the situation changes somewhat. At low concentrations (the dilute regime illustrated in Figure 2.7a), the polymer coils are in an expanded configuration and, on average, are separated from each other. Hence, if we were to measure the concentration profile across a section of solution, there would be clear variations, as illustrated schematically in Figure 2.7b. As the concentration increases to a value denoted by $c*$, the polymer coils become 'space-filling'. This is the threshold of the 'semi-dilute' state. The global polymer concentration is just equal to that which would be calculated

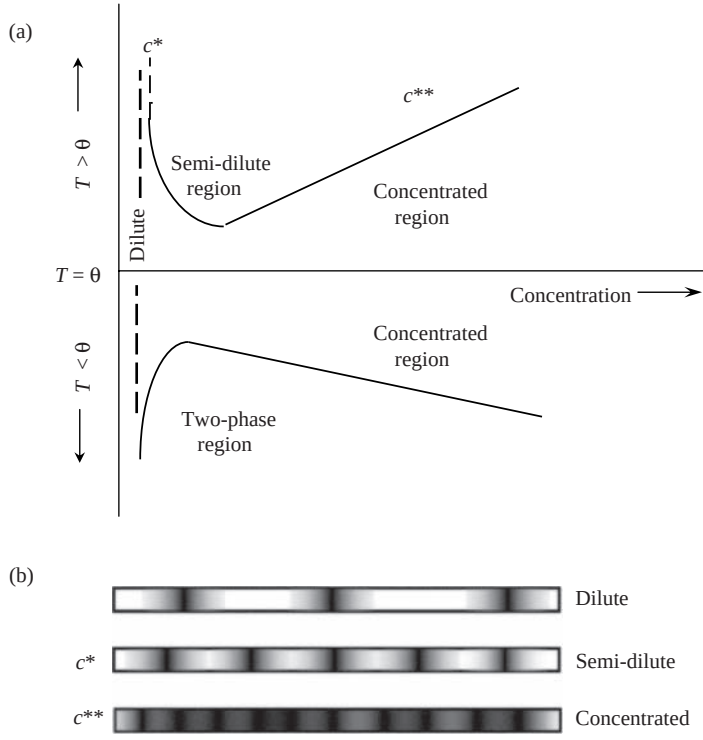


Figure 2.7 Schematics of (a) a polymer solution phase diagram and (b) the concentration variation across a polymer solution, showing the boundary from dilute to semi-dilute as the coils 'space-fill' at $c*$ and the approach to a uniform concentration at c^{**}

for a single coil, so by using Equation 2.6 we have the relationship

$$c^* \sim \frac{M_n}{N_A \langle r_c \rangle^{\frac{3}{2}}} = \frac{M_n}{N_A (c_\infty N)^{\frac{9}{5}} l^3} \tag{2.18}$$

As the solution concentration increases, the variation in local concentration becomes small as interpenetration increases and the polymer solution is said to be 'concentrated'. This is illustrated in the lowest part of Figure 2.7b. (c^{**} is the concentration where the individual coils are longer discernible and the chains are in their 'ideal' state). This boundary moves to higher concentrations as the temperature is increased above the θ-temperature. The excluded volume interactions in the concentrated state result in the osmotic pressure continuing to increase with increasing polymer concentration, in addition to decreasing the diffusivity of the molecules.

In both the semi-dilute and concentrated regimes, each polymer molecule is a component of a mesh due to the interpenetration of each molecule by it neighbours. The mesh size is referred to as the *correlation length* and decreases with increasing concentration until the dimension is of the order of the segment size in the melt state. Structural relaxation of the bulk system is effected by the diffusion of the molecules. When they are part of a entangled mesh, the net motion is by the wriggling or *reptation* [6] of each chain through the mesh. The model for this motion [5, 6] is of a chain moving through a tube. The dimensions of the tube cross-section are governed by the mesh size as the walls are formed by the surrounding molecules. Of course, as the concentration increases and the mesh size is reduced, the dynamics are slowed.

2.6 POLYMERS AT SURFACES

The starting point is to consider the interaction energy between the atoms or molecules making up this third component with those between the solvent and macromolecular species. If we use the concept of the Flory χ parameter, then we may assign a value for the interaction with the surface by considering the interaction energy between the polymer–solvent, polymer–surface and solvent–surface. So, if $\chi < \chi_{surf}$, the polymer will not absorb (where χ_{surf} is the polymer–surface value). Conversely, if $\chi > \chi_{surf}$, the polymer will adsorb. Detailed modelling has been carried out by Scheutjens and Fleer [8], who used the lattice model at a surface

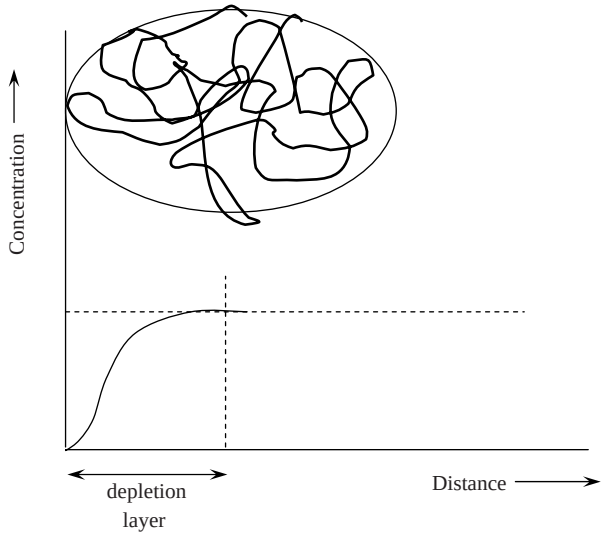


Figure 2.8 Illustration of the closest approach of a non-adsorbing polymer coil to a surface, showing the reduction in the local polymer concentration close to the surface from the average value in solution – the region is termed the *depletion layer*

and varied the χ parameter over the first few layers. This enabled predictions of concentration profiles to be made for both adsorbed homopolymers and adsorbed copolymers. The profile has also been modelled as a 'self-similar mesh' by deGennes [9, 10]. The details of the outer part of the concentration profile becomes of interest in the discussion of particles stabilized by adsorbed macromolecules. Figure 2.8 illustrates the concentration profile for a non-adsorbing polymer.

To obtain a uniform polymer concentration right up to the interface, the conformations in parts of the polymer would have to be reduced as that part of the coil close to the surface became more concentrated. This is energetically unfavourable without a competing attraction from the interface and the result is a *depletion layer* where the local concentration is lower than the global average within a distance of $\sim R_g$ away from the surface. When the enthalpy change for the adsorption, coupled with the increase in entropy of solvent molecules displaced from the surface, is greater than the decrease in entropy due to the restriction on in polymer conformation, the free energy is favourable for adsorption and the polymer will stick to the surface. Figure 2.9 illustrates the type of conformation that occurs for a polymer adsorbed from a θ- or better solvent. In a poor solvent, of course, the polymer will adsorbed in a dense layer

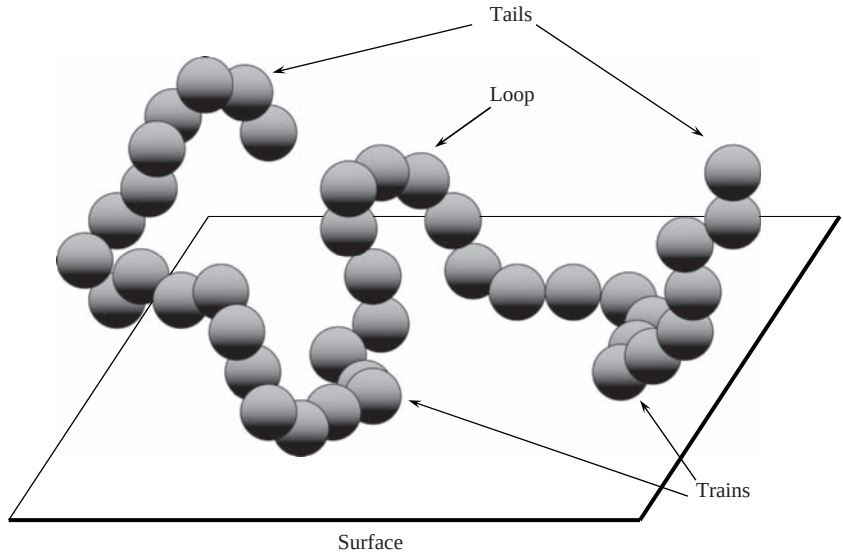


Figure 2.9 Three-dimensional representation of the conformation of a polymer adsorbed at an interface, showing the features of 'tails', 'loops' and 'trains'

on the surface. Figure 2.10 shows the concentration profile in the surface layer. Note that the tails project further into the solution phase than the loops and so the total concentration profile falls to that of the tails at the outer periphery.

Homopolymers are not usually added to colloidal systems to enhance the colloidal stability by adsorption. They are, however, frequently added

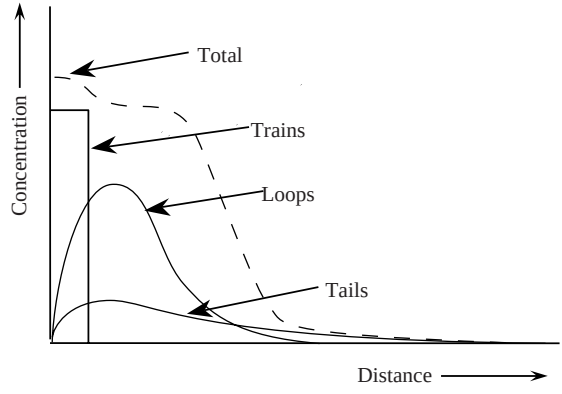


Figure 2.10 Illustration of the concentration profile of an adsorbed polymer

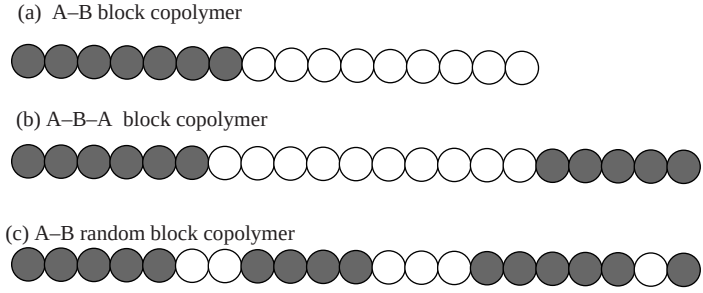


Figure 2.11 Some examples of block copolymers: (a) A–B, e.g. poly(propylene oxide)-*co*-poly(ethylene oxide); (b) A–B–A, e.g. poly(12-hydroxyoctadecanoic acid)-*co*-poly(ethylene oxide); (c) A–B random, e.g. poly(vinyl alcohol)

as *rheology modifiers* (*thickeners*). Copolymers are much better candidates to act as stabilizers. The molecular design is chosen to be amphipathic so that part of the molecule is in a 'worse than θ-solvent' whereas the rest is in a 'better than θ-solvent'. This provides strong attachment to the surface while maximizing the extension of the soluble moieties. A variety of copolymer configurations are available and are shown in Figure 2.11. Poly(vinyl alcohol) is prepared by the partial hydrolysis (e.g. $\sim$80%) of poly(vinyl acetate) and is a commonly used random block copolymer stabilizer. The surface configuration adopted by such random block copolymers will be similar to that sketched in Figure 2.9.

The adsorption isotherms measured for adsorbing polymers are usually of the high-affinity type so that when a low amount of polymer is added it is all adsorbed. Figure 2.12 illustrates the type of curve

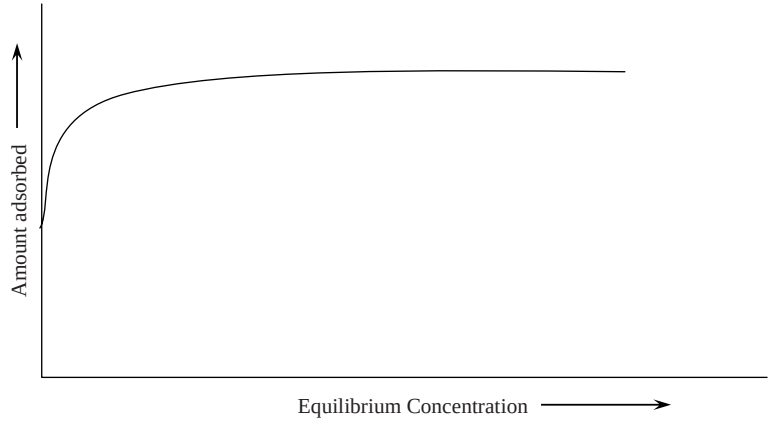


Figure 2.12 Illustration of the type of high-affinity isotherm found for most adsorbing polymers used for stabilizing colloidal particles

frequently obtained in this case. The determination of an isotherm is usually done by equilibrating aliquots of a dispersion with varying amounts of polymer, separating the particles and carrying out an assay for the polymer in solution. This has some practical difficulties, however. For example, the equilibration time can be fairly long – in some cases, it can be days. The lower molecular weight species diffuse more rapidly and can occupy the surface first. Subsequent displacement by larger molecules may then occur, but this is slow due to the 'multi-point' attachment of the chains with the added complication of slow surface motion to optimize the conformation. This behaviour makes the mixing process important as the desorption and reorganization of polymer are very slow. An added problem is that the equilibrium concentration of polymer in the solution phase, close to the onset of full coverage, is low, thus making the analysis difficult. When we use the dimensions of a polymer coil illustrated in Figure 2.1 and pack these together on a surface, monolayer coverage corresponds to a value of the order of $1\,\mathrm{mg\,m^{-2}}$. With the specific surface area for many colloidal systems being in the range $1–100\,\mathrm{m^2\,g^{-1}}$, it is clear from this that fairly small amounts of polymer are sufficient to saturate the surface of a dispersion.

2.7 POLYMER CHARACTERIZATION

The molecular weight is a key piece of information that we require for any polymer system, and there are several experimental options available to give us this information. It is always useful to employ two different methods as this will provide an insight into the polydispersity of the polymers that we are using. At this point, it is useful to review briefly the methods most frequently used.

2.7.1 Intrinsic Viscosity

It is straightforward to measure the viscosity of a polymer solution in a simple U-tube viscometer where the time is measured for a known volume of solution to flow through a capillary tube [11]. The viscosity of a polymer solution is described by the Huggins equation:

$$\eta = \eta_0 \left(1 + [\eta]c + k_{\mathrm{H}}[\eta]^2 c^2 + \cdots\right) \tag{2.19}$$

where c is the concentration of the polymer solution with a viscosity η, η_0 is the solvent viscosity, $[\eta]$ is known as the *intrinsic viscosity* and k_{H} is

the Huggin's coefficient. The intrinsic viscosity is directly related to the size and shape of the molecule in solution, and therefore to the molecular weight, and the Huggin's coefficient is a function of the pairwise interaction of the molecules. This equation could be extended to include higher order interactions, but as written is the low-concentration result. Equation 2.19 can be rearranged to give a linear expression:

$$\frac{\dfrac{\eta}{\eta_0} - 1}{c} = [\eta] + k_{\mathrm{H}}[\eta]^2 c \tag{2.20}$$

The term on the left-hand side is known as the *reduced viscosity*. A plot of the reduced viscosity versus the solution concentration gives a straight line with the intrinsic viscosity as the intercept. Experimentally, the problem is that at low concentrations the relative viscosity, (η/η_0), is very close to unity, so when we calculate the reduced viscosity we need data with high precision. For example, if our viscometer had a flow time of 120 s, we would like data to be reproducible to within 0.2 s. This requires a great deal of care, cleanliness and excellent temperature control as viscosity changes exponentially with temperature.

The molecular weight can be obtained from the intrinsic viscosity by using the Mark–Houwink equation:

$$[\eta] = KM_{\mathrm{v}}{}^{a} \tag{2.21}$$

where K and a are constants for a given polymer and solvent pair. The value of a is dependent on the quality of the solvent. In a θ-solvent $a = 0.5$ and in 'better than θ conditions' it can rise to 0.8. There is a comprehensive set of data available in the literature [12]. Equation 2.21 was originally derived using polymers with a narrow molecular weight fraction but of course, with the broad distribution with which we usually work, the value is an average. The viscosity is function of both the number density and the size in solution and the average is in between the values we would calculate on the basis of number and the average on the basis of weight.

2.7.2 Limiting Osmotic Pressure

The osmotic pressure – concentration relationship was given in Equation 2.17 and forms the basis of a useful method of determining the molecular

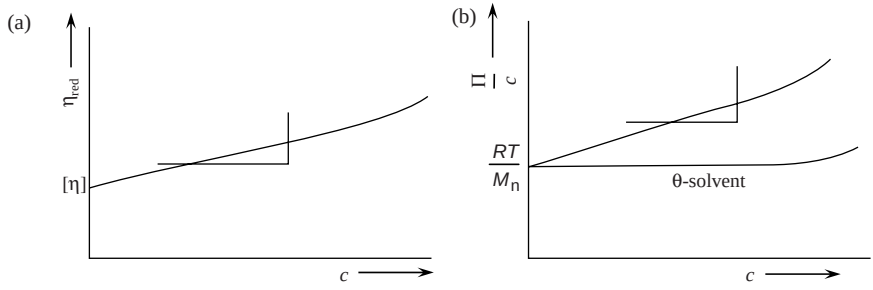


Figure 2.13 (a) Reduced viscosity plotted as a function of concentration for a polymer solution. The intercept yields the intrinsic viscosity $[\eta]$ and the slope is $k_H[\eta]^2$. (b) Reduced osmotic pressure plotted as a function of concentration for a polymer solution. The number-average molecular weight is obtained from the intercept and the second virial coefficient from the slope, with the latter reducing to zero as χ approaches 0.5

weight. The osmotic pressures of a series of solutions of different polymer content are measured and a linear plot of Π/c as a function of c should be obtained at low concentrations, as illustrated in Figure 2.13. The number-average molecular weight is obtained from the intercept and the value of the χ parameter from the slope.

The type of osmometer used is dependent on the molecular weight of the polymer. A membrane osmometer is used for $M_n > 10^4$ Da, whereas lower molecular weight polymers are characterized using a vapour pressure osmometer. With a membrane osmometer, pressure is applied to the solution to prevent the transfer of solvent across a microporous membrane. The vapour pressure instrument measures the change in temperature of an evaporating drop on the tip of a syringe. The higher the vapour pressure, the faster is evaporative cooling rate. The vapour pressure of a solution is a function of the concentration and the molecular weight of the solute. The equipment is calibrated with standard polymers of narrow molecular weight distributions.

2.7.3 Angular Light Scattering

When light passes through a medium and penetrates a region where the refractive index changes, light is scattered radially from that region. A dissolved polymer molecule is one such region and each molecule will scatter light. The scattering of electromagnetic radiation is discussed in some detail later in this volume. However, at this point we should

recognize that the amount of light scattered is a function of the number of scattering units per unit volume, their size and the difference in refractive index of the scattering unit and the surrounding medium.

The simplest approach is to treat the scattering molecules as 'point scatterers'. As the dimension in solution has R_g as less than 5% of the wavelength, this is a good approximation. We measure the relative intensity of light scattered at a given angle to give the *Rayleigh ratio*, $R(\Theta)$, as follows:

$$R(\Theta) = \frac{Ir^2}{I_0 v_s} \tag{2.22}$$

where I is the transmitted intensity, I_0 is the incident intensity, r is the pathlength from the cell to the detector and v_s is the scattering volume. It is convenient to measure the scattering at 90° and we can express the dependence of the Rayleigh ratio on the concentration of a dilute polymer solution as follows [13, 14]:

$$\frac{K_l c}{R(90)} = \frac{1}{M_w} + 2B_l c \tag{2.23}$$

with the optical constant

$$K_l = \frac{2\pi^2 {n_0}^2}{\lambda^4 N_A} \left(\frac{\partial n}{\partial c}\right)^2 \tag{2.24}$$

which is dependent on the incident wavelength, λ, the refractive index of the solvent, n_1, and rate of change of refractive index with polymer concentration; B_l is the second virial coefficient and is dependent on the 'quality of the solvent'. In this case, the weight-average molecular weight is obtained from the intercept of a plot of Equation 2.23 at zero concentration.

2.7.4 Gel Permeation Chromatography

When a polymer solution is passed slowly through a porous bed or swollen cross-linked gel, diffusion of the macromolecules means that they explore the pores as they pass through. This slows the passage of the macromolecules as when inside a pore, they are out of the flow and only pick up forward motion when they emerge. Smaller molecules can

explore small pores in addition to large ones, with the result that they spend more time inside the pores. The net result is that a separation occurs if the column is long enough. The larger macromolecules emerge first, with the smaller fraction emerging last as the column is eluted. The concentration of the polymer in the eluent is measured. A common and convenient method is to measure the refractive index of the solution as it passes through. This provides a continuous electronic read-out. The concentration can be calculated from a calibration curve of the refractive index dependence of solutions of the polymer being studied. The residence time on the column is calibrated by using narrow molecular weight distribution standards. This technique provides a molecular weight distribution and not just an average value. Also, the method may be scaled-up to provide a preparative route to small amounts of material with a narrow distribution of molecular weights.

Columns are available for aqueous or non-aqueous systems. Adsorption of the macromolecules on a chosen column material can prevent the technique from being used successfully and so the choice of an appropriate packing material and solvent system is important.

The characterization of copolymers such as A–B–A block or hydrophobically modified polymers can present special difficulties. Such polymers are synthesized so that part of the macromolecule is in a poor solvent. The result is that some self-assembly may occur and lead to the measurement of aggregates of macromolecules. Characterization of molecular weight needs to be carried out in a solvent system which will suppress such aggregation. It must then be recognized that the conformation in the solvent system used for the final application may need some additional experiments.

2.8 BIOPOLYMERS

Although many of the polymers that we use are based on synthesis from petrochemicals, we are at least as dependent on those from biological sources. From the biosphere we have three types of polymers available to us:

1. polyesters
2. polysaccharides
3. polypeptides.

These are listed in reverse order to their importance to us; for example, the last category includes DNA, which is key to our existence.

The description of each is not intended to be exhaustive but more to stimulate interest in where these polymers are currently important. In some cases, for example the cellulosics, we are making use of a polymer backbone that Nature has already produced, whereas in others the key process is the fermentation to produce the desired monomer feedstock. Some of the biopolymers that we use are in solution in water, for example the cellulose ethers that are used as 'thickeners' or as swollen gels, the alginates or gelatin used in foods, and there is increasing application of their fabrication as films, fibres and solid forms because of their biodegradability. Many of the solid structures are composite materials with disperse phases ranging from the micro- down to the nanoscale.

2.8.1 Polyesters

Bacteria can be used to prepare polyesters as a metabolic by-product. Fructose from soy or maize is the initial source in the fermentation broth using the bacterium *Ralstonia eutropha* and polyhydroxyalkanoate granules are then harvested from the culture in a batch process. Polyhydroxybutyrate is a thermoplastic produced in this manner. By using a transgenic version of *Arabidopsis thaliana* with genes from the bacterium *Alcaligenes eutrophus*, polyhydroxybutyrate granules were obtained from the plant cells [15]. Note that *Arabidopsis* is frequently used as a plant model and would not be the final 'bioreactor' for a process.

Poly-3-hydroxybutyrate is a hydroxyalkanoate with high potential as a biodegradable polymer matrix. The melting point is ~175 °C and with a tensile strength of 40 MPa, although brittle, it would make biodegradable plastic items. Colloidally dispersed pigments and fillers are readily added. One drawback for broad application is that the polymer can be hydrolysed and this hydrolysis is acid and base catalysed and the oxygen permeability is high.

The manufacture of poly(lactic acid) or poly(hydroxypropionic acid) starts from the production of hydroxypropionic acid from the fermentation of cane sugar or corn starch. This produces the L-enantiomorph, which is dimerized to give the lactide, which is then polymerized using a tin(II) Lewis acid catalyst to give the polymer. The polymer is produced commercially and can be used to make objects such as containers and spun to make fibres for fabric production. The reaction scheme is illustrated in Figure 2.14. The polymer is brittle due there being significant crystallinity. If a racemic mixture of the D- and L-enantiomorphs are

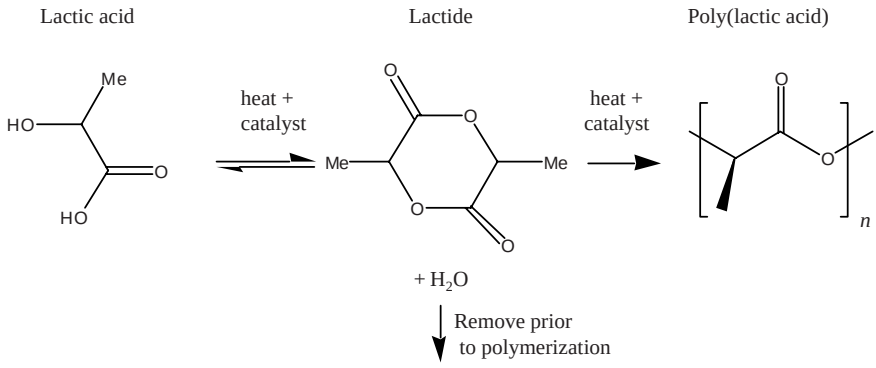


Figure 2.14 The synthetic route to poly(lactic acid)

used in the polymerization reaction, the polymer is amorphous as the stereoregularity leading to crystalline regions is lost.

2.8.2 Polysaccharides

The polysaccharides are a class of biopolymers which are abundant, cheaper than petrochemical polymers to produce and have a long history of use. They are extracted directly from plants, from trees (e.g. cellulose), grain or root crops (e.g. starch), seaweed (e.g. agarose) and from microbial fermentation (e.g. xanthan). They are all used as 'thickeners' due to their ability to produce very viscous solutions or weak gels at low concentrations and are used in products as diverse as paints and foods. This class of water-soluble polymers is often referred to as 'hydrocolloids'. The review by Chaplin [16] is an excellent compact resource with a very large number of references to relevant text. Two characteristics that they all have in common is their ability to form intermolecular hydrogen bonds and to form stiff expanded structures. We will now briefly review the common polysaccharides and we will group them from their common origins, reflecting their production rather than by their end use.

2.8.2.1 Polysaccharides from trees and plant stems – cellulosics

Cellulose forms the structural element of plants and is obtained in large amounts from wood pulp. The base polymer is a linear arrangement of β-D-pyranose units linked in the 1:4-positions via a shared oxygen.

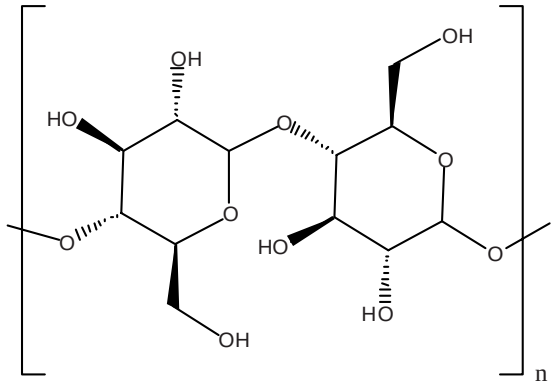


Figure 2.15 The repeat structural unit of cellulose showing the β-(1→4)-D-glucopyranose units

This is illustrated in Figure 2.15. The effectiveness of cellulose as a structural polymer is due to its insolubility in water. This is due to the extensive H-bonding, both inter- and intramolecular. This produces a crystalline sheet structure which forms the strong structural elements of plant stems. In order to produce water-soluble products, 20–30% of the hydroxyl groups are etherified to give products such as methylcellulose, ethylcellulose, hydroxyethylethylcellulose and hydroxypropylmethylcellulose. In some cases, anionic charge can be added with carboxyl groups, as is the case with carboxymethylcellulose.

Typical molecular weights are of the order of 100 kDa. The chain is relatively rigid with an expanded conformation, as indicated by a value of $c_\infty = 8$. This is due to the ring structure and also the side groups of the cellulose ethers. In the dilute state, carboxymethylcellulose at moderate to high pH forms a very open, almost rod-like structure due to the repulsion of the carboxy groups linked to the backbone. At higher concentrations, the total ion concentration rises, screening the charges progressively as the concentration increases, resulting in a coil conformation. The coils contract as the pH is lowered and the ionic strength is increased, as is the case for any polyelectrolyte. Although the etherification of a proportion of the hydroxyl groups reduces inter- and intramolecular H-bonding, allowing the molecules to become water soluble, the solubility does rely on H-bonding to the water molecules. As solution temperatures are increased, H-bonding is reduced and this allows the more hydrophobic regions on the chains to associate and strong gels are formed. Usually, a temperature >80 °C is required and the gels are thermally reversible.

At concentrations of a few percent by weight, the soluble cellulose derivatives form weak *pseudoplastic* gels and frequently also show time-dependent flow behaviour (see Chapter 9 for a definition of rheological terminology). Cellulose ethers are added to many products to make use of this rheological behaviour to modify the handling properties. Latex paints are a good example of this. In Section 1.2.1, we discussed the complex range of properties that we required from a decorative paint. The thixotropic behaviour of the final paint is an example of the strong synergistic effect of a cellulose ether solution and the latex particles. The cellulose ether solution shows some shear thinning behaviour but, when the latex particles are added, a soft solid is formed as the insoluble polymer induces *depletion flocculation* of the latex particles (see Section 3.9 in Chapter 3). The flocculate structures formed are progressively more solid-like as the particle volume fraction increases [17, 18].

The surface activity of the cellulose ethers is a function of the type and extent of the modification. Hydroxypropylmethylcellulose (HPMC) can effectively stabilize gas bubbles and has applications in bread making in recipes which are gluten-free. It has film-forming properties and is used in the pharmaceutical industry [19] as a gelatin substitute for the production of hard capsules which are accepted for vegetarian use. Hydroxypropylcellulose (HPC) has a lower critical solution temperature (LCST) of $\sim 45\,^{\circ}C$ and is used as a food additive, providing emulsion stability and rheological control. In addition, a solution of HPC has application as a tear substitute. Methylcellulose is used pharmaceutically as a dietary fibre supplement.

2.8.2.2 Polysaccharides from grain – glucan, amylose and amylopectin

Cereal crops provide us with a large amount of polysaccharides annually. β-Glucan is a polysaccharide that is present in the bran portion of the developed grains, with barley and oats being the richest. The structure is linear, consisting of D-glucopyranose units linked $1\rightarrow 4$ and $1\rightarrow 3$ with $\sim 2.5 \times 10^5$ units [16]. With such a high molecular weight, solutions are very viscous and markedly shear-thinning. This can give problems in the brewing industry. Nutritionally, β-glucan is recommended as a source of fibre and is thought to reduce cholesterol. Recent research also shows some promise in effecting the immune system response in a favourable manner.

Starch granules are recovered from the body of the grains and also from roots such as potato tubers. The grains contain large amounts of

the two polysaccharides amylose and amylopectin, which are released by cooking the granules. Amylose is the less abundant of the two molecules, has a lower molecular weight and consists of α-D-glucopyranose units linked in a 1→4 conformation, which gives the unbranched molecule a relatively stiff helical structure. A section of the helix is illustrated in Figure 2.16. The 'jet-cooking' of starch (high pressure at ~120 °C)

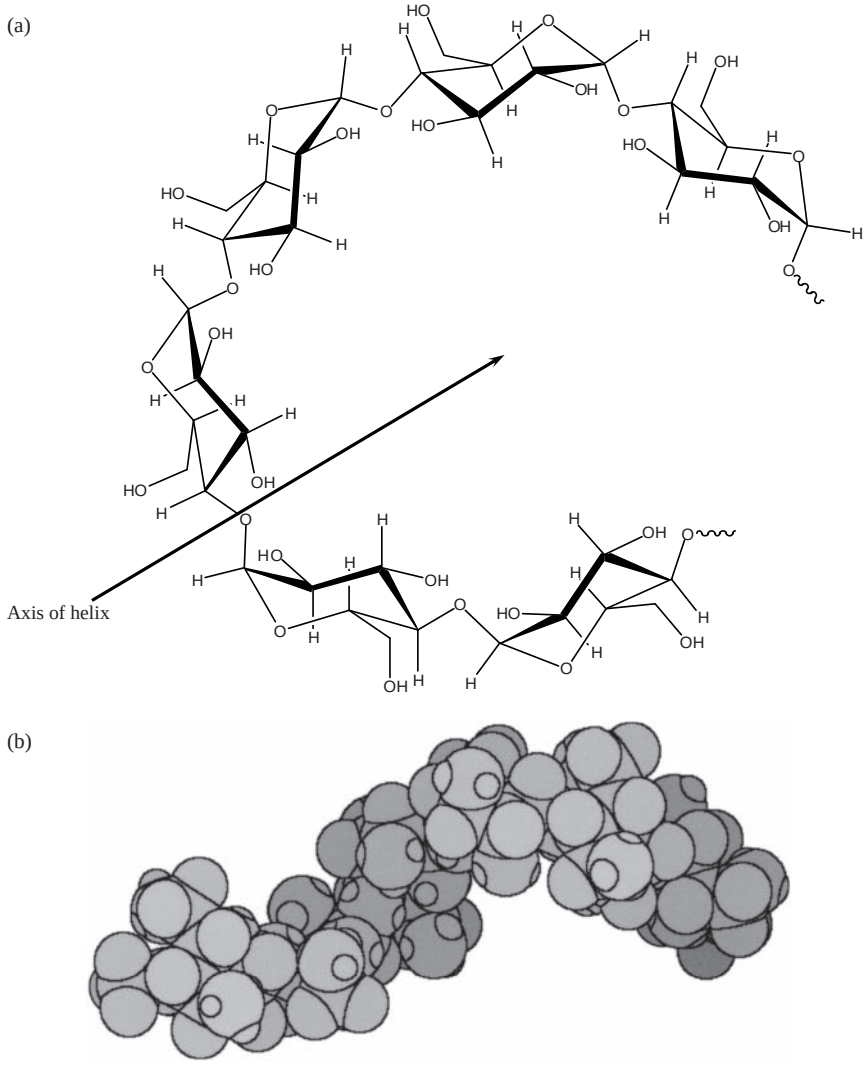


Figure 2.16 (a) A section of the helical chain structure of amylose. (b) The space-filling model of the same helical section provides a better indication of the structure

removes some of the H-bonded water and allows the amylose to form double-stranded crystallites. This is known as the retrogradation of starch and when carried out in the presence of fatty acids, such as octadecanoic (stearic) acid, gives biconcave disc-shaped particles similar in size and shape to erythrocytes. The inside portion of the single-stranded helix is moderately hydrophobic and can bind iodine ions; indeed, the blue colour produced in the titration of starch with iodine is due to the electron donor–acceptor interactions of iodide ions within this part of the molecule.

Amylopectin is the major component in starch granules and is a heavily branched version of the amylose chain with a branch at approximately every 25 glucopyranose units. The structure is similar to that of glycogen, which is found in animals but which has a higher degree of branching.

Starch is important nutritionally as a source of energy. We have enzymes that can hydrolyse about 30% of the starch that we eat. However, starch is also used as a thickener in colloidal dispersions, as an adhesive, in papermaking to help bind the cellulose fibres and as a mould-release agent in the confectionery industry. Genetic modification of plants has opened up the possibility of targeting functional improvements for specific products [20].

2.8.2.3 Polysaccharides from seaweeds – agar, carrageenan and alginates

Red seaweed is the commercial source of agar and the carrageenans. Agarose is the useful component of agar as it forms gels when a solution is cooled to $\sim 35\,^{\circ}C$, which can then be reheated to $> 85\,^{\circ}C$ without dissolution. It is widely utilized as a base for microbiological culture. It is a good base for this purpose as the gels are stable at different temperatures and the molecular structure prevents any microorganisms from using it nutritionally. It is also be used in the processed foods area as it can form gels when present in concentrations $> 0.5\%$. Because of its thermal stability, it is a useful gelatin substitute in countries with hot climates.

The molecule is linear, has a molar mass of ~ 120 kDa, is insoluble in cold water but is soluble at temperatures $> 90\,^{\circ}C$. The molecules self-assemble into complex double-helical units which can form large aggregates, producing strong gels. The chains consist of (1→3)-β-D-galactopyranose-(1→4)-3,6-anhydro-α-L-galactopyranose units.

The carrageenans are prepared by the alkaline extraction from red seaweed, with the most common form being κ-carrageenan. The

molecular structure is similar to that of agarose except that the D-galactopyranose unit has a sulfate group in the 4-position and the anhydro-galactopyranose is the D-form. Because the polymer has a high anionic content of strong acid groups, the gel formation is strengthened by interaction of cations. Potassium and calcium ions are the most effective for this purpose. The polymer is used as a rheological modifier in personal care and pharmaceutical products and also in foods, especially in milk products as the sulfate groups can interact with the protein stabilizers (casein).

Alginates are produced from the brown seaweeds and, like carrageenan, are polyelectrolytes although, in this case, the acid groups are weak carboxylic groups. The acid moieties are α-L-guluronate and β-D-mannuronate linked in the 1,4-positions. The acids can be link in blocks or can alternate. Alginates absorb water readily and form thermally stable gels. Calcium ions enhance the gelation and the calcium salt is the most frequently used form in commercial products. The polymer can be obtained in a range of molecular weights from 10^4 to 10^6 kDa. In their free acid forms, the molecules form helices linked via H-bonds. In the calcium form, the divalent calcium ions bind strongly, which strengthens the gels, but the guluronate sequences fold into a conformation which surrounds the calcium ions in an analogous fashion to eggs in an egg-box. The alginates are used as a rheological modifier in foods and pharmaceuticals. Calcium alginate has uses in medical dressings for wounds and burns. The strong gelling properties also make it useful for mould-making in areas such as casting from life, for example, in dentistry, prosthetics and sculpture.

2.8.2.4 Polysaccharide from bacterial fermentation – xanthan gum

Bacterial fermentation is employed to produce a polysaccharide with a backbone of β-(1→4)-D-glucopyranose glucan with polyelectrolyte side chains. The side chains are (β-D-mannopyranoside-6-acetate)-(α-D-glucuronopyranosyl)-(β-D-mannopyranosyl). The structure in solution is that of a double helix formed from two molecules. This produces a stiff rod-like unit and the fermentation process yields a moderately narrow range of molar masses. Xanthan is a useful pseudoplastic rheological modifier with a sharper response than many thickeners which have been produced by digestion of very large molecular structures. The increasing alignment of the rod-like molecules with increasing shear rates gives low viscosities at high rates, while the rapid molecular diffusion at the

cessation of flow is responsible for the rapid recovery of the viscosity. The value of the low shear viscosity can be enhanced by freeze–thaw cycles [21].

2.8.3 Protein Hydrocolloids

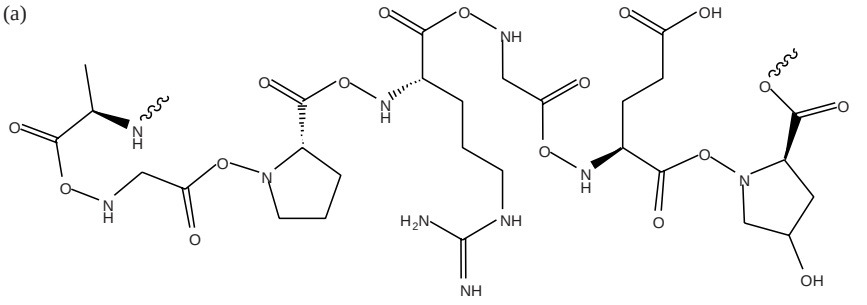
Perhaps the most important biopolymers are DNA and RNA and, although they may be of great interest to a scientist working with colloidal systems, they do not form part of the general 'tool-kit' for manipulating the properties. As a result, we shall not consider them further here.

2.8.3.1 Gelatin

Gelatin has been produced industrially since the early part of the nineteenth century by boiling animal skins and bone with acid or alkali to denature the collagen. Fish skins can also be used, but the bulk of the gelatin is produced from cattle with a lesser amount from pigs. It is widely used in the food industry as a gelling agent where the thermo-reversible nature of the gels formed produces attractive 'mouth-feel' properties, that is, the products have the correct psycho-rheological properties for successful consumer products. In the non-digital photographic field, gelatin is the common matrix for film and plates and stabilizes silver halides and chromophores. Colour films can have up to seven different layers laid down simultaneously at high speed from the melt. In this process, the dynamic wetting of substrate, the control of the viscosity and the gelation properties all have to be tightly controlled. It is used as a stabilizer for colloidal particles in many areas and was the stabilizer used by Michael Faraday when he synthesized nanoparticles of gold.

The backbone of gelatin is made up of amino acids, the dominant ones being glycine, proline (formally an imino acid) and 4-hydroxyproline in combination with alanine, glutamic acid and arginine. Glycine occurs most frequently, indeed, it makes up about one-third of the structural elements. Figure 2.17 illustrates a section of the gelatin molecule. The molecule in solution forms triple helices in the regions rich in 4-hydroxyproline and proline. The production process results in a mean molar mass of $\sim 95 \times 10^3$ Da, but there is a wide range of chain lengths making up this average. It is the combination of hydrophobic sections with hydrophilic sections containing acid or base functionality that gives the molecule it properties as an effective colloidal stabilizer.

(a)



(b)

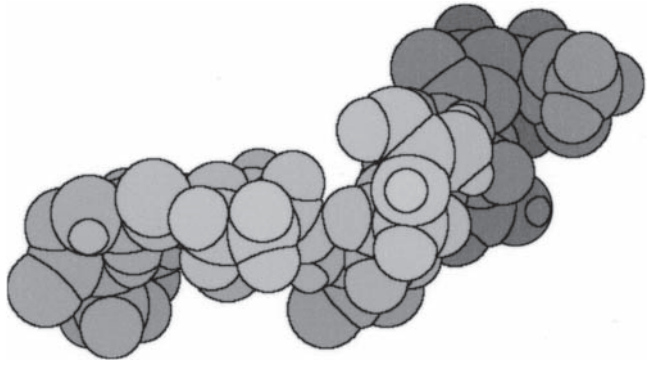


Figure 2.17 (a) The structure of a section of a gelatin molecule. (b) The space filling model illustrates the helical nature of the molecule and also shows the presence of significantly large side groups

2.9 SURFACTANTS IN SOLUTION

A *surfactant*, or *surface-active agent*, is a general term used to describe molecules that interact with an interface. These consist of two parts, one of which is highly soluble in one of the phases whereas the other is not. They are small, mobile molecules which are widely used in colloidal systems. For example, they are used as soaps, detergents, dispersants, wetting agents and germicides. Their structure consists of a hydrophobic tail which is usually a hydrocarbon (although fluorocarbon and dimethylsiloxane chains can also be used), with a polar hydrophilic head group which may be ionic or non-ionic. This type of molecular structure gives rise to the non-ideality of solutions of surfactants and their phase behaviour. Figure 2.18 shows some examples of different types of surfactant molecules. Surfactants are used in both aqueous and

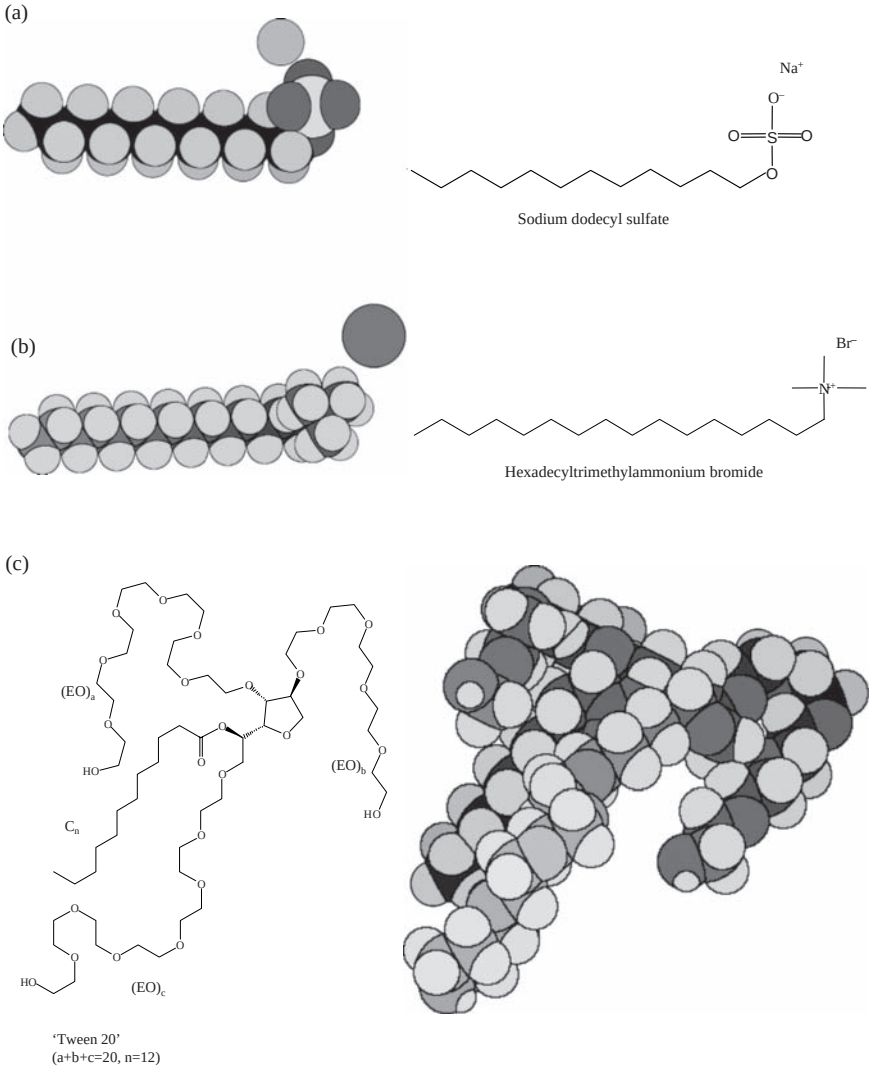


Figure 2.18 Some examples of different types of surfactants: (a) anionic; (b) cationic; (c) non-ionic. EO, ethylene oxide

non-aqueous systems and, although we usually think of the synthetic surfactants that are manufactured in large quantities, there are some very important naturally occurring types. For example, the surfactant in our lungs is vital for their operation, as are the bile salts produced by the pancreas, and act to disperse dietary fat into colloidal-sized droplets (or *chylomicra*) that pass into the bloodstream where they are utilized

by the body. The phospholipid lecithin is a constituent of cellular membranes. The fatty acids are also surfactants and the source of soaps, the manufacture of which consists basically of producing the sodium salt in a high-concentration phase which can be conveniently handled.

2.9.1 Dilute Solutions

We will mainly concern ourselves with aqueous solutions, but it should be kept in mind that analogous behaviour may be found in other solvents. As surfactant is added to water, the molecules dissolve. In most cases, an increase in temperature aids this process, but this is not universally true. In solution, the polar head groups are hydrated – they can take part in the H-bonding structure of the water. The hydrocarbon tails, on the other hand, cannot do this and we can visualize a discontinuity in the structure of the water around these tails. This has been referred to as a 'cage' and is cited as the origin of the 'hydrophobic effect'. What this means is that for the tail to be in solution, there is a increase in free energy relative to the reduction in the number of H-bonds, which we may think of as proportional to the area of the 'cage'. If several of these tails are brought together into an aggregate, the surface that is required to surround them would be less than the sum of the areas of the individual 'cages'. Furthermore, there will be an additional enthalpic term from the van der Waals attraction between the tails. Opposing this, of course, is the decrease in entropy associated with the clustering of the chains, although the entropic term for the water increases. With ionic head-groups, the electrostatics opposes their close approach, but counter-ion binding reduces this. In addition, water is also partially freed from the hydration sheath of the head groups. Any interface such as an air–water or oil–water interface also provides an opportunity for the tails to move out of the water structure.

With this picture in mind, let us consider the process of adding crystals of surfactant to water in increasing amounts and changing the temperature in each case. What we are doing, of course, is investigating the phase behaviour from a series of isopleths (varying temperature at fixed concentrations). Figure 2.19 sketches the type of phase diagram that we could expect. At low temperatures, the solubility is low with surfactant crystals in equilibrium with surfactant solution. There is a critical point known as the *Krafft point* [22, 23], and at temperatures higher than the value at this point (the *Krafft temperature*) the solubility appears to increase rapidly and the solution phase consists of surfactant aggregates or

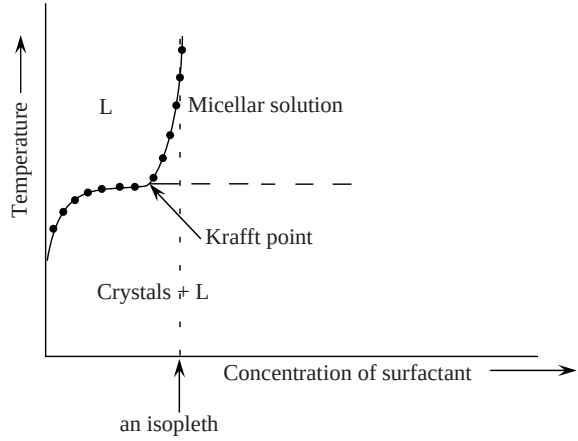


Figure 2.19 Schematic phase diagram for a dilute surfactant–water system. L represents a 'simple' solution. The Krafft point is the critical point and above the Krafft temperature the solution consists of surfactant micelles and surfactant molecules

micelles in addition to single molecules. Below the Krafft temperature, micelles are not formed. The concentration at which micelles are formed is known as the critical micelle concentration (cmc). The cmc varies with temperature above the Krafft temperature.

2.9.2 Micellization

Provided that we keep our solution above the Krafft temperature for our surfactant, there is clear evidence for a sharp self-assembly process occurring at a particular concentration from a sharp change in the physical properties of the surfactant solution. If the surface tension is measured for increasing concentrations of surfactant in solution, the former decreases steadily as increasing adsorption of surfactant molecules at the air–water interface disrupts the local H-bonding and makes the surface more 'oil-like'. The extent of the adsorption can be calculated from the Gibbs adsorption equation as discussed in Chapter 1. However, suddenly, the slope of the surface tension versus concentration curve decreases to almost zero. This indicates that the adsorption at the interface is now constant, although our solution concentration is increasing. One explanation would that might occur to us is that the surface is now saturated by a monolayer so that no more sites are available. However, then we find that other properties also change at the same concentration. For

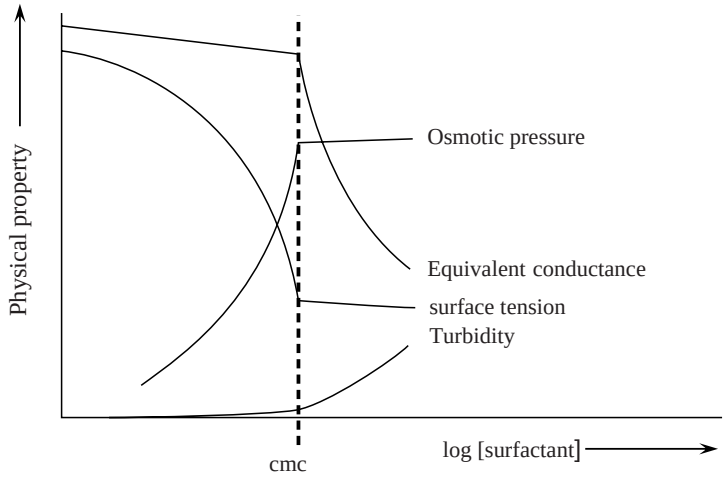


Figure 2.20 Schematic plot of the variation of certain physical properties of a surfactant solution at the critical micelle concentration

example, the rate of increase in osmotic pressure falls dramatically to a plateau. Light scattering experiments show a sharp increase in turbidity. These last two observations suggest that the increased amount of surfactant that we have added is not in the solution phase. Measurements of the equivalent conductance with increasing concentration show a marked decrease in slope after the critical concentration, indicating that we now have much less mobile charged units than we would expect from the individual molecules of surfactant. Figure 2.20 illustrates the type of change that we observe. Other measurements, by using NMR spectroscopy, also indicate marked changes at the same concentration. All of these experiments provide clear evidence of a phase change but with the formation of a sub-microscopic phase. When we measure the size of the units of the new phase, by light scattering or neutron scattering, for example, the size is approximately twice the length of the linear surfactant molecule.

As the surfactant concentration is increased to a level slightly above the cmc, we have spherical units – micelles – with a diameter twice that of the individual molecules, in addition to a constant concentration of single molecules. Attention was drawn above to the concept of the 'hydrophobic effect' and we may think about the micelle in these terms. Consider typical anionic surfactants such as sodium dodecanoate or sodium dodecyl sulfate. In both cases, there is a 12-carbon paraffin chain with a polar head group. The head groups will remain in the

water phase, with the tails in the spherical 'oily' phase. If the radius is larger than a stretched surfactant molecule, some of the heads would automatically have to be buried in the oil phase, which would be energetically expensive. At any given moment, some of the tails will be linear, and others will be bent to fill the volume to give a density similar to that of the bulk paraffin. With a CH_2—CH_2 distance of 0.127 nm, we easily calculate that there would be 60 or 80 molecules per micelle for the dodecanoate or sulfate molecule, respectively (the latter having a longer head group). If we take the diameter of a paraffin chain as 0.4 nm, we can compare the 'water contact area' for the appropriate number of chains with that of the outer surface of the micelle. Of course, we need to allow for the area occupied by the polar head groups on the surface (0.2 nm^2 per COOH group, for example). When this is done, we find that about two-thirds of the micellar surface is carbon chains in contact with water, although that is only one-sixth of the area that would be in contact if the molecules were separated. In addition to increasing the entropy of the water by freeing up the local structure, there will also be a contribution from the hydrocarbon because its motion is less constrained than when restricted by the water 'cage'. A detailed thermodynamic treatment would include contributions from the mutual repulsion of the head groups mitigated by the ion atmosphere around the micelle.

Of course, the question is whether there is a variation in the size of micelles around a mean value. To answer this, we need to consider the aggregation process in more detail. The aggregation is a progressive process which can be represented by the following multiple equilibria:

$$\begin{aligned} &S_1 + S_1 \rightleftharpoons S_2 \\ &S_1 + S_2 \rightleftharpoons S_3 \\ &\quad\quad\downarrow \\ &S_1 + S_{i-1} \rightleftharpoons S_i, \text{ and so on} \end{aligned} \tag{2.25}$$

which we can represent as follows:

$$iS_1 = S_i \tag{2.26}$$

where we have a single molecule, S_1, coming together to form a micellar aggregate, S_i. The equilibrium constant for this 'reaction' is given by

$$K_i = \frac{x_i}{(x_1)^i} \tag{2.27}$$

where x_1 and x_i are the mole fractions of the monomers and micelles. respectively. The equilibrium constant gives the standard free energy of

formation of the micelle containing i monomers as follows:

$$\Delta G^{o}(i) = -RT \ln K_i = -RT \ln x_i + i\,RT \ln x_i \tag{2.28}$$

that is

$$\frac{\Delta G^{\circ}(i)}{RT} = i \ln x_1 - \ln x_i \tag{2.29}$$

The standard free energy of formation of a micelle has contributions from the following:

(a) decrease as the hydrophobe–hydrophobe interactions replace the hydrophobe–water interactions;
(b) an increase as we form an interface between the micro-phase of the micelle and the surrounding water;
(c) an increase as we bring the hydrophilic head groups closer together by concentrating the hydrophobes in one place.

Everett [24] suggested the following equation for the standard free energy of formation based on those contributions:

$$\frac{\Delta G^{\circ}(i)}{RT} = -a\,(i-1) + b\,(i-1)^{\frac{2}{3}} + c\,(i-1)^{\frac{4}{3}} \tag{2.30}$$

where the coefficients a, b and have their origins in the contributions listed above and the values will vary with the chemical architecture of the system For example, a will become increasingly negative as the chain length of the hydrophobe increases and c would increase if the molecule had a charged hydrophilic group. To illustrate the implications of Equation 2.30, we can write it with suitable numerical values for these three parameters, as follows:

$$\frac{\Delta G^{\circ}(i)}{RT} = -18.8\,(i-1) + 19.55\,(i-1)^{\frac{2}{3}} + 1.25\,(i-1)^{\frac{4}{3}} \tag{2.31}$$

By combining Equation 2.31 with Equation 2.27, we can compute the ratio of the mole fractions of micelles to that of monomers as a function of i. By using single molecule concentrations at $\sim 10^{-2}$ M (which is a typical value for the cmc of a surfactant), we see from Figure 2.21 that there is a maximum in the curve at $i \approx 80$. However, the concentration

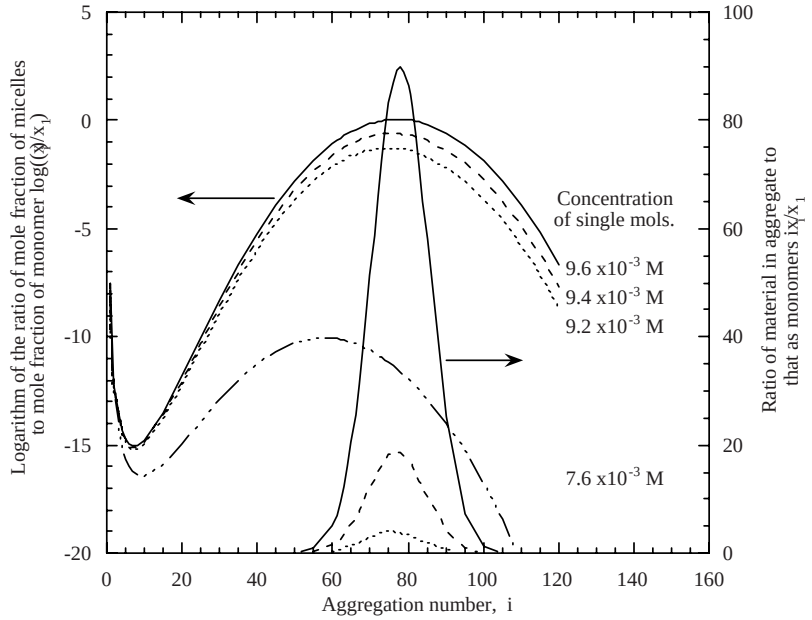


Figure 2.21 The ratio of monomer in micelles to single species in solution for a C_{12} surfactant

has to approach very close to the cmc before there is a significant proportion of material in aggregates [the numerical value of $\log(x_i/x_1)$ must approach 0]. If the concentration falls to around two-thirds of the cmc, the maximum disappears, indicating that aggregation would be absent.

As the concentration is increased to the cmc and slightly above, we find that the majority of the surfactant is present in aggregates and that these are of a preferred size. When we express the data as the ratio of the material in a micelle with an aggregation number i to the monomer level as a function of that aggregation number, it becomes clear that a preferred micellar size is obtained with a sharp distribution around the modal value. We can conclude that at concentrations slightly above the cmc we have a phase separation to give a monodisperse or uniform sized micro-phase. It is the balance between the opposing molecular interactions that define the preferred size and prevents the separating phase from growing indefinitely. With ionic surfactants, the counterion type and charge play a part and the micelle has to be modelled taking into account the potential distribution around the sphere [25]. The value of the cmc will be a function of the electrolyte concentration and temperature [26].

The micelles are not static, of course – the chains and head groups are mobile. Bound counter-ions of ionic head groups, which neutralize 30–50% of the charge, are not associated with a particular group but are mobile across the surface. Surfactant molecules can leave and rejoin micelles in a dynamic equilibrium, with the residence time being of the order of 1 μs. In addition, we should not think of a micelle as being a permanent entity but a 'unit' with a lifetime of the order of 1 ms. Even when there are oil-soluble molecules dissolved in the interior of the micelle (this is known as *solubilization*), the micelle is still not a permanent entity. The phenomenon of solubilization will only extend the lifetime. Fairly large molecules such as naphthalenic dyes can be solubilized by micelles and this process causes the micelle to increase slightly in size. A typical micelle would only solubilize about two naphthalene molecules, so the size increase is not great. Aliphatic alcohols with chain lengths of 6–12 carbon atoms also stabilize the micelle by reducing the interaction between head groups. Figure 2.22 shows the variation of the cmc with chain length for a number of different surfactants.

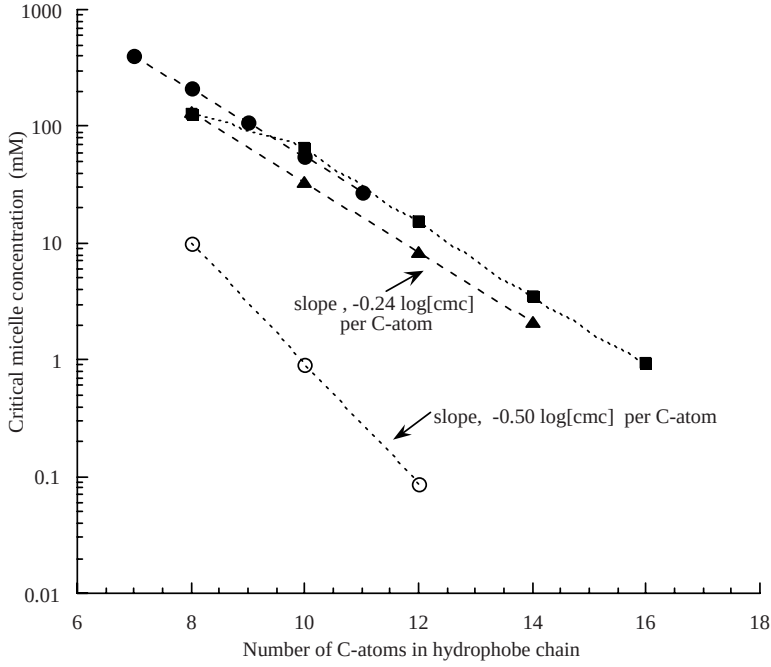


Figure 2.22 Variation of critical micelle concentration with surfactant chain length: ○, C_nEO_6; ■, C_nTAB; ▲, C_nSO_4; ● C_nCOONa. EO, ethylene oxide; TAB, trimethylammonium bromide

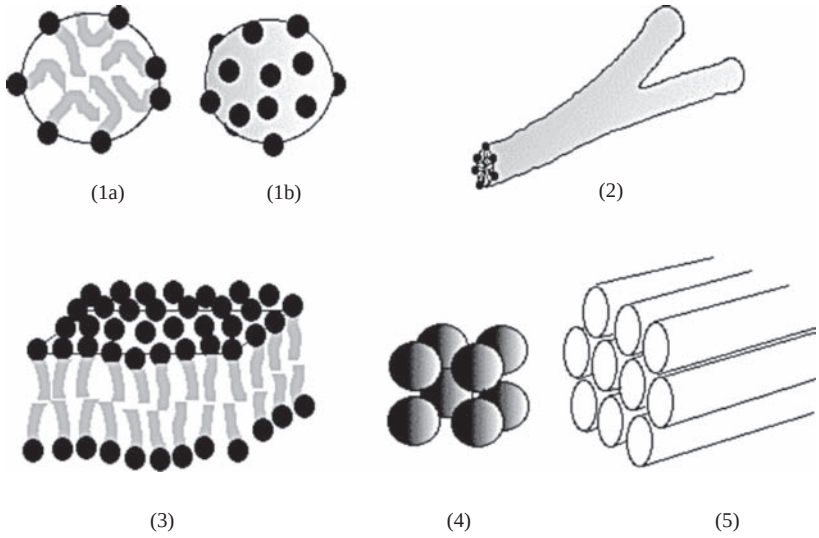


Figure 2.23 Examples of some of the (self-assembled) phase structures that can occur at temperatures above the Krafft temperature and with increasing concentrations above the cmc. (1) Spherical micelle ('a' as a cross-section); (2) 'worm-like' micelle; (3) lamellar phase; (4) cubic phase; (5) hexagonal phase

The above picture illustrates the phase behaviour at concentrations close to the cmc. At higher concentrations, the problem becomes more complex and the other phase structures appear. Figure 2.23 shows some of the structures that occur as the surfactant concentration increases.

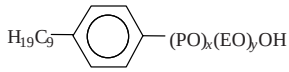
The most recently studied structure is the long, branched, rod-like structures known as 'worm-like' micelles, for which a large amount of literature has been produced in the past 15 years. The more organized structures, such as the liquid crystalline mesophases illustrated as (3), (4) and (5) in Figure 2.23, have been known since the early days of soap making when they were known as 'middle', 'viscous' and 'neat' phases, respectively. All of phases (2)–(5) are viscoelastic but with different textures. The more fluid ones can be useful for thickened detergent systems, with the lamellar phase giving us bar-soaps. Laughlin [23] has produced a particularly useful text describing the phase behaviour and how to recognize the various phases. The simplest way to observe the phase behaviour is to place a small amount of the solid surfactant on a microscope slide in contact with water. This is then viewed with a polarizing microscope. In the contact zone between the solid surfactant and the water, there is a progression of concentration from an isotropic solution through to the solid phase. The refractive index is different in

different directions for both the hexagonal and lamellar phases. This results in distinct optical patterns for each phase. With the cubic structure, the refractive indices are the same along each of the three axes and no patterns are observed. The sample may also be heated and a change from one phase to another observed. Usually, the phase behaviour is less sensitive to changes in temperature than it is to changes in concentration, that is, the phase boundaries do not change very much as the temperature is raised, with the major change being the loss of that phase as an isotropic solution is formed. With non-ionic surfactants, the formation of a second, coexisting liquid phase is observed as the *cloud point*. This is the result of the balancing of the hydrophobic effect of the tails and the head group hydration. The latter reduces as the H-bonding structure lessens with temperature and a second surfactant-rich phase is produced, giving the clouding phenomenon. The addition of electrolytes decreases the cloud point whereas the addition of co-solvents, such as ethanol or propanol, increases the temperature at which we observe the phenomenon. We need to operate close to the cloud point (around $\pm 10\,^{\circ}\mathrm{C}$) if we wish to use the surfactant efficiently as a dispersing aid [27]. This maximizes the interfacial responsiveness of the molecules.

Other structures can also be found. Di-chain surfactants, such as lecithin, for example, can form spherical bilayer structures. These may be made of just a single bilayer – rather like a cell – and are known as *vesicles*. Spherical structures made up of multiple bilayers can also be formed. We can summarize by noting that the richness of surfactant phase behaviour is due to the combination of packing constraints and the free energy changes associated with hydrophobe–water interactions, the oil–water interface and head group interactions, which include charge for ionic surfactants – much easier to study than to model!

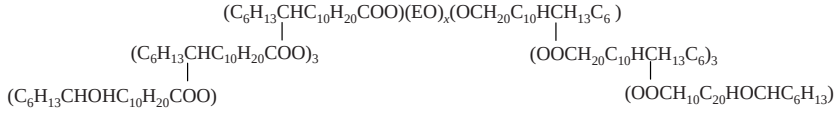
2.9.3 Macromolecular Surfactants

This group of surfactants consists of large, mostly non-ionic molecules with a molecular weight in the range 1×10^3–3×10^3 Da. A variety of structures are produced, with the various types and their uses given in a review by Hancock [27], which includes, in addition, much information on short chain non-ionics. Just like short-chain surfactants, the molecules consist of a hydrophobic part and a hydrophilic part. Poly(ethylene oxide) chains are a common choice for the hydrophile as this polymer has good water solubility at molecular weights $>10^3$ Da. On the other hand, poly(propylene oxide) has poor water solubility at this molecular

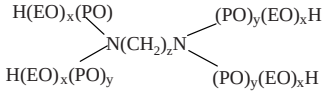


A-B block copolymer 'Synperonic' NPE 1800

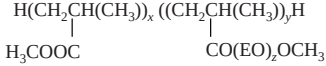
Linear A-B-A block copolymer 'Pluronic'



Branched A-linear B-A block copolymer poly(12-hydroxystearic acid)-co-poly(ethylene glycol)



Linear A-branched B block coploymer 'Tetronic'



Comb block copolymer [poly(methyl methacrylate)-*co*-poly(ethylene glycol)]

Figure 2.24 The structure of some macromolecular surfactants [27]. EO, ethylene oxide; PO, propylene oxide. Reprinted with permission from *Surfactants*, ed. Th Tadros, Macromolecular surfactants by R. I. Hancock, pp. 287–321, Academic Press, London, Copyright (1984) Elsevier

weight and is therefore a useful choice as a hydrophobe, either alone or in conjunction with other hydrophobic molecules. The structures of some commercially important macromolecular surfactants are shown in Figure 2.24. The structures can vary from simple linear block structures to sophisticated branched structures resembling brushes or combs.

2.9.4 Choices of Surfactants for Applications

This is a major problem to most workers in the colloids field, especially when the choice is not limited to those with government approval for the particular product. Hancock [27] gives a useful review of the uses of small and large non-ionic surfactants in a wide variety of applications. However, some general observations are appropriate at this point. Short-chain anionic surfactants are very widely used as stabilizers and wetting agents. These molecules are mobile, adsorb on hydrophobic surfaces

and form stable films. The sulfates and sulfonates are less vulnerable to the calcium ions in hard water than the potassium salts of fatty acids, where the calcium binds strongly to the carboxyl groups and precipitates the calcium salt as the familiar 'scum' seen when using such soaps with 'hard' water. These materials are widely used as stabilizers and cleaning agents. The cationics interact with biological membranes, which are invariably negatively charged and are useful as germicides. Coco-betaines and similar ionic materials have relatively low biological activity and can be found in personal care products such as shampoos. Short-chain non-ionics form more mobile, less rigid, films than anionics of similar chain length and are useful in low-foaming cleaning systems. The low rigidity of the surfactant films allows the thermal fluctuations to become large enough to aid collapse of the films. Rigid films damp out this motion and are therefore more stable. These surfactants are also useful as stabilizers in systems which have too high an electrolyte concentration for charge stabilization to work adequately.

Macromolecular surfactants do not make good wetting agents or emulsifiers. For this purpose, we require small molecules which diffuse rapidly and stabilize new interfaces by adsorption, building charge density and lowering interfacial tensions in the process. However, macromolecular surfactants make excellent stabilizers of emulsions or solid particles. The lower mobility of the molecules and the large moieties which are insoluble in the continuous phase mean that they are reluctant to leave the interface. Typically, the molecular weights of the soluble and insoluble parts of the molecules are similar, which maximizes the interfacial concentration of the stabilizing elements. We should also note that these large molecules can show some self-assembly tendencies, which can result in multilayer formation at the interfaces, which results in even greater stability. Poly(12-hydroxystearic acid)-*co*-poly(ethylene glycol) is a good example of this. The surfactant is an excellent stabilizer of high phase volume water-in-oil emulsions such as can be found in cosmetic moisturizing creams. The water droplets can be surrounded by a trilayer of surfactant in some emulsions. In addition, the macromolecular materials are usually less sensitive to temperature or electrolyte levels, which enables them to be used under a wide variety of conditions.

2.9.5 Proteins at Surfaces

Large protein molecules should also be thought of as macromolecular surfactants, albeit with special properties, as it is typical of biological

systems to have multiple roles. Proteins have a stabilizing function in naturally occurring systems, with milk being a good example in which β-casein species stabilize fat droplets. Of course, proteins are also widely employed as stabilizers in processed foods. However, it is not just in foodstuffs and pharmaceuticals that they have application. They have also been used as stabilizers for inorganic particles in paints, inks and photographic film, for example.

Adsorption on a hydrophobic surface occurs by the attachment of the hydrophobic regions and may or may not involve the opening out of the protein structure, that is, denaturing of the protein, as it adsorbs [28]. Some proteins are relatively unstructured (such as β-casein), whereas globular proteins have much more structure which is little changed by the adsorption process. The attachment occurs via a large number of segments and, like synthetic polymers, desorption is a very slow (even unlikely) process, although if the system is challenged by the addition of small, mobile surfactants, desorption can be induced. These are mobile enough to adsorb on surface sites from where a protein segment has temporarily desorbed, thus preventing reattachment. With an adsorption energy of $\sim 0.5 k_B T$ [28], we should visualize the adsorption at each site as being a dynamic process but, with attachment at several hundred sites, this will not lead to the loss of the whole molecule from the surface unless each site is immediately occupied by a competitor.

Just as adsorption equilibrium is a slow process with synthetic polymers, so it is with proteins. Rapid stirring, of course, limits the diffusion process to the movement across the laminar boundary layer of fluid very close to the surface, but the approach to the final conformation and packing density can be a slow process. In the final state, the stabilizing protein film provides a very robust form of stabilization in the form of electrosteric stabilization (see Chapter 3). There is an electrostatic component, although this is sensitive to pH changes as the major component is due to carboxyl groups. This is enhanced by a steric component which resists the local increase in concentration of the stabilizing moieties as two surfaces approach, in addition to moving the origin of the electrostatic component away from the surface. Finally, the protein films are viscoelastic and this damps out the thermal interfacial fluctuations, which produce coalescence of emulsions and foams.

Biodegradability is an additional advantage when using proteins as stabilizers, although this may be too rapid and toxicity is less likely to be a problem. The aliphatic alcohol ethoxylates are viewed as being acceptable in their biodegradability behaviour, unlike the nonylphenol

ethoxylates. The toxicity of both to fish and other species, however, can be a problem in some instances.

REFERENCES

1. Peebles, L. H. (1971) *Molecular Weight Distribution in Polymers*, Wiley-Interscience, New York.
2. Flory, P. J. (1953) *Principles of Polymer Chemistry*, Cornell University Press, Ithaca, NY.
3. Flory, P. J. (1969) *The Statistical Mechanics of Chain Molecules*, Wiley-Interscience, New York.
4. Yamakawa, H. (1971) *Modern Theory of Polymer Solutions*, Harper and Row, New York.
5. deGennes, P.-G. (1979) *Scaling Concepts in Polymer Physics*, Cornell University Press, Ithaca, NY.
6. Doi, M. and Edwards, S. F. (1986) *Dynamics of Polymer Solutions*, Oxford University Press, Oxford.
7. Flory, P. J. (1970) *Discuss. Faraday Soc.*, **49**, 7.
8. Scheutjens, J. M. H. M. and Fleer, G. J. (1979) *J. Phys. Chem.*, **83**, 1619.
9. deGennes, P.-G. (1980) *Macromolecules*, **13**, 1069.
10. deGennes, P.-G. (1987) *Adv. Colloid Interface Sci.*, **27**, 189.
11. Bodanecky, M. and Kovar, J. (1982) *Viscosity of Polymer Solutions*, Elsevier, Amsterdam.
12. Brandrup, J. and Immergut, E.H. (eds) (1989) *The Polymer Handbook*, 2nd edn, John Wiley & Sons, Inc., New York.
13. Kerker, M. (1969) *The Scattering of Light and Other Electromagnetic Radiation*, Academic Press, New York.
14. Huglin, M.B. (ed.) (1972) *Light Scattering from Polymer Solutions*, Academic Press, New York.
15. Poirier, Y., Dennis, D. E., Klomparens, K. and Somerville, C. (1992) *Science*, **256**, 520.
16. Chaplin, M., www.lsbu.ac.uk/water, last accessed May 2009.
17. Sperry, P. R. (1984) *J. Colloid Interface Sci.*, **99**, 97.
18. Faers, M. A. (2003) *Adv. Colloid Interface Sci.*, **106**, 23.
19. Gelderman, K. A., Thomlinson, S., Ross, G. D. and Gorter, A. (2004) *Trends Immunol.*, **25**, 158.
20. Jobling, S. (2004) *Curr. Opin. Plant Biol.*, **7**, 210.
21. Giannouli, P. and Morris, E. R. (2003) *Food Hydrocolloids*, **17**, 495.
22. Shinoda, K. (ed.) (1967) *Solvent Properties of Surfactant Solutions*, Marcel Dekker, New York,.
23. Laughlin, R. G. (1994) *The Aqueous Behaviour of Surfactants*, Academic Press, New York.
24. Everett, D. H. (1988) *Basic Principles of Colloid Science*, Royal Society of Chemistry, London.
25. Lindman, B., Wennerström, H. and Eicke, H. F. (eds) (1980) *Micelles*, Springer, Heidelberg.

26. Mukerjee, P. and Mysels, K. J. (1971) *Critical Micelle Concentrations of Aqueous Surfactant Systems*, NSRDA-NBS-36, US Government Printing Office, Washington, DC.
27. Hancock, R. I. (1984) Macromolecular surfactants, in *Surfactants* (ed. Th. Tadros), Academic Press, London, pp. 287–321.
28. Dickenson, E. (1992) *Introduction to Food Colloids*, Oxford University Press, Oxford.

3

Interactions Between Colloidal Particles

3.1 INTRODUCTION

In Chapter 1, we started our discussion of colloidal systems with the influence of size on the timescale of the motion of the primary components, that is, the particles. In this present chapter, we will extend that discussion to how the particles interact with each other and how these interactions result in the various structures found in colloidal systems. This means that the focus will be on describing the energy of the interactions. Usually, we will be thinking in terms of a potential energy and not be too concerned about kinetic energy. It is also usual to think in terms of Newtonian mechanics. We can visualize two particles close to each other, say with a centre-to-centre separation r, and ask how much work we would have to do to separate them to some large distance apart. This is the potential energy of the particle–particle interaction and it is termed the *pair potential*, $u(r)$. This is calculated from how much force is required to move the particles. So, if the force at distance r is $f(r)$, then we obtain the pair potential from the following:

$$u(r) = -\int_{r}^{\infty} f(r)\mathrm{d}r \tag{3.1a}$$

Colloids and Interfaces with Surfactants and Polymers, Second Edition James Goodwin

and of course

$$\frac{\mathrm{d}u(r)}{\mathrm{d}r} = -f(r) \tag{3.1b}$$

Therefore, the particles can be thought of as though they are interacting via 'colloidal springs', and it is the nature of these springs that we need to describe. Let us also recall that the spring constant (or modulus) is the rate of change of the force with the distance:

$$\frac{\mathrm{d}^2u(r)}{\mathrm{d}r^2} = y(r) \tag{3.2}$$

It is easier to deal with particles in an equilibrium state, but we should never forget the timescales as the systems that we use every day may be far from their equilibrium state. Formally, the interaction energy is a free energy which should include both the enthalpic and entropic contributions of *all* the components in the system, whether particles, solvent molecules, small ions, surfactant molecules or polymer molecules. This means that any expression for the interaction energy between any two particles would take into account an average contribution from all the other components. This is the *potential of mean force*. There are two approaches that are possible to estimate this: we may determine it from the equilibrium structure of the colloidal system, or we can produce a mathematical model of the system. The former route may be possible in some simple, idealized systems, but the latter is very difficult to achieve. Therefore, what we will do is to try to estimate the various contributions to the potential and then make the assumption that they can be added to give the total potential. In order to achieve this, we will use relatively simple models which have enough information to bring us close to where we wish to go and give an adequate description of what we observe. It is important to remember that these are simple models and we should not be too surprised if they only agree with experiment under limited conditions. Often, though, they will adequately serve our purpose.

There is an excellent text by Israelachvili [1] that gives a detailed outline of both intermolecular forces and the interactions between particles. In this chapter we shall just be working with the salient features and if the interested reader should need more fine details, the cited text will supply them. A number of contributions to the interaction between particles can be identified. Some of the interparticle forces work to bring particles into close contact, whereas others act to separate them. As we

shall see, it is the interplay between these that result in the final state of our dispersions. If we understand the origins of these forces, it becomes possible to modify them by control of the chemical environment or the chemical architecture of the components of the dispersion. Each of the major contributions to the net or total interparticle force can be described using physical models. These can become complex, but they are also centred around fairly simple initial models. This makes it at once easy to understand the underlying concepts even if the full manipulation becomes difficult, but then the application to the complicated systems that we wish to use can appear a daunting prospect. However, the general principles apply and, although an accurate predictive calculation may not be within our reach for a particular product, the choice of experiments will be much better focused.

3.2 INTERMOLECULAR ATTRACTION

The starting point for our discussion must be an examination of the forces that occur between molecules. Some of these interactions are strong and hence are long-lived, such as the covalent bond or the weaker, more transient, hydrogen bond. The origin of the attraction between particles does not lie in these interactions but in the weaker interactions that are often referred to as van der Waals interactions. These are the forces that account for the non-ideality of gases and account for the deviations from the simple behaviour described by the ideal gas equation:

$$\frac{PV}{T} = NR = nk_B \tag{3.3}$$

where P is the pressure, V is the volume, T is the temperature, n is the number of gas molecules in the volume, or N if we use the number of moles, and k_B is the Boltzmann constant. There are several interactions that can occur which are electrodynamic in origin, and the traditional description separates these into three distinct interactions. We will discuss the origins of these but only at a level sufficient to enable us to see the underlying mechanism for the attraction.

3.2.1 Keesom Interaction

It is not much of a surprise to find that the molecules of a gas such as hydrogen chloride have a strong permanent dipole moment due to the

polarization of the covalent bond. The dipoles tend to align and this will be the preferred arrangement. This dipole–dipole attraction is known as the *Keesom interaction* and we may write the interaction free energy at an intermolecular distance r as follows:

$$u(r)_{\mathrm{K}} = -\frac{C_{\mathrm{K}}}{r^6} \tag{3.4}$$

where C_{K} is a constant which depends on the particular type of molecule being considered. For example, if we have two identical molecules of dipole moment μ:

$$C_{\mathrm{K}} \propto \mu^4 \tag{3.5}$$

Because of the marked dipole alignment, the rotational motion of the molecules is restricted and we should be thinking of this as a 'long-time' interaction, that is, a low-frequency interaction.

3.2.2 Debye Interaction

This is the type of interaction that occurs between a polar and a non-polar molecule. The dipole on the polar molecule polarizes the electron cloud of the non-polar molecule. The molecular rotation is still occurring and therefore we could think here of frequencies associated with the interaction as those in the microwave region. The interaction free energy for this dipole–induced dipole interaction can be described by a similar expression to that used for the dipole–dipole interaction (Equation 3.4). In this case, the polarizability of the non-polar molecule is a key feature. It should also be recognized that even for two polar molecules with different dipole moments, the net dipoles will be affected by interaction with the adjacent molecules. The equation given illustrates this for molecules 1 and 2:

$$u(r) = -\frac{C_{\mathrm{D}}}{r^6} \tag{3.6}$$

and we note that the interaction constant, C_{D}, for a system consisting of two different molecules has the following dependence:

$$C_{\mathrm{D}} \propto \left(\alpha_2 {\mu_1}^2 + \alpha_1 {\mu_2}^2\right) \tag{3.7}$$

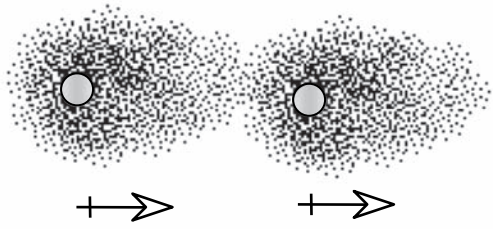


Figure 3.1 The coupling of oscillating atomic dipoles which form the basis of the London (or dispersion) force between atoms

We are using an example with two permanent dipoles that will orient to maximize the interaction and the timescales will again be long.

3.2.3 London or Dispersion Interaction

This describes the interaction that results in attraction between non-polar molecules. It is due to the movement of the electron cloud around the atomic nucleus resulting in a dipole that fluctuates. When two atoms come into close proximity, the temporary dipoles become aligned, that is, the fluctuations become coupled and this is a preferred (or lower) energy state. A schematic illustration is shown in Figure 3.1.

The range of the interaction is similar to the two discussed previously, but the timescale of the fluctuations is now that of the electronic transitions and we should think towards the visible–ultraviolet part of the electromagnetic spectrum. The interaction energy can be written as

$$u(r) = -\frac{C_L}{r^6} \tag{3.8}$$

Now, the London constant, C_L, for two similar atoms is proportional to the ionization energy of the outer electrons, $h\nu_I$, where h is Planck's constant, and the polarizability, α, as follows:

$$C_L \propto h\nu_I\alpha^2 \tag{3.9}$$

and for two different types of molecule:

$$C_L \propto h\frac{\nu_{I1}\nu_{I2}}{\nu_{I1} + \nu_{I2}}\alpha_1\alpha_2 \tag{3.10}$$

It turns out that this is a particularly important type of intermolecular interaction, because it is much larger than the Keesom or Debye contributions in nearly all cases of colloidal materials. Water is a notable exception, with the dispersion interaction contributing only one-quarter of the total. An important feature of this interaction is that there is only a weak tendency to change the orientation of neighbouring molecules. In the solid phase this is important as the fluctuations can still couple without requiring motion of the molecules to maximize their alignment, as would be the case for molecules with permanent dipoles.

Let us now consider the implications of this attractive interaction with the *n*-alkane series. These have particular relevance to many colloidal systems as many surfactant systems have a linear aliphatic chain as one element of their composition. This element has a tendency to self-assemble at interfaces where the chains come into close proximity. The cohesive energy of the solid hydrocarbon is estimated from the measured latent heats of melting and vaporization. Here, we will use the data given by Israelachvili [1] and use his simple model calculation. Figure 3.2 shows the value of the molar cohesive energy as $-CH_2-$ groups are added to the chain. The slope gives the value per additional $-CH_2-$ as $6.8\,\mathrm{kJ\,mol^{-1}}$. This is equal to $\sim 2.75 k_B T$ at 25 °C, showing that these non-polar molecules are fairly 'sticky'. Taking each chain surrounded by six nearest neighbours and summing the interactions over the neighbouring groups up and down the chain, and also around, gave a value of $\sim 6.9\,\mathrm{kJ\,mol^{-1}}$.

This simple result is a very important illustration of how the dispersion forces can apparently be additive to a good approximation. This will lead us to consider the behaviour of the relatively large groups of molecules that make up colloidal particles. However good this approximation appears to be in some simple cases, it should be remembered that the interaction of each molecule is affected by all of the neighbouring fields so that pairwise additivity is only an approximation.

3.2.4 A More Generalized Approach

All of the three interactions describe above are based on the attraction between dipoles and have the same distance dependence. Therefore, we can expect that we could handle these in one general description. The London constant in Equations 3.9 and 3.10 is written as dependent on the ionization of the outer electrons. Other electronic transitions also take place so that contributions at other frequencies can also occur.

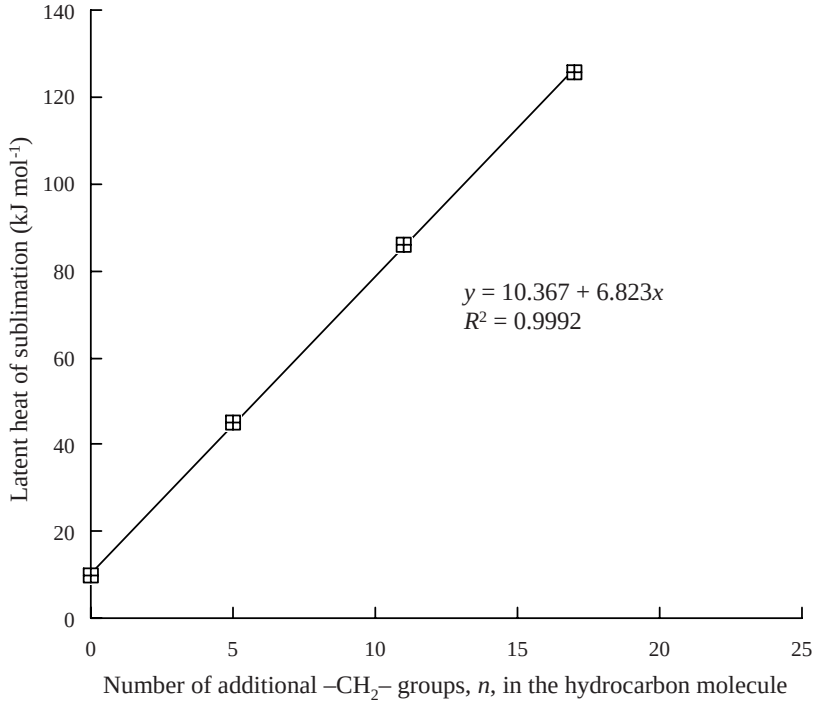


Figure 3.2 The latent heat of sublimation for a series of linear alkanes, $CH_3{-}(CH_2)_n{-}H$ [1]. This figure shows the mathematical relationship for the straight line drawn through the experimental points. *R* is known as the *correlation coefficient* and provides a measure of the 'quality' of the fit'. In fact, R^2 (the *coefficient of determination*) is also used because it is more sensitive to changes. This varies between –1 and +1, with values very close to –1 or +1 pointing to a very tight 'fit' of the data [1].

Therefore, we must think in terms of the full range of frequencies ranging from those of a few hertz up to the ultraviolet region at $\sim 10^{16}$ Hz. Furthermore, we have only considered pairs of atoms or molecules interacting in the absence of any intervening medium. The *dielectric constant* or *relative permittivity* of the medium is the important factor here. If a sphere is placed in a medium of the same permittivity and an external field is applied to the system, the sphere will not be polarized. If it has a dielectric constant which is either larger or smaller than the medium, then it will be polarized. This effect is the basis for electrorheological fluids whose solid–liquid properties are readily and rapidly controlled by the application of an electric field. In order to achieve the high fields necessary for strong effects, the choice is to use a low

dielectric constant for the medium with a higher value for the material making up the particles. It is the difference in dielectric constant which is important and we should think in terms of an 'excess polarizability' [2]. It is important to recognize that for the case of permanent dipoles there is an alignment contribution to the polarizability in addition to the electronic component. It is only the latter, of course, that we need to consider with non-polar molecules but the full spectral range would need to be included for full prediction. The dielectric behaviour as a function of frequency to give $\varepsilon(\nu)$ is a tractable experiment at lower frequencies and the refractive index, $n(\nu)$, is a viable measure of the dielectric behaviour at the higher end of the spectrum [recall that $\varepsilon(\nu) = n^2(\nu)$].

The general description of the interaction was given by McLachlan [3] as a summation of over the range of interaction frequencies for molecules 1 and 2 interacting in a medium 3, as follows:

$$u(r) = -\left\{ \frac{3k_B T \alpha_1(0)\alpha_2(0)}{[4\pi\varepsilon_3(0)\varepsilon_0]^2 r^6} + \frac{6k_B T}{(4\pi\varepsilon_0)^2 r^6} \sum_{n=1}^{\infty} \frac{\alpha_1(i\nu)\alpha_2(i\nu)}{\varepsilon_3^2(i\nu)} \right\} \tag{3.11}$$

The summation is carried out over all frequencies so that all the interactions are captured. The first term on the right-hand side of the above equation is the 'zero frequency' term, values for which can usually be found in data tables. The second term on the right-hand side contains the contributions for all other frequencies, with the steps in terms of $k_B T$:

$$n = \frac{\nu_n h/2\pi}{k_B T} \tag{3.12}$$

However, this term uses the frequency dependence of the polarizability at complex frequencies, $i\nu$. Here, i indicates the *quadrature* or *imaginary* component of the frequency, which is the dissipative component as we are looking at the interaction of the oscillating electromagnetic field with the molecules, and not the transmitted component of the field.

3.3 NOTES ON COMPLEX NUMBER MANIPULATION

We need to use complex number notation whenever we wish to describe the behaviour of oscillating fields. This is something that we have to do

frequently in physical science. For example, we will use this in Chapter 9 when dealing with rheology, where, for example, we describe what happens when we apply an oscillating force mechanically to a concentrated colloidal system. There, we use the algebra to separate the elastic storage of work done from that dissipated by viscous flow. (The response is similar to what we would obtain from oscillating a damped spring such as that used in an automotive suspension unit.) Another common example is that of an AC electric circuit with a capacitor and a resistor in series. At zero frequency (DC voltage), we can store electrical energy in the capacitor and we would have little interest in the resistor. As we increase the frequency of the AC voltage, we store less in the capacitor and dissipate energy in the resistor. Let us come now to examples closer to the subject of this chapter and consider electromagnetic radiation. In the usual science courses, we learn about spectra and how the applied radiation interacts with matter. We are familiar with the fact that light can pass through a solution, but we may find the intensity reduced at some particular frequencies, perhaps in the UV, the IR and the microwave regions. However, at the same time we talk of a refractive index. What we mean here is what we would measure by transmission at some frequency. It is, in fact, a complex number in which we should include dissipative terms in addition to storage, just like the AC circuit. The same applies to dielectric constants. We are referring to the static (low-frequency plateau) value. We can carry out the measurements at say 10^3 Hz and at higher frequencies we find interesting behaviour as polar molecules respond and give dissipative contributions as the timescales become too short for their motion.

Initially, the use of complex number algebra appears much more difficult to many scientific workers than is actually the case. The greatest barrier is the nomenclature, such as when we say the number $P = x + iy$, where $i^2 = \sqrt{-1}$ and that x is the *real* part of P and y is the *imaginary* part! All we have to remember is that the ordinary rules of algebraic manipulation apply and that whenever we arrive at i^2 we write –1 and also remember that $1/i = -i$. Even this looks a little intimidating without some practice. It is useful just to think of the notation as a simple way of writing down how we are dividing P, the numerical value that we measure, into two contributions, x and y. Figure 3.3 shows this in graphical form. P is given in both Cartesian coordinates and polar coordinates.

To illustrate how this works, let us consider the polarizability, which is a function of frequency. We can write the total polarizability as the sum of any dipole (permanent or induced) alignment term, $\alpha_a(\nu)$, and the electronic polarizability of a molecule, $\alpha_e(\nu)$, which is responsible for

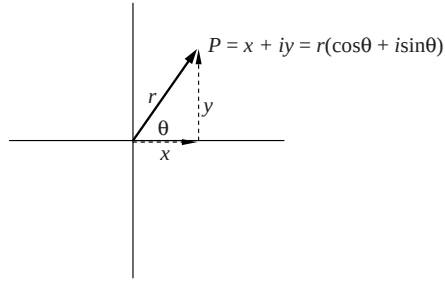


Figure 3.3 Illustration of complex number notation

the dispersion force [1, 4] as follows:

$$\alpha(\nu) = \alpha_{\mathrm{a}}(\nu) + \alpha_{\mathrm{e}}(\nu) \tag{3.13a}$$

which may be written in more detail as

$$\alpha(\nu) = \frac{\mu^2}{3k_{\mathrm{B}}T\,(1-i\nu/\nu_{\mathrm{rotation}})} + \frac{\alpha_{\mathrm{e}}}{1-(\nu/\nu_{\mathrm{I}})^2} \tag{3.13b}$$

where ν_{rotation} is the rotational frequency of the dipolar molecule and ν_{I} is the ionization frequency (so here we assuming that there is only one absorption frequency, ν_{I}).

The quadrature (or imaginary) component is

$$\alpha(i\nu) = \frac{\mu^2}{3k_{\mathrm{B}}T\left(1-i^2\nu/\nu_{\mathrm{rotation}}\right)} + \frac{\alpha_{\mathrm{e}}}{1-(i\nu/\nu_{\mathrm{I}})^2} \tag{3.13c}$$

and for the range of n frequencies we obtain

$$\alpha\,(i\nu_n) = \frac{\mu^2}{3k_{\mathrm{B}}T\,(1+\nu_n/\nu_{\mathrm{rotation}})} + \frac{\alpha_{\mathrm{e}}}{1+(\nu_n/\nu_{\mathrm{I}})^2} \tag{3.13d}$$

Therefore, the utilization of Equation 3.11 should appear less daunting as the frequencies are just those where we observe adsorption. Consider now the situation where we have two similar non-polar molecules interacting in a gas. Methane would be a good example. Equation 3.11

becomes ($\varepsilon_3 = 1$ as the interaction is across a vacuum)

$$u(r) = -\left[\frac{3k_B T\alpha_{CH_4}{}^2(0)}{(4\pi\varepsilon_0)^2 r^6} + \frac{6k_B T}{(4\pi\varepsilon_0)^2 r^6}\sum_{n=1}^{\infty}\alpha_{CH_4}{}^2(i\nu)\right] \tag{3.14}$$

We can replace the summation by an integration. Using Equation 3.12, we have

$$dn = (h/2\pi k_B T)\, d\nu \tag{3.15}$$

and so

$$u(r) = -\left[\frac{3k_B T\alpha_{CH_4}{}^2(0)}{(4\pi\varepsilon_0)^2} + \frac{3h}{(4\pi\varepsilon_0)^2\pi}\int_{\nu_1}^{\infty}\alpha_{CH_4}{}^2(i\nu)d\nu\right]\frac{1}{r^6} \tag{3.16}$$

In Equation 3.16, when frequencies approach the visible range, the second term on the right-hand side becomes dominant. At frequencies below the microwave region, only the first term on the right-hand side is important and is, of course, a constant. Methane has no permanent dipole so the frequency-dependent polarizability from Equation 3.13d is

$$\alpha_{CH_4}(i\nu) = \frac{\alpha_{eCH_4}}{1 + (\nu/\nu_I)} \tag{3.17}$$

Substituting this into Equation 3.16 and integrating yields the London dispersion result as we are just using the ionization frequency:

$$u(r) \approx -\left[\frac{3h}{(4\pi\varepsilon_0)^2}\alpha_{eCH_4}{}^2\nu_I\right]\frac{1}{r^6} \tag{3.18}$$

Hence we may write the interaction energy with the London constant as follows:

$$u(r) = -\frac{C_L}{r^6} \tag{3.19}$$

The electronic polarizability for methane is $2.89 \times 10^{-40}\,C\,m^2\,J^{-1}$ and the ionization frequency is $3.05 \times 10^{15}\,s^{-1}$, which give a value of $C_L = 1.02 \times 10^{-77}\,J\,m^6$.

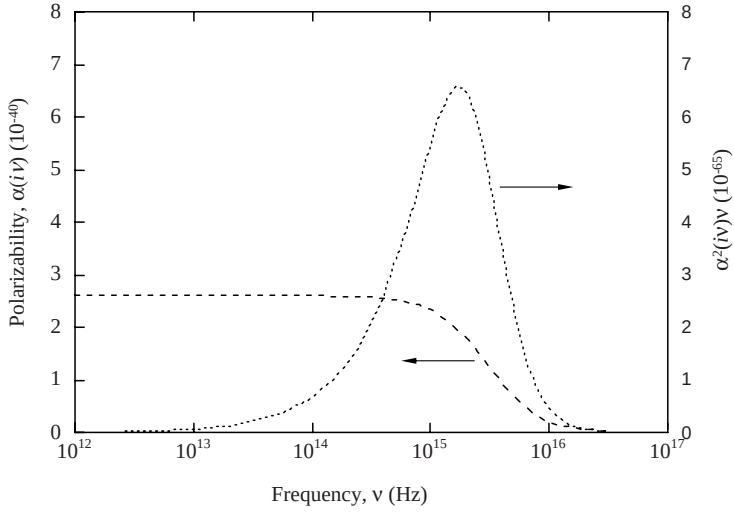


Figure 3.4 Variation of the frequency dependence of the quadrature component of the polarizability of methane

A plot of $\alpha_{CH_4}{}^2(i\nu)\nu$ as a function of ν is given in Figure 3.4 and this shows the spectral region which dominates the attraction between the methane molecules. Note how it is centred in the ultraviolet region where we are accustomed to seeing electronic excitations.

3.4 DISPERSION FORCES BETWEEN PARTICLES

We now have some understanding of how atoms and molecules attract each other. In addition, we have indicated that considering the dispersion forces to be additive can be a reasonable approximation as we start to work with larger molecules. This clearly has importance when we are going to consider surfactants and polymers, but it may seem surprising that we can extend this approximation to describing the interaction between two particles. This approach laid the foundation for our understanding of why particles coagulate and was pioneered by workers such as de Boer [5], Hamaker [6], Derjaguin [7] and Langbein [8] using London's approximation based on the interaction at a frequency around the ionization frequency. This is where we will begin.

The starting point is to consider the interaction between a single molecule and a slab of material and then extend that to two slabs interacting. As we are assuming that the energies are additive, we simply

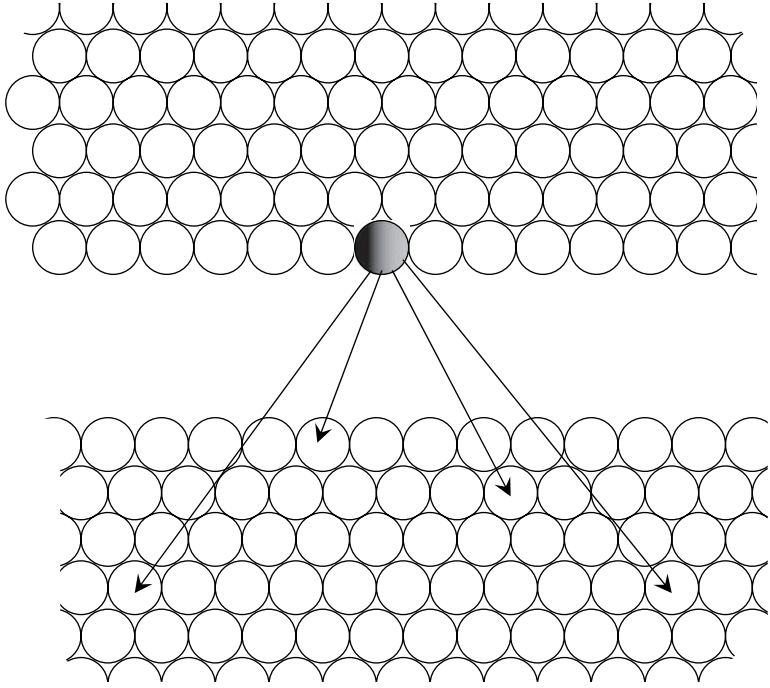


Figure 3.5 The summation of the atomic dipolar interactions between two slabs of material

use Equation 3.19 and add up the interactions between our reference molecule and all the molecules in the slab of material (illustrated in Figure 3.5) Therefore:

$$\text{interaction energy} = \text{sum for molecule with each in lower slab}$$

Next, we add this up for all the molecules in the upper slab, so we now have:

$$\text{interaction} = \text{sum for all moleculess in upper slab} \times \text{sum for molecule with each in lower slab}$$

The problem is that to calculate r for each interaction we would need to know the detailed structure. It is simpler to use the number density of molecules in the slab, ρ_N, and to integrate over the volumes. This 'semi-continuum' approach is in the spirit of the additivity assumption and would only become a problem at very close separations. This results

in the energy as follows:

$$\text{interaction energy} = C_L {\rho_N}^2 \times (\text{a geometric term})$$

and we have a general approach which has the characteristic of the material, that is, the electronic polarizability, the ionization frequency and the square of the product of the density and the molar mass as a front factor to a term based on the shape of the material and the separation distance between the two bodies. Depending on the shape of the two particles, solving the integrals can be non-trivial, but it is essentially a problem of geometry and calculus and not one of physics. Examples of the algebraic manipulations are given in the literature [1, 9, 10], although the purpose of this text is just to illustrate the underlying mechanisms of colloid science. So, for example, for two slabs which we can consider to be infinitely thick, the energy of unit area of one slab interacting with the whole of the other slab is

$$V_A = -\frac{\pi C_L {\rho_N}^2}{12H^2} \tag{3.20}$$

where V_A is used to represent the dispersion energy between two slabs a distance H apart and is the energy per unit area of surface. The numerator is the material property and the denominator arises from the geometry. It is common to express the material properties as a single material constant, the *Hamaker constant*:

$$V_A = -\frac{A_{11}}{12\pi H^2} \tag{3.21}$$

where $A_{11} = \pi^2 C_L {\rho_{N1}}^2$. The subscript 11 indicates the interaction is between two slabs of the same material. So, is we had a slab of material 1 interacting with material 2, the Hamaker constant notation would indicate this, namely:

$$A_{12} = \pi^2 C_{L12} \rho_{N1} \rho_{N2} \tag{3.22}$$

If the 'semi-infinite' slabs are replaced by plates of finite thickness, t, Equation 3.21 has a slightly more complicated form:

$$V_A = -\frac{A_{11}}{12\pi}\left(\frac{1}{H^2} + \frac{1}{(H+2t)^2} - \frac{2}{(H+t)^2}\right) \tag{3.23}$$

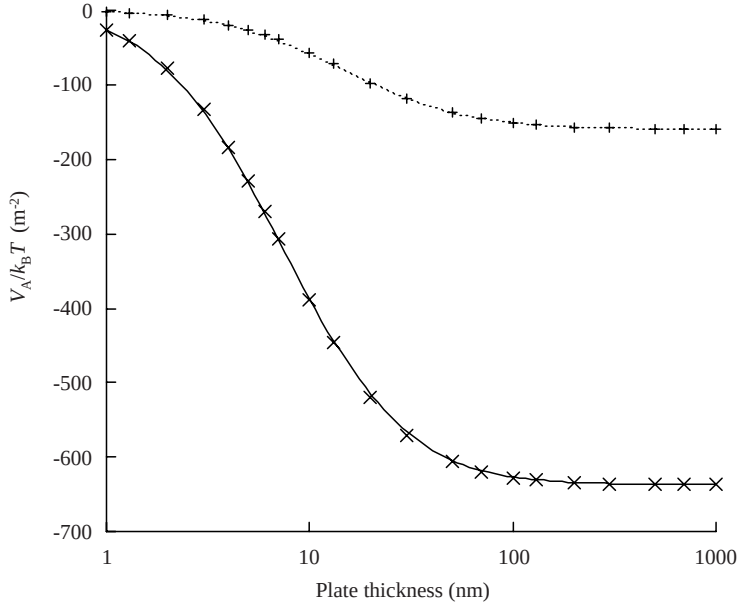


Figure 3.6 Variation of the energy of attraction with the thickness of two plates, using Equation 3.23: ×, separation = 10 nm; +, separation = 20 nm

Equation 3.23 was used to calculate the attractive energy for two small plates as a function of thickness; a value of $A_{11} = 10^{-19}$ J was used and the plate area used was 0.1 μm^2. These values would be similar to those found for clay particles. It is clear from Figure 3.6 that the finite thickness of the plate is only important for thin plates but also that the variation of the interaction with distance is greater for thin plates.

However, we still have the energy of attraction equal to the product of a material constant and a geometric term that has been derived from the integration over the particle volumes. This is the case whatever the shape of the particles. For example, if we have two identical spheres of radius a and centre-to-centre distance r, then when the spheres are close together the attractive energy is

$$V_A(r) = -\frac{A_{11}a}{12(r-2a)} \tag{3.24a}$$

or

$$V_A(H) = -\frac{A_{11}a}{12H} \tag{3.24b}$$

with $H = (r - 2a)$ and $H \ll a$, H being the closest distance between the surfaces of the spheres. This is the total energy of interaction and not an energy per unit area of surface. The more general expression for two spheres of radii a_1 and a_2 is [6]

$$V_A(r) = -\frac{A_{11}}{6}\left\{\frac{2a_1a_2}{r^2-(a_1+a_2)^2} + \frac{2a_1a_2}{r^2-(a_1-a_2)^2} + \ln\left[\frac{r^2-(a_1+a_2)^2}{r^2-(a_1-a_2)^2}\right]\right\} \tag{3.25a}$$

and if the spheres are the same radii:

$$V_A(H) = -\frac{A_{11}}{12H}\left[1 + \frac{H}{2a+H} + \frac{H}{a}\ln\left(\frac{H}{2a+H}\right)\right] \tag{3.25b}$$

The problem in calculating the dispersion interaction is the determination of the geometric term, which can become fairly complex for relatively simple geometric forms (see Figure 3.7). Some of these, such as parallel cylinders, crossed cylinders and a sphere and a plate (the last two result in the same geometric factor), are given in the literature [1, 9–11]. For example, when a spherical particle is close to a plate, we have

$$V_A(H) = -\frac{A_{12}a}{6H} \tag{3.26}$$

Inspection of Equations 3.24b and 3.26 shows that the interaction of a sphere with a thick plate has a geometric factor that is twice that of two similar sized spheres.

It should be noted that in Equation 3.26 we are considering that the material of the sphere is different from that of the plate – hence the notation A_{12}. This is an appropriate point to include consideration of the additional problem of when the particles are immersed in a continuous medium and what effect that will have compared with the case of two particles *in vacuo*. The effect of the particles being immersed in a medium is to reduce the net interaction. If, for example, the Hamaker constant of the medium approaches that of the particles, the attractive interaction approaches zero. We must think in terms of a *net* or *effective* Hamaker constant. There are approximate estimates that we can use to estimate the effective Hamaker constant (for example, see Hunter [9]). If we have particles of type 1 and 2 immersed in a medium 3, we can

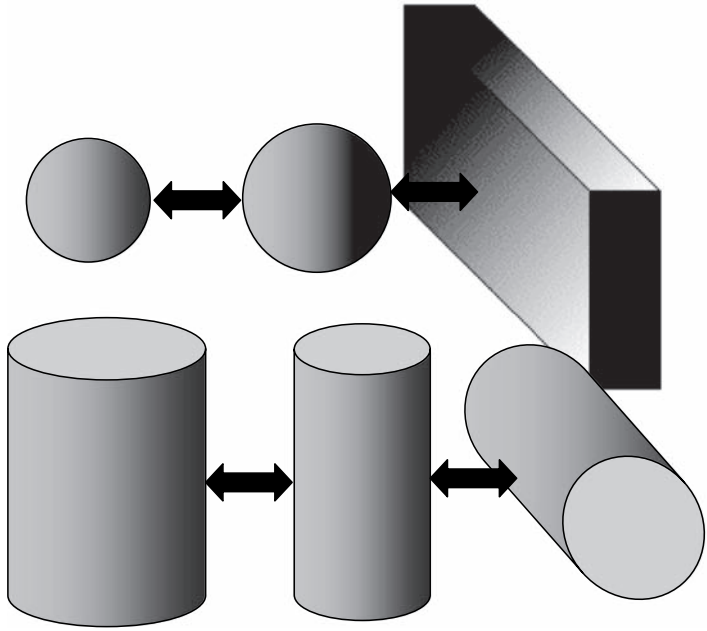


Figure 3.7 Illustrations of interactions between different geometric forms: sphere–sphere, sphere–plate, parallel cylinder and crossed cylinder. Note that the last is equivalent to the sphere–plate interaction if both cylinders have the same radius

write the effective Hamaker constant as the sum of the particle–particle and medium–medium interactions for identical volumes across vacuum minus the cross terms:

$$A_{132} = A_{12} + A_{33} - A_{13} - A_{23} \tag{3.27a}$$

If the two particles are the same material, this, of course, becomes

$$A_{13} = A_{11} + A_{33} - 2A_{13} \tag{3.27b}$$

Now, if we make the assumption that

$$A_{jk}^2 = A_{jj} A_{kk}$$

then we may write Equations 3.27a and 3.27b as

$$A_{132} \approx \left(A_{11}{}^{\frac{1}{2}} - A_{33}{}^{\frac{1}{2}} \right) \left(A_{22}{}^{\frac{1}{2}} - A_{33}{}^{\frac{1}{2}} \right) \tag{3.28a}$$

and

$$A_{131} \approx \left(A_{11}{}^{\frac{1}{2}} - A_{33}{}^{\frac{1}{2}}\right)^2 \tag{3.28b}$$

As an example of how these equations are relatively simple to use, let us consider the situation where we have oil droplets (say *n*-tetradecane), which are 2 μm in diameter, dispersed in water with the dispersion in a polystyrene (ps) container. Figure 3.8 shows the interactions between the droplets and the droplets and the container wall at close approach. In this figure the interactions were calculated using Equations 3.24 and 3.26 with the effective Hamaker constants from Equations 3.28 for the case of sphere–sphere and sphere–plate. This figure illustrates that first there is a strong attraction in both cases, even at distances of tens of nanometres, and second that the stronger attraction is between the oil droplets and the wall of the container. This is partially because the sphere–plate interaction is twice that of two similar spherical particles, but also because the net Hamaker constant for the interaction is larger. Although both are simple hydrocarbons, polystyrene is a denser molecule with lots of aromatic character giving different electronic polarizabilities.

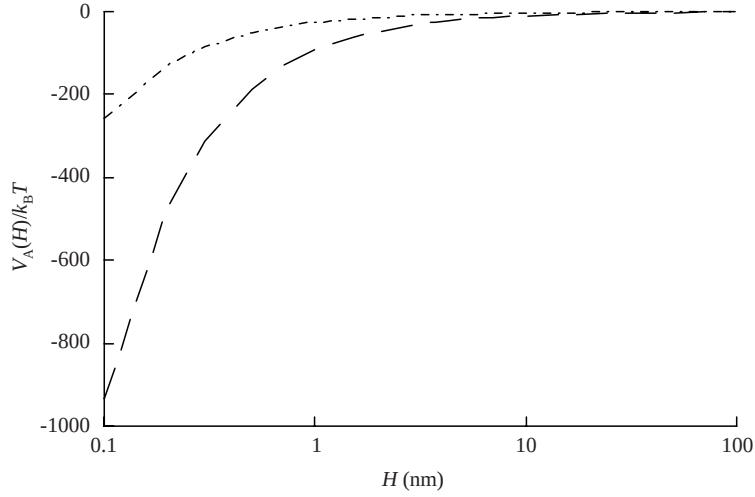


Figure 3.8 Attractive interactions between 2 μm spherical tetradecane droplets in water, both with each other and with the polystyrene container wall. (·····) $C_{14}H_{30}$ sphere–sphere; (– – –) $C_{14}H_{30}$ sphere–wall. $A_{C_{14}} = 5.2 \times 10^{-20}$; $A_{H_2O} = 3.7 \times 10^{-20}$; $A_{PS} = 6.6 \times 10^{-20}$

3.5 RETARDED DISPERSION FORCES

The range of the separation axis shown in Figure 3.8 is limited for several reasons. At close separations, the detailed surface structure becomes a limiting factor. We have assumed an atomically smooth surface. This can only be the case for a limited number of conditions such as fluid surfaces or solids such as carefully cleaved mica. Even with these surfaces there is a limit as the surfaces approach. This occurs as the electron orbital overlap becomes significant. At this point, the free energy becomes dominated by a very short-range repulsion. This is known as *Born repulsion* and is found as the inverse r^{12} dependence in the well-known Lennard–Jones–Devonshire equation [12]. The former limitation, however, is of even greater importance in most practical systems as we have often intentionally adsorbed material to mitigate the interaction and the outer periphery of any adsorbed layers will not be as well defined as the surface of the adsorbing particle. (The properties of the layer should also be taken into account in estimating the attraction between the surfaces.) The upper limit in Figure 3.8 is partially due to the restrictions for the use of the simplified form of Equation 3.24 instead of the full Hamaker expressions, Equation 3.25. However, there is another feature that we must discuss at this point. As the distance between the particles increases, the correlation of the oscillating atomic dipoles becomes poorer and the free energy decreases at a greater rate than we would have predicted due to the timescale of the field propagation compared with the timescale of the oscillation. This is termed *retardation* of the dispersion interaction.

The field propagation time is equal to the separation distance between the surfaces divided by the velocity of light in the intervening material. So, for example, if we consider two surfaces separated by 30 nm, the propagation time from one surface to the other and then back would be $\sim 10^{-16}$ s. Now, the frequency of the radiation where the strongest interactions are occurring is usually in the near-UV region and so the propagation time represents the time taken for a significant part of an oscillation to occur. In other words, the oscillations will no longer be exactly 'in phase' (i.e. the dipoles are no longer completely aligned) and the attraction is weakened. This weakening will become most marked for the higher frequency contributions. Hence we may conclude that we may have a good approximate description of the attraction between particles at close separation but at separations >30 nm the values calculated will overestimate the interaction. The indication is also that there should be a frequency dependence of this effect and a different approach might be sought.

3.6 THE GENERAL OR LIFSHITZ THEORY OF DISPERSION FORCES BETWEEN PARTICLES

There is an additional problem with particles that are dispersed in a medium such as an electrolyte solution, namely that the low-frequency contributions (i.e. in the extreme case the interaction of permanent dipoles) are heavily screened by the ionic environment and the low-frequency interaction falls off more rapidly with distance than would be the case in a vacuum. This indicates that we can expect to run into difficulties with the Hamaker model and should see if the more general approach that was introduced above will help us.

The starting point is to consider two slabs of material, 1 and 2, interacting across a medium 3. We need to think of each of these materials as slabs of a dielectric material with static dielectric constants of $\varepsilon_1(0)$, $\varepsilon_2(0)$ and $\varepsilon_3(0)$. The interaction energy is a material property (the Hamaker constant) multiplied by a geometric factor. Equation 3.22 defined the Hamaker constant for us and so for the two slabs acting across medium 3 we have

$$A_{132} = \pi^2 C_{\mathrm{L}132} \rho_{\mathrm{N}1} \rho_{\mathrm{N}2} \tag{3.29}$$

The London term, $C_{\mathrm{L}132}$, is given by Equation 3.11 (McLachlan's equation), which includes the polarizabilities of the molecules making up materials 1 and 2. The bulk polarisability of slab 1 in the medium 3 can be written in terms of the molecular properties of the material in 1 and the measurable macroscopic dielectric properties, as follows:

$$\rho_{\mathrm{N}1}\alpha_1(i\nu) = 2\varepsilon_0\varepsilon_3(i\nu)\frac{\varepsilon_1(i\nu) - \varepsilon_3(i\nu)}{\varepsilon_1(i\nu) + \varepsilon_3(i\nu)} \tag{3.30}$$

The frequency dependent Hamaker function is then [1]

$$A_{132} \approx \frac{3k_{\mathrm{B}}T}{4}\left[\frac{\varepsilon_1(0) - \varepsilon_3(0)}{\varepsilon_1(0) + \varepsilon_3(0)}\right]\left[\frac{\varepsilon_2(0) - \varepsilon_3(0)}{\varepsilon_2(0) + \varepsilon_3(0)}\right] + \frac{3h}{4\pi}\int_{\nu_1}^{\infty}\left[\frac{\varepsilon_1(i\nu) - \varepsilon_3(i\nu)}{\varepsilon_1(i\nu) + \varepsilon_3(i\nu)}\right]\left[\frac{\varepsilon_2(i\nu) - \varepsilon_3(i\nu)}{\varepsilon_2(i\nu) + \varepsilon_3(i\nu)}\right]\mathrm{d}\nu \tag{3.31}$$

The above equation indicates that we should know the dielectric properties over the whole spectral range. There are some interesting implications of Equation 3.31. Let us consider a small particle, 1, dispersed in

medium 3 interacting with a larger particle, 2. If, for example, $\varepsilon_1(i\nu) < \varepsilon_3(i\nu) < \varepsilon_2(i\nu)$, then A_{132} would be negative. This means that a particle of material 1 would be rejected from the medium 3. If the particle of material 2 were liquid, particle 1 could be engulfed. Engulfment is an important process in biological systems, for example it is part of our body's defence mechanism where white cells engulf and remove foreign cells such as bacteria in a process known as *phagacytosis*. After engulfment, particle 1 is in medium 2 and interacting with the surface of 3. In this case, $\varepsilon_3(i\nu)$ is interchanged with the $\varepsilon_2(i\nu)$ in Equation 3.31. Both terms in the parentheses are negative and A_{123} is positive.

Calculation of the Hamaker 'constant' requires extensive measurements from dielectric spectroscopy. The task is clearly a non-trivial one and the full dielectric data are available for only a very few systems. Polystyrene particles dispersed in aqueous media are a popular model colloidal system and Parsegian and Weiss [13] have carried out the full calculation for this system and compared the results with the Hamaker treatment. Good agreement was obtained, but this does not mean that the more complicated Lifshitz analysis should be neglected. It naturally includes retardation effects in addition to changes to the intervening media between particles.

Higher frequency dielectric information is found from the frequency-dependent refractive index data [1, 11]:

$$\varepsilon(i\nu) = 1 + \frac{n^2(0) - 1}{1 + \nu/\nu_e} \tag{3.32}$$

where ν_e is the frequency of the dominant adsorption in the UV region and $n(0)$ is the low-frequency value of the refractive index in the visible range. For the dispersion interaction, we can neglect interactions in the microwave region and below. Some approximations for Equation 3.31 are available (e.g. Israelachvili [1] and Russel *et al* [10]). For example, if we have two particles of similar material, 1, interacting in a medium 3, the effective Hamaker 'constant', that is, the distance-dependent value of A which we will write as $A_{131}(H)$, is

$$\begin{aligned} A_{131}(H) = {} & \frac{3k_B T}{4}\left[\frac{\varepsilon_1(0) - \varepsilon_3(0)}{\varepsilon_1(0) + \varepsilon_3(0)}\right]^2 \\ & + \frac{3h\nu_e}{16\sqrt{2}} \frac{\left[n_1^2(0) - n_2^2(0)\right]^2}{\left[n_1^2(0) + n_2^2(0)\right]^{\frac{3}{2}}} F\left(H/\lambda_e\right) \end{aligned} \tag{3.33}$$

where $F\,(H/\lambda_e)$ is a function of the distance between the surfaces relative to the wavelength at the UV adsorption peak. The function is equal to 1 at small distances of separation and reduces fairly rapidly with distance H to a value

$$\frac{H}{\lambda_e} \to \infty; \quad F(H/\lambda_e)\frac{2^{\frac{5}{2}}}{\pi n_3(0)\left[n_1^2(0)+n_3^2(0)\right]^{\frac{1}{2}}}\frac{H}{\lambda_e} \tag{3.34}$$

and there is an interpolation equation available [10] which can be used to calculate intermediate values:

$$F\,(H/\lambda_e) \approx \left(1+\left\{\frac{H}{\lambda_e}\frac{\pi n_3(0)\left[n_1^2(0)+n_3^2(0)\right]^{\frac{1}{2}}}{2^{\frac{5}{2}}}\right\}^{\frac{3}{2}}\right)^{-\frac{2}{3}} \tag{3.35}$$

Hence our computational procedure for estimating the interaction between two spherical particles of the same material dispersed in a medium would be to use Equation 3.25 with an effective Hamaker constant estimated from Equation 3.33 with Equation 3.35 at each separation distance. Although the process may appear complicated, it is relatively easy to program a computer to handle it. However, before we can do this, we need to take another look at Equation 3.33. The first term in Equation 3.33 uses the static dielectric constants. This means that we are looking at the contributions from permanent dipoles. Now, in media that are electrolytes, whether made up of aqueous or organic components, the dipole interactions become heavily screened and the first term approaches zero.

3.7 SUMMARY AND CALCULATION GUIDE

So far in this chapter, we have seen how the van der Waals forces acting between molecules arise from the dipolar behaviour of the molecules even when there are no permanent dipoles. The latter are the London dispersion forces. This early treatment assumed that the interaction occurred at a single frequency. The forces may be summed to account for the interaction between particles, as was carried out by Hamaker and de Boer. The interaction was shown to be the product of a material property, a geometric term for the shape of the bodies and a distance dependence.

More recently, the analysis has been carried out by treating the particles as macroscopic dielectric materials. This is known as the Lifshitz treatment and has some strong advantages, as follows:

(a) The pairwise summation is no longer carried out and the macroscopic behaviour automatically includes multi-molecular interactions.
(b) The interaction over the whole spectrum is included so that all the dipolar interactions are inherent in the computation.
(c) The retardation of the interaction that shows as a reduction in the interaction free energy as the separation between the bodies is increased can be included as an integral part of the calculation. Moreover, the retardation naturally is shown to reduce the high-frequency part of the spectral response.
(d) Any intervening medium is included naturally in the calculation and it is no longer necessary to include combining relationships to provide an approximation to the Hamaker constant.
(e) When the intervening medium is an electrolyte, the low-frequency interaction (Keesom and Debye terms) are screened out.
(f) The equations indicate that two bodies of the same material will always attract each other, but also that it is possible to have a repulsive interaction with the right combination of dielectric properties.

3.8 CALCULATION STRATEGY

In principle, the full Lifshitz calculation requires a detailed knowledge of the full electromagnetic spectral response and these data are not often available in sufficient detail. However, the approximations given above will usually be sufficient. There are several decisions that we have to make, based on the following questions:

(a) For what purpose are we going to use the results?
(b) Do we have a simple geometry or is the approximation to a simple geometry a limiting factor?
(c) Do we need to know the distance dependence of the interaction or are we only really concerned with the effects when the surfaces are relatively close?
(d) Do we have the required material properties to hand?

The answers to each of these questions will set our strategy and so let us consider each in more detail.

It is only if we are carrying out very precise experiments to determine the attractive interaction using model systems that we would need to tackle the full calculation. Often, we can use the value of the Hamaker constant as opposed to the distance-dependent Hamaker function. In these cases, the attractive energy equations for close distances would also suffice (for example, Equations 3.24 and 3.26). The value of A could be calculated from Equation 3.33 with the function $F(H/\lambda_e) = 1$. In addition, there are values for a wide variety of materials listed in the literature [1, 10, 14, 15]. However, it would be preferable to utilize Equation 3.33 rather than the combining rules, as the values of $\varepsilon(0)$ and $n(0)$ are fairly easy to find in the literature [16, 17]. Examples of this situation include problems of estimating adhesion of particles to surfaces, separation of particles during dispersion, estimation of the aggregation rates of particles and estimation of the rheology of concentrated aggregated systems.

If our materials are produced by techniques such as milling or high-speed dispersion, the particles may be small, compact aggregates of primary crystallites or fragments of crystals. Plate-like surfaces may not be flat but be 'stepped'. In cases like these, the simple approximations may be sufficient but the representation at close distances can easily be poor. The surface detail makes less difference at larger distances. We expect small liquid droplets to be spherical but we should be aware that the value of the interfacial tension is important in conjunction with the droplet size. In addition, at close distances of approach the surfaces are likely to flatten and increase the interaction area, thereby increasing the attraction.

At this point, it will be useful to look at an example calculation. For this purpose, we will take a poly(methyl methacrylate) latex particle in a simple aliphatic hydrocarbon, namely dodecane. This has been widely used as a model colloid by numerous workers. The example is with a particle size of 500 nm (i.e. a radius of 250 nm). The material properties that we need are the following:

$\varepsilon_1(0) = 3.6,\ \varepsilon_3(0) = 2.01;$

$n_1(0) = 1.492,\ n_3(0) = 1.411;$

$\lambda_e = 99.9$ nm (from the speed of light/frequency of the absorption peak in the UV region).

Equation 3.33 is used to give the Hamaker constant [i.e. putting $F(H/\lambda_e = 1]$ and the Hamaker function to show the effects of retardation.

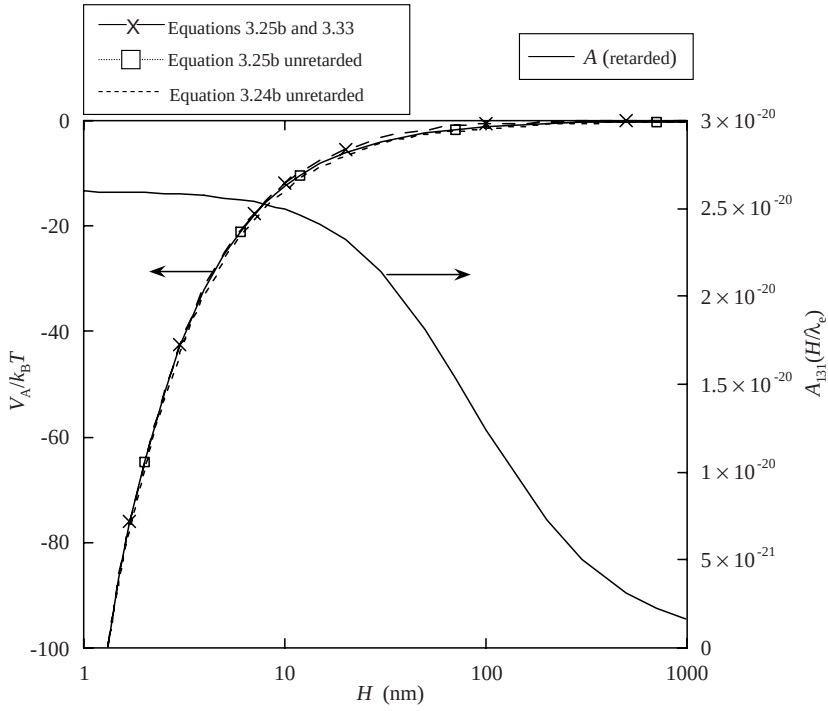


Figure 3.9 The attraction between two 2500 nm radius poly(methyl methacrylate) spheres in dodecane. The variation of the combined Hamaker constant with distance is also shown

In addition, we use the full Hamaker expression for identical spheres (Equation 3.25b), and also the approximation for close approach (Equation 3.24b). The results are shown in Figure 3.9 along with the plot for the Hamaker function.

The first point to note is that there is little retardation until a distance >10% of the wavelength of peak absorption is used. The drop has become significant at distances greater than 30 nm. Although the retardation is large at distances in excess of 500 nm, the attraction is becoming less than the thermal energy at ~25 °C and we would be unable to discern the interaction, let alone the detail. The conclusion that we may draw from this is that the neglect of retardation and the use of the simplest equation for the interaction (Equation 3.24b) will be adequate for most purposes for this system. If, for example, we had the task of dispersing these particles, the amount of work that we would put in would be the value of V_A at the closest distance of approach. Alternatively, if our task is to prevent the particles from sticking firmly together, we have

to supply another interaction which will have a greater potential energy but opposite in sign to the values at close distances.

However, before we draw the conclusion that we can always take the simpler route, let us look at another system. This time we will use water as the suspending medium with oil droplets as the particles. The oil phase here is tetradecane, which represents a light fuel oil and has a chain length commonly found in Nature. This chain length is the main fraction of coconut oil, for example. The particles are assumed to be $2\,\mu$m and all the same size for this calculation. The material properties for this case are the following:

$$\varepsilon_1(0) = 2.03,\ \varepsilon_3(0) = 80;$$
$$n_1(0) = 1.418,\ n_3(0) = 1.333;$$
$$\lambda_e = 103.4 \text{ nm}.$$

In the example calculation, we use Equations 3.33 and 3.25b as we did in the previous example. Here, though, we are going to compare the attraction between the oil drops in 'pure' water and 'seawater'. In the 'pure' water the only ions are OH^- and H_3O^+ with concentrations of 10^{-7} mol dm^{-3} (i.e. no dissolved carbon dioxide). 'Seawater' is used to give such a high electrolyte concentration that we can assume that the low-frequency contributions to the attraction are completely screened out. We shall see later in this chapter how the ionic strength affects the range of electrical fields and that at high ionic strengths the range is only a nanometre or two.

The results are shown in Figure 3.10, and indicate that there is a large contribution due to the low-frequency term, that is, the first term on the right-hand side of Equation 3.33. This means that the oil droplets are 'less sticky' in seawater, although an attraction of $\sim 100 k_B T$ is still a strong attraction. In most cases when we have particles dispersed in an aqueous medium, there is a significant ionic content. This may be from the addition of ionic surfactants and electrolytes (acids, bases or salts) and from ions associated with the particle surface due to surface charge. As we shall see later, the concentration is always at its greatest between pairs of particles and the neglect of the low-frequency contribution is an adequate approximation. The result now becomes similar to the case where the particles are dispersed in low-polarity media where the low-frequency contributions amount to $<5\%$ of the Hamaker constant.

Finally, we should note that the case of the interaction between different particles is also quite tractable in approximate form where we are

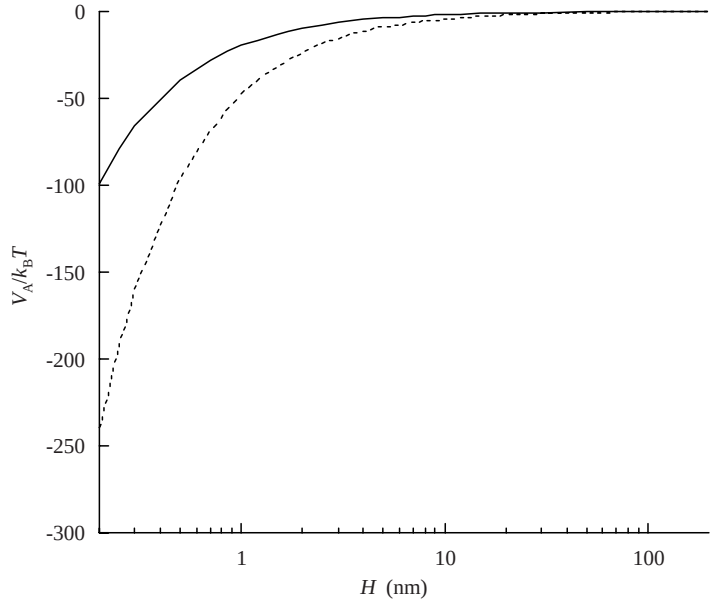


Figure 3.10 The attraction between two 2 μm tetradecane droplets in 'pure water' (- - -) and 'seawater' (—) (using Equations 3.25b and 3.33), where the low-frequency term is screened

assuming a dominant absorption peak in the UV region. This is in a similar spirit to the earlier London model. However, the approximation we are using makes it easy to include a wide frequency variation so that retardation and the variation with dipolar materials are naturally included. The value of λ_e changes only slightly from material to material and a value of 100 nm will normally suffice, although with aromatic materials such as benzene, toluene or polystyrene 150 nm would give a better approximation. The equation for two particles of different material, 1 and 2, acting across a medium 3, is as follows [1]:

$$A_{132} \approx \frac{3k_BT}{4}\left[\frac{\varepsilon_1(0)-\varepsilon_3(0)}{\varepsilon_1(0)+\varepsilon_3(0)}\right]\left[\frac{\varepsilon_2(0)-\varepsilon_3(0)}{\varepsilon_2(0)+\varepsilon_3(0)}\right] + \frac{3hc}{\lambda_e \times 3\sqrt{2}} \times \left(\frac{\left[n_1^2(0)-n_3^2(0)\right]\left[n_2^2(0)-n_3^2(0)\right]}{\left[n_1^2(0)+n_3^2(0)\right]^{\frac{1}{2}}\left[n_2^2(0)+n_3^2(0)\right]^{\frac{1}{2}}\left\{\left[n_1^2(0)+n_3^2(0)\right]^{\frac{1}{2}}+\left[n_2^2(0)+n_3^2(0)\right]^{\frac{1}{2}}\right\}}\right) \quad (3.36)$$

If the particles were spherical, Equation 3.25a would be used to calculate the attraction. Although the equations look unwieldy, they are really fairly straightforward to calculate. Again, in aqueous systems the low-frequency contributions could be neglected and the following equation would provide an adequate approximation:

$$A_{132} \approx \frac{3hc}{\lambda_e \times 3\sqrt{2}} \times \left(\frac{\left[n_1^2(0) - n_3^2(0)\right]\left[n_2^2(0) - n_3^2(0)\right]}{\left[n_1^2(0)+n_3^2(0)\right]^{\frac{1}{2}}\left[n_2^2(0)+n_3^2(0)\right]^{\frac{1}{2}}\left\{\left[n_1^2(0)+n_3^2(0)\right]^{\frac{1}{2}} + \left[n_2^2(0)+n_3^2(0)\right]^{\frac{1}{2}}\right\}} \right) \tag{3.37}$$

Metal particles such as gold, silver and platinum are important technologically as they are widely used in the form of inks for printing electronic circuit boards and integrated circuits. As they are conductors, the value of $\varepsilon_{Au}(0) = \infty$. The characteristic frequency is the plasma frequency (i.e. $60\,\text{nm} < \lambda_e < 100\,\text{nm}$) and the value of the Hamaker constant for the metal particles in water is found to be about 4×10^{-19} J, that is, that metal particles are very much 'stickier' than particles made up of organic materials.

A final point that we must remember is that this and any other continuum approach are limited in terms of the minimum separation of surfaces. At very small distances, the molecular structures of both the surfaces and the species of the continuous phase become important. Even at a few tenths of a nanometre, we cannot expect the approximations to be reliable. At this dimension, we are dealing with the size of hydrated ions and solvent molecules such as aromatic or aliphatic hydrocarbons, and also surface irregularities on what we might regard as 'smooth' surfaces. The problem of the dimensions of the intervening solution species leads us naturally to consider the next interaction that we will study.

3.9 THE DEPLETION INTERACTION

This is usually an attractive interaction that can cause aggregation of particles in concentrated dispersions, although under some concentration conditions a weak repulsion may occur. It is most frequently considered to occur in dispersions to which non-adsorbing polymer has been added. This would typically be the situation where a polymer thickener had been

added to a dispersion to control its rheological properties. A decorative paint system is a good example of this. Polymer thickeners are added to prevent sedimentation of pigment, and also to give a thicker film and so increase covering power. The particulate components of the paint (pigments and latex particles) are covered with adsorbed surfactants that restrict the adsorption of the soluble polymers. What we observe is that the thickening power of the polymer is much greater when particles are present in significant concentrations than when they are absent. However, the depletion interaction may also cause particles in the presence of concentrated surfactant micellar phases, or if there are two widely differing size populations, to phase separate.

The range of the interaction is no greater than the dimension of the solution species that is causing the effect, so we can consider it to be a short-range interaction. That is, the range is much less than the dimensions of the particles that we are considering to be the 'main population'. Although the greatest mass may be in the large particles that we are dispersing, the greatest number is going to be in the small component that is causing the depletion interaction. For example, a few percent of a soluble, non-adsorbing polymer thickener could produce a marked interaction in a paint dispersion of 60–70% solids. The interaction was first suggested by Asakura and Oosawa [18, 19] and has since been further developed [20, 21]. The concept is based on the idea of a 'depletion layer' close to the surface of a particle. Within this layer, the concentration of the non-adsorbing species (polymer molecules, for example) is lower than the average. This does not mean that there is a complete absence of material in the layer but just that the centre of mass of the solution species cannot approach closer than its radius of gyration. This is shown in Figure 3.11, with R_g being the radius of gyration.

As a starting point, think initially in terms of the non-adsorbing species being hard spheres and then the concentration of material will follow the profile sketched in the lower section of Figure 3.11a. Of course, if the material were a polymer in a good solvent, the increase in concentration would be slower with the value calculated from the product of the profile sketched and the concentration distribution within the coil. The mechanism for the interaction is a phase separation of the large and small particles due to an effective decrease in concentration of the smaller particles as medium in the 'depletion layers' is squeezed out by the aggregation of the large particles. The argument then is as follows:

1. There is a spherical shell around each particle which the centre of a small sphere cannot occupy, that is, we should exclude this volume when we calculate the number density of small spheres, ρ_{p1}.

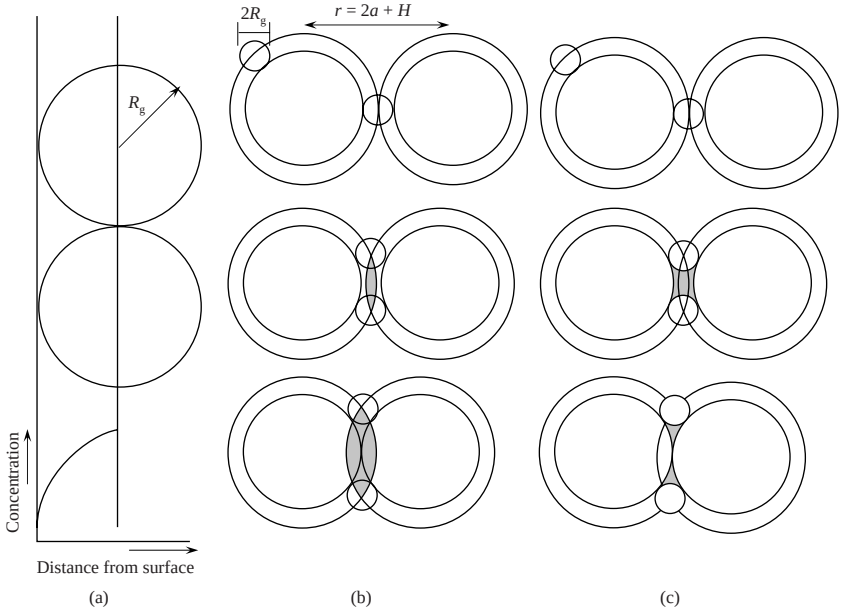


Figure 3.11 The origin of the depletion interaction. Part (a) illustrates how the polymer concentration close to the surface follows the polymer concentration across the coil in combination with the coil shape; (b) illustrates the volume excluded to polymer coils as two particles approach [18–21]; (c) illustrates the toroidal model of the excluded volume showing how the excluded volume passes through a maximum at intermediate distances of approach [22]

2. The osmotic pressure from the small spheres is $\rho_{p1} k_B T$.
3. As the large particles come together, the overlapping part of the depletion layers is now a volume which is freed from occupation by the small particles.
4. The osmotic pressure acts to push the large particles together as there is an osmotic pressure difference between the volume excluded from the polymer and the bulk.
5. The pressure multiplied by the interaction area is the force, so we may construct the potential energy by integrating this force over the separation distance. This gives the depletion contribution to the pair potential, V_d, simply as the osmotic pressure of the polymer solution multiplied by the overlap volume, v_o:

$$V_d = \left(\rho_{p1} k_B T\right) v_o \tag{3.38}$$

Figure 3.11b indicates the volume of liquid that is not accessible to a small sphere as the larger spheres come together (shaded portions). We

may calculate this excluded volume from simple geometry. The lens is simply twice the volume of a spherical cap. The centre-to-centre separation of the large particles is r with the surface-to-surface separation H, the radius of the large particles is a and that of the small particles is R_g. The volume of the cap of a sphere is given by

$$\text{volume of spherical cap of height } h = \frac{\pi}{3}h^2\left[3\left(a + R_g\right) - h\right] \quad (3.39a)$$

where

$$h = \left(a + R_g\right) - \frac{r}{2} \quad (3.39b)$$

Now, we may write the depletion interaction between two spheres of equal size as

$$V_d = -\left(\rho_{p1}k_BT\right)2\frac{\pi}{3}\left[\left(a + R_g\right) - \frac{r}{2}\right]^2\left[2\left(a + R_g\right) + \frac{r}{2}\right] \quad (3.40)$$

The negative sign indicates that the interaction is attractive. Multiplying this out, we obtain

$$V_d = -\left(\rho_{p1}k_BT\right)\frac{4\pi}{3}\left(a + R_g\right)^3\left[1 - \frac{3}{4}\left(\frac{r}{a + R_g}\right) + \frac{1}{16}\left(\frac{r}{a + R_g}\right)^3\right]() \quad (3.41a)$$

and if we write this in terms of the separation between surfaces, $H = r - 2a$, we have

$$V_d = -\left(\rho_{p1}k_BT\right)\frac{4\pi}{3}\left(a + R_g\right)^3\left[1 - \frac{3}{4}\left(\frac{2a + H}{a + R_g}\right) + \frac{1}{16}\left(\frac{2a + H}{a + R_g}\right)^3\right] \quad (3.41b)$$

As we are assuming that the small particles are hard spheres, we may write the relative depletion potential in terms of the volume fraction of the small particles, φ_{p1}:

$$\frac{V_d}{\varphi_{p1}k_BT} = -\left(\frac{a + R_g}{R_g}\right)^3\left[1 - \frac{3}{4}\left(\frac{2a + H}{a + R_g}\right) + \frac{1}{16}\left(\frac{2a + H}{a + R_g}\right)^3\right] \quad (3.42)$$

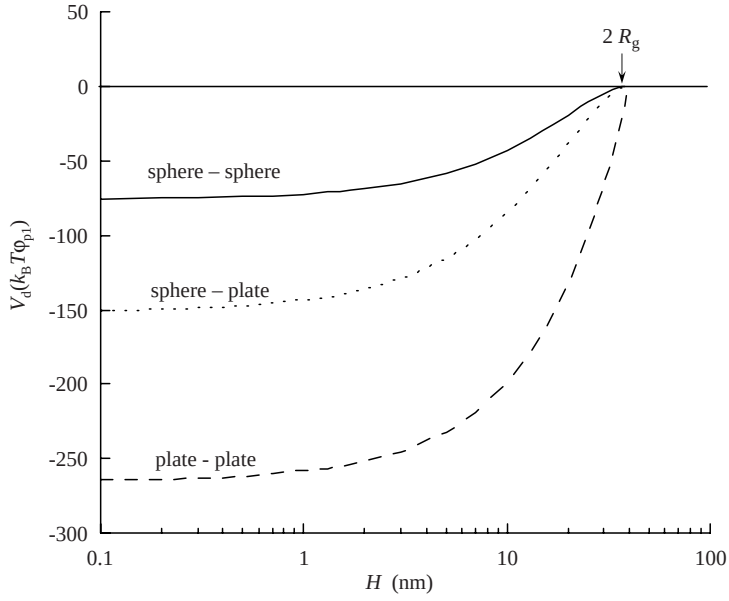


Figure 3.12 Relative depletion potential as a function of the separation distance between surfaces. R_g, 20 nm; a, 10^3 nm; plate overlap 150 × 150 nm

This geometric function is plotted in Figure 3.12. The interaction is zero until a distance equal to the diameter of the small particle is reached between the large particles. A volume fraction of 0.05 would result in an attraction of $\sim 4k_BT$ in this example.

The interaction for other geometries is straightforward. For a sphere and a plate, we just have the osmotic pressure multiplied by the interaction area (the circular area of the cap of the sphere defined by the overlap of the depletion layers), integrated as we move the two surfaces closer together. The overlap distance is $2R_g - H$. This gives us the volume of a spherical cap and not a lens as was the case with two spheres. For plates, we have the osmotic pressure multiplied by the overlap are of the two plates to give the force. When this is integrated over distance as the plates are moved together, we have the volume of the two depletion layers multiplied by the osmotic pressure. Hence for two plates the interaction increases linearly as the plates approach each other with the overlap again being $2R_g - H$. Hence when we write the expressions for this model of the depletion potential for these geometries, we have:

for a sphere with a plate:

$$\frac{V_d}{\varphi_{p1} k_B T} = -\frac{1}{4}\left(2 - \frac{H}{R_g}\right)^2 \left(1 + \frac{3a}{R_g} + \frac{H}{R_g}\right) \tag{3.43}$$

and for two plates:

$$\frac{V_d}{\varphi_{p1} k_B T} = -\frac{3A_o}{4\pi}\left(\frac{2}{R_g^2} - \frac{H}{R_g^3}\right) \tag{3.44}$$

where A_o is the overlapping area of the two plates. These expressions are plotted in Figure 3.12 using some typical values for colloidal systems.

At this point, we to have a simple model of the depletion potential, but we must remember that it is a *simplified* model and we should consider its limitations. The effect is of most interest with the case of solution polymers which have been added to dispersions, often as a rheology modifier or thickener. To exploit the effect, we need take heed of the following:

(a) The polymer should be in a good solvent, that is $\chi < 0.5$.
(b) The polymer should be non-adsorbing or, if it adsorbs, then the concentration must be sufficient to be greater than that required to saturate the surface, that is, the concentration would ensure that we are on the plateau of an adsorption isotherm.
(c) We must remember that a polymer coil is not a hard sphere but is dynamic and with a concentration that decreases towards the periphery of the coil.
(d) The polymer concentration should be relatively high so that the osmotic pressure is high. This will invariably mean that we will have to use a more complicated expression for the osmotic pressure than the van't Hoff limit, that is, $\Pi = \rho_{p1} k_B T + O(\rho^2{}_{p1}) + \ldots$.

The last two points are important. The softness of a polymer coil can be expected to change the interaction profile. This possibility will be enhanced if the large particles are covered with a polymer layer. However, trying to model variations would not be very useful compared with the variation that must occur due to the variation in coil size due to the polydispersity of the polymer. The molecular weight distribution of the

majority of polymer thickeners is large and the variation in the value of R_g is, of course, directly related to this spread.

The osmotic pressure of the polymer solution can be tackled relatively easily. The simplest approach might be to determine it experimentally, although there are expressions in the literature which can be used. However, at this point we should look carefully at the model. When we do this, other factors become apparent. If we consider a single polymer coil in isolation, the dimension is a function of the molecular weight and the polymer–solvent interaction, that is, the Flory–Huggins χ parameter. We are not dealing with an isolated coil in pure solvent, however, and whatever else there is in solution can be expected to have an impact. Electrolytes increase to χ values towards 0.5 for water-soluble polymers such as poly(ethylene oxide). Hence we expect the polymer coil to shrink as the salt concentration is increased. Increasing temperature will also do this in aqueous systems as the H-bonding levels are reduced. The addition of low molar mass alcohols will produce a similar result. Surfactants can interact with polymer chains and change their spatial characteristics. The presence of other polymer molecules is particularly interesting. Of course, we have already drawn attention to the osmotic pressure change for the calculation of the depletion potential, but we should also remember that the dimensions of the polymer coils are also dependent on the polymer concentration. This is easy to visualize if we recall that at a concentration c^*, the polymer coils would just fill the fluid volume. This means that at the peripheries the local polymer concentration becomes higher than that of an isolated coil. The diffusional motion increases this effect, resulting in a redistribution of the segments to give a more compact conformation. In other words, the value of R_g is reduced. Here we have two competing effects: the increase in polymer concentration increases the depletion potential as the osmotic pressure is increased, but the dimension of the depletion layer is reduced, thereby decreasing the strength of the interaction. The result is that we can find a maximum attraction at an intermediate polymer concentration.

A re-examination of Figure 3.11 should also indicate another problem with the model that can alter the profile of the potential curve. As two large particles approach at distances $H < 2R_g$, we can see that there is a volume of liquid in between which does not contain the small spheres – the origin of the depletion effect. Note that we are not referring to the overlap volume of the depletion layers! That volume is just a mathematical consequence of integrating the interaction area with separation distance. The 'real' excluded liquid volume is a more complicated

toroidal shape. However, this volume also varies as the interparticle distance varies. For an isolated particle pair in a large volume of polymer solution, there would be no change in the osmotic pressure. In a concentrated system we must be careful. This volume variation exists with every interaction. At volume fractions of >0.3 with weakly aggregating systems such as these, there can be six or more interactions per particle and this must be multiplied by the number of particles (divided by 2 so that we do not double count, of course). What are the implications of this? We usually work in a region where the osmotic pressure is high and this also means that it is varying sharply with concentration. Figure 3.11c shows clearly that the excluded volume passes through a maximum at intermediate separations. This means that in a concentrated system, the osmotic pressure is at a maximum at those intermediate separations also and not a constant as the model above assumes. The result is that the maximum attractive force is not at contact but a little way prior to that point [22]. This is important for properties such as the rheology where, for example, the elasticity is a rate of change of the force with distance, that is, the curvature of the potential well.

REFERENCES

1. Israelachvili, J. (1992) *Intermolecular and Surface Forces*, Academic Press, London.
2. Landau, L. D. and Lifschitz, E. M. (1984) *Electrodynamics of Continuous Media*, Vol. **8**, Pergamon Press, Oxford.
3. McLachlan, A. D. (1963) *Proc. R. Soc. London, Ser. A*, **202**, 224.
4. Von Hippel, A. R. (1958) *Dielectric Materials and Applications*, John Wiley & Sons, Inc., New York.
5. de Boer, J. H. (1936) *Trans. Faraday Soc.*, **32**, 10.
6. Hamaker, H. C. (1937) *Physica*, **4**, 1058.
7. Derjaguin, B. V. (1934) *Kolloid Z.*, **69**, 155.
8. Langbein, D. (1974) *Theory of van der Waals Attraction*, Springer Tracts in Modern Physics, Vol. **72**, Springer, Berlin.
9. Hunter, R. J. (1987) *Foundations of Colloid Science*, Vol. **1**, Oxford University Press, Oxford.
10. Russel, W. B., Saville, D. A. and Schowalter, W. R. (1989) *Colloidal Dispersions*, Cambridge University Press, Cambridge.
11. Mahanty, J. and Ninham, B. W. (1976) *Dispersion Forces*, Academic Press, London.
12. Atkins, P. W. (1982) *Physical Chemistry*, 2nd edn, Oxford University Press, Oxford.
13. Parsegian, V. A. and Weiss, G. H. (1981) *J Colloid Interface Sci.*, **81**, 285.
14. Gregory, J. (1973) *Adv. Colloid Interface Sci.*, **2**, 396.
15. Hough, D. B. and White, L. R. (1980) *Adv. Colloid Interface Sci.*, **14**, 3.
16. Weast, R. C. and Astle, M.J. (eds) (1981) *Handbook of Chemistry and Physics*, CRC Press, Boca Raton, FL.

17. Brandrup, J. and Immergut, E. H. (eds) (1975) *Polymer Handbook*, 2nd edn, John Wiley & Sons, Inc., New York.
18. Asakura, S. and Oosawa, F. (1954) *J. Chem. Phys.*, **22**, 1255.
19. Asakura, S. and Oosawa, F. (1958) *J. Polym. Sci.*, **33**, 183.
20. Vrij, A. (1956) *Pure Appl. Chem.*, **48**, 471.
21. Joanny, J. F., Liebler, L. and de Gennes, P.-G. (1979) *J. Polym. Sci. Polym. Phys. Ed.*, **17**, 1073.
22. Goodwin, J. W., Hughes, R. W., Kwaamba, H. M. and Reynolds, P. A. (2000) *Colloids Surf. A*, **161**, 361.

4

Forces of Repulsion

4.1 INTRODUCTION

In Chapter 3, we discussed the attractive interactions between colloidal particles. There are other contributions to the pair potential which prevent particles from adhering to each other. These are the forces of repulsion and this chapter will describe the most important of these.

4.2 ELECTROSTATIC INTERACTIONS

Many interfaces in aqueous systems carry an electrical charge. Like charges repel each other. In the simplest case, this may be described by Coulomb's law; however, we shall see that the ionic content of the system modifies this. The repulsion, however, can be strong and prevent interfaces from coming together. Let us consider a soap film, as an example. In this case, a surfactant is strongly adsorbed at the interface with hydrocarbon tails on the vapour side and ionic head groups on the water side of each interface. Sodium or potassium fatty acid salts make the classic soaps with the number of carbon atoms ranging from 14 to 18, depending on the source. When a foam is produced, the water drains due to gravity and the close-packed arrays of ionic end-groups approach each other. At some point, the forces come into balance with the magnitude of the electrostatic repulsion controlling the final thickness of the foam films. This repulsion occurs between any 'like-charged' interfaces and is an important factor in determining the colloidal behaviour of aqueous

Colloids and Interfaces with Surfactants and Polymers, Second Edition James Goodwin

systems and also those of intermediate polarity (those with a relative permittivity ≥ 10).

The model that we will use is mainly a continuum model in which we will assume any charged surface to be uniformly charged and therefore structureless. Our colloidal systems will also be electrically neutral so that any surface charge is always balanced by an equal amount of charge opposite in sign in the continuous phase. The distribution of the solution charge close to the charged surface, however, is where we put our efforts. The two layers of charge are known as the e*lectrical double layer*. Ions in solution of the same sign as the charged surface are referred to as *co-ions* and those of opposite sign are *counter-ions*. There will always be some ions present in our aqueous solvent phase, of course, such as H_3O^+, OH^- and species such as carbonate and silicate, even when we have not intentionally added electrolytes. Ions which interact specifically with the surface to vary the value of the surface charge are known as *potential-determining ions*; H_3O^+ is an example of such an ion with fatty acid soap films, for example, as the carboxylic acid is a weak acid and the degree of dissociation would be different at pH 2 and pH 10. Electrolytes whose ions play no direct part in the charging mechanism are known as *indifferent electrolytes*. However, the specificity of ion–surface interaction is of major importance in controlling the behaviour of colloidal systems and is still a rich field of study.

4.3 THE ORIGINS OF SURFACE CHARGE

There are several mechanisms by which an interface may acquire a charge. In many cases, we have some control over the value of the charge. Each mechanism is introduced below with some examples, but it is important to be aware that and there can be more than one mechanism operating in our practical systems.

4.3.1 Adsorbed Layers

Surfactants are often added as a component to disperse systems, for example, as a wetting aid for the dispersion of dry powders, an emulsifier for liquid systems, and as a stabilizer for the end product. Ionic surfactants are relatively cheap and are frequently added. The final conformation at the interface is for the hydrophobic tails to be at the interface with the ionic groups outermost. The surface density is always high in

a surfactant-stabilized system. Proteins also adsorb on hydrophobic surfaces and provide a hydrophilic charged outer layer. An example is the fat droplets in milk, which are stabilized by casein. The charge is negative as it results from the carboxyl groups on the protein.

Natural rubber latex is a dispersion of poly(*cis*-isoprene) with a negatively charged surface from the adsorbed protein on the surface. Reduction of the pH prevents the dissociation of the carboxyl groups and coagulation results. Synthetic rubber latexes [poly(styrene–butadiene) rubbers] have some charged groups chemically bound to the polymer but invariably have a high concentration of anionic surfactants adsorbed on the surface of the particles. The surfactant has two functions: first, as an emulsifier for the monomer prior to synthesis, and second, to provide colloid stability to the final product.

Poly(acrylic acid) can be used with oxides in order to provide a coated surface with a negative charge at pH values where the oxide surface may have a low positive charge. Inorganic 'macro-ions' can also be used for this purpose. China clay (kaolinite) is an important additive to paper to provide a glossy surface and it can be stabilized with poly(acrylic acid) adsorbed on the positively charged edges. Calgon, a phosphate which adsorbs on the faces, is also used as a stabilizer for kaolinite. Titanium dioxide pigment particles are often coated with a layer of alumina and poly(acrylic acid) will adsorb on the surface at neutral pH.

4.3.2 Ionogenic Surfaces

The charge on this type of surface is controlled by the ionization of chemical moieties at the surface. Carboxyl groups that are chemically bound to the polymer of synthetic latexes provide an example of this. The charge is a function of pH as the degree of dissociation is a function of pH. Although the pK_a of an isolated $-COOH$ group is pH $\approx$ 4, this is not the situation with a surface with many groups in close proximity. The dissociation of one group makes it more difficult for the immediate neighbours to dissociate. This is the *polyelectrolyte effect* and means that the surface has a variable pK_a and pH values as high as 9 may be required to ensure dissociation of all the surface groups.

The surfaces of uncoated oxide particles have surface hydroxyl groups. At high pH, these can ionize to give $-O^-$ and at low pH the lone pair on the oxygen can hold a proton to give $-OH_2^+$. For example, if we consider the surface of a titanium dioxide particle, we have oxygen atoms on the

Table 4.1 Isoelectric points of oxides

Oxide	Formula	iep
α-Alumina	α-Al_2O_3	9.1
Haematite	α-Fe_2O_3	6.7
Magnesium oxide	MgO	12.5
Rutile	TiO_2	3.5
Silica	SiO_2	2.0
Zinc oxide	ZnO	10.3
Zircon	ZrO_2	5.7

surface of the crystal giving an amphoteric surface according to

$$\text{surface } TiOH_2^+ + H_2O \xleftarrow{H_3O^+} \text{surface } TiOH \xrightarrow{OH^-} \text{surface } TiO^- + H_2O$$

Hence this type of surface not only shows a variation in the magnitude of the surface charge with pH but also a variation of the sign. At an intermediate pH value, the charge can be reduced to zero and the pH at which this occurs is termed the *isoelectric point* (iep) or the *point of zero charge* (pzc). Frequently, the two terms are used interchangeably but they are not always the same. For a simple oxide surface, the former term is to be preferred as we expect the charge to be uniformly zero across the surface. With more complex materials, where there is more than one type of surface present, we may have a situation where both surfaces carry a charge but of opposite sign. Then, at some pH value we can have zero net charge as the two values balance each other. This would then be the pzc. China clay (kaolinite) is a layer lattice aluminosilicate, where the edges are positive with negative basal surfaces. In this case, the iep of the edges is at pH $\approx$ 8 with the face iep at pH $\approx$ 2 and a pzc for the particle at pH $\approx$ 6. Table 4.1 lists the isoelectric points for some oxides.

4.3.3 Isomorphous Substitution

This is a common occurrence in clay minerals. The basic structure of a clay particle is an aluminosilicate layer lattice. As the clay is formed, it crystallizes with a layer of silicon atoms tetrahedrally coordinated to oxygen atoms (an SiO_2 layer). The next layer of the lattice is aluminium with octahedrally coordinated oxygens (an Al_2O_3 layer), some of which are shared with the tetrahedral silica layer. This layer structure is

repeated throughout the crystal. If it is a 1:1 layer lattice structure, there are alternating layers of silica and alumina (e.g. kaolinite). Alternatively, it can crystallize in a 2:1 structure, with the alumina layer sandwiched between two silica layers (e.g. montmorillonite). During the crystallization process, an occasional silicon atom can be substituted by an aluminium atom and more frequently an aluminium atom in by a magnesium atom. A *small* amount of such substitution does not produce too much distortion of the lattice to stop it growing and it retains the same structure – hence the term *isomorphous substitution*. Thus, the number of oxygen atoms is the same. As anions, they are larger than the cations in the structure and their packing is the dominant factor. However, the difference in valency of the magnesium, compared with the aluminium anions, for example, means that some of the oxygen anion coordination will be incomplete, thus resulting in a negative charge. This charge is balanced by soluble cation species at the surface of the crystal, with the result that the basal surfaces of clay minerals carry a significant negative charge. In the dry state, the counter-ions to this surface charge are located on the surface, whereas in the hydrated state they are in solution near the surface. The 2:1 layer lattice clays have a unit cell $\sim$1 nm thick and, on hydration, water penetrates between the layers and the negative charges of the surfaces repel and cause the clays to swell. The swelling can result in a separation between (original) unit cells of a greater dimension than that of the (original) unit cell. This means that the expansion of the clay can be very marked. This phenomenon causes major difficulties in building on land with high clay contents.

4.3.4 Differential Solution of Surface Ions

When the colloidal particles are made up of sparingly soluble salts, dissolution occurs until the concentration of the ionic components in solution corresponds to the solubility product of the compounds. Silver halides are a much studied example of this class of material, with silver bromide dispersions, for example, having a long history as a commercially important component of photographic emulsions. For many years, silver iodide was used as a model colloid for academic studies. It is readily produced by mixing silver nitrate and potassium iodide solutions. Under the correct conditions, a dispersion is formed with the particles confined to a narrow size distribution and with a cubic shape. Now, the solubility product $K_s = [Ag^+][I^-] = 10^{-16}$. If the precipitation is produced under conditions which have an excess of I^-, say 1×10^{-4} M, we form a

dispersion with negatively charged particles. Using $p[I^-]$ as $-\log_{10}[I^-]$, we have a solution with $p[I^-] = 4$ and hence $p[Ag^+] = 12$ and the solution of Ag^+ ions is suppressed with the surface of the particle consisting of I^- species. At $p[I^-] = 10.5$, that is, $p[Ag^+] = 5.5$, the surface populations of the two ion types are equal and the pzc is reached as the net charge is zero. Here, the potential-determining ions are Ag^+ and I^-. We might have expected that the pzc would have occurred when there were equal numbers of Ag^+ and I^- species present, that is, at $p[I^-] = 8$; however, the solvation of the smaller cation is greater and the adsorption is in favour of the large anion.

On the surface of the silver iodide crystal, at or close to the pzc, we have a large number of charged species. This means that as the surface is charged up, only a relatively small increase in the number of anions or cations is required and we may represent the surface potential as we would for an electrode surface by the Nernst equation [1]:

$$\psi_0 = \frac{RT}{F_c} \ln \left(\frac{[Ag^+]}{[Ag^+]_{pzc}} \right) \tag{4.1}$$

where F_c is the Faraday constant. This gives a surface potential of about ± 60 mV for a factor of ± 10 in the silver ion concentration from that corresponding to the pzc.

Other colloidal crystallites can also show *Nernstian behaviour* and some examples were given by Hunter [1]. However, it should be pointed out that this behaviour is far from universal for colloidal particles. For example, with amphoteric surfaces such as oxides, the surface charge population close to the pzc is very small and, as the surface charges up, the activity of the species at the surface changes and we no longer find Nernstian behaviour. There are a variety of models available for different types of surfaces and a good summary can be found in Hunter [1].

4.3.5 The Structure of the Electrical Double Layer

When we have a colloidal particle with a charged surface, there is always an equal and opposite charge in the solution. The structure of the solution part of the double layer must now be considered. We will start with modelling a flat surface and then consider curved interfaces. The earliest treatment was due to Helmholtz and described a simple model of a uniformly charged surface with ions of opposite charge, the

counter-ions, in a uniform layer adjacent to the surface and treated as point charges. However, a more complex picture superseded this model. The first point that we need to consider is that the solvated counter-ions have a finite size, that they can interact laterally and that there may be specific chemical interactions with the surface. That is, we must be aware that other than straightforward electrostatic forces may be involved. The second point to remember is that the diffusive motion of the ions will oppose the tendency to concentrate the counter-ions in the interfacial region and result in a diffuse array. It is in studying this diffuse region that we will put most of our effort here after considering the first layer of ions.

4.3.5.1 The Stern plane

This is the inner region of the solution part of the double layer. It is usually modelled using the *Langmuir isotherm*, which describes the formation of a monolayer. We should be familiar with this isotherm in the context of gas adsorption, where the adsorption is a function of both the pressure and the fraction of the sites that are occupied. In the present case, the adsorption energy is made up of the electrostatic attraction and any other 'specific' interactions. Hence we have a monolayer consisting mainly of counter-ions at the surface whose population is a function of the electrostatic potential plus specific chemical interactions, in addition to the ionic content of the aqueous phase. (Here, the ionic strength is analogous to pressure in the gas phase adsorption.) Another problem is that in colloidal systems, the molecular structure of the surface means that the discrete nature of the charges needs to be considered at this level. This acts to increase the site occupation so that, if we were to estimate the adsorption energy, a lower energy would be required to attain the same level of counter-ion density at the surface than would be the case if the surface charge is considered to be 'smeared out' uniformly over the surface.

The picture that is emerging now is a complicated one and a detailed analysis is outside the scope of this text. Good descriptions of the current ideas are given in the books by Hunter [1] and Lyklema [2] along with the major references to the original research papers. For our present purposes, we will consider our surface to be uniformly charged with a surface charge density of σ_0 C m^{-2} and adjacent to this is a layer of counter-ions. It is to the outer edge of this layer that we must now turn our attention. At this plane, the adsorbed ions have changed the charge

to σ_δ and the potential relative to ground (i.e. at a very large, and so effectively infinite, distance from the surface) is ψ_δ. This is just the work done in bringing a point charge from infinity to this plane. From here we will be treating the ions as point charges. Now, unfortunately, this is a difficult quantity to measure on a routine basis but we can relatively easily estimate a potential from electrokinetic measurements where we have a motion between the fluid and the interface. This potential is known as the zeta (ζ)-potential and is the potential where the centre of the first layer of solvated ions that are moving relative to the surface is located. This is termed the 'shear plane', but it is not at a well-defined distance from the surface so we have the situation where we can determine a potential although at a position that is slightly uncertain, but is at ~0.5 nm or so from the surface. This distance would be about three times the radius of a solvated ion, but the utility of the ζ-potential is that it reflects the value of ψ_δ which can differ in sign as well as markedly in magnitude from the potential at the surface ψ_0. A sketch of our emerging model of the double layer is given in Figure 4.1.

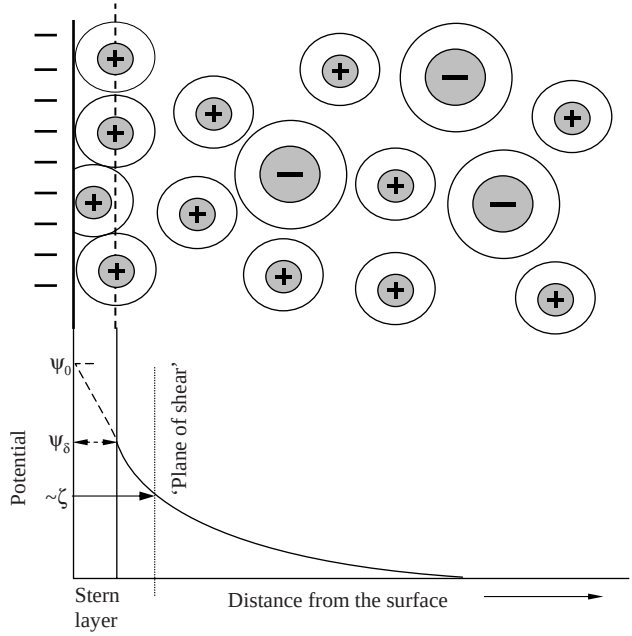


Figure 4.1 Illustration of a simple model for the electrical double layer at the surface of a colloidal particle

The population of the Stern layer is a function of ion type. Multivalent ions bind more strongly than monovalent ions. The solvation and polarizability are also important factors. These characteristics lead to the specificity of the various ions. So, for example, with a positively charged surface we observe a very marked difference in the binding of the halide anions in the series from F^- to I^-. The addition of multivalent counter-ions to a colloidal dispersion can result in a densely populated Stern layer and the reversal of the sign of the charge relative to the particle surface.

4.3.5.2 The diffuse double layer

This is the part of the electrical double layer outside the Stern layer where the concentration of ions is determined by the competing effects of electrostatics and diffusion. What we are going to do is to calculate the distribution of ions and from this, the potential as a function of distance. Ultimately, this will enable us to model the electrostatic repulsive energy as we estimate the work done in bringing one surface up to a second surface.

The model that we are going to use is known as the *Gouy–Chapman model* and is named from the two physicists who developed it independently. We start by defining the potential at a distance x from our surface as $\psi(x)$. As we are considering a flat surface which is uniformly charged, there will only be a dependence in the x-direction, with the potential being constant in planes parallel to our surface. In addition, the ions are treated as though they were point charges and the number of ions of type i per unit volume in the bulk electrolyte is n_{i0}. We make use of the Boltzmann distribution to estimate the ion density at x:

$$n_i = n_{i0} \exp\left[-\frac{z_i e \psi(x)}{k_B T}\right] \tag{4.2}$$

where e is the charge on the electron and z_i is the valency of the ion i. We need to be careful about signs here and remember that as the sign of the charge on the counter-ion is always opposite to that of the surface, the exponent will always be negative for the counter-ion concentration and positive for the co-ion concentration. Equation 4.2 describes the situation where the counter-ion concentration increases close to the interface – adsorbed – whereas the co-ion concentration is reduced close to the interface – negatively adsorbed. Figure 4.2 shows how the local ion concentration profiles behave according to Equation 4.2. The number of

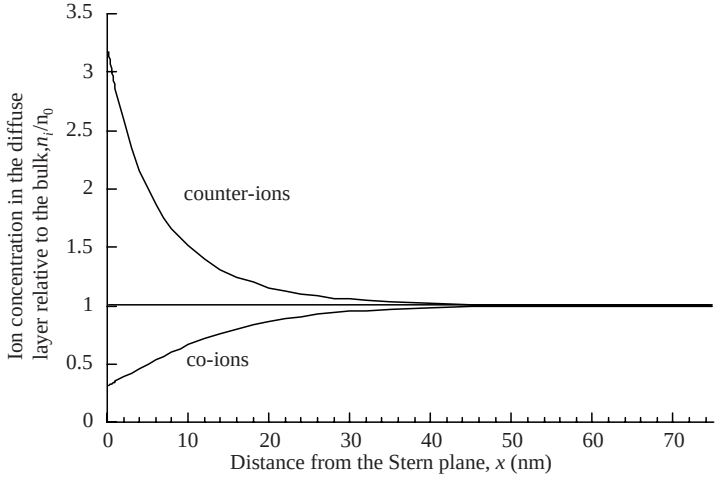


Figure 4.2 Local ion concentration profiles: $\psi(0) = -30$ mV, in 10^{-3} mol dm^{-3} NaCl

charges per unit volume, that is, the space charge density, ρ, is given by

$$\rho = \sum_i n_i z_i e \tag{4.3}$$

In order to make the algebra more transparent, we can take the case of a symmetrical electrolyte where $z_- = -z_+$. Using Equation 4.2, we can write the net space charge density as

$$\begin{aligned} \rho &= n_0 z e \left\{ \exp\left[-\frac{ze\psi(x)}{k_B T} \right] - \exp\left[\frac{ze\psi(x)}{k_B T} \right] \right\} \\ &= -2n_0 z e \sinh\left[\frac{ze\psi(x)}{k_B T} \right] \end{aligned} \tag{4.4}$$

recalling the identity $\sinh(x) = (e^x - e^{-x})/2$.

Now, Poisson's equation relates the space charge density to the curvature of the potential, as follows:

$$\nabla^2 \psi(x) = -\frac{\rho}{\varepsilon_0 \varepsilon_r} \tag{4.5}$$

where the Laplace operator $\nabla^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2$, whereas for a planar surface $\nabla^2 = d^2/dx^2$. Substituting for ρ in Equation 4.4

gives the *Poisson–Boltzmann equation*, as follows:

$$\nabla^2\psi(x) = \frac{2n_0ze}{\varepsilon_0\varepsilon_r}\sinh\left[\frac{ze\psi(x)}{k_BT}\right] \tag{4.6}$$

Rewriting this for the planar interface:

$$\frac{d^2\psi}{dx^2} = \frac{2n_0ze}{\varepsilon_0\varepsilon_r}\sinh\left[\frac{ze\psi(x)}{k_BT}\right] \tag{4.7}$$

Prior to seeking solutions to this equation, let us specify the boundary conditions that we have:

$$\begin{array}{ll} x = \delta & x = \infty \\ \psi(x) = \psi_\delta & \psi(x) = 0 \\ & \dfrac{d\psi(x)}{dx} = 0 \end{array} \tag{4.8}$$

As an initial proposition, let us look at the case for small potentials where $ze\psi(x)/k_BT < 1$. For a univalent electrolyte we have $z = 1$ and then we have the condition that $\psi(x) < 25$ mV. When x is small, $\sinh(x) \sim x$ and Equation 4.7 becomes

$$\frac{d^2\psi}{dx^2} \approx \frac{2n_0(ze)^2}{\varepsilon_0\varepsilon_r k_BT}\psi(x) \tag{4.9}$$

This is known as the *Debye–Hückel approximation*. Equation 4.9 is a linear homogeneous second-order differential equation whose solution is in terms of exponentials, and with the boundary conditions given above we have the potential as a function of distance away from the Stern layer as

$$\psi(x) \approx \psi_\delta \exp(-\kappa x) \tag{4.10a}$$

where for any electrolyte

$$\kappa = \left[\frac{\sum_i (ez_i)^2 n_{0i}}{\varepsilon_0\varepsilon_r k_BT}\right]^{\frac{1}{2}} \tag{4.10b}$$

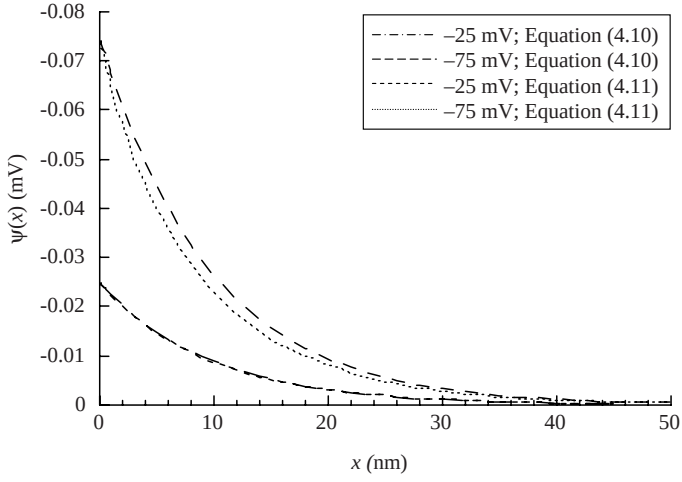


Figure 4.3 Decrease of electrical potential with distance for a planar surface in 10^{-3} mol dm^{-3} NaCl

where κ is the Debye–Hückel parameter which we find in the theory of electrolytes and which controls the rate of decay of potential with distance away from a surface.

We will look in more detail shortly at the behaviour of κ as the electrolyte concentration and type are changed, but first we should consider the solution of Equation 4.7 at higher potentials. With a little manipulation (see Hunter [1] and Russel *et al.* [3], for example), Equation 4.7 may be integrated to give

$$\psi(x) = \frac{2k_{\mathrm{B}}T}{ze} \ln \left[\frac{1 + \exp(-\kappa x) \tanh\left(\dfrac{ze\psi_{\delta}}{4k_{\mathrm{B}}T}\right)}{1 - \exp(-\kappa x) \tanh\left(\dfrac{ze\psi_{\delta}}{4k_{\mathrm{B}}T}\right)} \right] \tag{4.11}$$

recalling that the identity $\tanh(x) = (\mathrm{e}^{x} - \mathrm{e}^{-x})/(\mathrm{e}^{x} + \mathrm{e}^{-x})$ and so $= (\mathrm{e}^{2x} - 1)/(\mathrm{e}^{2x} + 1)$. A comparison of the results for surfaces with potentials of -25 and -75 mV is shown in Figure 4.3, calculated from Equations 4.10 and 4.11. The 1:1 electrolyte concentration in this case was 10^{-3} M.

Let us consider the case where we have a high potential [$\tanh(x) \approx 1$ as $x \gg 1$]. At long distances from the surface, x is large and $\exp(-\kappa x)$ is small, so the logarithms can be expanded as a series and only the leading

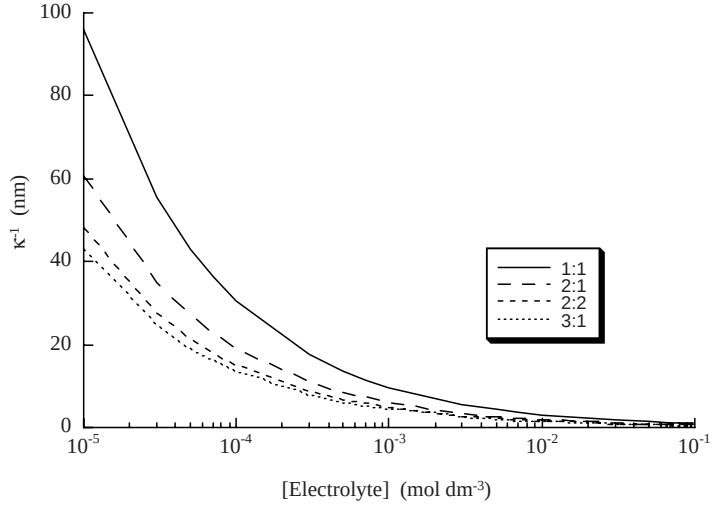


Figure 4.4 The Debye–Hückel decay length as a function of the concentration and type of electrolyte

terms used. Hence at long distances we have

$$\psi(x) \approx \frac{4k_B T}{ze} \tanh\left(\frac{ze\psi_\delta}{4k_B T}\right) \exp(-\kappa x) \tag{4.12}$$

Therefore, we have the exponential fall of potential at large distances and, as the tanh function $\rightarrow 1$ for high potentials, the behaviour appears to be that of a surface with a reduced potential $ze\psi/k_B T \approx 4$. Therefore, at long distances the surface appears to have a potential of 100 mV regardless of how much higher it may be. This is important as we are often interested in the interactions of highly charged surfaces as they approach at separations corresponding to $\kappa x \gg 1$.

The Debye–Hückel parameter characterizes the decay of the potential with distance from the surface. Equation 4.10b shows that is a function of both the concentration of the bulk electrolyte and the valency of the ions. Figure 4.4 gives a plot for various electrolyte types. Thus, when $\kappa x = 1$ the potential will have fallen to 37% of the value at the start of the diffuse layer. For the potential to have fallen to ~1% of the Stern layer value, the distance will have to be as far away from the surface as ~4.5 'decay lengths'. Hence, even at moderate electrolyte concentrations, we can see that the decay of the potential occurs at distances comparable to the dimensions of many colloidal particles.

When the ζ-potential was introduced, it was defined as the potential at the 'shear plane' and was taken as an approximation to the potential of the Stern layer, that is, $\psi_\delta \approx \zeta$, although the location of the shear plane is about a hydrated ion diameter further from the surface. Because it is often easy to measure, it is normally the ζ-potential that we will use for calculating diffuse layer interactions. It is straightforward to estimate the charge at the Stern layer which is, of course, equal but opposite in sign to diffuse layer charge. We can use Equation 4.5 to estimate the charge at a distance δ from the surface:

$$\sigma_\delta = -\int_\delta^\infty \rho \mathrm{d}x = \varepsilon_0 \varepsilon_\mathrm{r} \int_\delta^\infty \frac{\mathrm{d}^2\psi(x)}{\mathrm{d}x^2}\mathrm{d}x \tag{4.13}$$

so that

$$\sigma_\delta = \varepsilon_0 \varepsilon_\mathrm{r} \left[\frac{\mathrm{d}\psi(x)}{\mathrm{d}x}\right]_\delta^\infty \tag{4.14}$$

Our boundary condition is that at infinity the slope of the potential–distance curve is zero and so we have

$$\sigma_\delta = -\varepsilon_0 \varepsilon_\mathrm{r} \frac{\mathrm{d}\psi_\delta}{\mathrm{d}x} \tag{4.15}$$

Equation (4.10a) gives us the distance dependence of the potential for small potentials, with the slope at δ given by

$$\left.\frac{\mathrm{d}\psi(x)}{\mathrm{d}x}\right|_\delta = -\kappa\psi_\delta \exp(-\kappa\delta) = -\kappa\psi_\delta \text{ as } \kappa\delta \to 0 \tag{4.16}$$

so that

$$\sigma_\delta = \kappa\varepsilon_0\varepsilon_\mathrm{r}\psi_\delta \approx \kappa\varepsilon_0\varepsilon_\mathrm{r}\zeta \tag{4.17}$$

For higher potentials, we can use the same route but using Equation 4.7 instead of Equation 4.10 to yield [1]

$$\left.\frac{\mathrm{d}\psi(x)}{\mathrm{d}x}\right|_\delta = -\frac{2\kappa k_\mathrm{B}T}{ze}\sinh\left(\frac{ze\psi_\delta}{k_\mathrm{B}T}\right) \tag{4.18}$$

so that

$$\sigma_\delta \approx \frac{2\kappa\varepsilon_0\varepsilon_r k_B T}{ze} \sinh\left(\frac{ze\zeta}{2k_B T}\right) \tag{4.19}$$

4.3.5.3 The diffuse layer and spherical particles

The solution of the Poisson–Boltzmann equation is not available in an analytical form except when the Debye–Hückel approximation for low potentials is used. As there is spherical symmetry, it is usual to use spherical polar coordinates instead of Cartesian coordinates so that the equivalent form of Equation 4.7 for a symmetrical electrolyte with r as the distance from the centre of the particle is

$$\frac{1}{r^2}\frac{\partial}{\partial r} r^2 \frac{\partial \psi(r)}{\partial r} = \frac{2zen_0}{\varepsilon_0\varepsilon_r} \sinh\left[\frac{ze\psi(r)}{k_B T}\right] \tag{4.20}$$

With the Debye–Hückel approximation, we obtain the diffuse layer potential as a function of distance (for a particle of radius a):

$$\psi(r) = \psi_\delta \frac{a}{r} \exp\left[-\kappa\,(r-a)\right] \tag{4.21}$$

Here, we are assuming δ to be very much smaller than a. We can also estimate the Stern layer charge along similar lines to the flat surface, to give

$$\sigma_\delta = \frac{\varepsilon_0\varepsilon_r\psi_\delta\,(1+\kappa a)}{a} \tag{4.22}$$

In addition, the charge at the 'shear plane' can be estimated if the ζ-potential is substituted for the Stern potential.

For potentials in excess of 25 mV, we need numerical solutions to the Poisson–Boltzmann equation. Loeb *et al.* [4] have provided extensive tables from such numerical calculations and these authors, and others, have also supplied some approximate analytical expressions (see, for example, Hunter [1] and Oshima *et al.* [5]). As an example, a useful expression for the charge of the Stern layer, approximating it to that at

the 'shear plane', would be [4]

$$\sigma_\delta \approx \frac{\varepsilon_0 \varepsilon_r \kappa k_B T}{ze}\left[2\sinh\left(\frac{ze\zeta}{2k_B T}\right) + \frac{4}{\kappa a}\tanh\left(\frac{ze\zeta}{4k_B T}\right)\right] \qquad (4.23)$$

4.4 THE INTERACTION BETWEEN DIFFUSE DOUBLE LAYERS

As two charged surfaces approach each other, they interact. If the sign of the charge is the same, they usually repel each other; alternatively, if the charges are opposite, they attract. In a vacuum, the repulsion is described by Coulomb's law but in a liquid medium the interaction is screened by the ion atmosphere. The algebra for two planar surfaces is easiest to follow. The model used to calculate the repulsion uses the ion concentration between the surfaces to give an osmotic pressure difference between the surfaces and the bulk electrolyte, that is, the *excess osmotic pressure*. The potential distribution in the intervening space is assumed to be the sum of the potentials due to the two surfaces. This is illustrated in Figure 4.5, where we plot the potential distribution between two flat surfaces with a small overlap of diffuse layers.

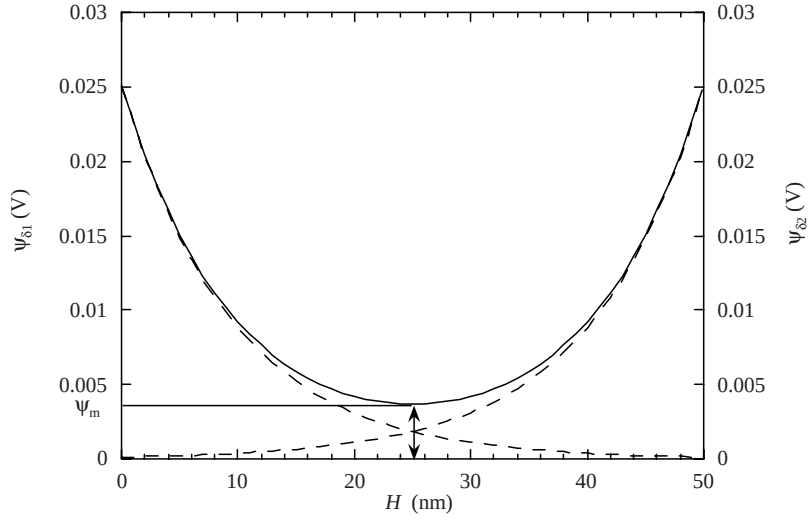


Figure 4.5 The potential profile between two planar surfaces at 25 mV, in 10^{-3} mol dm^{-3} NaCl

4.4.1 The Interaction Between Similar Flat Plates

The simplest route is to estimate the excess osmotic pressure at the mid-plane position between two surfaces a separated by a distance h, that is, at $h/2$ where $\mathrm{d}\psi(x)/\mathrm{d}x = 0$ and $\psi(x) = \psi_\mathrm{m}$. Once we have the excess osmotic pressure, the potential energy per unit area of surface is calculated from integration of the force (osmotic pressure) that we overcome in bringing the two surfaces from an infinite distance apart to position H.

The difference in the ion concentration at the mid-plane and the bulk electrolyte gives us the osmotic pressure, $\Pi(H)$:

$$\Pi\,(H) = k_\mathrm{B}T\,(n_+ + n_- - 2n_0) \tag{4.24}$$

Using the Boltzmann distribution of ions, Equation 4.2, we can rewrite Equation 4.24 for a symmetrical electrolyte as

$$\Pi\,(H)\,2n_0k_\mathrm{B}T\left[\cosh\left(\frac{ze\psi_\mathrm{m}}{k_\mathrm{B}T}\right) - 1\right] \tag{4.25}$$

[recalling that $\cosh(x) = (\mathrm{e}^x + \mathrm{e}^{-x})/2$]. For small values of x, we can expand $\cosh(x) \approx 1 + (x^2/2)$ and so for small potentials we have

$$\Pi(H) \approx \frac{\kappa^2\varepsilon_0\varepsilon_\mathrm{r}}{2}\psi_\mathrm{m}^2 \tag{4.26}$$

Substituting for the mid-plane potential $\psi_\mathrm{m} = 2\psi(H/2)$, we now have

$$\Pi(H) \approx \frac{\kappa^2\varepsilon_0\varepsilon_\mathrm{r}}{2}\left[2\psi_\delta\exp\left(-\kappa H/2\right)\right]^2 = 2\kappa^2\varepsilon_0\varepsilon_\mathrm{r}\psi_\delta^2\exp\left(-\kappa H\right) \tag{4.27}$$

and the potential energy is

$$V_\mathrm{R} = -\int_\infty^H \Pi\,(H)\,\mathrm{d}H \tag{4.28}$$

which, with the boundary condition that $\Pi(H) \to 0$ as $H \to \infty$, gives

$$V_\mathrm{R} = 2\kappa\varepsilon_0\varepsilon_\mathrm{r}\psi_\delta^2\exp\left(-\kappa H\right) \tag{4.29a}$$

namely

$$V_R \approx 2\kappa\varepsilon_0\varepsilon_r\zeta^2 \exp(-\kappa H) \tag{4.29b}$$

The same route is taken for systems with higher potentials where we cannot justify the simplifying Debye–Hückel approximation. If we look at weak overlap of the diffuse layers so that $\kappa H > 1$, we can still simply add the local potentials that we estimate from the isolated surfaces (Equation 4.12). Following the same route as used above for the low-potential case, we can derive the resulting expression for V_R:

$$V_R = \frac{64 n_0 k_B T}{\kappa} \tanh^2\left(\frac{ze\psi_\delta}{4k_B T}\right) \exp(-\kappa H) \tag{4.30}$$

4.4.2 The Interaction Between Dissimilar Flat Plates

Of course, the interaction will not always be between two identical surfaces. For two flat plates with potentials $\psi_{\delta 1}$ and $\psi_{\delta 2}$, we have in the Debye–Hückel limit [1]:

$$V_R = \frac{\varepsilon_0\varepsilon_r\kappa}{2}\left[\left(\psi_{\delta 1}^2 + \psi_{\delta 2}^2\right)\left[1 - \coth(\kappa H)\right] + 2\psi_{\delta 1}\psi_{\delta 2}\,\mathrm{cosech}(\kappa H)\right] \tag{4.31}$$

Some results using Equation 4.31 are plotted in Figure 4.6 for two surfaces with potentials of -25 and -15 mV. Also plotted for reference is the result for two surfaces of equal potential at -20 mV. We see a maximum in the interaction energy at $\kappa H \approx 0.5$. At close separations, the potential energy decreases and eventually becomes attractive. At the maximum of the curve, the electrostatic interaction force is zero and is attractive in the above example when $\kappa H < 0.5$. This is an exciting result as it indicates why some materials with a similar sign charge may stick together or *heterocoagulate*. For example, we could picture a protein-coated surface of a cell close to its isoelectric point tending to stick to clay particles just due to electrostatic interactions. Of course, dispersion forces are also usually attractive and will aid the process.

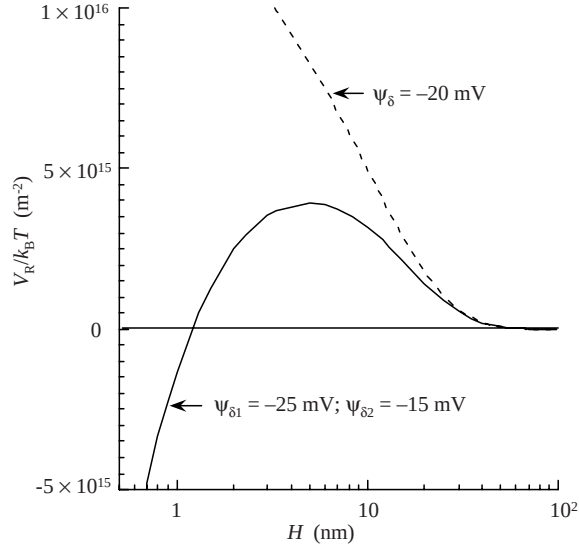


Figure 4.6 Potential energy of interaction for two flat plates

4.4.3 Constant Charge or Constant Potential?

At this point, we should note that the charge at the Stern plane is related to the local potential gradient (Equation 4.15). If we examine Figure 4.5 and keep the potential at the surfaces as fixed points, when we bring the two plates close together the slope of the potential decreases due to the summation of the potential distributions. This means that the local charge changes (decreases) as a result of the constant potential. Hence the result expressed by Equations 4.29 and 4.30 are strictly for that constant potential condition. In fact, we obtain an identical result with both the Debye–Hückel approximation and weak overlap condition if we keep the charge constant. At higher potentials and greater overlap, this is not the case and we have different results if we work at constant potential or constant charge. Where appropriate, the inclusion of ψ or σ in the subscript will be used to indicate if the equation relates to constant potential or constant charge. The choice of the appropriate equation would seem to be easily related to the type of surface that we have. For example, a surface with a charge due to isomorphous substitution in a lattice should be a good candidate for a constant charge surface. On the other hand, a surface where the charge is controlled by the concentration of potential-determining ions in solution might best be

treated as a constant-potential surface. In practice, the choice is rarely clear cut. However, care should be exercised when using the constant-charge expressions at small separations as the approximations become inadequate, predicting much too strong a repulsion at $\kappa H < 1$.

4.5 THE INTERACTION BETWEEN TWO SPHERES

A similar approach to the repulsion between flat plates is taken in estimating the repulsion between two spheres. The problem of solving the potential distribution around a sphere is carried over to the problem of overlapping diffuse layers. However, there are a number of approximate expressions available (for example, see Hunter [1], Russel *et al.* [3] and Verwey and Overbeek [6]). Here, we give the approximate equations with an indication of where they can be used rather than discuss the details of the derivations.

4.5.1 Identical Spheres at $\kappa a < 5$

Here, the diffuse layer is becoming of a similar magnitude to the particle radius. With the Debye–Hückel approximation for the potential around each sphere and simply summing the two, that is, weak overlap, we have

$$V_{\psi \mathrm{R}} = 2\pi \varepsilon_0 \varepsilon_\mathrm{r} a \psi_\delta^2 \exp(-\kappa H) \tag{4.32}$$

As with the case of flat plates we have $V_{\sigma \mathrm{R}} = V_{\psi \mathrm{R}}$. This solution is best for weak overlap so that at $\kappa H > 2$ it is satisfactory over a wide range of κa. As $\kappa a \to 1$ or less, it is useful at close distance of approach [7].

4.5.2 Identical Spheres at $\kappa a > 10$

In this case, the diffuse layer is thin compared with the particle radius. The results for low potentials, with H being the closest distance between the surfaces of the particles (i.e. $H = r - 2a$ with r the centre-centre distance) are as follows:

$$V_{\psi \mathrm{R}} = 2\pi \varepsilon_0 \varepsilon_\mathrm{r} a \psi_\delta^2 \ln[1 + \exp(-\kappa H)] \tag{4.33a}$$

$$V_{\sigma \mathrm{R}} = -2\pi \varepsilon_0 \varepsilon_\mathrm{r} a \psi_\delta^2 \ln[1 - \exp(-\kappa H)] \tag{4.33b}$$

The constant-potential expression, Equation 4.33a, works well for all separations and is acceptable down to $\kappa a > 2$ if at close approach such that $\kappa H < 2$ [3]. The constant-charge expression, Equation 4.33b, should be used with caution, especially at close approach as a large overestimate can be obtained for the repulsive potential.

4.5.3 Spheres of Radii a_1 and a_2 at $\kappa a > 10$

The equivalent form of Equations 4.33a for spheres of any radii was given by Hogg *et al.* [8] as

$$V_{\psi R} = \frac{\pi \varepsilon_0 \varepsilon_r a_1 a_2}{a_1 + a_2} \left(\psi_{\delta 1}^2 + \psi_{\delta 2}^2\right) \left\{ \frac{2\psi_{\delta 1}\psi_{\delta 2}}{\psi_{\delta 1}^2 + \psi_{\delta 2}^2} \ln \left[\frac{1 + \exp(-\kappa H)}{1 - \exp(-\kappa H)} \right] + \ln\left[1 - \exp(-2\kappa H)\right] \right\} \tag{4.34}$$

4.5.4 A Sphere and a Plate

The result is that the potential energy is twice that estimated for two similar spheres, so that, for example, at $\kappa a < 5$ at weak overlap we can write

$$V_{\psi R} = 4\pi \varepsilon_0 \varepsilon_0 a \psi_\delta^2 \exp(-\kappa H) \tag{4.35}$$

and for close approach and large κa

$$V_{\psi R} = 4\pi \varepsilon_0 \varepsilon_r a \psi_\delta^2 \ln\left[1 + \exp(-\kappa H)\right] \tag{4.36}$$

4.6 THE EFFECT OF PARTICLE CONCENTRATION

Up to this point, we have only considered the interaction of two isolated colloidal particles. It is reasonable to neglect the presence of other particles if the time-average separation of the particles is very much larger than the range of the diffuse layer. Once the separation becomes similar in magnitude to the range of the diffuse layer, it becomes untenable to consider the osmotic pressure between a pair of particles relative to a 'bulk value'. The background electrolyte now contains other particles with

their counter-ions. The simplest approach is to consider the additional concentration of ions due to the particle counter-ions as an addition to the ions of the background electrolyte. Of course, each charged particle is a 'macro-ion', but the number is very much smaller than the number of the corresponding counter-ions, hence the particle contribution may be ignored without introducing a large error. However, once the volume fraction of the particles becomes high, the volume occupied by the particles must be allowed for as this volume is excluded to the ions. Russel *et al.* [3] gave a convenient expression for κ for a symmetrical electrolyte with ions of valency z:

$$\kappa = \left(\frac{e^2}{\varepsilon_r \varepsilon_0 k_B T} \frac{2z^2 n_0 - \dfrac{3\sigma_0 z\varphi}{ae}}{1-\varphi} \right)^{\frac{1}{2}} \tag{4.37}$$

The $1-\varphi$ term in the denominator corrects the ion concentration for the volume occupied by particles. σ_0 is the surface charge density, which is often measurable by titration. If the Stern layer charge were available, it would be better to use that value as the strongly bound counter-ions are effectively removed from osmotic activity. The ζ-potential could be used to give an estimate of σ_δ ; $(\sigma_0/e)\left(4\pi a^2\right)$ is the number of charges per particle of radius a and surface charge density of σ_0 and so $(\sigma_0/ze)\left(4\pi a^2\right)$ is the number of counter-ions associated with each particle. $3\varphi/4\pi a^3$ is the number of particles per unit volume. The product of these two terms and z^2 gives the counter-ion contribution to the equation. Now, the counter-ions are assumed to have the same activity as would an ion in bulk solution, so we are not making any correction for the structuring effect of the charged particles. To do so would require a complex statistical mechanical analysis. However, the approximation given in Equation 4.37 is a useful correction for many concentrated dispersions. This is most important for small charged particles where we have added little extra electrolyte.

As an illustration of the problem, we may consider two suspensions of charged spherical particles in a dilute solution of sodium chloride. Polymer latex particles would be a good example here as they can be produced with fairly high surface charge densities and also in many different sizes. Figure 4.7 shows how the range of the diffuse layer can vary with particle concentration with a plots of the Debye length, κ^{-1}, against the volume fraction, φ.

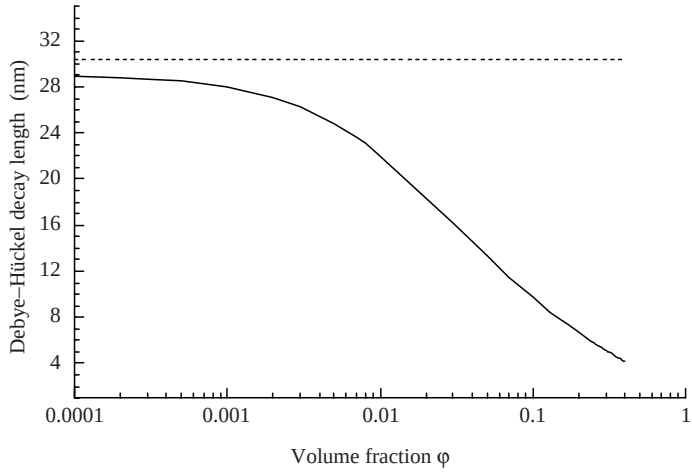


Figure 4.7 The Debye–Hückel parameter as a function of φ: $a = 200$ nm; [NaCl] $= 10^{-4}$ mol dm^{-3}; $\sigma_0 = -10$ μC cm^{-2}. The dashed line represents the limiting value for zero volume fraction

There is a factor of nearly 10 change in the Debye length, which, because of the square root relationship with ion concentration, means that the background electrolyte is completely swamped by the counter-ion concentration as the dispersion becomes concentrated. At volume fractions of ~0.35 we are at similar concentrations to many coatings. The effect on the electrostatic repulsive potential is illustrated in Figure 4.8, where curves are plotted for three concentrations of particles. The value of 10 μC cm^{-2} represents a higher value than that found in many systems, which are often in the range 0.1−1 μC cm^{-2}. However, even at these lower values the we obtain values of 12 nm $< \kappa^{-1} <$ 20 nm at high solids contents, and so with the exponential functionality in the repulsive energy expression is still a significant correction.

The addition of ionic surfactants to dispersions is common. They are added to aid dispersion and to provide charge to prevent aggregation. The density of surfactant molecules on the surface of the particles usually corresponds to a monolayer and so the surface charge density is high. However, the Stern charge is lower as many counter-ions are strongly bound, but it is also important to recognize that the ionic surfactants are also electrolytes. The problem is slightly more complex, however, as the surfactant molecules form micelles, which limits their numbers in free solution and the micelles also bind the counter-ions to the amphiphilic molecule. However, the diffuse layer counter-ions should still be included

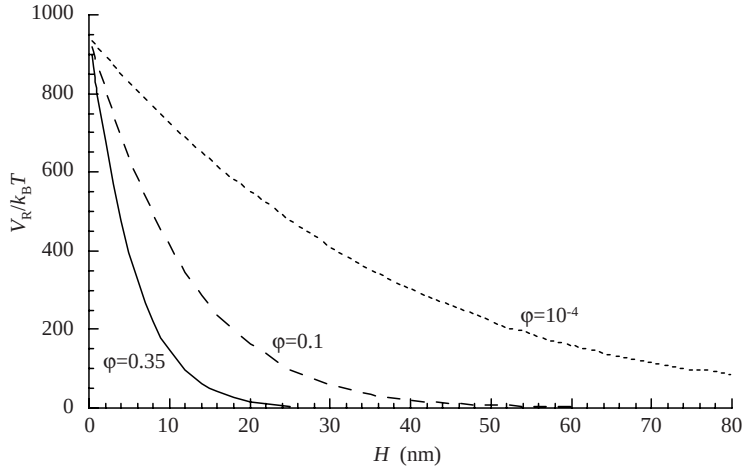


Figure 4.8 Electrostatic repulsion for three different concentrations of particles (using Equation 4.33a): $a = 200$ nm; [NaCl] $= 10^{-3}$ mol dm^{-3}; $\psi_\delta = -80$ mV; $\sigma_\delta = -10$ μC cm^{-2}

and the micelles can act as small particles, and we will need to keep this in mind when we consider the behaviour of concentrated dispersions.

4.7 STERIC INTERACTIONS

In many colloidal systems, we add non-ionic materials which adsorb on the surfaces of the particles. The purpose is to prevent aggregation by keeping the surfaces apart. For this to be effective, the molecules must be firmly attached, but must also extend into solution so that the surfaces cannot approach too closely, and this gives the term 'steric interaction'. The types of molecules that are effective for this purpose are non-ionic surface-active block copolymers. (Charged species such as proteins and polyelectrolytes are also used, but we will delay discussion of these until we discuss the full picture of colloid stability.) Simple homopolymers are rarely used for this purpose as they need to be chemically grafted to the surface to prevent desorption if they are lyophilic, and if they are lyophobic they will be difficult to get into solution and would form a dense, ineffective layer. In this section, we will concentrate on materials that are most effective as stabilizers, as this is where the important applications lie.

So, the picture emerging is of molecules with one or more strong anchoring moieties but with other moieties which are in a good solvent. This means that we can design effective systems that can operate as stabilizers in high electrolyte concentrations or in non-aqueous (or, better, low-polarity) media. The simple linear non-ionic surfactants can be used in addition to larger molecules. The model that we are going to work with is relatively simple, but it is important to recognize that the problem is many faceted and necessarily complex if a full treatment is to be attempted. In outline, what we will discuss is the approach of two surfaces that are fully covered by strongly attached polymer (or non-ionic surfactant) layers. As the surfaces approach, the outermost parts of the layers start to mix as the molecules interpenetrate. The result is an increase in free energy. At close approach, we may picture an additional constraint as the surface–surface separation becomes less than the dimension of a single layer. This would result in a further increase in free energy as the chains are 'compressed', that is, there are even fewer possible configurations available to each molecule, resulting in a marked decrease in entropy. However, this is not a problem that we shall have to deal with when considering good steric stabilizers, as we shall see that the repulsion becomes very large before such a close approach is attained. Hence only the outermost parts of the layers that interpenetrate in systems are of practical importance, although we could construct systems where this would not be the case.

Before we can begin to construct our model, we must have some information on the composition of the layers attached to the surfaces. Figure 4.9 illustrates some of the scenarios that we must consider. The schematic in Figure 4.9a illustrates a surface covered with terminally anchored polymer molecules. The local polymer concentration, ρ_2, shows a maximum a small distance away from the surface with the concentration decreasing further out. Remember that the coils are dynamic and we are trying to obtain the closest lateral spacing. The profile is distorted from that which we could expect from an isolated coil, first, by the proximity of the particle surface, and second, by the crowding of the neighbouring coils. Hence the extension away from the surface should be slightly greater than the dimension of an isolated coil in solution. If we consider a block or graft copolymer strongly attached to the surface, we may expect the profile to be slightly denser as the adsorbing sections of the molecule force the soluble moieties closer together. The extension away from the surface is a maximum for the longest tails where only one end is attached. Loops are pulled to the surface at both ends and tend to have a shorter extension. The adsorbed sections ('trains'), of course, are

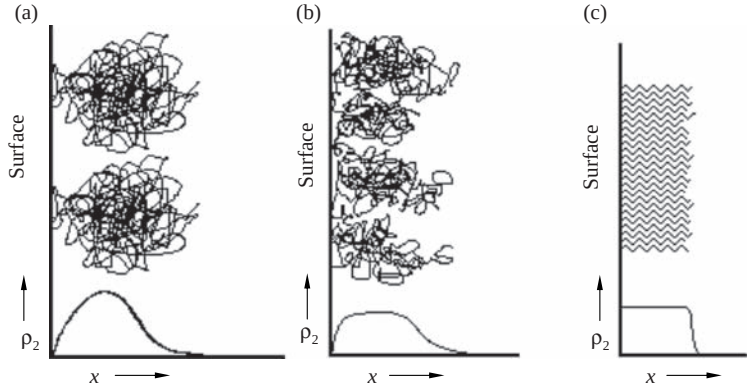


Figure 4.9 Schematics of the concentration profiles of adsorbed layers: (a) terminally anchored homopolymer; (b) A–B block copolymer; (c) polymer 'brush'

held close to the surface. This is illustrated in Figure 4.9b, where the profile is shown as relatively uniform until the outer portion of the tails are reached. Lastly, in Figure 4.9c, we illustrate an adsorbed surfactant layer which has a uniform concentration profile with a short-range reducing profile due to the variation in chain length. The polydispersity of most surfactants tends to be less than we find with most polymer systems. The polydispersity of the latter means that the outer profile of the layer can be expected to fall more slowly than we might expect from the periphery of an isolated chain.

In summary, we would like to have three pieces of information:

- the amount of material attached to unit area of surface;
- an adsorbed layer 'thickness';
- the profile of the outer part of the layer.

The experimental estimation of each of these features has been discussed in Chapter 2. In this section, we will restrict our discussion to the interaction between two coated particles. Furthermore, the model will be restricted to that of layers of uniform concentration, as this will simplify the algebraic description and make the route more transparent. This represent a less serious limitation than might be apparent at first sight, as frequently effective steric stabilization is provided by densely packed layers of non-ionic surfactants or branched polymer chains forming part of a block copolymer. The interaction calculated using this model will result in a steeper pair-potential than may actually be the case, but the

softness is nearly always the result of stabilizer polydispersity and this is not easy to model effectively. We will return to this point at the end of the discussion.

The starting point for our model is the recognition that the steric interaction is zero until the outer periphery of the stabilizer layer on one particle comes into contact with that of an adjacent particle. Closer approach means that the polymer or surfactant chains start to mix. The chain concentration in the overlap volume is higher than that in either of the isolated layers, and this results in an osmotic pressure acting to separate the particles. If the thickness of the layer is δ, the interaction occurs when the particle surfaces are at separations $<2\delta$. This is the interaction that we will model.

Should the particle separation become $<\delta$, a second effect should be modelled. The surface of the second particle further restricts the possible conformation of the chains attached to the opposing surface, giving rise to an additional contribution to the interaction energy. This *volume restriction* term will be neglected here as in any successful steric stabilizer this degree of overlap of layers will never occur as only a small degree of overlap is sufficient to produce potentials much greater than $k_B T$.

In Figure 4.10, we see the overlap of adsorbed layers of uniform concentration for several geometries for layers which have a thickness δ and with a particle surface-to-surface separation distance H. The concentration of stabilizer in each layer is c_2 and so the concentration in the overlap region is $2c_2$. This means there is a higher osmotic pressure in the region where the two layers are mixing than before they came into contact. We will define this as the excess osmotic pressure, Π_E, which acts to separate the particles. In order to calculate the force, we must

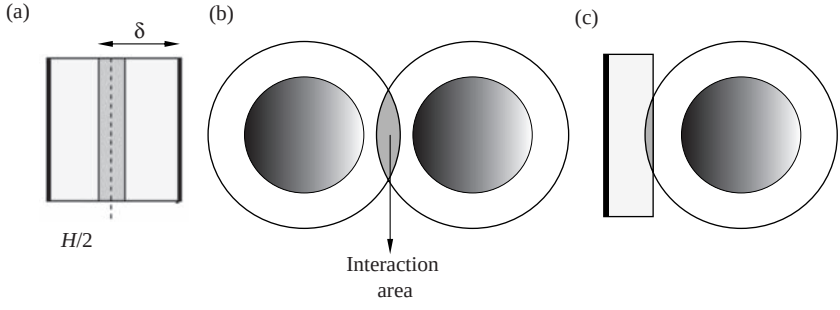


Figure 4.10 Schematics of steric interactions – overlap of adsorbed layers for different geometries: (a) plate–plate; (b) sphere–sphere; (c) sphere–plate

define the area over which the pressure is acting. For the examples illustrated in Figure 4.10, this is unit area for the plates and the area at the base of the spherical cap which is located at the mid-plane for the other two interactions. The interaction energy between the particles (which originates from the free energy of mixing the polymers in the solvent) is then difference in the free energy of two particles at overlap relative to the value just at contact. The value can be calculated by integrating the force with the distance between the particles:

$$V_s = \int_{2\delta}^{H} -\Pi_E A dx \tag{4.38}$$

Of course, the integral of Adx gives the overlap volume, which we may define as v_o.

The Flory–Huggins model of polymers in solution gives the osmotic pressure of a polymer solution as a function of concentration (as discussed in Chapter 2):

$$\Pi = \frac{RTc_2}{M_2} + RTB_2c_2^2 + \cdots \tag{4.39}$$

where B_2 is the second virial coefficient:

$$B_2 = \left(\frac{\bar{v}_2}{M_2}\right)^2 \frac{1}{\bar{v}_1}(0.5 - \chi) \tag{4.40}$$

with the subscripts 1 and 2 referring to the solvent and polymer species, respectively, as earlier. We can now use Equation 4.39 to calculate the energy difference as

$$V_s = -(2v_o\Pi_{x=2\delta} - v_o\Pi_{x=H}) \tag{4.41}$$

and so

$$V_S = \left[\left(\frac{RT \times 2c_2}{M_2} + RTB_2(2c_2)^2\right) - 2\left(\frac{RTc_2}{M_2} + RTB_2c_2^2\right)\right] v_o \tag{4.42}$$

which gives

$$V_S = 2RTB_2c_2^2v_o \tag{4.43a}$$

$$V_S = \frac{2RT}{\bar{v}_1}\left(\frac{c_2\bar{v}_2}{M_2}\right)^2 (0.5 - \chi)\, v_o \tag{4.43b}$$

It should be noted in Equation 4.43b that the term in parentheses that is squared is just the volume fraction of polymer in the layer. For the three geometries illustrated in Figure 4.10, we have the overlap volumes as follows:
plate–plate:

$$v_o = 2\delta - H \tag{4.44a}$$

sphere–sphere:

$$v_o = \frac{2\pi}{3}(\delta - H/2)^2 (3a + 2\delta + H/2) \tag{4.44b}$$

sphere–plate:

$$v_o = \frac{4\pi}{3}(\delta - H/2)^2 (3a + \delta + H) \tag{4.44c}$$

These are the same expressions as given by Hunter [1] and Napper [9]. Other models are also available, such as those given by Russel *et al.* [3], Dolan and Edwards [10] and deGennes [11], but the above is sufficient for our purpose here.

The expressions in Equations 4.43a and 4.43b indicates what we can expect for sterically stabilized particles. The major features are as follows:

(a) There is a marked change with temperature.
(b) The repulsion increases rapidly as the polymer concentration in the layer is increased.
(c) There is an increase in repulsion as the separation decreases.

An increase in temperature can be expected to give an increased repulsion with non-aqueous systems directly from the RT term, but also from the decrease in the χ parameter as the solubility improves. However,

especially in aqueous systems, this is not always the case as $\chi \rightarrow 0.5$ as $T \rightarrow$ LCST (where LCST is the *lower consolute solution temperature*, above which phase separation of the polymer takes place). As this occurs, repulsion approaches zero, the adsorbed layer collapses and further temperature increases result in attraction.

As shown in Equations 4.43a and 4.43b, the repulsion increases with the square of the polymer concentration in the adsorbed layer. A more rapid increase than this would have been included if we had not restricted ourselves to the simple Flory–Huggins treatment. In principle, we could use the osmotic pressure obtained experimentally from polymer solutions of similar concentration. However, that would not necessarily provide a better description as polymer attached to a surface is not exactly the same as similar molecules that are free in solution. In addition, the model assumes that higher polymer concentration is localized to the overlap volume, which is an oversimplification. Some rearrangement must occur, resulting in a larger but poorly defined interaction volume. This 'local dilution' would easily counteract the improvement of a more complex expression for the osmotic pressure due to the latter's marked dependence on concentration. At this point, it is relevant to recall the approximations made in developing the Flory–Huggins model. This was based on a simple mixing concept, using a lattice model with the molar volumes of the solvent and the chain segments being equal. The χ parameter was expressed as an enthalpic term. Now, the χ parameter depends on both the temperature and the polymer concentration as there is also an entropic contribution [9]. The value of χ increases as both the temperature and polymer concentration increase. This would reduce magnitude of the interaction energy as the surfaces approach.

The greatest problem, however, is concentration detail in the early overlap regime. This is the point at which polydispersity is really controlling the softness of the interaction and a detailed profile of the outermost layer concentration is needed. Although there are some models for the concentration profiles of adsorbed layers and some experimental measurements, this information is poorest at the outermost extremity of the layers. This is the region that is of greatest interest in systems with good steric stabilizers as it is the curvature here that we require when we need to estimate interparticle forces or rheological properties.

In order to illustrate how the system could be modelled, let us consider the simplified model for the concentration profile shown in Figure 4.11. Here, there is a uniform polymer concentration extending from the surface to a distance δ_1. The polymer concentration then decreases monotonically until it is negligible, at δ_m. A practical system which would

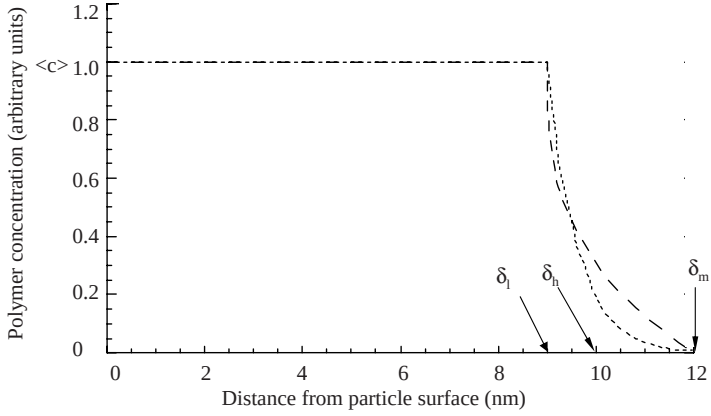


Figure 4.11 Illustration of the simplified model for adsorbed polymer layers, using the relationships shown in Equations 4.45a–c to give the concentration decrease of the outer layer: (····) polymer concentration, exponential decrease (Equation 4.45b); (– – –) polymer concentration, cube root decrease (Equation 4.45c)

have a profile somewhat similar to this would be a particle stabilized with poly(12-hydroxystearic acid), a useful stabilizing moiety for particles dispersed in hydrocarbons. This heavily branched stabilizer is predominantly a hexamer but with pentamers and septamers also present. The profiles shown in Figure 4.11 are given by the following equations:

$$x < \delta_1; \quad c = \langle c \rangle \tag{4.45a}$$

$$\delta_1 < x < \delta_m; \quad c = \exp\left[-5\left(\frac{x - \delta_m}{\delta_m - \delta_1}\right)\right] \tag{4.45b}$$

$$c = 1 - \left(\frac{x - \delta_m}{\delta_m - \delta_1}\right)^{\frac{1}{3}} \tag{4.45c}$$

It should be stressed that both rates of decay are purely illustrative. The exponential decay rate was chosen to give a very low concentration at δ_m. A hydrodynamic thickness of $\delta_h = 10$ nm is also marked on the figure and serves to illustrate that we might expect repulsion prior to a surface–surface separation of $2\delta_h$. We can use Equations 4.45a–c in conjunction with Equations 4.44a–c to describe the concentration variation prior to the integration that is used to give the energy. Figure 4.12 illustrates the result for two plates covered with polymer 'brushes' whose concentration profile could be that shown as the exponential decay in Figure 4.11. It is interesting that not only is the decay very

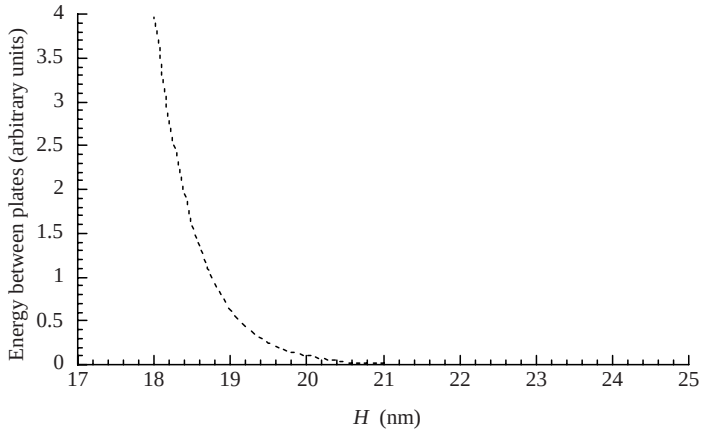


Figure 4.12 Illustration of the interaction between two plates with polymer 'brush' layers, displaying an exponentially decaying concentration of the outer periphery of each layer

rapid as the interaction 'follows' the square of the concentration but also that we do not see a significant repulsion until the surface–surface separation approaches $2\delta_h$. After this, it rises rapidly over the next 1–2 nm.

4.8 CALCULATION STRATEGY

Although the above discussion illustrates that a knowledge of the details of the concentration profile is desirable, it is difficult to obtain this information experimentally. Techniques such as neutron scattering lose their sensitivity at the low local concentrations at the periphery. In many practical circumstances, we prepare colloidal systems with the most effective steric stabilizers that we can. Thus, we can frequently use a uniform densely packed profile and the value of δ_h can be measured. The interaction energy will increase very rapidly as soon as the two layers come into contact. This situation can be approximated to

$$H > 2\delta_h \qquad V_S = 0 \tag{4.46a}$$

$$H < 2\delta_h \qquad V_S = \infty \tag{4.46b}$$

In the case of non-ionic surfactants, the stretched surfactant chain length would give a suitable distance for the repulsive wall. On combining this

with the attractive potentials, we can get an indication as to whether or not the attraction has been reduced to a few $k_B T$ or less so that the system can readily be handled. Only occasionally will we require more detailed information on the potential and then we have to model the concentration profile of the stabilizer layer. However, in aqueous systems or those of intermediate polarity, the electrostatic repulsion will control the pair-potential profile; the ζ-potential is the potential at a distance from the surface corresponding to δ_h.

As an illustration, we may take the problem of the interaction of terminally anchored polymer chains, which has been analysed in some detail by both Dolan and Edwards [10] and deGennes [12]. The expression using the latter's work for the force between two planar surfaces is as follows:

$$F_T(H) = k_B T \Gamma^{\frac{3}{2}} \left[\left(\frac{H}{2\delta} \right)^{-\frac{9}{4}} - \left(\frac{H}{2\delta} \right)^{\frac{3}{4}} \right] - \frac{A}{6\pi H^3} \tag{4.47}$$

where Γ is the number of chains per unit area of surface. As an illustrative calculation, we may take a δ value of 10 nm, a surface density of polymer of one chain every 100 nm^2 and a Hamaker constant for the system of $A \approx 10^{-20}$ J. The results obtained are shown in Figure 4.13 and illustrate the very rapid rise in repulsion at 2δ as the two 'brush' layers begin to

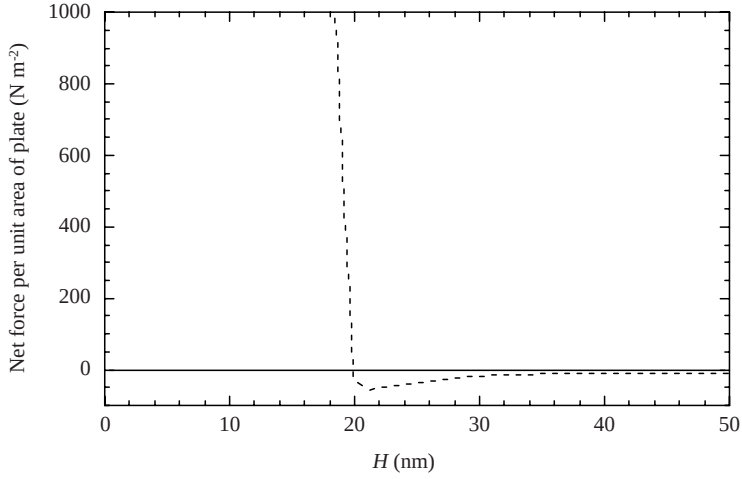


Figure 4.13 Illustration of the net force between two plates: $\delta = 10$ nm; $A = 1 \times 10^{-20}$ J

overlap. Of course, the polydispersity is not taken into account. The very steep wall is of less interest than the details of the minimum at the start of the near vertical rise, and this is where chain polydispersity and weak electrostatics are likely to control the profile.

REFERENCES

1. Hunter, R. J. (1987) *Foundations of Colloid Science*, Vol. **1**, Oxford University Press, Oxford.
2. Lyklema, J. (1991) *Fundamentals of Interface and Colloid Science*, Academic Press, London.
3. Russel, W. B., Saville, D. A. and Schowalter, W. R. (1989) *Colloidal Dispersions*, Cambridge University Press, Cambridge.
4. Loeb, A. L., Overbeek, J. Th. G. and Weirsema, P. H. (1961) *The Electrical Double Layer Around a Spherical Colloid Particle*, MIT Press, Cambridge, MA.
5. Ohshima, H., Healy, T. W. and White, L. R. (1982) *J. Colloid Interface Sci.*, **89**, 446.
6. Verwey, E. J. W. and Overbeek, J. Th. G. (1948) *Theory of Stability of Lyophobic Colloids*, Elsevier, Amsterdam.
7. Glendinning, A. B. and Russel, W. B. (1983) *J. Colloid Interface Sci.*, **93**, 93.
8. Hogg, R., Healy, T. W. and Fursteneau, D. W. (1966) *Trans. Faraday Soc.*, **62**, 1683.
9. Napper, D. H. (1983) *Polymeric Stabilization of Colloidal Dispersions*, Academic Press, London.
10. Dolan, A. K. and Edwards, S. F. (1974) *Proc. R. Soc. London, Ser. A*, **337**, 509.
11. deGennes, P.-G. (1987) *Adv. Colloid Interface Sci.*, **27**, 189.
12. deGennes, P.-G. (1982) *Macromolecules*, **15**, 492.

5

The Stability of Dispersions

5.1 INTRODUCTION

The term 'stability' has two distinct meanings in the context of colloidal dispersions. In many practical situations, it is taken to mean that there are no signs of phase separation over a period of time. If particles in a dispersion showed a tendency to sediment or cream over a period of storage, that dispersion would be termed 'unstable'. However, we also use the term in another context to mean that the particles have no tendency to aggregate. With large or dense particles, these two usages can be contradictory. With very small particles, this may not be the case. Hence we will refer to *colloid stability* when we mean that the particles do not aggregate and *mechanical stability* when they do not sediment. The key to understanding colloid stability is the pair potential. In Chapters 3 and 4 we developed expressions describing the distance-dependence of the various components of the potential energy of interaction of two particles. The total potential energy of interaction, that is, the *pair potential*, is calculated as the sum of each of these components.

5.2 THE STABILITY OF CHARGE-STABILIZED COLLOIDS – THE DLVO THEORY

The linear combination of the dispersion force contribution to the pair potential with the electrostatic repulsion gave the first comprehensive model framework for the stability of colloidal dispersions. The work

Colloids and Interfaces with Surfactants and Polymers, Second Edition James Goodwin

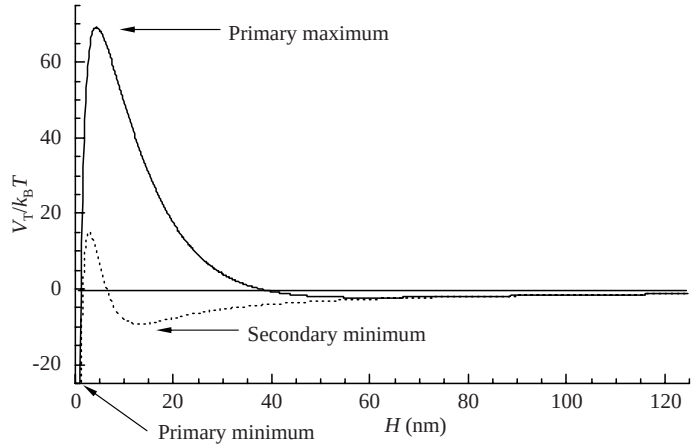


Figure 5.1 The pair-potential calculated for the rutile form of titanium dioxide at concentrations of 1 mM (—) and 10 mM (- - - -) of sodium chloride

was due to Derjaguin and Landau [1] and independently Verwey and Overbeek [2], hence the term *DLVO theory* – and we may summarize the total interaction energy as follows:

$$V_T = V_A + V_R \tag{5.1}$$

So, for example, we may use Equation 3.25b for V_A and Equation 4.33a for V_R to describe the interaction between two titanium dioxide particles as a function of separation at different electrolyte concentrations. The results of such a calculation are illustrated in Figure 5.1 for 100 nm radius rutile particles with a ζ-potential of -45 mV. The calculation is for an isolated pair of particles and so we are considering very dilute dispersions. As V_A and V_R have different dependences on particle separation, the combined curve has a complicated structure. The general features are as follows:

1. A *primary minimum*, V_{pmin}, where the dispersion interaction is much larger than the electrostatic term. The details here are uncertain, however, as details of the molecular structures of both the surface and adsorbed ions are of a similar scale to the separation.
2. The *primary maximum*, V_{max}, occurs at a distance slightly further away where the electrostatic interaction is dominant.

3. After this, the interaction energy decreases to the *secondary minimum*, V_{smin}, as the attractive interaction again becomes slightly larger than the repulsion.

The primary minimum indicates that the aggregated state is the lowest energy condition and this is where we would expect the particles to reside. The primary maximum opposes the close approach of the particles and is an activation energy that must be exceeded for aggregation to occur. The motion of the particles is governed by the thermal energy and, of course, we may describe the energy distribution by the Boltzmann equation. This is important as we can see that we are predicting a *kinetic* stability as the rate of aggregation should be proportional to $\exp(-V_{\text{T}}/k_{\text{B}}T)$. When $V_{\text{T}} \gg k_{\text{B}}T$, the particles will be in a colloidally stable state. In our example, $V_{\text{smin}} \approx k_{\text{B}}T$ and therefore we may expect many of the particles to be in close proximity for much of the time. In other words, as the net attractive energy is only slightly larger than the average thermal energy, any aggregation in the secondary minimum is reversible. We will return to discuss the kinetics of the aggregation of particles later in this chapter.

In Chapters 3 and 4, we discussed the various contributions in some detail. If there is a net repulsive interaction between two particles such that $V_{\text{T}} \gg k_{\text{B}}T$, the particles will not aggregate. If there is a strong attraction so that $-V_{\text{T}} \gg k_{\text{B}}T$, there will be strong aggregation or *coagulation* of the particles with the particle surfaces coming into close proximity. In the case of fluid particles, this is the precursor to *coalescence*. There are other situations where the particles are attracted at long distances but do not come into close contact. Here we will use the term *flocculation* to describe the aggregation. Many authors use the terms interchangeably but it is useful to separate them. For example, particles can be flocculated in a shallow attractive well but still have a large energy barrier to overcome before they can come into close contact. Some examples here are aggregation in the secondary minimum predicted for some electrostatically stabilized dispersions or aggregation due to the depletion forces produced when a non-adsorbing polymer is added to a dispersion. In both of these situations, the well is $-1 > V_{\text{min}}/k_{\text{B}}T > -100$ and the aggregates may be redispersed with moderate shearing forces, unlike the coagulated state. The term *bridging flocculation* is used to describe the aggregation of particles by the adsorption of polymer on two particles simultaneously, thereby 'tying them' together.

5.3 MECHANISMS OF AGGREGATION

5.3.1 Electrostatically Stabilized Dispersions

One of the great early successes of the DLVO formulation of the pair potential for electrostatically stabilized particles was the ability to predict their coagulation on the addition of electrolytes [3]. The addition of an indifferent electrolyte reduces the range of the repulsive component and the maximum in the potential energy curve is reduced. An indifferent electrolyte is one which does not contain a potential-determining ion, such as Ag^+ with silver halide particles or H_3O^+ with oxide particles. At some point, the value of the maximum approaches zero and there is no barrier to particles coming into close contact due to the dispersion forces. Of course, the concentration and type of the counter-ions do have an effect other than compression of the diffuse layer. The adsorption in the Stern layer can also change, which we can readily observe experimentally via changes in the ζ-potential. The ζ-potential is frequently assumed to equate to the Stern potential, ψ_δ, and is therefore used in calculations of the pair potential. The variation in the tendency to adsorb in a series of ions of the same valency is known as the *Hofmeister series* and the specificity of these ions also shows up in other colloidal aspects such as micellization [3]. The series for monovalent cations and anions is as follows:

$$Cs^+ > Rb^+ > K^+ > Na^+ > Li^+$$
$$CNS^- > I^- > Br^- > Cl^- > F^- > NO_3^- > ClO_4^-$$

With the cations, the small size of the lithium ion results in a higher charge density, which, in turn, means that it is more strongly hydrated. The larger anions are more easily polarized and this increases the adsorption. We see a stronger effect of the ion type with positive particles and anionic counter-ions than with negative particles and cationic counter-ions.

The valency of the counter-ion is extremely important. The coagulating power of an ion increases dramatically with its valency, as encompassed in the Shultz–Hardy rule [3–5], which states that the coagulating power varies as z^6, where z is the counter-ion valency. Experimentally, this is not always the case and a lower exponent can be found. Figure 5.2 illustrates the conditions for the loss of stability. The data shown are for polystyrene latex particles with a diameter of 1 μm and were calculated

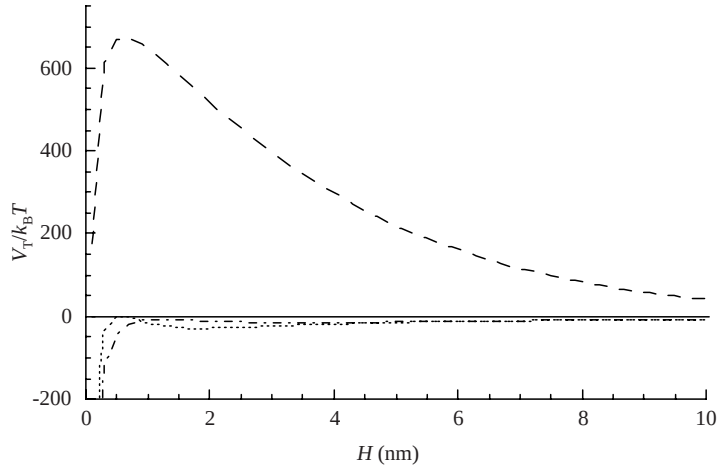


Figure 5.2 The pair potential calculated for polystyrene latex particles with a radius of 500 nm as a function of the ζ-potential and sodium chloride concentration: (·····) −20 mV, 10 mM; (····) −30 mV, 400 mM; - - - -) −50 mV, 10 mM

for different combinations of ζ-potential and electrolyte concentration resulting in the conditions for instability being met. The unstable particles are those where the conditions pertain that $V_{max} = 0$, so at this point

$$\frac{V_R}{V_A} = -1 \tag{5.2a}$$

and also

$$\frac{\partial V_T}{\partial H} = 0 = \frac{\partial V_R}{\partial H} = \frac{\partial V_A}{\partial H}, \quad \text{with } \kappa \to \kappa_{ccc} \tag{5.2b}$$

If we take the equations for two similar flat plates, that is, Equations 3.21 and 4.30 and differentiate these, we obtain

$$0 = -\kappa_{ccc} V_R - \frac{2}{H} V_A \tag{5.3a}$$

or

$$\kappa_{ccc} H = 2 \tag{5.3b}$$

and so putting this value back into the expression for V_{max}, we have, from Equations 3.21 and 4.30,

$$\frac{64nk_BT}{\kappa_{ccc}}\tanh\left(\frac{ze\psi_\delta}{k_BT}\right)\exp(-2) = \frac{A_{121}\kappa_{ccc}^2}{48\pi} \tag{5.4}$$

The expression for κ was given earlier by Equation 4.10a. Substitution of this into Equation 5.4 gives the following result for the critical coagulation concentration (ccc):

$$\text{ccc} \propto \frac{1}{z^6} \tag{5.5}$$

However, what often occurs as the electrolyte concentration is increased is that the potential, ψ_δ, falls to a low value and then we find experimentally that we have the weaker dependence

$$\text{ccc} \propto \frac{1}{z^2} \tag{5.6}$$

What we observe experimentally, is that as we approach the critical coagulation concentration, we start to see aggregates in suspension. Hence we can determine the value simply with a series of test-tubes containing a range of electrolyte concentrations, to each of which we add a known mount of dispersion. After a few minutes, aggregates will be seen at concentrations above the ccc. This experiment can be made slightly more precise if a spectrophotometric determination of turbidity is carried out. The procedure here is as follows:

1. Mix a known volume of the dispersion with the a known volume of different electrolyte concentrations.
2. Allow these to stand for a fixed time; 30 min would be suitable.
3. Centrifuge the dispersions at a low *g*-value so that only the aggregates are removed.
4. Measure the turbidity of the supernatant.

A curve similar to the schematic plot in Figure 5.3 will be obtained.

It is important to be aware that the coagulating power of multivalent cations can be more complex than we might expect initially. Trivalent cations such as Al^{3+} and La^{3+} only exist in this form at low pH values. As Matijévic pointed out [6], as the pH approaches values of 7 and

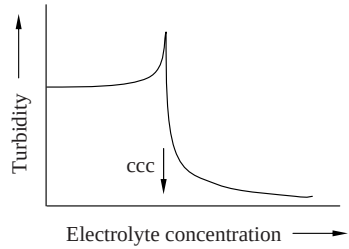


Figure 5.3 The turbidity of the supernatant from a dispersion as a function of concentration of added electrolyte

above, these ions exist as large, complex, highly charged species. Hence we should begin to think of these counter-ions as small counter-charged nanoparticles which can cause hetero-coagulation, and this is why they are such effective coagulants.

5.4 HETERO-COAGULATION AND HETERO-FLOCCULATION

This is an important aspect of the aggregation of colloidal particles in many practical situations. We will discuss this in broader terms here than is often done. Thus, under this heading we include aspects of the interaction of particles which can carry a charge opposite in sign from each other and the effects of polyelectrolytes with a charge different in sign from the particles. It was pointed out in the previous paragraph that ions such as Al^{3+} are only in that form at values of pH $< \sim 4$. At higher pH values, large, hydrolysed complex ions are formed with high charge densities. These can be very strongly adsorbed on the surface of negatively charged particles. Clearly, if present in high concentration, the surface charge can be reversed in sign in exactly the same manner as occurs with high concentrations of the simple multivalent ions and the system would be potentially 'restabilized'. Hence, when used as coagulants, the concentrations are kept low. Neutralization of the surface charge is one mechanism for aggregation, just as it is with the simple multivalent ion species. The diffuse layer compression is, of course, occurring simultaneously as discussed in the previous section, but the strong binding has a very large effect in the reduction of the Stern potential, ψ_δ. However, complete neutralization of the charge is not a prerequisite for aggregation to occur. There are two further possibilities. First, the

small macro-ions can act as bridges at low concentrations. Second, the adsorption can produce 'patches' on the surface of different charge so that on close approach local attraction can occur. This is just 'charge patch' flocculation, as suggested by Gregory [7] as a mechanism for the aggregation of negatively charged particles by cationic polymers. For this to work, the background electrolyte concentration must be sufficiently high that an approaching particle surface can 'recognize' that local difference with the resulting attraction. This means that the patches have to be of a similar magnitude to the decay distance of the diffuse layer potential, κ^{-1}.

5.4.1 Polymeric Flocculants

Cationic polyelectrolytes, such as a positively charged polyacrylamide-based copolymer, can form bridges, causing hetero-flocculation. The radius of gyration of a polyelectrolyte is much larger than if the polymer were uncharged, as the charged groups repel each other, forcing the coil to take up an expanded configuration. The dimension of the coil is a function of electrolyte concentration as the screening of the charges varies with ionic strength and in many cases also with changes in pH. Bridging by polymers (or by other particles) is most effective if the particles are small and concentrated enough that their average separation is of a similar magnitude to the maximum dimension of the coagulant species. This will not be the case with larger colloidal particles and then the charge patch mechanism becomes the likely route. The critical factor here is that the coagulating polymer species will have sufficient time to be adsorbed 'flat' on the surface before an encounter with another particle can occur. A schematic of this situation is shown in Figure 5.4.

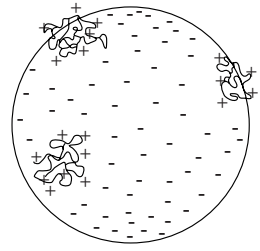


Figure 5.4 Schematic of 'charge patch' formation by the adsorption of a cationic polyelectrolyte on an anionic particle

Both bridging and charge patch flocculation work well when the surface coverage by adsorbed polymer is less than 50%. As greater coverage is attained, there is an increasing tendency for a stable system to be engineered. Polymeric flocculants are normally used at concentrations of the order of parts per million (ppm).

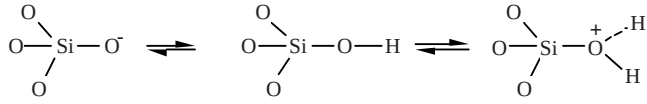
Anionic polyelectrolytes, such as poly(acrylic acid) and partially hydrolysed polyacrylamide, can also be utilized as polymeric flocculants. However, many colloidal systems of practical importance are also anionic. In this case, divalent ions such as Ca^{2+} can be added and act as ion bridges to bind the polymer to the surfaces and promote bridging flocculation. Calcium ions bind very strongly to carboxyl groups and this can help to attach the anionic polymer to material coated with proteinaceous material. Ion bridges between two separately coated particles can also be formed, hence producing a flocculated system.

Depletion flocculation should also be considered at this point of the discussion. This mechanism, however, is not useful in terms of separation processes as the attractive well is too shallow. The flocculated state is readily redispersed and so the particles could not be successfully separated, for example by pressure filtration. In fact, this reversibility is a feature that can be used to advantage in some situations, as we shall see when we discuss mechanical stability.

5.4.2 Particulate Hetero-coagulation

The charge patch flocculation mode described above is just one example of hetero-coagulation where surfaces with charges of different sign attract one another. The dispersion force contribution is also present so that the total attractive force is large. An example of hetero-coagulation, which is of immense practical importance, is to be found in the colloidal behaviour of clays. Kaolinite is an aluminosilicate which has a 1:1 layer lattice structure. The clay crystal is made up of alternating layers of silica tetrahedra and alumina octahedra connected via shared oxygen atoms. The basal planes are silica and the crystals are irregular hexagons which have an axial ratio of approximately 10:1. As with all oxides, the silica surfaces are amphoteric so that the surface hydroxyl groups can either ionize or bind a proton, as shown in Scheme 5.1.

The point of zero charge (pzc) occurs at a low pH for silica and hence the surfaces are negatively charged over most of the pH range. However, there is an additional factor to be considered in that during the formation of the clay, some of the silicon atoms were replaced by aluminium



Scheme 5.1

and some of the aluminium in the next layer down by magnesium. The crystal structure, although distorted, is basically the same as the cations are smaller than the oxygen anions. This is known as *isomorphous substitution* and the result is that there is an excess negative charge due to the replacement of ions by ones of a lower valency. This charge is balanced by positive ions in the Stern and diffuse layers. The surfaces are therefore strongly negatively charged. The isomorphous substitution is only present in the outermost crystal layers and is probably one of the factors that limits the growth of the crystals. The crystal edges have exposed alumina layers and the pzc of this edge surface is close to that of alumina. The result is that the edges are positive at pH < 7 and become increasingly negative at pH > 7. Hence, at pH values < 7 the clay particles aggregate in an edge–face structure similar to a 'house of cards'. Figure 5.5 illustrates this mode of hetero-coagulation.

Another situation in which hetero-coagulation can occur is in mixed particle systems, especially when at least one of the components is an oxide or a particle coated with a layer which has an isoelectric point at a pH at a different value from the other component. In this situation, the

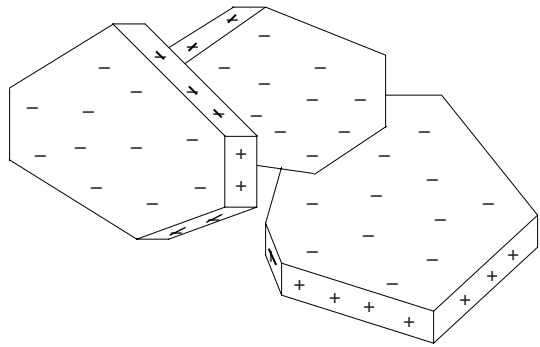


Figure 5.5 Schematic representation of 'edge–face' coagulation found with kaolinite at pH < 7 when the edges carry a positive charge and the faces are negative. The particles form a very open structure, referred to as 'house of cards' aggregation structure

van der Waals attraction is aided by the attraction of oppositely charged surfaces over some of the pH range. Readjustment of the pH of the dispersion after aggregation has taken place will not result in restabilization of the aggregated particles due to the large amount of energy required to overcome the deep primary minimum attractive energy. The exception would be the particles coated with a protein layer. In this case, the expansion of the hydrophilic layer due to the internal charge repulsion may be sufficient to aid the 'repeptization'.

The mixing of particle systems of different surface charges is important. To illustrate this, we can take the example of the mixing of a system of large negatively charged particles with some small positively charged particles. As the particle number varies as the cube of the size, the small particles are likely to be present in greater numbers than the large particles in the final mixture unless very small amounts are used. If the small particles are added to the large particles, the probable result will be a bridging aggregation with complete separation of the large particles from suspension. If the large particles are added to the small with good mixing, a system of large particles coated with small particles can be achieved. The order and mode of mixing in a hetero-coagulative system are just as important as they are with the mixing of adsorbing polymers into a dispersion.

5.4.3 Aggregate Structure

We have already seen how the hetero-coagulation of a clay leads to an open structure. This type of structure can space-fill, that is, occupy the whole volume available to the dispersion at low volume fractions. This has important implications for applications involving aggregated systems. For example, the rheological properties change dramatically so that handling can become difficult. Filtration may start as an easy separation process once a system is aggregated, but the final 'dewatering' is limited so subsequent drying can be a slow and expensive process. A relatively weak aggregation would be a preferable situation here so that collapse of the filter cake to a high solids density could be achieved. In the case of ceramics, the rheology of the open structures is excellent for shape formation with minimal elastic recovery after yield has occurred at moderate to high stress. However, the open structure results in considerable shrinkage on drying and firing. These two examples illustrate clearly that the details of aggregate formation have great practical importance.

If the potential barrier to aggregation is reduced to zero, we have the situation where particle collisions are 'sticky'. This mode of aggregation is termed *diffusion-limited aggregation*. The simplest visualization of this is to consider the particles with a hard-sphere potential which has been modified to give a very narrow but deep square attractive well close to the surface. Particles then collide and stick as they diffuse. The aggregates grow in size with an open-dendritic or fractal type of structure. Computer models generate this type of open branched structure and some careful experiments have confirmed these models. A fuller description may be found in Russel *et al.* [5]. As these aggregates grow by accretion of 'sticky' particles, they grow into each other and span the available space. This point is known as the *percolation threshold*. At higher concentrations, denser structures result and these are more difficult to define by a single parameter such as the 'fractal dimension'. In addition, although these structures are academically interesting, they are invariably modified in general applications. Systems are mixed during the addition of a coagulant and further *shear processed* during subsequent handling. The shear forces on these very large and fragile structures compact them to relatively high densities. For example, systems of monodisperse spherical particles can be compacted to random packing densities (i.e. $\varphi \approx 0.64$) in monodisperse spherical aggregates by shearing the coagulating system [8]. The point to remember here is that the strongly aggregated structures are *metastable* structures. The lowest energy configuration would be a very dense unit with the maximum number and/or area of contacts. However, the fractal structure of a dilute strongly aggregated system would be very long lived in the absence of external forces as $V_{\text{pmin}} \gg k_{\text{B}}T$ and densification purely by diffusive motion would be imperceptibly slow.

The structures that we observe then are a combination of processing and the strength of the attractive interaction. The attractive well can be controlled by the addition of material to the surface prior to coagulation. Non-ionic surfactants and polymers are excellent candidates for this, providing a steric barrier and limiting the aggregation to a weak flocculation. Excellent experimental studies of aggregate structures formed with weakly flocculated systems have been reported by Russel *et al.* [5] and Sperry [9]. A series of different sized polymer latexes were mixed with different molecular weight water-soluble polymers to give depletion-flocculated systems. Phase separation occurred at a value of $V_{\text{min}} \approx -2k_{\text{B}}T$. Stronger attractions resulted in rigid branched aggregates, whereas weaker attractions gave fluid-like aggregates more akin

to a separated fluid phase. Hence, at some combination of attractive potential and thermal energy, equilibrium structures were attainable, but if $V_{\min} \gg -k_B T$, non-equilibrium structures are formed. For phase separation to occur, the time for the break-up of a pair of particles must be significantly longer than the time for more particles to collide with the doublet. However, if it is much longer than the time over which we are prepared to study the system, we will not wait long enough to see an equilibrium structure. The diffusion coefficient, D, of a particle is given by the Stokes–Einstein equation:

$$D = \frac{k_B T}{6\pi \eta_0 a} \tag{5.7}$$

where a is the particle radius and η_0 is the viscosity of the medium. The diffusion constant has units of a flux, that is, $m^2\ s^{-1}$, and we can calculate the characteristic diffusional time as the time it takes for a particle to diffuse through a mean distance of a. Thus, the Einstein–Smoluchowski equation gives the time as

$$t_D = \frac{a^2}{D} = \frac{6\pi \eta_0 a^3}{k_B T} \tag{5.8}$$

However, if there is an attractive potential, the break-up rate will be slowed as only a fraction of the particles will have sufficient thermal energy at any time to escape. The diffusion time must therefore be modified by the Boltzmann factor to reflect this energy distribution:

$$t_D = \frac{6\pi \eta_0 a^3}{k_B T} \exp\left(-\frac{V_{\min}}{k_B T}\right) \tag{5.9}$$

Figure 5.6 indicates how t_D increases with the attractive potential for particles of different size.

5.4.4 Slow Structural Changes

The usual laboratory timescale (our timescale) is 1 ms to 1 ks, but times which are very much longer than this cannot be neglected as products can be in storage for several months. There are two points to note here. The first is that if new, 'sticky' particles arrive at a doublet before the

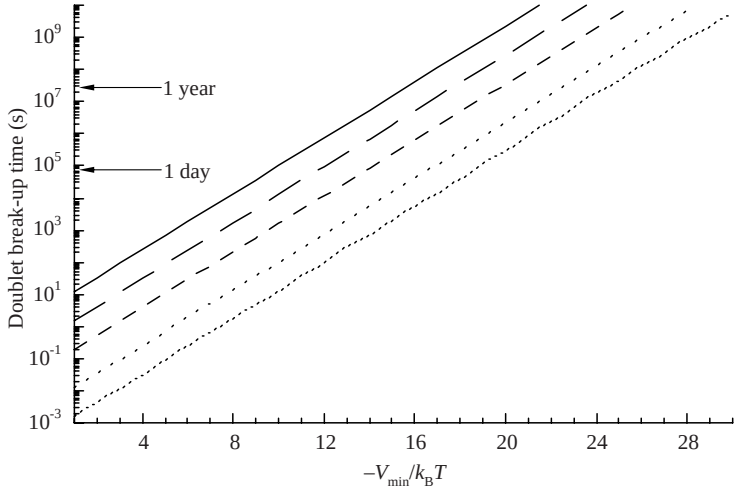


Figure 5.6 The time taken for doublets to separate as a function of particle radius and the depth of the attractive potential minimum: (····) $a = 50$ nm; (- - - -) $a = 100$ nm; (– – –) $a = 250$ nm; (— —) $a = 500$ nm; (—) $a = 1000$ nm

doublet can break up, we can expect to observe phase separation. We see how that break-up time increases rapidly with both particle size and attractive potential. The second point is more subtle as we can find changes occurring slowly with time. These are also related to the diffusive process. Consider a particular system with monodisperse spherical particles of 500 nm radius and weakly flocculated with an attractive potential of $V_{\min} \approx -10k_BT$. At a volume fraction of ~0.3, we have a flocculated phase and experiments indicate that this is close to space filling (see, for example, Goodwin and Hughes [10]). However, this is not the densest packing, which would be closer to face-centred-cubic. From Figure 5.6, we see that the break-up time for a pair is ~2 h. What occurs is that particles in the lower density zones diffuse to higher density zones, that is, they increase the number of nearest neighbours (coordination number) to move into a lower energy state. Migration in the opposite direction is clearly less favourable as each extra contact increases the total $V_{\min}$ by $10k_BT$. The final result is a slow change in local density followed by a collapse of the structure as it suddenly consolidates. This can take place over times from hours to weeks depending on the volume fraction (as this controls the local mean particle density or coordination number). We see an apparently stable system, which, after a long incubation time, sediments rapidly with tracks being formed as the medium is forced up to

the top of the bed. Consolidation of suspension concentrates on storage is a common problem in systems of large dense particles.

5.5 THE RATE OF COAGULATION

5.5.1 Diffusion-limited Aggregation

The initial rate of aggregation was first analysed quantitatively by von Smoluchowski, who modelled the system as that of diffusing spherical particles which stick on collision contact but the pair potential is zero up to that contact. Detailed descriptions of the model for the diffusion behaviour can be found in Hunter [4] and Russel *et al.* [5]. A simplified approach is sufficient for our purposes here.

Figure 5.7 illustrates the geometry of the model. If r is the centre-to-centre distance between the reference sphere and an approaching particle, contact occurs when $r = 2a$ for a monodisperse system. The central particle is treated as a 'sink' so that no account is taken of any growth of the reference unit. This is satisfactory simplification as the description is of the *initial* coagulation rate. As both particles are diffusing, the net diffusion coefficient is equal to $2D\,\mathrm{m^2\,s^{-1}}$ [2]. The net velocity of an incoming particle is therefore $2D/a\,\mathrm{m\,s^{-1}}$. The surface area of our

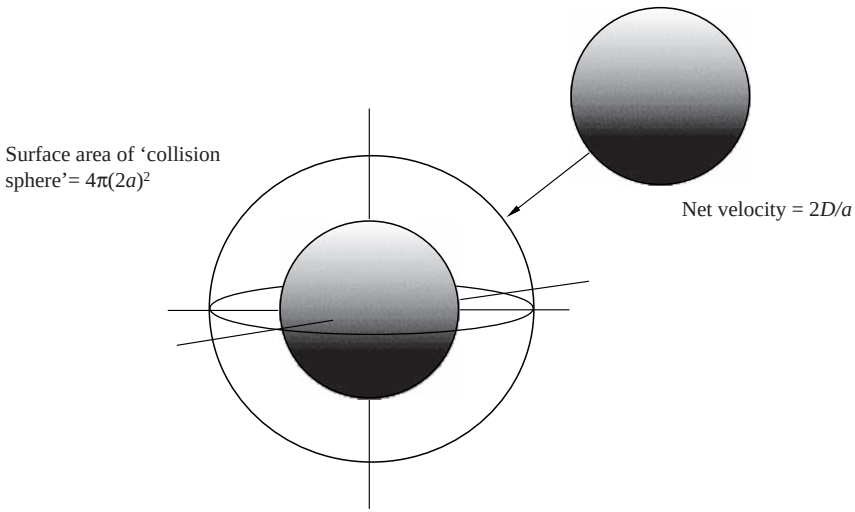


Figure 5.7 Illustration of a Brownian collision of two identical particles of radius a with a diffusion constant of D

'collision sphere' is $4\pi(2a)^2$. The flux through the collision sphere, if there are initially N_p particles per unit volume, is

$$J_B = N_p \frac{2D}{a} 4\pi(2a)^2 \tag{5.10}$$

As this process is occurring with each particle, the collision frequency due to Brownian diffusion is

$$c_B = \frac{N_p J_B}{2} \tag{5.11}$$

where the factor of 2 is introduced to prevent double counting. Substituting for D from Equation 5.7 into Equation 5.11 gives

$$c_B = N_p^2 \frac{8k_B T}{3\eta_0} \tag{5.12a}$$

As each collision results in coagulation, the initial coagulation rate that we should observe in a quiescent system is given by

$$-\frac{dN_p}{dt} = N_p^2 \frac{8k_B T}{3\eta_0} \tag{5.12b}$$

where the half-life of the rapid aggregation process is determined for this second-order rate equations as

$$t_{\frac{1}{2}} = \frac{3\eta_0}{4k_B T N_p} \tag{5.13}$$

However, the formation and aggregation of doublets and larger multiplets should be included [3] and this results in Equation 5.13 providing a poor estimate this far into the process.

5.5.2 The Effect of the Continuous Phase

The fast aggregation rate is inversely proportional to the viscosity of the suspending medium. The rate appears to be directly proportional to the temperature but the rate increases more rapidly than this because, in addition, the viscosity of a liquid decreases exponentially with increasing temperature. Clearly, a high viscosity reduces the rate of aggregation, but

there is an additional effect of the particle being in a liquid medium. The range of the hydrodynamic disturbance around a moving particle scales with the particle radius. This means that as two particles approach, the hydrodynamic interactions start to become significant as $r < \sim 3a$. As $r \rightarrow 2a$, the reduction in the rate of approach becomes large with the liquid between the two particles having to drain out of the intervening space before the particles can come into contact. The hydrodynamic interaction acts in a similar fashion to a repulsion in that it slows the approach and hence extends the time taken for the formation of doublets. This problem has been analysed by Derjaguin [11] and Spielman [12], with the latter calculating the reduction in the initial aggregation rate as a function of the viscosity of the medium.

5.5.3 Potential-limited Aggregation

The von Smoluchowski rate given in Equation 5.12b will overestimate the rate if there is a repulsion between the particles. Fuchs tackled this problem and the analysis is discussed in detail in Hunter [4] and Russel *et al.* [5]. The pair potential slows the approach of two particles. At any distance, the fraction of particles with thermal energy in excess of the potential at that distance is given by the Boltzmann factor: $\exp(-V_T/k_B T)$. The flux through successive spherical shells as the particles approach is slowed from the simple collision case and only a fraction of the particles that encounter one another can approach close enough to stick. The fraction of the encounters that stick is $1/W$, where W is known as the *stability ratio*. We can express W as the ratio of the two fluxes as follows [4]:

$$W = 2a \int_{2a}^{\infty} \exp(V_T/k_B T) \frac{dr}{r^2} \tag{5.14}$$

Reerink and Overbeek [13] pointed out that the maximum in the pair potential was the dominant factor in restricting the approach of particles, as in the slow coagulation regime the electrolyte concentration is such that the diffuse layer is very compressed. They showed that a useful approximation to the integral Equation 5.14 was

$$W \approx \frac{1}{2\kappa a} \exp(V_{max}/k_B T) \tag{5.15}$$

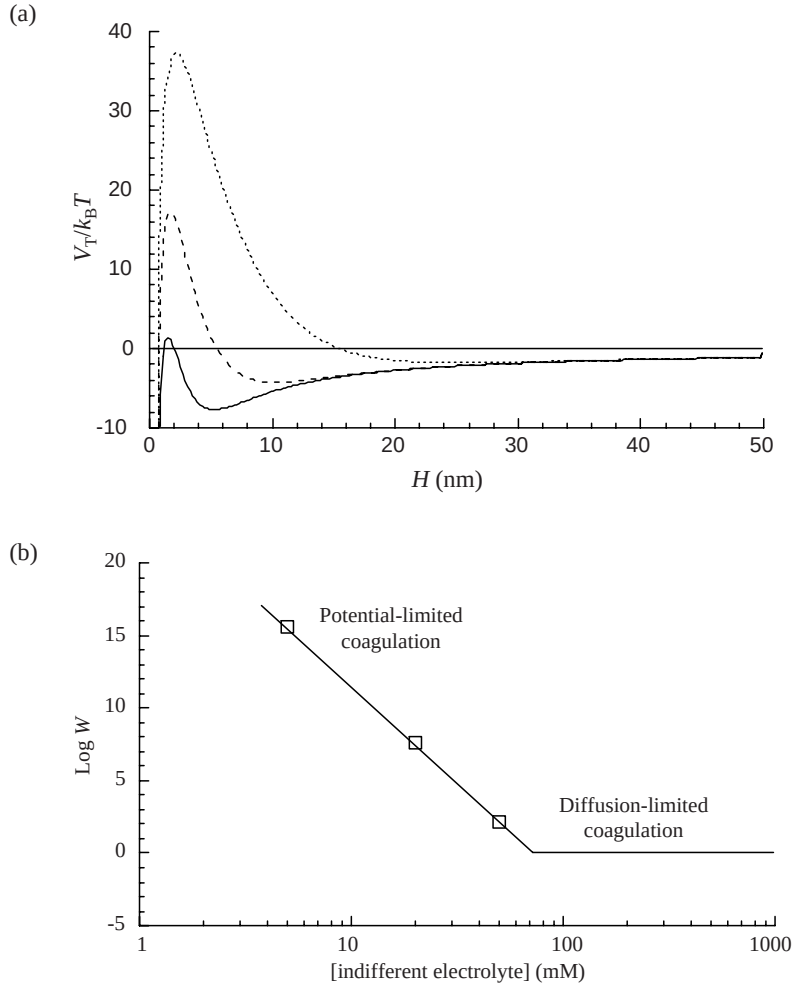


Figure 5.8 (a) The pair potential calculated for silver bromide with a particle radius of 50 nm and a ζ-potential of -50 mV: (····) $\kappa a = 11.6$, $V_{max} = 37.3k_BT$; (- - - -) $\kappa a = 23.2$, $V_{max} = 17.4k_BT$; (—) $\kappa a = 36.7$, $V_{max} = 1.5k_BT$. (b) The stability ratio calculated for silver bromide as a function of the concentration of a 1:1 indifferent electrolyte (using Equation 5.15)

Figure 5.8a shows the pair potentials calculated for silver bromide particles with a ζ-potential of -50 mV and a particle radius of 50 nm. Equation 5.15 was employed to estimate the stability ratio for the same particles with the pair potentials in Figure 5.8a and the result is shown in Figure 5.8b. However, it should be stressed that the height of the primary maximum relative to the secondary minimum was used in the

calculation, as this would be the energy required for the particles to move from a secondary minimum flocculation to a primary minimum coagulated state. It is clear from Figure 5.8b that the stability ratio changes very rapidly with electrolyte concentration. The result is that we will usually just notice a change from a stable to a rapidly aggregated dispersion under normal laboratory conditions. Therefore, although it might appear that the electrolyte concentration of sample is low enough for the stability to be adequate, the slow coagulation process is the type of problem that occurs during storage.

5.6 AGGREGATION IN FLOWING DISPERSIONS

There are two effect s of a flow field on the aggregation behaviour of dispersions that we need to consider. The first is that the collision frequency is increased and this effect was first analysed by von Smoluchowski (see, for example, Hunter [4] and Russel *et al.* [5]). This just has an effect on the rate of aggregation, but the hydrodynamic contribution to the interaction should also be considered [5]. This latter effect can cause a stable dispersion to aggregate under some conditions. Aggregation under quiescent conditions is termed *orthokinetic* coagulation, where the collisions are due solely to Brownian motion. In a strong flow field, where the collisions are controlled by the shear field, the aggregation is termed *perikinetic* coagulation.

5.6.1 The Effect of Flow on Collision Frequency

This discussion will be restricted to laminar flow. In a high-speed mixer, the flow is turbulent with chaotic vortices so that the particles are subjected to a wide and unpredictable range of hydrodynamic forces and a much more complex treatment would be required.

In a mixed or flowing system, there is a variation of fluid velocity with position in the fluid. This known as a *shear field*. With simple shear, the fluid is moving at a uniform velocity in, say the x–y plane, with a change in velocity in the z-direction. This is sketched in Figure 5.9.

In the absence of Brownian motion, a particle will move at the velocity of the liquid at the plane coincident with the centre of the particle, v_p, which is a distance z_p from the reference x–y plane:

$$v_p = z_p \frac{dv}{dz} \tag{5.16}$$

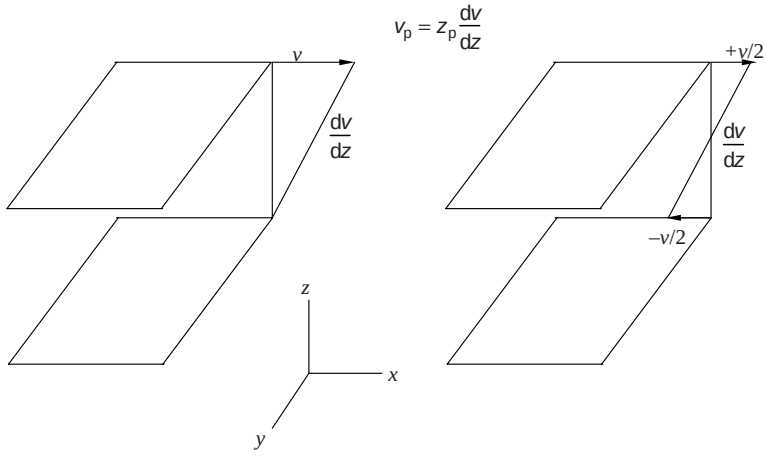


Figure 5.9 Illustration of a simple shear field, where both figures are mathematically equivalent with the same velocity gradient. The right-hand representation is used in this case as it emphasizes the symmetry of the collision process

The geometry is illustrated in Figure 5.10. Any particle whose centre passes through the 'collision disc' with a radius of $2a$ collides with the reference particle located on the x-axis. The problem is now straightforward as it reduces to calculating the fluid flux through the 'collision disc', and as we know that the particle concentration is N_p, we have the number of collisions with a reference particle. The fluid flux through the area element dz is

$$f_z = \frac{dv}{dz} z dz \times 2y_z \tag{5.17a}$$

$$f_z = \frac{dv}{dz} [2a \cos(\theta/2)] [-a \sin(\theta/2)] [2a \sin(\theta/2)] \tag{5.17b}$$

Therefore, the total flux of particles through the disc is

$$f_p = N_p \times 2\frac{dv}{dz} (2a^3) \int_{\pi}^{0} [-\sin^2(\theta/2) \cos(\theta/2)] d\theta \tag{5.18}$$

As the collisions occur with all the particles, the total collision frequency due to flow, c_f, is the product of f_p and N_p, although we have to

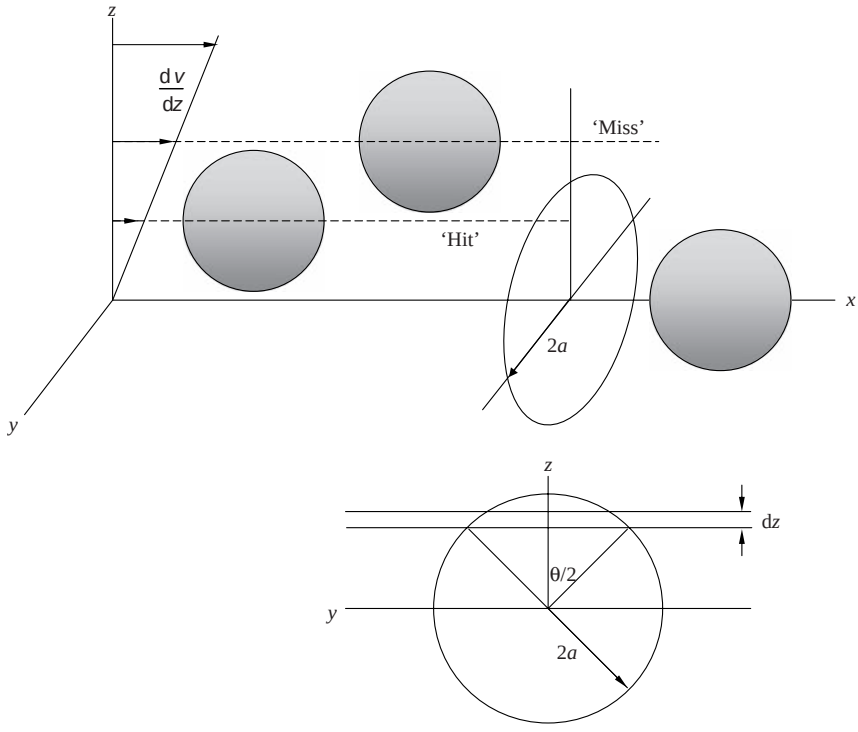


Figure 5.10 Illustration of collision trajectories for particles under simple shear flow. Note that there is symmetry on each side of the x–y plane

divide by two to prevent 'double counting':

$$c_f = \frac{16}{3} N_p^2 a^3 \left(\frac{dv}{dz}\right) \tag{5.19}$$

However, as the particles approach in the shear field, the hydrodynamic interactions cause the colliding pair to rotate and with the combination of the slowing of the approach due to liquid drainage [5, 12], that is, lubrication stresses and Brownian motion, not all collisions will lead to aggregation. Equation 4.59 must be reduced by a factor α to account for this:

$$c_f = \alpha \frac{16}{3} N_p^2 a^3 \left(\frac{dv}{dz}\right) \tag{5.20}$$

This 'collision efficiency factor' is of order 1 and a typical value would be $\alpha \approx 0.8$. It is interesting to compare the Brownian collision frequency

(c_B) from Equation 5.12a with that due to flow in Equation 5.20:

$$\frac{c_f}{c_B} = \frac{2\alpha\eta_0 a^3}{k_B T}\frac{dv}{dz} \tag{5.21}$$

If the particles are dispersed in water at a temperature of 25 °C, the ratio in Equation 5.21 becomes

$$\frac{c_f}{c_B} \approx 4 \times 10^{17} a^3 \frac{dv}{dz} \tag{5.22}$$

When we stir liquid in a beaker with a rod, the velocity gradient is in the range 1–10 s^{-1}, with a mechanical stirrer 100 s^{-1} would be a reasonable value, and at the tip of a turbine in a large reactor one or two orders of magnitude higher could be possible. Hence the particle radius must be $a < 1$ μm if even slow mixing can be disregarded.

This description of the collision process does not include the details of the collision trajectories, which are governed by the hydrodynamic interactions. A more detailed description should also includes the effects of interparticle repulsion and attraction and the contact time [3].

5.6.2 The Effect of Flow on the Interaction Force

The above description of the collision frequency due to von Smoluchowski is useful if we are looking at particle coagulation rates in the absence of a potential barrier. When we consider the case of potential-limited aggregation, we have to consider the contribution due to the hydrodynamic forces acting on the colliding pair. An excellent analysis of this situation has been carried out by Schowalter and is given in Russel *et al.* [5]. To illustrate the underlying concept, we will consider a simple model.

Figure 5.11 shows the forces acting on a collision doublet in simple shear. The figure shows the trajectory with the points at which the maximum compression and tension occur, that is, at $\theta = 45°$. The particles have the same radius a and the reference particle is at $z = 0$. The velocity of the streamline coincident with centre of the colliding particle at the orientation giving the force maximum is

$$v = \frac{dv}{dz} 2a \sin(45) \tag{5.23}$$

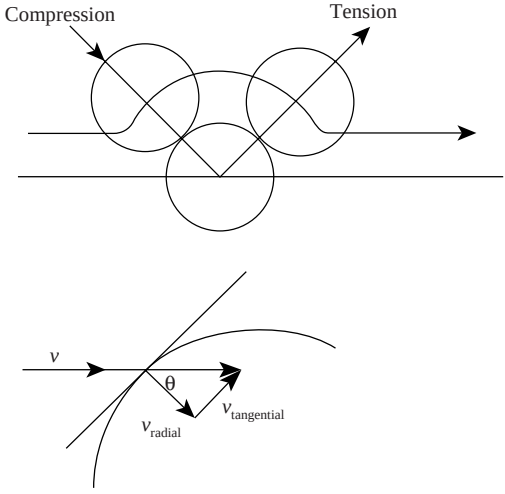


Figure 5.11 Illustration of the geometry of a colliding pair of particles with the maximum compression and tension at 45° to the shear plane

Now, the radial component of the Stokes drag force on the particle is

$$F_h = 6\pi\eta_0 a v_{radial} = 6\pi\eta_0 a \cos(45) \tag{5.24}$$

and so we may write

$$F_h = \pm\frac{dv}{dz}6\pi\eta_0 \times 2a^2 \sin(45)\cos(45) \tag{5.25a}$$

$$F_h = \pm\frac{dv}{dz}6\pi\eta_0 a^2 \tag{5.25b}$$

where the ± sign indicates the compression (+) or tension (−). It should be noted at this point that the trajectory sketched in Figure 5.11 is a simplification. The trajectory would be altered by the colloidal forces on close approach. The details would depend on whether there was a net attraction or repulsion, but the maxima and minima would still be at the positions shown so Equation 5.25b is satisfactory for our purposes.

It is possible to use Equation 5.25b to indicate where the stability or instability boundaries are for particular dispersions. To carry this out, all we have to do is to equate the interparticle force at the maximum and minimum points on the force–distance curve. As an example, we shall consider a polystyrene latex system at a sodium chloride concentration

of 50 mM. For the calculation, we shall choose a particle radius of 500 nm. The pair potential can be calculated using Equation 4.33a for the repulsion and Equation 3.24b for the attraction. The interparticle force is given by

$$F_{\mathrm{T}} = -\frac{\mathrm{d}V_{\mathrm{T}}}{\mathrm{d}H} \tag{5.26a}$$

$$F_{\mathrm{T}} = 2\pi\varepsilon_{\mathrm{r}}\varepsilon_0\kappa a\psi_\delta \frac{\exp(-\kappa H)}{1+\exp(-\kappa H)} - \frac{aA_{11}}{12H^2} \tag{5.26b}$$

The value of κa is 368 and so we have a situation where the interparticle forces are changing in a region very close to the particle surface; hence, hydrodynamics control the trajectories until the particles are at very close distances. The force–distance curve for this system with a ζ-potential of -40 mV is shown in Figure 5.12, where $\zeta \cong \psi_\delta$ is assumed. The stability boundaries are calculated in terms of the value of the shear rate required the change the aggregation state at different values of the ζ-potential. The results are shown in Figure 5.13.

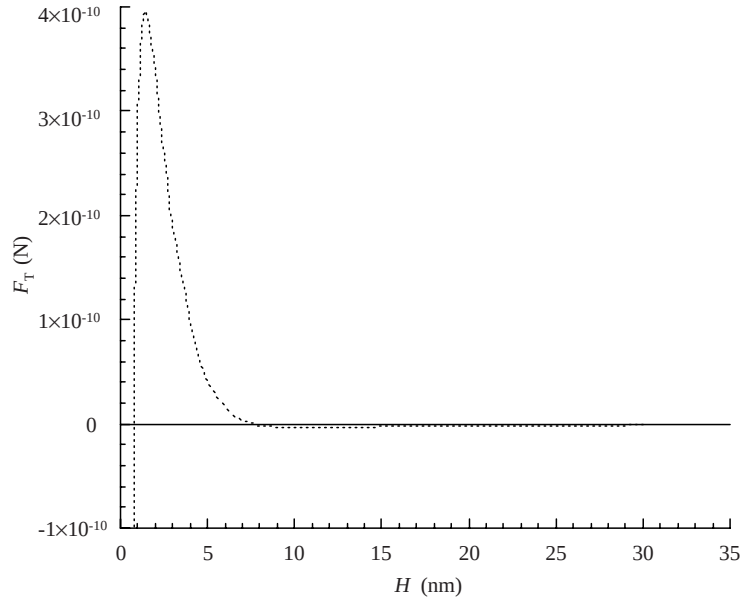


Figure 5.12 The interparticle force plotted as a function of surface separation for two polystyrene latex particles: $a = 500$ nm; ζ-potential $= -40$ mV; 1:1 electrolyte concentration $= 50$ mM; $F_{\max} = 3.96 \times 10^{-10}$ N; $F_{\mathrm{smin}} = -3.8 \times 10^{-11}$ N

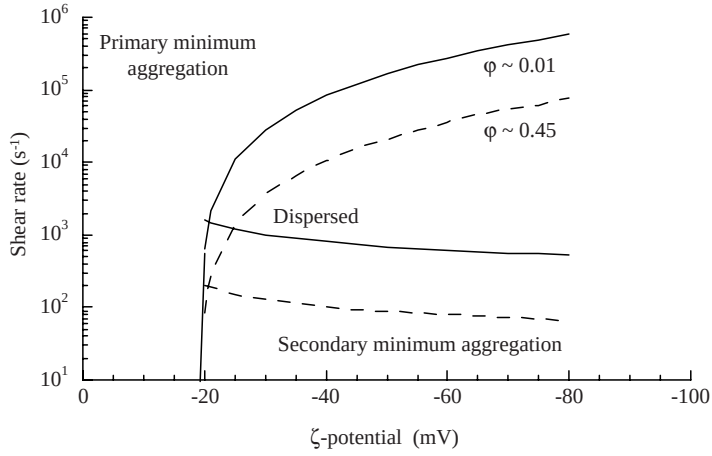


Figure 5.13 The stability map for polystyrene latexes as a function of shear rate and ζ-potential. The particle radius was taken as 500 nm, with the calculations being carried out for a 50 mM solution of a 1:1 electrolyte

There are several features of note in this stability diagram. At ζ-potentials less than -20 mV, the dispersion is coagulated at all the values of the shear rates that are plotted. With a small increase in potential above this value, the dispersion is aggregated, but it is flocculated and not coagulated at low shear rates. However, at shear rates of the order of $10^5\ s^{-1}$, the hydrodynamic forces are sufficient to cause the dispersion to form doublets which are coagulated. It is also interesting that the shear forces on this particle size, ionic strength and diffuse layer potential combination will only break down the doublets flocculated in the secondary minimum when shear rates of the order of $10^3\ s^{-1}$ are reached. Although shear rates of this magnitude are readily attainable in a viscometer, they would represent a very high stirring rate with a paddle stirrer in a beaker on the laboratory bench, but are easily achieved in pumps and large reactors with turbine mixers. Equations 5.25a and 5.25b shows that the shear forces increase as the square of the particle radius and so the stability boundaries drop rapidly with increasing particle size as the colloidal forces change more slowly with radius than the hydrodynamic forces. Hence particles with a radius of 3–4 μm are much more sensitive to shear-induced aggregation than particles of 0.3–0.4 μm in radius. It is also important to note that at electrolyte concentrations which are becoming high enough to be approaching a slow coagulation condition, the ζ-potential is unlikely to be as high as ± 50 mV. This means that we are most interested in the steep part of the coagulation boundary where

relatively low shear rates can be important. For example, localized high electrolyte concentrations can occur during the addition of a solution to a dispersion, even when the systems is stirred, so attention must be paid to both the mixing and addition rates.

In many situations, the volume fraction of the dispersion is maintained as high as possible while still maintaining adequate handling properties, that is, good heat and mass transfer properties. Hence, when the hydrodynamic forces on a pair of particles are considered, it is really the 'multibody' hydrodynamic forces that are critical. As a first-order approach, we can take the 'effective medium' or mean field approach and use the viscosity of the dispersion to calculate the local force. To illustrate this, we shall use the high-shear result for the viscosity of a dispersion of monodisperse hard spheres. The details of the derivation are considered in the discussion on the rheology of dispersions later in this volume, so here we just give the result obtained as follows:

$$\eta = \eta_0 \left(1 - \frac{\varphi}{0.605}\right)^{-1.513} \tag{5.27}$$

The stability boundaries shown in Figure 5.13 with the dashed lines were calculated using a volume fraction of 0.45. Clearly, the boundaries drop in proportion to the viscous forces as expected, so it is easier to break up flocculated pairs with an applied shear field. Therefore, we observe a larger fraction of the stable area that is occupied by particles that we would term 'dispersed' occurring at lower shear rates than with more dilute dispersions. It should be noted, however, that the shear rate axis is logarithmic and that the range of stability under shear at moderate ζ-potentials is actually reduced with increasing solid content. It is also important to observe that the boundary is moving away from the -20 mV threshold more rapidly, indicating more shear sensitivity. Indeed, we now see that close to the -20 mV threshold for rapid coagulation that flocculated particles will become coagulated at only moderate shear rates with there being no opportunity to separate them. At first sight, these observations may seem to be of minor significance, but we frequently have concentrated dispersions to which we wish to add stabilizers, such as a surfactant, for example, and of course mix the system thoroughly during the addition of the stabilizer solution. Once coagulated, separating the particles to allow the stabilizers to adsorb and function correctly may be impossible.

At this point of the discussion, the question naturally arises of the break-up of coagulated particles, as Equations 5.25a and 5.25b give an

estimation of the hydrodynamic forces available. This is, in part, the same problem as we see in the dispersion of dry powders into a liquid medium or the break-up of liquid drops in emulsification, although both of these processes have additional factors. In the case of the 'redispersion' of coagulated particles, the problem is to estimate the depth of the primary minimum as the shear force has to be sufficiently large to overcome the attraction and pull the particles apart to a distance equivalent to the maximum in the curve. Here, the problem comes down to details of molecular dimensions, so that the hydration of bound ions and their dimensions, for example, must control the approach. We could take a cut-off distance of, say, 0.5 nm as a typical minimum separation and use that to calculate a restabilization boundary. The force–distance curve is very steep at this point, so the calculation would not be of great practical use. Another problem is, that on this scale, the details of the surface topography, distribution of charges and so on become important. Finally, experimentation either to confirm the result or to provide a basis for estimation becomes difficult as the shear rates required would invariably involve turbulent flow fields, which are chaotic and difficult to describe.

REFERENCES

1. Derjaguin, B. V. and Landau, L., (1941) *Acta Physicochim.*, **14**, 633.
2. Verwey, E. J. W. and Overbeek, J. Th. G. (1948) *Theory of the Stability of Lyophobic Colloids*, Elsevier, Amsterdam.
3. Overbeek, J. Th. G. (1980) *Pure Appl. Chem.*, **52**, 1151.
4. Hunter, R. J. (1987) *Foundations of Colloid Science*, Vol. **1**, Oxford University Press, Oxford.
5. Russel, W. B., Saville, D. A. and Schowalter, W. (1989) *Colloidal Dispersions*, Cambridge University Press, Cambridge.
6. Matijévic, E. (1973) *J. Colloid Interface Sci.*, **43**, 217.
7. Gregory, J. (1996) Polymer adsorption and flocculation, in *Industrial Water Soluble Polymers* (ed. C. A. Finch), Royal Society of Chemistry, Cambridge, pp. 62–75.
8. Goodwin, J. W. and Mercer-Chalmers, J. (1998) Flow induced aggregation of colloidal dispersions, in *Modern Aspects of Colloidal Dispersions* (eds R. H. Ottewill and A. R. Rennie), Kluwer, Dordrecht, pp. 61–75.
9. (a) Sperry, P. R. (1982) *J. Colloid Interface Sci.*, **87**, 37; Sperry, P. R. (1984) *J. Colloid Interface Sci.*, **99**, 97.
10. Goodwin, J. W. and Hughes, R. W. (2000) *Rheology for Chemists – An Introduction*, Royal Society of Chemistry, Cambridge.
11. Derjaguin, B. V. (1966) *Discuss. Faraday Soc.*, **42**, 317.
12. Spielman, L. A. (1970) *J. Colloid Interface Sci.*, **33**, 562.
13. Reerink, H. and Overbeek, J. Th. G. (1954) *Discuss. Faraday Soc.*, **18**, 74.

6

The Wetting of Surfaces by Liquids

6.1 INTRODUCTION

At first sight, the wetting of a macroscopic surface by a liquid appears to be unrelated to the normal colloid discussion. It is, however, intimately related to the detailed surface chemistry and physics that determine the behaviour of particles. The same intermolecular forces are involved and the adsorption of macromolecules or surfactants is usually handled as though it was occurring on a macroscopic surface.

There are a large number of situations in which the manner and rate at which a surface is wetted by a liquid is of major importance. Many colloidal materials are obtained as dry powders which are then dispersed in a liquid medium. Pigments and fillers immediately come to mind. The extraction of oil from porous rock by replacing it through pumping water into the well is another case. We could pick many others from commercially important areas such as the enhancement of mineral ores via flotation to the more apparently mundane such as washing clothes and maintaining visibility through windows in wet conditions. Hence this is an area of study which fits naturally into problems of the application of colloidal systems, as we shall see in this chapter.

In Chapter 1, the concept of surface tension was introduced in terms of the Helmholtz free energy per unit are of surface in (Equation 1.15),

Colloids and Interfaces with Surfactants and Polymers, Second Edition James Goodwin

which can be restated as follows:

$$\gamma_s = \frac{F_s}{A_s} - \sum_i \mu_{is} n_{is} \tag{6.1}$$

were the subscript s denotes the surface, F is the free energy and A the area, with μ and n being the chemical potential and number of moles of each species in equilibrium, respectively. This formal definition helps to focus our attention on the details of the intermolecular forces, the molecular composition of the interfacial region and where we draw the 'Gibbs dividing surface' as our definition of the actual location of the surface. However, the experimental observations that we make, and the interpretation that we discuss in this chapter, treat surfaces as macroscopic with well-defined locations and analyse the force balances with classical mechanics. It is possible to do a great deal with this approach, but we must always keep in mind adsorption processes and that the surfaces should be in equilibrium unless we are specifically looking at rate processes.

6.2 THE CONTACT ANGLE

When a small amount of liquid is placed on the surface of a solid, it forms a drop which covers a limited area of the surface. This is illustrated in Figure 6.1, in which the contact angle, θ, is the angle between the tangent to the liquid surface at the contact line and the solid surface [1]. The limiting condition is that $0^\circ < \theta < 180^\circ$. If we take the condition where the contact angle approaches zero, we have a surface that is completely wetted by the liquid. Water on *clean* glass is an example of this. The glass is hydrophilic due to the silica surface, which is made up of a large number of oxygen atoms and surface silanol groups

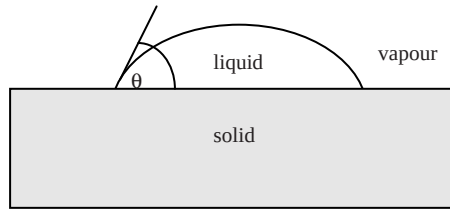


Figure 6.1 The contact angle, θ, formed by a drop of liquid on a solid surface

that can hydrogen bond with the water surface. Mercury on a polytetrafluoroethylene (PTFE) surface forms drops with a contact angle of about 150° and this can be considered to be complete non-wetting of the surface. If the value of contact angle is $\theta \geq 90°$, the droplet does not spread readily. If the volume in the drop is changed, we can observe that the line at the three-phase junction, the *wetting line*, moves with some difficulty. If liquid is withdrawn, the angle is reduced from the value that we see when liquid is added. These experiments have to be carried out slowly and carefully so that we are as close to an equilibrium value as possible. The angle obtained as we have just finished expanding the drop is known as the *advancing contact angle*, θ_A, and the angle that is observed when the liquid has just been withdrawn is the *receding contact angle*, θ_R. These angles change if they are measured while the wetting line is in motion. They are then the *dynamic contact angles* and merit a discussion in some detail due to their importance in many coating and printing processes.

6.2.1 The Young Equation

In the illustration of the liquid drop shown in Figure 6.1, the contact line will move until equilibrium is established. We may describe this situation in terms of the force balance in the plane of the surface shown in Figure 6.2a and expressed as the vector addition

$$\gamma_{vs} = \gamma_{ls} + \gamma_{vl} \cos\theta \tag{6.2}$$

where γ_{12} is the *interfacial* tension between phases 1 and 2 [e.g. vapour (v) and solid (s)] (it is only referred to as the surface tension if it refers to the liquid–saturated vapour interface). It should be noted that

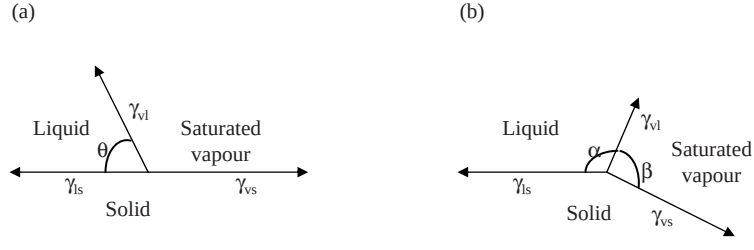


Figure 6.2 The balance of interfacial tensions at the vapour–liquid–solid contact line shown for (a) a flat surface and (b) a 'non-flat' surface

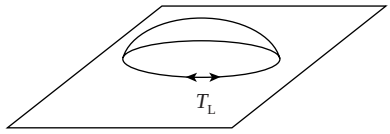


Figure 6.3 Schematic of the line tension, T_L, acting along the contact line of a liquid drop

Equation 6.2 refers to a drop on a flat surface. If the surface is not flat, then the balance shown in Figure 6.2b gives [2, 3]

$$\alpha + \beta \neq 180: \quad \gamma_{vs} \geq \gamma_{ls} + \gamma_{lv}\cos\beta; \quad \gamma_{ls} \geq \gamma_{vs} + \gamma_{lv}\cos\alpha \tag{6.3}$$

However, the contact line is not straight – it is curved and the radius of curvature is important, as this results in what is known as the *line tension*, T_L, which acts in opposition to the expansion of the drop (as illustrated in Figure 6.3). Equation 6.2 should be written for a drop of radius r as follows:

$$\gamma_{vs} = \gamma_{ls} + \gamma_{vl}\cos\theta + \frac{T_L}{r} \tag{6.4}$$

The line tension is not very large, being typically of the order of $T_L \approx 10^{-11}$ N, with the value $T_L/r > 1\,\text{mN}\,\text{m}^{-1}$ being required if it is to make a significant contribution. Therefore, we need only use Equation 6.4 for very small systems such as capillary condensation in porous solids.

6.3 METHODS FOR THE MEASUREMENT OF CONTACT ANGLE

Several methods are available for measuring the contact angle and suitable equipment may be purchased 'off the shelf'. The technique chosen will depend on the surface. For example, if it is a crystal face, one of the first two methods described below would be suitable using drops or bubbles. If it is a thin element which has had some form of surface treatment, one of the other methods may be easier. In all cases, the roughness of the surface is an issue which must be addressed.

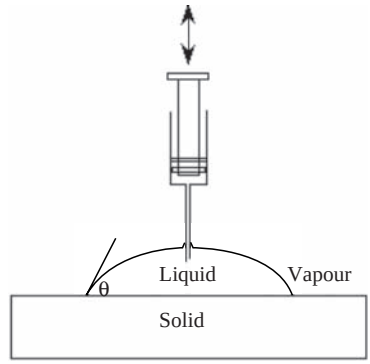


Figure 6.4 Schematic of the sessile drop experiment. In this approach, liquid is added or removed via the syringe and the drop profile is photographed

6.3.1 Sessile Drop

In this approach, a drop of liquid is placed on the surface for which the contact angle is required. A syringe is used to add liquid – either to add liquid to give the advancing angle or to withdraw liquid to give the value of the receding angle (Figure 6.4). A low-power microscope coupled to a computerized image analysis system is the most convenient method for the evaluation. The simplest method is to draw a tangent to the image on the screen at the contact point of the drop and the surface. Rather better, however, is to digitize the image of the curved surface, fit the points to a polynomial and compute the tangent to the curve at the contact point. The drop and experimental surface should be enclosed in a cell so that the drop is surrounded by saturated vapour. This allows the surface of the solid to reach an adsorption equilibrium and to limit evaporation of the droplet. The drop must be sufficiently large that the syringe needle does not affect the curvature of the drop surface close to the contact line.

6.3.2 Captive Bubble

In this experiment, the solid is immersed in the liquid and a bubble is produced from a gas-tight syringe and brought into contact with the surface (Figure 6.5). Again, a low-power microscope coupled to a camera and computerized imaging system is the most convenient method of determining the angle. In this case, expansion of the bubble enables

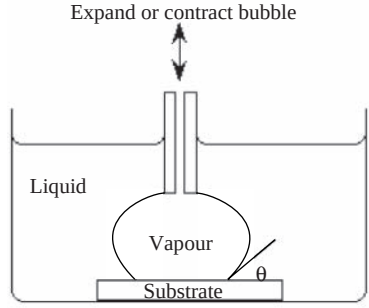


Figure 6.5 Schematic of the captive bubble experiment. Expanding the bubble gives θ_R and contraction gives θ_A

the receding angle to be recorded, while contraction gives the advancing angle. This method has an advantage in that a saturated vapour is readily achieved.

6.3.3 Wilhelmy Plate

This is carried out by employing the same equipment that can be used to determine the surface tension of a liquid in which mode the plate must be fully wetted and a contact angle of $\sim 0^\circ$ is assumed. However, if the surface tension of the liquid is known, the data can be used to calculate the contact angle. The experimental arrangement is shown in Figure 6.6.

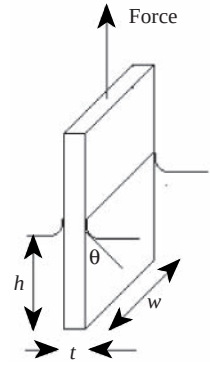


Figure 6.6 Schematic of the Wilhelmy plate experiment, where h is the depth of immersion of a plate of width w and thickness t. The net force is measured with a balance

The change in force recorded by the balance on partial immersion of the plate is equal to the surface tension force minus the buoyancy of the immersed section of the plate:

$$\text{force} = 2\,(w + t)\,\gamma_{\text{vl}} \cos\theta - wth\rho_{\text{l}} \tag{6.5}$$

where w is the width of the plate of thickness t and so $2(w + t)$ is the length of the contact line, h is the depth immersed and ρ_{l} is the liquid density. The advancing and receding angles are obtained by altering the depth of immersion.

6.3.4 Tilted Plate Methods

There are two methods which use tilted plates, as shown in Figure 6.7. When a drop is place on an inclined plane and the tilt of the plane is increased until the drop just starts to move, the advancing and receding angles may be obtained in one experiment (Figure 6.7a). This sounds easy, but it requires a continuous recording of the image to determine the point at which to take the measurement and the technique is particularly prone to vibrations, leading to significant experimental scatter. The second technique, illustrated in Figure 6.7b, also uses a tilted plate and makes use of the observation that the meniscus will appear flat when

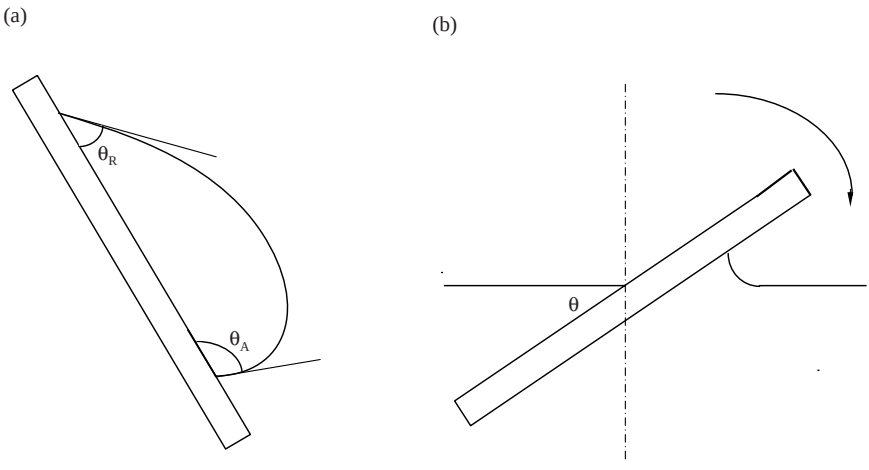


Figure 6.7 Schematic of the two tilted plate methods for measuring the contact angle

the plate makes an angle of $(180 - \theta)^\circ$ with the surface. Advancing and receding angles can only be obtained by changing the immersion depth of the plate.

6.4 CONTACT ANGLE HYSTERESIS

The value of the contact angle obtained experimentally is found to depend on whether the liquid is advancing over the surface or receding and $\theta_A \geq \theta_R$. We may also observe ageing effects. The first thing to consider is the nature of the surface. Although we may form a fresh surface, the nature of that surface rapidly changes owing to the adsorption of material from the atmosphere. This may be physical adsorption of water or hydrocarbons or chemical adsorption to give an oxide layer on metals or water as in the formation of surface silanol groups on freshly prepared silica. Recall from Chapter 3 that the surface tension tells us about the surface energy, and for 'atomic liquids' this can be calculated from the London dispersion forces, while the hydrogen bonding in liquids such as water makes the dominant contribution. With solids, things can be more complex as there can be residual strains present, but solids have much higher surface energies (and tensions) than liquids. Nevertheless, the adsorption on the surface can make a very large difference. Israelachvili [4] quotes data for cleaved mica, which illustrates this well. For mica, freshly cleaved under high vacuum, the value is $4.5\,J\,m^{-2}$ and this falls to $0.3\,J\,m^{-2}$ when it is cleaved under normal laboratory conditions and water can adsorb. A thin liquid layer can form on the solid surface and we might think that the contact angle would become very small, as the liquid would in effect be in contact with itself. However, the wetting film has to be thick for this to be the case. With thin films such as monolayers, the underlying intermolecular forces are still evident. Of course, when we make measurements of advancing angles, there may only be a monolayer or so of liquid adsorbed on the solid, but when we reverse the process by withdrawing liquid and rapidly measure the receding angle, a much thicker film may be present which only thins slowly.

However, there are other reasons for the observed hysteresis. Surface heterogeneity can result from polycrystallinity or variations in local composition. When the wetting line moves over a heterogeneous surface, it will stick as it passes a boundary from a lower to a higher energy region. This will hold back the contact line, thus increasing the advancing angle and reducing the receding angle. Surface roughness is

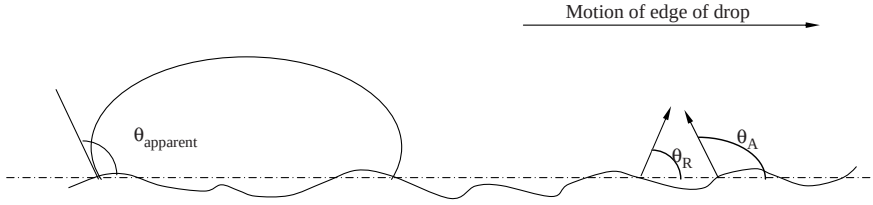


Figure 6.8 Schematic of a drop on a rough surface. Here, we measure an 'apparent' contact angle and hence the differences between advancing and receding angles, even though the true contact angle does not alter ($\sim 90^\circ$ in the figure)

an important cause of hysteresis. The schematic shown in Figure 6.8 illustrates the effect. When the angle is measured, the surface is taken as a plane through the undulations, although the liquid must form an angle to the local surface, which may be at an angle to the viewing plane. As shown in Figure 6.8, if the surface is at an angle α to the measurement or viewing plane, the receding angle will apparently be reduced by α and the advancing angle increased by α at the points shown. What will happen is that the drop will move in 'jumps' from one metastable configuration to the next. Wenzel [5] defined a 'roughness factor', R_f, in terms of the ratio of the actual surface area to the apparent geometric area for which the apparent contact angle was θ', so Young's equation now becomes

$$R_f (\gamma_{vs} - \gamma_{ls}) = \gamma_{vl} \cos \theta' \tag{6.6a}$$

and so

$$R_f = \frac{\cos \theta'}{\cos \theta} \tag{6.6b}$$

The dimensional scale of the surface heterogeneity or roughness is important. A dimension of at least 100 nm appears to be required. This means, for example, that optically smooth surfaces will not show much hysteresis due to roughness.

6.5 SPREADING

The surface tension of a pure liquid was defined as the surface energy of that liquid. That is, it is the energy required to generate unit area of a new surface. The *work of cohesion*, W_{ll}, of a liquid is the work that

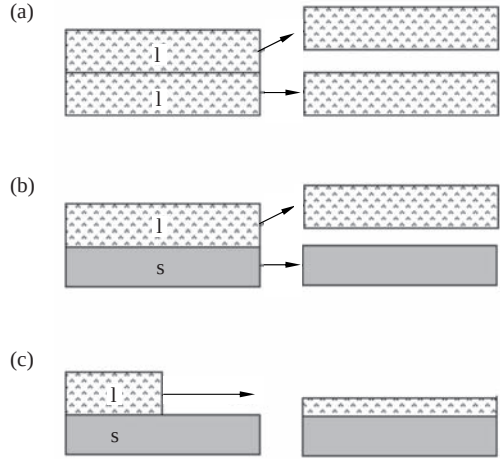


Figure 6.9 Schematics of the processes of (a) cohesion, (b) adhesion and (c) spreading

would be required to separate two 'slabs' of the liquid as shown in Figure 6.9a:

$$W_{ll} = 2\gamma_{vl} \tag{6.7}$$

The *work of adhesion*, W_{sl}, is the work required to separate two different materials (shown in Figure 6.9b). It is calculated as the sum of the energies of the new surfaces that are formed, minus that of the interface that is lost, as follows:

$$W_{ls} = \gamma_{sv} + \gamma_{lv} - \gamma_{sl} \tag{6.8}$$

Figure 6.9c shows spreading in similar terms. We may define the *equilibrium spreading coefficient*, S, in terms of the work of adhesion and work of cohesion. We have to do work to create the new liquid surfaces and then this is offset by the work of adhesion to the liquid to the vapour/solid surface:

$$S_{vls} = W_{ls} - W_{ll} \tag{6.9a}$$

$$S_{vls} = \gamma_{vs} - \gamma_{vl} - \gamma_{ls} \tag{6.9b}$$

We may use Young's equation (Equation 6.2) to put both the work of adhesion and the spreading coefficient in terms of readily measured

quantities such as the contact angle:

$$W_{ls} = \gamma_{vl}(\cos\theta + 1) \tag{6.10}$$

This is known as the Dupré equation and the spreading coefficient is given by

$$S_{vls} = \gamma_{vl}(\cos\theta - 1) \tag{6.11}$$

and, of course the contact angle can be written in terms of the ratio of the work of adhesion and the work of cohesion, as follows:

$$\cos\theta = 2\frac{W_{ls}}{W_{vl}} - 1 \tag{6.12}$$

It is instructive to note the limits:

$$\begin{aligned} &W_{ls} = W_{vl}; \quad \cos\theta = 1 \text{ and so } \theta = 0 \\ &W_{ls} = \frac{W_{vl}}{2}; \quad \cos\theta = 0 \text{ and so } \theta = 90 \\ &W_{ls} \ll W_{vl}; \quad \cos\theta \to -1 \text{ and so } \theta \to 180 \end{aligned} \tag{6.13}$$

Of course, this is no surprise as the work of cohesion and adhesion are calculated from the surface energies, which are a reflection of the intermolecular forces. For non-polar liquids, the London dispersion forces provide a good estimate of the surface energy via the appropriate Hamaker constants. Girifalco and Good [6] and Fowkes [7] observed that as the geometric mean gave a reasonable approximation for the combined Hamaker constant, the geometric mean of the work of cohesion and hence surface tensions could be used to calculate the interfacial tension. So, for two liquids with surface tensions equal to γ_1 and γ_2, the interfacial tension, γ_{12}, from Equation 6.8 is as follows:

$$\gamma_{12} = \frac{W_{11}}{2} + \frac{W_{22}}{2} - W_{12} \tag{6.14}$$

or

$$\gamma_{12} = \gamma_1 + \gamma_2 - W_{12} \tag{6.15}$$

Now

$$W_{12} \approx \sqrt{W_{11}W_{22}} = 2\sqrt{\gamma_1^d \gamma_2^d} \tag{6.16}$$

and

$$\gamma_{12} \approx \gamma_1 + \gamma_2 - 2\sqrt{\gamma_1^{\mathrm{d}}\gamma_2^{\mathrm{d}}} \tag{6.17}$$

where the superscript d is used to indicate the dispersion force contribution to the surface tension. For example, the surface tension of water has a dispersion force contribution plus a larger H-bonding contribution, with the result that $\gamma_{H_2O} = \gamma_{H_2O}^{\mathrm{d}} + \gamma_{H_2O}^{\mathrm{H}} = 22 + 50.5 = 72.5\,\mathrm{mN\,m^{-1}}$. So, using Equation 6.17 for octadecane on water, we can calculate the interfacial tension at $20\,^{\circ}\mathrm{C} \approx 28 + 73 - 2(28 \times 22)^{0.5} \approx 51\,\mathrm{mN\,m^{-1}}$ as the interaction across the interface that is due to the dispersion forces.

When we are considering liquids wetting a solid and we combine Equation 6.10 with Equation 6.16, we obtain

$$\gamma_{\mathrm{vl}}(\cos\theta + 1) = 2\sqrt{\gamma_{\mathrm{vl}}^{\mathrm{d}}\gamma_{\mathrm{s}}^{\mathrm{d}}} \tag{6.18}$$

Note that the surface tension of the solid is not modified by the vapour from the liquid in this expression and so we must be careful in the application of Equation 6.18. Rearrangement of the latter gives the following:

$$\gamma_{\mathrm{s}}^{\mathrm{d}} = \frac{\gamma_{\mathrm{vl}}^{\mathrm{d}}(\cos\theta - 1)^2}{4} \tag{6.19}$$

So, we may use measurements of the contact angle and the dispersion contribution to the surface tension of the liquid–vapour interface to estimate the solid surface tension. Good results have been obtained for non-polar materials. In many cases, the vapour from the liquid adsorbs on the surface of the solid and the surface pressure of this wetting film, Π_{slv}, needs to be considered. Therefore, we should write

$$\gamma_{\mathrm{s}} = \gamma_{\mathrm{sl}} + \Pi_{\mathrm{slv}} \tag{6.20}$$

The surface pressure is most significant when the contact angle approaches zero and the surface is completely wetted by the liquid which spreads over the surface, giving stable films at all thicknesses. When the contact angle is greater than zero, the thinning film produced as the liquid spreads becomes unstable and breaks up into droplets.

The *critical surface tension of wetting*, γ_{C}, can be defined by making use of Equation 6.19 as $\theta \to 0$; then $\gamma_{\mathrm{s}}^{\mathrm{d}} \to \gamma_{\mathrm{vl}}^{\mathrm{d}} = \gamma_{\mathrm{C}}$. A plot of $\cos\theta$

versus $(\gamma_{vl}^{d})^{-0.5}$ should be linear and the value of the surface tension found by extrapolating the line to $\cos\theta = 1$ gives an estimate of the surface tension of the solid. This is known as a Zisman plot [8] and has been used to determine the surface energy of solids which have a low surface energy. However, this is only an approximation as the surface pressure would have to be zero, and that would not be the case as θ approaches zero.

6.5.1 The Adsorption of Surfactants

In many practical situations, surfactants are present as stabilizers, dispersants or wetting aids. Often, we will be concerned with the wetting of solids with aqueous solutions, but that will not always be the case. The adsorption of the surfactant at each interface must be considered. Young's equation is a useful starting point and if it is written as

$$\cos\theta = \frac{\gamma_{vs} - \gamma_{ls}}{\gamma_{vl}} \tag{6.21}$$

it is immediately apparent that if either of the surface tension values at the liquid–solid or liquid–vapour interfaces decreases, then $\cos\theta$ will increase and so wetting is improved, as this means that the contact angle moves towards zero. As the surfactant is in the liquid phase, adsorption at the solid–vapour interface may not be complete. For example, this would be the case if we were coating a using a blade or roller technique. In other situations, this may not be the case as the solid may be in the liquid phase prior to the gas phase being introduced. The separation of minerals by *froth flotation* is an example here. In this process, the crushed mineral and its matrix are mixed with water and a surfactant solution is added (the 'collector'). Preferential adsorption on the mineral fragments occurs and air is pumped into the solution to produce a froth. Additional surface-active materials ('frothers') are used to reduce the liquid–vapour tension to enhance the frothing response. These have to be carefully chosen so that they do not change the contact angle too much. The contact angle at the mineral–liquid–vapour line is increased ($\cos\theta \to 0$), so that the desired mineral particles become attached to bubbles and are removed in the froth. There is a complex situation here as there is more than one type of solid surface present and also more than one surfactant being added with specific interactions with the surfaces occurring. As a simple example, let us consider the effect of a cationic surfactant such

as hexadecyltrimethylammonium bromide (C_{16}TAB) at the glass (fused silica)–water–air interface. The silica surface is negatively charged at pH $\geq$ 2 and the cationic head group produces strong adsorption of the surfactant at low solution concentrations. A monolayer of surfactant at the liquid–solid interface is formed by the time a solution concentration of $\sim 5 \times 10^{-5}$ M is reached. The contact angle of water on clean glass is $<10^{\circ}$ but with a monolayer of C_{16}TAB it is $\sim 90^{\circ}$. The cmc of the surfactant is $\sim 10^{-3}$ M and at that concentration a bilayer of surfactant is adsorbed and the contact angle is again reduced to $<10^{\circ}$. At 5×10^{-5} M, the adsorption of the surfactant at the air–liquid interface is low and has reduced the surface tension by only about 10%. The large change in contact angle is due to the solid surface, which now consists of a surface of densely packed alkane chains causing the values of both γ_{vl} and γ_{ls} to decrease dramatically, although the former falls more (to $\sim 27\,\mathrm{mN\,m^{-1}}$) than the latter. Wetting is again favoured by the ionogenic surface formed by the bilayer in combination with the reduced tension of the air–water surface as the cmc is approached.

6.6 CURVED SURFACES

The surface tension acts in the plane of the surface, so if we have a small volume of liquid a spherical drop is formed. There is a pressure drop across the interface and we may calculate this by balancing the work done in attempting to reduce the surface area with that generated by the internal pressure. The change in the surface energy of a sphere of radius r is given by

$$\gamma_{\mathrm{lv}} \mathrm{d}A(r) = \gamma_{\mathrm{lv}} 8\pi r \,\mathrm{d}r \tag{6.22}$$

At equilibrium, this is balanced by the change in the pressure in the drop, thus

$$4\pi r^2 \Delta P \mathrm{d}r = 8\pi \gamma_{\mathrm{lv}} r \,\mathrm{d}r \tag{6.23}$$

The result is the Young–Laplace equation [1]:

$$\Delta P = \frac{2\gamma_{\mathrm{lv}}}{r} \tag{6.24}$$

where r is the radius of the spherical drop. Equation 6.24 applies to any curved interface and it may be generalized for two different radii of curvature measured orthogonally, as follows:

$$\Delta P = \gamma_{lv}\left(\frac{1}{r_1} + \frac{1}{r_2}\right) \tag{6.25}$$

Note that the convention is that the radius of curvature is measured in the liquid phase so that the sign convention is that it is positive for a drop and negative for a bubble. This means that the vapour pressure of a liquid in a small drop is higher than for a flat surface, but it is lower in a bubble than for the flat surface. A consequence of this is that large drops grow at the expense of small drops as the material evaporates from the small and condenses on the larger drops.

6.7 CAPILLARITY

The behaviour of liquids in capillaries and pores is an important aspect of the wetting of surfaces. The wetting of rock in oil wells is one such example. A single capillary is a suitable starting point for the discussion. Figure 6.10a shows a capillary with a circular cross-section of radius r in which the liquid has risen to a height h. The liquid rises until the surface force which causes the liquid to spread on the surface is balanced by the gravitational force on the column of liquid. The pressure drop across the interface (the *capillary pressure*) is given by the Young–Laplace equation

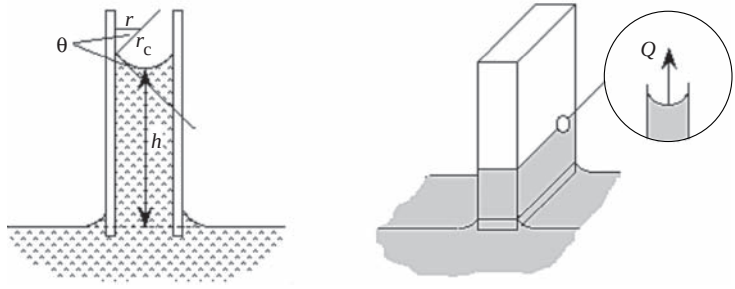


Figure 6.10 (a) Schematic of the capillary rise in a tube where the radius is r and the radius of curvature of the meniscus is r_c. (b) Schematic of the capillary rise in a porous block, where the volumetric flow rate in a pore is Q

for a circular capillary:

$$\Delta P = hg\,(\rho_{\mathrm{l}} - \rho_{\mathrm{v}}) = \frac{2\gamma_{\mathrm{lv}}}{r} \tag{6.26}$$

In nearly all cases $\rho_{\mathrm{l}} \gg \rho_{\mathrm{v}}$. When the contact angle is greater than zero, the radius of curvature of the meniscus, r_{c} in Figure 6.10a, is $r/\cos\theta$ and Equation 6.26 then becomes

$$h\rho_{\mathrm{l}}g = \frac{2\gamma_{\mathrm{vl}}\cos\theta}{r} \tag{6.27}$$

This is an equilibrium measurement and provides another route to the determination of the contact angle. It should be noted that we have an inverse dependence of h on the radius of curvature. Therefore, if we have a shape in which the radius of curvature changes, we should expect to see a greater rise with the sharpest curvature. It is instructive to look at the behaviour of water in a small, square, glass optical cell. The meniscus is highest in the inside corners where there is a small radius of curvature. If the same cell is then stood in a dish of water, we can see that the wetting line on the outside is lowest at the outer corners (negative radius of curvature).

The imbibition of liquid into a capillary (or pore) is illustrated in Figure 6.10b.The rate is often important in many situations. We start by using the Poiseuille equation for the flow of liquid in a tube:

$$Q = \frac{\Delta P \pi r^4}{8L\eta} \tag{6.28}$$

where L is the wetted length of a capillary of radius r, η is the viscosity of the liquid and Q is the volume flow rate:

$$Q = \pi r^2 \frac{\mathrm{d}L}{\mathrm{d}t} \tag{6.29}$$

We may now insert Equations 6.26 and 6.29 into Equation 6.28 to give the Washburn equation:

$$\frac{\mathrm{d}L}{\mathrm{d}t} = \frac{r\gamma_{\mathrm{vl}}\cos\theta}{4l\eta} \tag{6.30}$$

Integration of this equation gives the distance the liquid penetrates in a given time:

$$\int_0^{L(t)} L\mathrm{d}L = \frac{r\gamma_{\mathrm{vl}}\cos\theta}{4\eta}\int_0^t \mathrm{d}t \tag{6.31a}$$

$$L(t) = \left(\frac{r\gamma_{\mathrm{vl}}\cos\theta}{4\eta}t\right)^{0.5} \tag{6.31b}$$

This result is for a single pore but it is readily adapted for a porous solid. If we have a porous block of material, for example some rock, the pores will vary in diameter and path. The effect can be included in one factor, the '*tortuosity factor*', K:

$$L(t) = \left(\frac{r\gamma_{\mathrm{vl}}\cos\theta}{4\eta K^2}t\right)^{0.5} \tag{6.32}$$

For a given porous system, r/K^2 is a constant. If the penetration is measured using two or three different liquids each of which has a value of $\cos\theta \approx 1$, but with different viscosities, the value of r/K^2 can be determined using Equation 6.32.

The penetration of oils, water or surfactant solutions is of great importance both in the drilling for oil and gas and in their recovery from mature wells. Some of the clay-based formations are soft, porous and easily wetted by water so that controlling the size of the well-bore can be difficult. The 'breakthrough' of water into the recovery well from the well bore used to pump in replacement water is also a major problem. Polymers and surfactants can play an important role in controlling the contact angles and viscosity in these porous systems.

6.7.1 Dispersion of Powders

Another common situation concerning the penetration of liquid into a porous matrix is the dispersion of dry powders in liquids. Many colloidal systems are dried and subsequently redispersed. The dry powders are easy to store and transport and are also easy to handle in batch applications. Pigments, fillers and some foodstuffs are commonly handled in this way. When dried and ground, it is useful if the aggregates

are macroscopic so that dust problems are minimized, but large aggregates have to be wetted and broken up. The particles are held together by dispersion forces and in the dry state these are very strong as there is no intervening medium to mediate them, and also the particles are in intimate contact in a deep attractive well. The first requirement is that the liquid must wet the particles. If the contact angle is low, the medium will be imbibed and the Laplace pressure is given by Equation 6.25 (suitably modified by the tortuosity factor K). The rate of penetration is given by the Washburn equation, but note that as the aggregate is immersed in the fluid, the air inside is trapped and will give a back-pressure and slow the penetration. As the pressure is higher than atmospheric, the gas will dissolve and allow the penetration to continue. However, this will be slower as the gas has to diffuse out of the aggregate through the imbibed liquid. This is clearly a slower process for large than for small aggregates. As an example, we can take a pore radius of 0.1 μm, a surface tension for the wetting liquid of $40\,\mathrm{mN\,m^{-1}}$ and a contact angle of 60°. If $K \approx 1$, the Laplace pressure is 4 atm. Even if the pore penetration continues until the internal pressure reaches this value, it will be insufficient to break up the aggregate, but will slow the wetting of the aggregates. Of course, we normally use mechanical action, that is, shear forces, to aid the break-up, by milling, for example. If the deformation rates (shear, extension or both) are high and the concentration of aggregates is also high, there may be enough force to overcome the interparticle attraction. The hydrodynamic stress is then given by:

$$\frac{F_\mathrm{h}}{{a_\mathrm{f}}^2} \propto \eta(\varphi)\dot{\gamma} \tag{6.33}$$

where F_h is the hydrodynamic force produced on an aggregate with a radius a_f by a shear rate $\dot{\gamma}\,\mathrm{s^{-1}}$ and $\eta(\varphi)$ is the viscosity of the concentrated suspension of aggregates. Deformation rates of 10^5–$10^6\,\mathrm{s^{-1}}$ are high, and much higher values are difficult to achieve. The value of the dispersion viscosity is of great importance here as stresses equivalent to several atmospheres would be required to overcome the dispersion forces. This is why it is much easier to disperse powders as a paste and subsequently dilute rather than add all the dispersion medium initially.

The wetting of the aggregate, however, is still an essential part as solvation of the surface allows repulsion to occur in addition to reducing the attraction. Repulsion may be the development of electrical

charge or it may be the solvation of polymeric or protein layers. Both of these result in repulsion aiding break-up of the aggregates and also stabilizing the dispersed units to prevent reagglomeration. This introduces another feature of capillary behaviour in porous systems, namely the tendency for vapours to condense in small pores, which will aid the wetting process.

6.7.2 The Kelvin Equation

The Young–Laplace equation indicates the pressure difference across a curved interface. The Kelvin equation enables us to calculate the change in vapour pressure due to the curvature of the interface. A clear derivation was given by Hunter [9] so just the outline will be given here. For a pure liquid, at the curved interface at equilibrium the chemical potentials in the liquid and vapour phases are equal and the pressure drop is given by the Young–Laplace equation, as follows. With $\mu_l = \mu_v$,

$$P_l - P_v = \frac{2\gamma_{lv}}{r} \tag{6.34}$$

We may make a small displacement from equilibrium, and then we have $d\mu_l = d\mu_v$ and

$$dP_l - dP_v = d\left(\frac{2\gamma_{lv}}{r}\right) \tag{6.35}$$

where the changes in chemical potentials are

$$d\mu_v = \overline{V}_v dP_v - \overline{S}_v dT \tag{6.36a}$$

$$d\mu_l = \overline{V}_l dP_l - \overline{S}_l dT \tag{6.36b}$$

where the bars indicate the molar volumes and entropies. At constant temperature and as we are close to equilibrium, we may write

$$P_l = \frac{\overline{V}_v}{\overline{V}_l} P_v \tag{6.37}$$

Substituting this into Equation 6.34 yields

$$\left(\frac{\overline{V}_v}{\overline{V}_l} - 1\right) dP_v = d\left(\frac{2\gamma_{lv}}{r}\right) \tag{6.38}$$

The boundary conditions that we need for the integration of Equation 6.38 are $r \to \infty$, $P_v \to P^0$ (the vapour pressure of the liquid with a flat surface). Then, by taking note that as

$$\frac{\overline{V}_v}{\overline{V}_l} \gg 1 \text{ and } \overline{V}_v \approx \frac{RT}{P_v}, \text{ then } \int_{P^0}^{P_v} \left(\frac{RT}{\overline{V}_l}\frac{dP_v}{P_v}\right) \approx \int_{\infty}^{r} d\left(\frac{2\gamma_{lv}}{r}\right) \tag{6.39}$$

the result for the vapour–liquid surface is given by

$$\ln\left(\frac{P_v}{P^0}\right) \approx \frac{\overline{V}_l}{RT}\frac{2\gamma_{lv}}{r} \tag{6.40}$$

If the cross-section is oval instead of circular, then Equation 6.40 is written in terms of the major and minor radii, r_1 and r_2:

$$\ln\left(\frac{P_v}{P^0}\right) \approx \frac{\gamma_{lv}\overline{V}_l}{RT}\left(\frac{1}{r_1} + \frac{1}{r_2}\right) \tag{6.41}$$

If we consider a liquid drop, the radius of curvature (in the liquid) is positive and Equation 6.40 indicates that as r decreases, $P_v/P^0 > 1$, that is, the vapour pressure of the liquid in the drop becomes increasingly larger than that of the flat surface. This is in contrast to the case of a bubble, where the radius of curvature (in the liquid) is negative and then decreasing the radius means that the vapour pressure becomes lower although the actual pressure (the capillary pressure) is higher (as given by the Young–Laplace equation). The vapour pressure above the liquid in a capillary or a pore is also lower and this gives rise to the phenomenon of *capillary condensation*. In this case, vapours start to condense in the finest pores or crevices where there is a highly curved region prior to the saturation vapour pressure being reached.

This effect of condensation in regions with a sharp curvature can also be a problem with dry powders stored in a humid atmosphere. At the contact points between the particles (usually agglomerates), the curvature is such that capillary condensation occurs in the contact zones. The

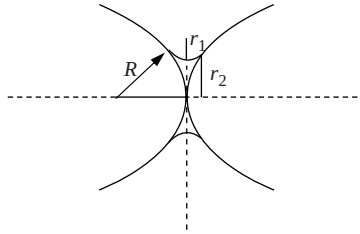


Figure 6.11 Illustration of liquid condensation in the contact region between particles in a powder to give a liquid bridge, where R is the radius of curvature of the particle surface and r_1 and r_2 are the radii of curvature of the meniscus (note that r_2 is negative in this figure)

force exerted by the liquid can be large enough to make the powder particles stick firmly and we recognize this as 'caking'. It is a useful exercise to calculate the force, assuming that the particles are smooth spheres completely wetted by the liquid phase [9]. The geometry is illustrated in Figure 6.11, where the radius of the particles is R and the radii of curvature of the liquid wetting the particles are r_1 and $-r_2$ as shown in the figure. The Laplace pressure is

$$\Delta P = \gamma_{\mathrm{lv}} \left(\frac{1}{r_1} - \frac{1}{r_2} \right) \tag{6.25}$$

This gives the force F_{cc}, holding the surfaces together, as follows:

$$F_{\mathrm{cc}} = \Delta P \pi {r_2}^2 = \pi \gamma_{\mathrm{lv}} \left(\frac{{r_2}^2}{r_1} - r_2 \right) \tag{6.42}$$

Now, as $r_2 > r_1$ and noting from Figure 6.11 that $2R \approx {r_2}^2/r_1$:

$$F_{\mathrm{cc}} \approx 2\pi R \gamma_{\mathrm{lv}} \tag{6.43}$$

6.7.3 Solid Particle Solubility

It is interesting that the Kelvin equation can also be used to describe the change in solubility of a solid particle in a liquid as the particle size is reduced. The solubility increases as the size is reduced in the same manner as the vapour pressure of the liquid in a drop increases as the size is reduced. This means that larger particles grow by the

transfer of material from smaller particles. This is known as *Ostwald ripening* and the analogous form of the Kelvin equation is

$$\ln\left[\frac{s(r)}{s(r \to \infty)}\right] \approx \frac{2\gamma_{ls}\overline{V}_s}{rRT} \tag{6.44}$$

where $s(r)$ is the solubility of a solid particle with a radius r. However, the equivalent of capillary condensation also occurs where two crystals make contact (in a sedimented bed, for example) and then they fuse together and make redispersion impossible.

6.8 TEMPERATURE EFFECTS

As a material is heated, we observe melting and then boiling of the liquid phase to produce the vapour. In some cases, we just observe sublimation of the solid. The surface tension of a material is due to the intermolecular forces in the material as we have already discussed. The surface tension then reflects the decreasing attraction and decreases with increase in temperature. This effect is larger in the liquid phase than the corresponding solid and occurs for both simple molecular liquids where the intermolecular forces are dispersion forces and for those with other contributions in addition. For example, the decrease in hydrogen bonding in water as the temperature is increased towards 100 °C is familiar to us.

When we consider the equilibrium expressed by the Young equation (Equation 6.2), we note that the change in the tension at the liquid–vapour interface will be greater than the difference in the changes in the tensions at the solid–vapour and the solid–liquid interfaces. For equilibrium to be maintained, the contact angle must decrease, that is, $\cos\theta$ increases. This means that the surface that we are considering is more easily wetted at higher temperatures. Adamson [1] provides a more extensive discussion on the effect of temperature on contact angle. We shall make use of the change to provide an estimate of the enthalpy of the wetting process.

6.8.1 Heat of Wetting

The free energy change on completely immersing a solid in a liquid can be estimated from the difference in tensions as we replace the solid–vapour

interface by a solid–liquid interface:

$$-\Delta G_{\mathrm{w}} = \gamma_{\mathrm{sv}} - \gamma_{\mathrm{sl}} \tag{6.45}$$

We must keep in mind that Equation 6.45 implies that the solid is uniform throughout and we can consider it to be in an equilibrium state. Thus, we are assuming that the tensions are a good approximation to the surface free energy of the solid. We associate the temperature dependence with the change in entropy, so we may take the approach of Harkins and Jura [10] in their work using heats of immersion:

$$\Delta H_{\mathrm{w}} = \Delta G_{\mathrm{w}} + T\Delta S_{\mathrm{w}} \tag{6.46}$$

and using Equation 6.45 with the variation of each tension with temperature, we have

$$\Delta H_{\mathrm{w}} = \left(\gamma_{\mathrm{sl}} - T\frac{\mathrm{d}\gamma_{\mathrm{sl}}}{\mathrm{d}T}\right) - \left(\gamma_{\mathrm{sv}} - T\frac{\mathrm{d}\gamma_{\mathrm{sv}}}{\mathrm{d}T}\right) \tag{6.47}$$

Grouping the like terms:

$$\Delta H_{\mathrm{w}} = -\left[(\gamma_{\mathrm{sv}} - \gamma_{\mathrm{sl}}) - T\left(\frac{\mathrm{d}\gamma_{\mathrm{sv}}}{\mathrm{d}T} - \frac{\mathrm{d}\gamma_{\mathrm{sl}}}{\mathrm{d}T}\right)\right] \tag{6.48}$$

We may use the Young equation to put this in terms of the contact angle and the liquid–vapour tension:

$$\Delta H_{\mathrm{w}} = -\left[\gamma_{\mathrm{lv}}\cos\theta - T\,\frac{\mathrm{d}\,(\gamma_{\mathrm{lv}}\cos\theta)}{\mathrm{d}T}\right] \tag{6.49}$$

This enables us to estimate the heat of wetting simply by measuring the changes in contact angle and the surface tension of the liquid as a function of temperature. In many cases, this is a simpler experiment than using calorimetry as the heat change is so small. However, the precision of the contact angle measurements is only a few degrees, and we must recognize the approximation in equating the surface tension of the solid with the surface free energy.

6.9 DYNAMIC CONTACT ANGLES

Thus far, the discussion in this chapter has been restricted to static processes, which, if not in equilibrium, are in a metastable state. During the coating of surfaces, often at very high rates, the dynamics are important. The first issue to consider is the tensions at the liquid interfaces. If surfactants have been included in the formulation of a coating, the kinetics of the adsorption process at both the liquid–solid and the liquid–vapour interfaces is important. As the timescale for diffusion to the interface is approached by the timescale of the creation of the new interfaces, adsorption will not be completed. As the coating rates are increased, the tensions approach those due to the major liquid component. In practice, this is not often a factor, however, as the surfactant concentrations are frequently around the critical micelle concentrations so as to ensure sufficient material to provide the colloidal stability of the particulates in the system.

A more universal problem is that of the hydrodynamic or viscous stress. The viscous force acts in the same direction as the solid–liquid surface tension, that is, in the same direction as the surface is moving or in the opposite direction to the motion of the coating implement, whether a blade, brush or roller. Figure 6.12 provides a simplified schematic for the coating of a moving tape using a blade to determine the thickness of the coat. At rest we can use the Young equation (Equation 6.2) for the

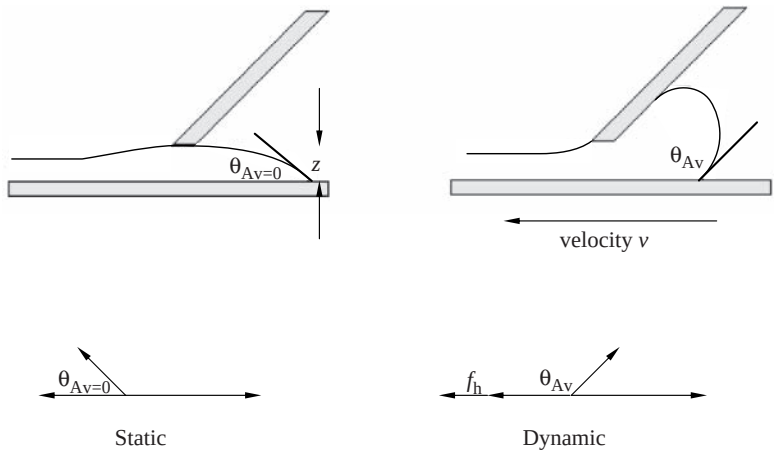


Figure 6.12 Schematic representation of a coating process, using a blade to determine the film thickness

force balance by using the advancing contact angle, as follows:

$$\gamma_{sv} = \gamma_{sl} + \gamma_{lv} \cos \theta_{Av \to 0} \tag{6.2}$$

and when the coating is being applied we must include the viscous force term in the new force balance:

$$\gamma_{sv} = \gamma_{sl} + \gamma_{lv} \cos \theta_{Av} + f_h \tag{6.50}$$

where f_h is the hydrodynamic force per unit length of interface moving at a velocity v with θ_{Av} being the dynamic advancing contact angle at that velocity. Rearrangement gives the dynamic contact angle as

$$\theta_{Av} = \arccos\left(\cos \theta_{Av \to 0} - \frac{f_h}{\gamma_{lv}}\right) \tag{6.51}$$

Assuming the coating is a Newtonian fluid, the shear stress ($N\,m^{-2}$) is equal to the product of the viscosity and the shear rate. The latter is calculated from the tape velocity and the gap, z. Hence we have

$$\sigma = \eta\dot{\gamma} \text{ and } \dot{\gamma} = \frac{v}{z} \tag{6.52}$$

We need to characterize the area over which the stress is acting to give the force, and per unit length of wetting line we may use the distance into the liquid film where flow is occurring as nz, where $n \geq 1$. A value of $n \approx 1$ will simplify the calculations and is a reasonable approximation, so the viscous force contribution is

$$f_h = \eta \frac{v}{z} nz \approx \eta v \tag{6.53}$$

The dynamic advancing contact angle is now

$$\theta_{Av} \approx \arccos\left(\cos \theta_{Av \to 0} - \frac{\eta v}{\gamma_{lv}}\right) \tag{6.54}$$

Note that for the equivalent relationship for a receding angle the viscous contribution must be added:

$$\theta_{Rv} \approx \arccos\left(\cos \theta_{Rv \to 0} + \frac{\eta v}{\gamma_{lv}}\right) \tag{6.55}$$

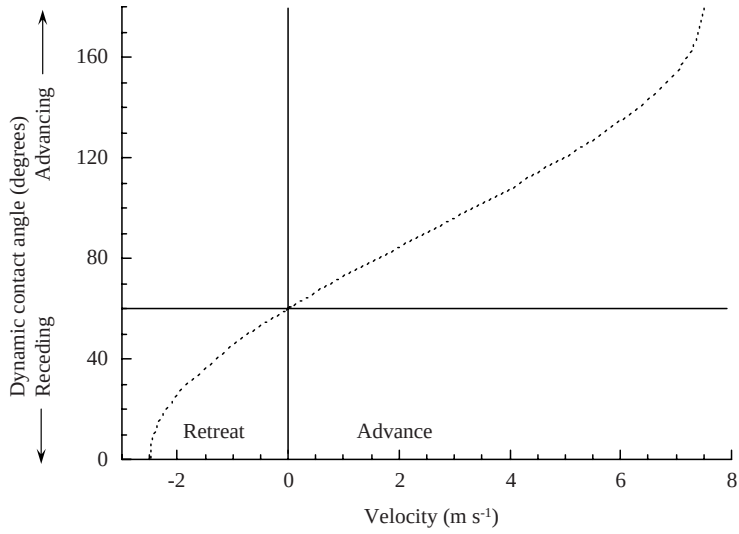


Figure 6.13 The dynamic contact angle plotted as a function of the front velocity, where the calculations were carried out for a solution with a static surface tension of 50 mN m^{-1} and a viscosity of 0.1 Pa s

Let us take an example to illustrate the process. If the static contact angle is 60° (no hysteresis), the surface tension of the coating is 50 mN m^{-1} and the viscosity is 0.1 N m^{-2} s, we have the effect of tape velocity shown in Figure 6.13. We see from Figure 6.13 that with this modelled coating system operating at a velocity of ~7.5 s^{-1}, the dynamic advancing contact angle is ~180°. When this occurs, air is entrained and the resulting coating is poor with patches of bare surface. The implication for the receding angle curve is that dewetting becomes ineffective at velocities of ~2.5 m s^{-1} with this model. This would be important in situations such as drawing a fibre out of a coating fluid. Even when a surface is completely wetted by the coating, air entrainment still occurs if the velocity is high enough.

The results of an another example calculation, which uses the same simple model, is shown in Figure 6.14. In this case, the entrainment is predicted at a fairly low velocity of ~50 cm s^{-1} as the viscosity is much higher at 0.25 Pa s.

Many coatings consist of concentrated dispersions and the viscosity deformation rate curve is non-linear with the material showing shear thinning. The shape of the curve will be modified, with a more rapid rise at low velocities as the stress is proportionately larger. The curve shape

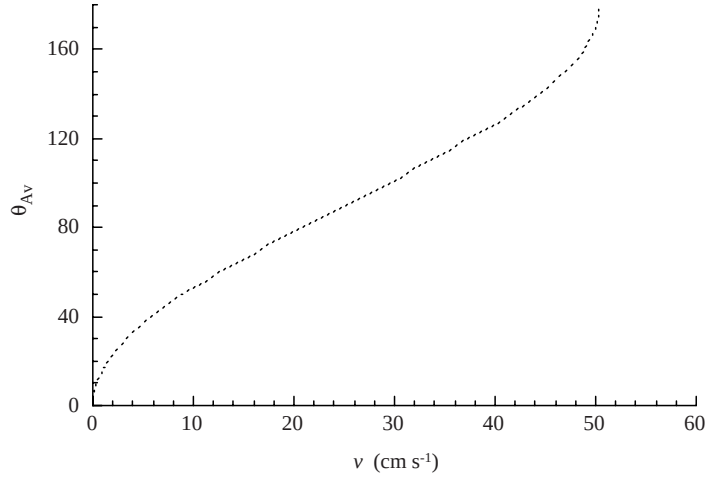


Figure 6.14 The dynamic contact angle calculated for a moving surface with a static contact angle of zero, a surface tension of 63 mN m^{-1} and a viscosity of 0.25 Pa s. Air entrainment occurs as the velocity reaches 50 cm s^{-1}

will also depend on the precise flow behaviour in the coating system and the model given above would need to be modified for the particular flow pattern and the change in that flow with coating velocity. However, it illustrates the general behaviour adequately.

REFERENCES

1. Adamson, A. W. (1990) *Physical Chemistry of Surfaces*, 5th edn, Wiley-Interscience, New York.
2. Blake, T. D. (1981), in *Surfactants* (ed. Th. F. Tadros), Academic Press, London, 1981, pp. 221–275.
3. Gibbs, J. W. (1928) *The Collected Works of J. Willard Gibbs*, Vol. **1**, Longman, New York.
4. Israelachvili, J. N. (1991) *Intermolecular and Surface Forces*, 2nd edn, Academic Press, San Diego.
5. Wenzel, R. N. (1936) *Ind. Eng. Chem.*, **28**, 988.
6. Girifalco, L. A. and Good, R. J. (1957) *J. Phys. Chem.*, **61**, 904.
7. Fowkes, F. M. (1964) *Ind. Eng. Chem.*, **56**, 40.
8. Zisman, W. A. (1964) *Adv. Chem. Ser.*, **43**, 1.
9. Hunter, R. J., (1987) *Foundations of Colloid Science*, Vol. **1**, Oxford University Press, Oxford.
10. Harkins, W. D. and Jura, G. (1944) *J. Am. Chem. Soc.* **66**, 1362.

7

Emulsions and Microemulsions

7.1 INTRODUCTION

Dispersions of one liquid in another immiscible liquid are of great importance in many applications, ranging through many diverse fields such as foods and pharmaceuticals to the oil industry. Many of the general features of colloidal dispersions are also common to emulsions, such as the stabilizing mechanisms that counteract the attraction due to dispersion forces and the tendency of large particles to settle or *cream* due to the effects of gravity. The fluid nature of both the particle and the interface results in special features that we do not observe with solid particles, although some of them are shared with the disperse phase in foams. Not everything that is commonly referred to as an emulsion is in fact a liquid-in-liquid dispersion. For example, the photographic 'emulsion' that is coated on a film matrix is a dispersion of solid particles (silver halide and dye crystals, depending on the film type) and has a fluid continuous phase when coated, but neither phase is a liquid when in the camera. Emulsion polymers used in decorative paints provide another example. The monomers are dispersed as an emulsion prior to polymerization, but once polymerized the particles are solid although their glass transition temperature is well below room temperature so that they can fuse to make a film. The polymer particles, however, never completely lose their individual identity and dissolve in one another.

Colloids and Interfaces with Surfactants and Polymers, Second Edition James Goodwin

Most common emulsions have a particle size distribution at the upper end of the colloid size range, that is, with a radius in the 1–10 μm range. The problem here is that in their formation a great deal of energy is required to create the large amount of new interface that is produced with particles with radii of one or two orders of magnitude smaller. By using a mixed surfactant system, say an ionic surfactant with a co-surfactant such as an alcohol, droplets with a radius of ~100 nm can be produced, and such systems are frequently referred to *miniemulsions*. Just like their larger brethren, these systems are inherently colloidally unstable. However, if the surfactant and co-surfactant contents are increased to high levels, a thermodynamically stable hetero-phase system may be formed with droplet radii in the 10 nm range. These systems are called *microemulsions* and are a completely different type of colloidal dispersion from the mini- and normal emulsions. Therefore, although both types are discussed in this chapter, they will be considered separately.

Many of the considerations that we apply to dispersions of one liquid in another apply to the dispersion of any fluid in a liquid; for example, dispersions of a gas in a liquid are foams and they are very similar to emulsions. Not all foams have the very high phase volumes that we are familiar with in the foams formed by the detergents used in the hand-washing of dishes or in shampoos. A low phase volume foam is termed a 'low-quality' foam and a high phase volume foam (say at $\varphi > 0.75$) is termed a 'high-quality' foam. When we are referring to a liquid disperse phase, we would use the terms 'emulsion' and 'cream' for the equivalent low and high concentrations, respectively. The compressibility of a gas phase is much greater than that of a liquid disperse phase and the solubility is also invariably much greater. Both of these factors provide additional layers of complexity. However, we will not deal with foams specifically in this text but should keep them in mind when we think of emulsions.

7.2 EMULSIFICATION

When we make an emulsion, we disperse one liquid phase in a finely divided form in another liquid phase which is immiscible with the first. Hence we may disperse water in an oil (a W/O emulsion), oil in an aqueous phase (an O/W emulsion) or one oil in another (an O/O emulsion). The last type is not very common but W/O and O/W emulsions are very widely used. It is also possible to turn one into the other – a process known as *phase inversion*. For example, milk is an O/W emulsion which

is phase inverted to give a W/O emulsion that we recognize as butter. However, the first question that we need to address is, 'how do we disperse on system in the other?'. During the emulsification process, we have to create a large amount of new interface. The free energy change is estimated from the product of the interfacial tension and the areal change. There is also a positive entropy change as we are dispersing one phase, so we may write

$$\Delta G_{\text{formation}} = \gamma_{12}\Delta A - T\Delta S \tag{7.1}$$

where γ_{12} refers to the interfacial tension of liquid 1 against liquid 2, with the latter being the continuous phase. We are not dispersing the disperse phase into very small drops, so the entropy change is small:

$$\gamma_{12}\Delta A \gg T\Delta S \tag{7.2}$$

Hence the free energy change is positive, informing us immediately that emulsification is not a spontaneous process and that we have to provide the energy input. In part, this may be in terms of heating, but it is mainly in terms of mechanical energy by using very high shear rates. Equation 7.1 tells us that the resulting droplets would be unstable, so that we must find a means of preventing them from reforming into two distinct phases. The free energy change only tells us about the difference in the initial and final states, so once formed we may provide a kinetic stability via the forces of repulsion that we employ with rigid particles. However, the details of the route are also important in the formation process. When shear forces are produced by intense flow fields, the droplets are progressively broken down from large macroscopic drops to small drops in the colloidal domain. In order to disrupt a droplet, a force has to be applied to it, hydrodynamically in the case we are considering here. The pressure gradient must be large enough to exceed that due to the Laplace pressure:

$$\frac{\mathrm{d}P}{\mathrm{d}r} \geq \frac{\Delta P}{a} = \frac{2\gamma_{12}}{a^2} \tag{7.3}$$

Of course, the area produced by dispersing unit volume of phase 1 in phase 2 is simply $3/a$ as the droplets are spherical and so we readily note by comparing Equation 7.1 with Equation 7.3 that the latter is where most of our work will come from.

The work required to break up the droplets may be applied in various ways. In a high-speed disperser or blender, large shear and elongational forces are applied to the droplets. The droplets elongate and break in the middle. As they separate, a small or 'satellite' drop is formed as the neck breaks. The newly formed drops must be rapidly stabilized as they tend to coalesce during subsequent collisions. When very high flow velocities are used, turbulent flow is usually produced. In turbulence, the flow is chaotic with localized eddies with high energy dissipation. The inertial forces are large and are the cause of the instabilities in the flow field and hence the density of continuous phase is important with higher density leading to turbulence at lower flow rate. The size of the local eddies can be defined in terms of the *Kolmogorov scale* (from Kolmogorov's analysis), so we may write the droplet size as [1]

$$a \sim \dot{E}^{-0.4} \gamma_{12}^{0.6} \rho_2^{-0.2} \tag{7.4}$$

where $\dot{E} = \mathrm{d}E/\mathrm{d}t$ is the rate of energy dissipation per unit volume. An ultrasonic probe may also be used to produce emulsions. In this case, the local cavitation is used to break up the droplets.

In all cases, however, the processes are not very efficient and there is normally significant heating [2]. This can aid the break-up as the timescale of the disruption process is also important. The viscosity of both phases decreases with increase in temperature and an important factor is the timescale of the extension of the droplets during break-up (slower with higher viscosity) compared with the timescale associated with the flow field (inversely proportional to the flow rate). The result is that larger drops are produced from disperse phases with higher viscosities. Of course, a higher viscosity of the continuous phase means higher shear forces on the drops and so faster deformation leads to a smaller drop size. However, the turbulence will not be as marked, the flow having a larger eddy size, and so the beneficial effect of increasing the continuous phase viscosity may by reduced.

7.2.1 Emulsion Type

The control of which phase is to be the disperse phase is of major importance. The primary controlling factors are the type of surfactant that is used and the temperature of the emulsification process, although the volume ratio and the viscosity ratio of the two phases can be important. The higher viscosity phase tends to be the continuous phase, with the

stability of the disperse phase being enhanced due to the slower drainage of the thin film produced as two drops collide. The choice of surfactant is normally the controlling factor. There may be restricted choices when the emulsion is intended for a specific use such as food or personal care products. However, there are some general guidelines which can be useful.

The simplest of these is known as *Bancroft's rule*, which states that the continuous phase should be the phase in which the surfactant is the most soluble. A good example with which to illustrate this is the water–cyclohexane system studied by Shinoda and co-workers [3, 4], in which a non-ionic surfactant (a nonylbenzene ethoxylate at 5 wt%) was used. We will consider what they observed at a phase ratio of 1:1. At room temperature, the surfactant is below its *cloud point* and most soluble in the water phase, with the ethylene oxide groups strongly hydrated and fitting into the H-bonding structure of the water. The emulsification process results in an O/W emulsion being formed. At 80 °C, the surfactant is more soluble in the cyclohexane, where it forms inverted micelles with the ethylene oxide groups oriented towards the centre of the micelles. This occurs as the hydration of the ethylene oxide moieties is reduced. An interesting question to address is, 'what happens when an emulsion is formed at, say, room temperature and it is then heated to 80 °C?'. The result is that the emulsion can invert and the temperature at which this occurs is the *phase inversion temperature* (PIT). In this case, the PIT was ~55 C and in this region the surfactant produced a concentrated lamellar surfactant phase as the solubility was not high in either phase, with some cyclohexane droplets and some water droplets present. The PITs occur at the cloud point of the non-ionic surfactants in water, whereas with ionic surfactants it can occur at the Krafft point. The phase-inversion can be followed by observing changes in the physical properties of the system, such as conductivity or viscosity.

Attempts have been made to classify surfactants numerically in terms of their chemical structure so that the selection of emulsifiers would be more straightforward. This was attempted first by Griffin [5] for non-ionic surfactants and the concepts extended by Davies and Rideal [6]. The *hydrophilic–lipophilic balance* (HLB) number is defined in terms of numerical values assigned to the chemical groupings in the surfactant, as follows:

$$\text{HLB} = 7 + \sum(\text{hydrophilic group numbers}) - \sum(\text{lipophilic group numbers}) \tag{7.5}$$

Table 7.1 HLB numbers for various chemical groups. Data from Davies and Rideal [6]

Type	Chemical group	Group number[a]
Lipophilic	—CH—	0.475
	=CH—	0.475
	—CH_2—	0.475
	—CH_3	0.475
Hydrophilic	—SO_4Na	38.7
	—COOK	21.1
	—COONa	19.1
	—SO_3Na	11
	=N—	9.4
	Ester (sorbitan ring)	6.8
	Ester (free)	2.4
	—COOH	2.1
	—OH (free)	1.9
	—O—	1.3
	—OH (sorbitan ring)	0.5

[a]Example of *derived* group numbers: ethylene oxide (—CH_2—CH_2—O—), = 1.3 − 2(0.475) = 0.35 − *hydrophilic*; propylene oxide [—CH_2—$CH(CH_3)$—O—] = 1.3 − 3(0.475) = −0.125 − *lipophilic*.

The group numbers assigned by Davies and Rideal [6] are given in Table 7.1. It is interesting that whereas the ethylene oxide group is hydrophilic, the propylene oxide group is hydrophobic. This is confirmed by the effectiveness of the ABA block copolymers PEO–PPO–PEO as effective stabilizers. Hence we may estimate the HLB number for our surfactant systems from Equation 7.5. For example:

- Sodium dodecyl sulfate, $CH_3(CH_2)_{11}OSO_3Na$, has an HLB number of 40.
- Dodecyl hexaethylene glycol monoether, $CH_3(CH_2)_{11}O(CH_2CH_2O)_5CH_2CH_2OH$ or $C_{12}E_6$, has an HLB number of 5.3.
- Glycerol monostearate, $HOCH_2CH(OH)CH_2OOC(CH_2)_{16}CH_3$, has an HLB number of 3.7.

When two surfactants are mixed together, the HLB of the mixture may be estimated from the HLB number of each component multiplied by the mass fraction of the component.

Griffin [5] used the HLB numbers to classify surfactants for particular uses and his classification is given in Table 7.2. Hence glycerol monostearate would be a suitable choice to produce a W/O emulsion and sodium dodecyl sulfate could be used to produce an O/W emulsion.

Table 7.2 Classification of surfactant applications by the HLB range of the surfactant, according to Griffin [5] (note that the original classification was for non-ionic systems)

Application	HLB range
W/O emulsifier	*3–6*
Wetting agent	7–9
O/W emulsifier	*8–18+*
Detergent	13–16
Solubilizer	15–18

The HLB classification is a useful initial guide, but it is only an indication. A better solution is to use the *Hildebrand solubility parameter*. This can be related to the van der Waals forces between the components [7]; however, the HLB numbers are widely used. It should also be noted that the HLB value for the non-ionic surfactants correlates well with the cloud point temperature of the surfactant and so in turn should provide a guide to the value of the PIT.

7.3 STABILITY OF EMULSIONS

The aggregation of emulsion droplets is a function of the same parameters as were discussed for the aggregation of solid particles. Hence it is important to know the value of the combined Hamaker constant, the magnitude of any electrical charge and the state of steric stabilizer layers. The fluid nature of the particle adds complexity as the interface can deform under the influence of both attractive forces and shear. This means that the factors such as the density of the packing of steric layers or charges are slightly variable but in addition that the location of the stabilizing moieties can also change rapidly as lateral diffusion in the stabilizer layer can be rapid.

There are further additional factors that we have to consider. Perhaps the most important of these is that aggregation is frequently not the final state. The coalescence of particles to form larger particles, and finally two distinct phases, leads to a larger decrease in the free energy than is the case with aggregation alone. Of course, aggregation does not always have to lead to coalescence and with stable dispersions of fluid particles we will have to consider the details of the trilayer film formed by the two droplet surface films separated by a thin film of the liquid disperse

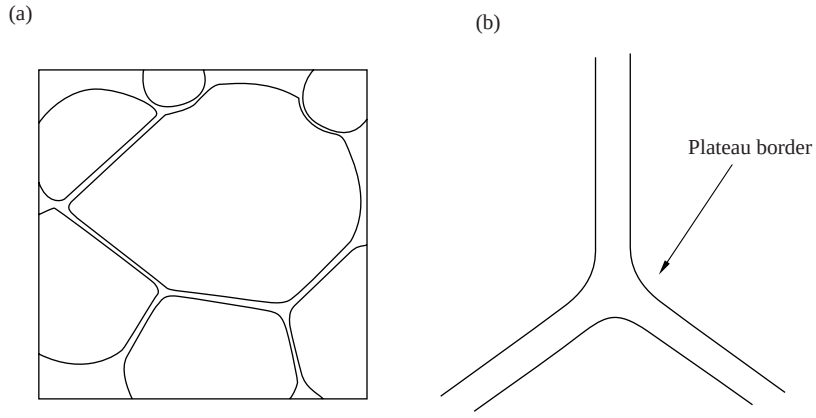


Figure 7.1 Schematics of (a) a section of a concentrated emulsion illustrating the polyhedral shapes of the droplets and (b) the junction between three droplets, showing the thin-film region with parallel faces and the region of varying curvature known as the *plateau border*

phase. Although two-body interactions can occur during collisions as the result of Brownian motion or collision due to a shear field, we will usually have coalescence occurring in the concentrated state. This may be produced by aggregation and/or by gravitational separation. The latter is just sedimentation or creaming if the disperse phase is less dense than the continuous phase. Monodisperse emulsions are very unusual and the wide particles size distribution of most emulsions means that droplets at $\varphi < 0.75$ are spherical, albeit with flattened interaction zones if in the aggregated state. Above this concentration we have creams and the droplets become increasingly deformed as the disperse phase volume is increased. The smaller droplets are more rigid than the larger droplets and the latter deform first, giving polyhedral shapes with some surfaces concave due to the presence of particles with smaller radii of curvature. This is illustrated in Figure 7.1a.

Many common emulsions have the majority of their mass of material in droplets with radii of curvature in the range 0.5–10 μm. It is relatively easy to deform droplets of these dimensions and so there are flat parallel areas between the particles. The interparticle forces acting across the thin films determine the mean separation distance. These are the dispersion forces, electrostatic forces and/or steric forces discussed in detail in Chapters 3 and 4. Such forces are formulated from static or equilibrium models, but when considering the stability to coalescence of emulsions, (and many of the points are also directly applicable to foams), the dynamics are a major factor. When two fluid droplets come together,

whether during a Brownian collision or a shear-induced collision or due to the gravitational force, the following processes occur. The rate of close approach is determined by the balance of those forces bringing the particles together, the interparticle forces and the viscous force from the fluid drainage, as we would expect for rigid particles. In addition, we have to consider the changes in the interaction region, that is, at close separations the interaction zones flatten and this increases both the strength of the interactions and the viscous resistance to the drainage as the potential contact area is increased. As the interaction zone grows, the area of the interface increases. The rapid stretching of the interface results in a locally higher interfacial tension and this opposes the stretching. The difference in the dynamic surface tension and the static value resists the expansion and acts as a driving force to move the adsorbed stabilizing material into the interaction zone. This flow also causes fluid to be drawn into the thinning film region by viscous drag and, in addition, opposes the drainage. These effects are grouped as the *Marangoni effect*. There is an additional effect, known as the *Gibbs effect*, which is governed by the decrease in the amount of stabilizer adsorbed at the expanded interface, which increases the stiffness (dilatational elastic modulus) of the film. In thin films, the local mass concentration of stabilizer can be too low to re-establish the initial equilibrium. These effects are usually grouped together and referred to as the *Gibbs–Marangoni effect*, and act to resist deformation of the interfaces in the thin-film region. However, we must also think in terms of local thermally (diffusion) driven 'ripple-like' disturbances. The Gibbs–Marangoni effect will act to damp these out, but they are a main mechanism by which coalescence occurs. To understand this, we must visualize the thin intervening film between two fluid particles where local wave effects cause local fluctuations in the separation. The stiffer the film, the smaller are the fluctuations. The force which maintains the film, known as the *disjoining pressure*, is the sum of the attractive and the repulsive forces. The elastic modulus is the rate of change of the force with distance and again is the sum of the attractive and repulsive terms. We may write this as follows:

$$\text{modulus} = \frac{\partial F_{\mathrm{R}}}{\partial H} - \frac{\partial F_{\mathrm{A}}}{\partial H} \tag{7.6}$$

It is clear from Equation 7.6 that if $\partial F_{\mathrm{R}}/\partial H < \partial F_{\mathrm{A}}/\partial H$ the stiffness decreases and the magnitude of the ripples (local thinning) increases and coalescence will occur.

We may summarize the factors that lead for a stable emulsion as follows:

1. The stabilizer on the surface of the droplets should produce sufficient repulsion to maintain a film of the continuous phase between droplet surfaces during collision or in a concentrated system such as a sediment/cream.
2. The stabilizer should be strongly adsorbed at the interface so that it will not desorb as droplets come together.
3. The dilatational modulus of the stabilizer layer is high so that the thermal disturbances are damped and do not grow to a magnitude comparable to the dimensions between adjacent drop surfaces.
4. The solubility of the disperse phase in the continuous phase should be as low as possible to minimize diffusive transport from one droplet to another in the process known as *Ostwald ripening*.

7.3.1 Ostwald Ripening of Emulsions

This is name given to the process whereby the droplet size distribution in an emulsion progressively shifts towards larger sizes. The origin of the effect is the Laplace pressure acting to increase the pressure inside the drop with decreasing radius. As a result of this pressure, molecules of dispersed phase diffuse from the high- to the low-pressure regions, that is, the small droplets dissolve and the larger droplets grow as the material is transferred by diffusion through the continuous phase.

The Laplace pressure is given by Equation 6.24 for droplets in a vapour and for a drop of radius r_1 in a liquid we have

$$P_{r_1} = P_{r=\infty} + \frac{2\gamma_{12}}{r_1} \tag{7.7}$$

where γ_{12} is the interfacial tension. Clearly, if there are two droplets with different radii, the difference in pressure between them is

$$P_{r_2} - P_{r_1} = 2\gamma_{12}\left(\frac{1}{r_2} - \frac{1}{r_1}\right) \tag{7.8}$$

and this acts to equalize the radii at a rate governed by the diffusion of the disperse phase through the continuous phase. In Chapter 6 we also discussed the Kelvin equation (Equation 6.40), which describes the

change in vapour pressure with drop radius. The analogous situation for a drop in a liquid (or a solid particle in a liquid, for that matter) is the solubility of the disperse phase in the continuous phase. Hence, if c_r is the concentration of the disperse phase in the continuous phase as a result of the applied pressure P_r, we can write the analogue of Equation 6.40 for the solubility of the disperse phase as

$$\ln\left(\frac{c_{r_1}}{c_{r=\infty}}\right) = \frac{\bar{V}_1}{RT}\frac{2\gamma_{12}}{r_1} \tag{7.9}$$

and for two drops we therefore have

$$\ln\left(\frac{c_{r_1}}{c_{r_2}}\right) = \frac{M_1}{\rho_1 RT}2\gamma_{12}\left(\frac{1}{r_2} - \frac{1}{r_1}\right) \tag{7.10}$$

where M_1 and ρ_1 are the molar mass and density of the disperse phase, respectively. Hence Ostwald ripening is most marked if the distribution of drop sizes is broad and the interfacial tension is high.

The data [8] shown in Figure 7.2 illustrate the relative growth in the mean particle size of an O/W emulsion after 10 days. The stabilizer was sodium dodecyl sulfate and the initial mean droplet size was ~1 μm. The figure also shows clearly how the relative growth rate follows the solubility of the alkane in water.

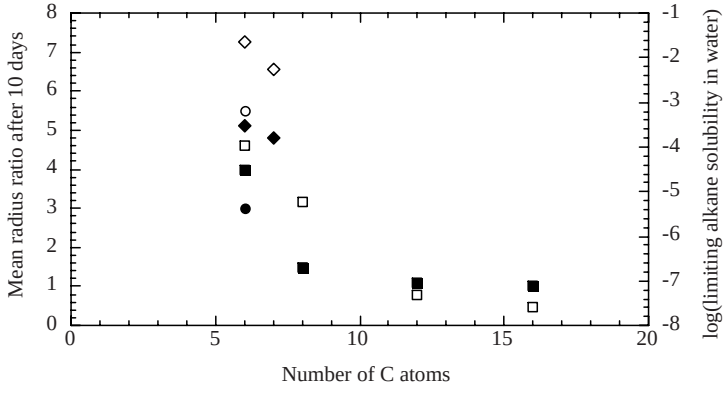


Figure 7.2 The Ostwald ripening behaviour of oil–water emulsions stabilized with sodium dodecyl sulfate, from data reported by Winsor [8]: ■, □, *n*-alkanes; ●, ○, cyclohexane; ♦, ◇, aromatic alkanes (where the solid symbols refer to data for the radius ratios and the open symbols refer to the alkane solubilities)

The prevention of Ostwald ripening is usually desirable as ageing a formulation is not an attractive proposition. The solubility of the disperse phase in the continuous phase is a factor which affects the rate rather than the ultimate state. The interfacial tension may be manipulated somewhat, but this is also likely to change the initial size distribution. The addition of small amounts of a third component, which is soluble in the disperse phase but has an extremely low solubility in the continuous phase, is an effective strategy to limit the size drift. In this case, the concentration of the third component is the same in all the particles after emulsification. As ripening commences, the concentration in the small particles rises and that in the large particles decreases, and this opposes and ultimately limits the ripening process. (The situation in the droplets is directly analogous to an osmometer, where the continuous phase is acting as the semi-permeable membrane.) Now, if the surfactant has a very low solubility in the continuous phase while being sufficiently surface active to act as a stabilizer, of course, it will act as this third component. The reduction in the specific surface area that occurs as the ripening proceeds increases the concentration of surfactant dissolved in the small droplets. It is a common practice to dissolve surfactants in what will become the disperse phase.

7.4 MICROEMULSIONS

The interfacial tension is an important parameter in the control of droplet size in emulsification. This may be reduced by the addition of co-surfactants such as long-chain alcohols. Droplet sizes of ~100 nm may be produced in this way and the resulting emulsions are referred to as 'miniemulsions'. These differ little from emulsions with larger drop sizes in terms of stability, although the rate of creaming/sedimentation is reduced and this may no longer be a problem in a formulated product. If interfacial tensions are reduced to very low levels, however, a further decrease in drop size down to ~10 nm can occur. The systems are produced at moderate to high levels of surfactant and are transparent due to the small size of the disperse phase. The viscosity is usually low, unlike liquid crystal phases, and the stability is different from what we regard as 'normal emulsions'. Such systems are a thermodynamically stable phase and are termed *microemulsions*. Formation is spontaneous, requiring little or no input of mechanical energy beyond the gentle mixing of the components. This inherent stability means that the systems are completely different from 'normal emulsions' and have very little in

common with them except that they are colloidal systems consisting of oil, water and surfactants.

Microemulsions have been used in many formulations [9, 10]. Cleaning systems range from dry-cleaning fluids to self-polishing floor and car waxes. The hard wax (carnauba wax or synthetic polymers) is in the oil phase and when the coating dries the wax particles form the film. As the dimensions are so much smaller than the wavelength of light, the surface asperities are too small to result in individual reflections. Other examples of applications are in cutting oils, pesticides and flavours for foods. There are some examples of chemical reactions being carried out in microemulsions, the most successful being the polymerization of acrylamides to give very high molecular weight products [11]. Interest in the use of microemulsions in tertiary oil recovery increases and decreases in parallel with oil prices. The attraction is the spontaneous formation of the microemulsion so that oil trapped in the pores of depleted oil wells (there can be 75% of the original amount remaining) can be flushed out. The large amounts of surfactant required is the cost-limiting factor. Microemulsion fuels using vegetable oils from sustainable sources have also be formulated [10].

7.4.1 Phase Behaviour

The phase diagrams are fairly complex as we have at least three components (and often four), namely the oil, water (electrolyte solution) and surfactant (and often a co-surfactant). We must also remember that temperature and pressure will be important in many situations. A diagrammatic representation of a six-component phase map is not possible – just four require a tetrahedron – but the usual practice is to reduce the map to a three-component form. This may be by working at a constant surfactant–co-surfactant mix in addition to constant temperature and pressure. Alternatively, the surfactant concentration in the electrolyte solution can be kept constant and the co-surfactant concentration can be the third variable. This enables us to utilize the conventional triangular phase diagram plot. Temperature variations are then slices across a parallel-sided triangular prism. The plotting of the phase diagrams is a lengthy endeavour as many samples are required in order to define the boundaries precisely. The process may be speeded up using a titration technique [8] in which two of the components are adjusted sequentially. A schematic of a typical phase map is given in Figure 7.3.

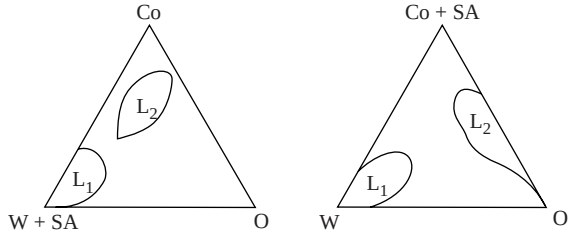


Figure 7.3 Schematic ternary phase diagrams showing the oil-in-water microemulsion region, L_1, and the water-in-oil microemulsion region, L_2: O, oil; W, water; SA, surfactant; Co, co-surfactant

A wide range of surfactants can produce microemulsions. Anionic surfactants, such as sodium dodecyl sulfate and potassium oleate, require co-surfactants such as aliphatic alcohols of shorter chain length. As an alternative to the alcohols, some non-ionic surfactants can be used in which a poly(ethylene glycol) moiety replaces the simple alcohol, although it should be noted that the properties are slightly different [10]. Some di-chain ionic surfactants, such as Aerosol OT [sodium di(2-ethylhexyl)sulfosuccinate] and didocecylammonium bromide, do not require a co-surfactant to form microemulsions with oil and water. The same is true for some non-ionic surfactants. This occurs in the phase inversion boundary region between W/O and O/W systems.

Figure 7.4 shows schematically the type of microemulsion systems that were classified by Winsor [8], where we see microemulsion phases in equilibrium with excess water or oil. The composition of the three-component system is shown as the intersection of the broken lines. The ends of the 'tie-lines' gives the composition of the two phases, while the 'lever rule' gives the amounts of each. Any change to the system that changes the hydrophobic balance of the surfactant may enable us to move from one type of system to the other. For example, at a constant temperature the change of surfactant to one with a longer chain length would take us from left to right in the figure. A similar result can be achieved by changing the polarity of the water phase (by the addition of short-chain alcohols, for example), increasing the temperature for a non-ionic surfactant or increasing the electrolyte concentration for an ionic species. It is the same process as described earlier for the phase inversion temperature for the non-ionic system studied by Shinoda and Kunieda [4] where the H-bonding in the water is reduced. This decreases the solubility of the hydrophile moiety and increases the solubility of the hydrophobe. The head group of an ionic surfactant becomes less soluble

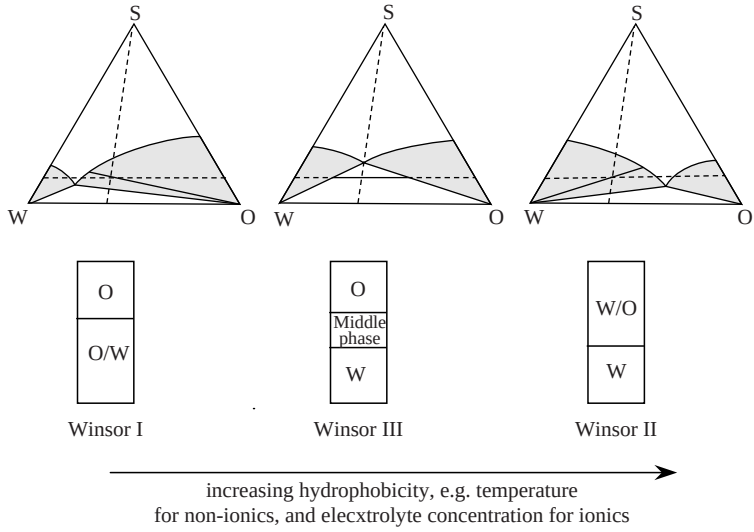


Figure 7.4 Schematic phase diagrams of various types of microemulsion systems, as classified by Winsor [8]: O, oil; W, water; S surfactant. The overall compositions are given by the intersections of the dashed lines, the shaded areas represent the microemulsion plus the pure component regions, and the ratios of the microemulsion phases and the pure liquid phases are given by the 'tie-lines' and the 'lever rule'

with increasing added electrolyte. This is readily observed as the cmc decreases with added salt as the ion binding is increased and the mutual repulsion decreased (i.e. we are tending to 'salt out' the molecule). Hence, just as for macroemulsions, we have the continuous phase as the one in which the surfactant is most soluble, as described by Bancroft's rule. The curvature of the surface is governed by the packing of the molecules at the interface and this, of course, includes solvation and electrostatic interactions. For the formation of such small droplets as we find in microemulsions (5–50 nm), the interfacial tension has to be 'ultra-low', with typical values being in the range 10^{-4}–10^{-2} mN m^{-1}.

The Winsor III system (see Figure 7.4) is particularly interesting. The concentration of the system is outside the two-phase regions marked in the phase diagram and is in a multi-phase region. The HLB of the surfactant is such that we have equal solubility and no preferred curvature. This means that the interfacial tension is not just 'ultra-low' but is approaching zero. The rigidity (elasticity) of the interface is therefore low and the interface readily deformed by thermal fluctuations so that curvatures in favour of both the water and the oil directions are

probable. A 'bicontinuous' or 'sponge' microemulsion phase is produced, with the domain size typical of the microemulsion size domain, that is, ~10 nm. If the rigidity of the film is increased, by the use of long-chain alcohols as co-surfactants, for example, the bicontinuous middle phase has a lamellar structure with oil swelling the hydrophobe region.

7.4.2 Characterization of Microemulsions

The first problem that we must solve is 'which type of system do we have?'. Conductivity provides an easy check here. If the conductivity is low, the system is a W/O microemulsion. Conversely, if it is high, then it is an O/W system. A bicontinuous phase would also have a high conductivity and we would need other information to separate this from the O/W microemulsion.

Transmission electron microscopy, coupled with a freeze–fracture technique, is especially useful in studying the bicontinuous phase [12]. In this, we use liquid nitrogen to cool a small sample so rapidly that the water is turned to 'amorphous ice'. If we were to put the sample straight into liquid nitrogen, a gaseous layer round the sample acts as an insulator – therefore, we place the sample in a volume of liquid propane that is held at liquid nitrogen temperature. The cooled sample is placed in a high vacuum, fractured and then coated with tungsten to give a shadow – a replica is then formed from a platinum–carbon coating, which can be viewed in an electron microscope.

Dynamic light scattering can be used to follow the diffusive motion of the individual droplets. However, the samples are relatively concentrated and the diffusion is slowed by particle–particle interactions. NMR spectroscopy can be used to determine the molecular diffusion – in this case, a technique known as pulsed-gradient spin-echo NMR is used.

Small-angle neutron scattering is particularly powerful as the use of D_2O with deuterated oils and surfactants and/or co-surfactants enables detailed information on the droplet structures to be deduced. Small-angle X-ray scattering may be useful but it is not possible to vary the contrast to bring out detail with this technique, as is the case with neutron scattering. In addition, the scattering is much weaker with X-rays as the elements are all low atomic number elements and the X-rays are scattered by the electrons. Neutrons are scattered by the atomic nuclei and give stronger scattering.

7.4.3 Stability of Microemulsions

It was stated above that microemulsions are a *thermodynamically stable* phase and are therefore unlike macroemulsions. The modelling of these systems is a more recent development than that of the stability of other colloidal systems and is still not fully complete. Most practical systems contain many components and usually these will contain impurities that will also be surface active. Nevertheless, it is instructive to consider the main ideas involved in modelling the stability, although we will not follow the fine detail of the analysis.

The foundations were laid in the work of Overbeek [13] and Ruckenstein and Krishnan [14]. They discussed the free energy change in the formation of a microemulsion consisting of droplets, and we will not broaden the discussion to the details of the bi-continuous or 'sponge' phase. Following the above authors, we may split the free energy of formation into several components:

$$\Delta G_{\text{disp}} = \Delta G_{\text{I}} + \Delta G_{\text{E}} - T\Delta S_{\text{disp}} \tag{7.11}$$

where ΔG_{I} is the free energy associated with creating the interfacial area of the droplets. The electrostatic components associated with interactions produced by charged surfactant molecules is the ΔG_{E} term and are a function of the curvature, that is, droplet radius, and also of the relative interfacial concentrations of the charged surfactant and any uncharged co-surfactant. The entropic term is the contribution of the fine droplets in a random dispersion. To obtain a thermodynamically stable system, the free energy of the microemulsion phase must be lower than that of the original mixture of macroscopic phases. Hence the value of ΔG_{disp} must be negative. To be rather more specific, we can write

$$\Delta G_{\text{I}} \propto \gamma_{\text{ow}} A \tag{7.12}$$

where A is the area of the interface that has been created with an interfacial tension γ_{ow}, and ΔG_{I} is positive. As the droplets are small, the value of A is necessarily large and so the interfacial tension has to be 'ultra-low'. The electrical work needed to get two drops to their mean position is calculated from the electrostatic potential multiplied by the charge:

$$\Delta G_{\text{E}} \propto \int \psi \, \mathrm{d}\sigma \tag{7.13}$$

This again is positive as particles of similar charge repel, and confining many charged particles into the space available to the macroscopic phases will require the input of energy. The entropic term for n particles occupying a space is given by nk_BT, so

$$T\Delta S_{\text{disp}} \propto k_B T \ln \frac{\varphi}{\varphi_m} \tag{7.14}$$

where the volume fraction φ is that of the disperse phase, and the maximum volume fraction that could be occupied by hard spheres is φ_m. The droplets in a microemulsion are spherical and monodisperse, and in a static system we could imagine that we might pack them together to a face-centred cubic or hexagonal close-packed structure which would give a maximum volume fraction of 0.74. The dense random packing of hard spheres is lower, at a maximum volume fraction of $\sim$0.62. The thermal motion of the droplets means that each has a larger excluded volume than we might predict for a static system and a phase transition occurs at a volume fraction of >0.5 and even lower when particles are charged (this point is discussed in more detail in Chapter 9). We can see immediately from Equation 7.14 that for $\varphi < \varphi_m$ the entropic term is negative and, for a stable microemulsion system to be formed, this term must be larger than those given by Equations 7.12 and 7.13. The work carried out by Ruckenstein and Krishnan [14] included adsorption isotherms of the surfactant and co-surfactant at the interface, and this modifies both the interfacial tension and charge terms.

Although the discussion has referred to 'droplets' and 'hard spheres', it is important to keep in mind that the radius of a microemulsion droplet is of the same order of magnitude as the molecular length of the surfactants used, noting that any co-surfactants are invariably shorter molecules. Hence, as the droplet size is on the same scale as the molecules, the interfacial structure is much less sharp, with significant penetration of the two phases into the interfacial region. The molecular shapes and the packing together of these shapes in a curved surface [8, 10, 15] are a critical consideration. In an O/W microemulsion, the surfactant chains are on the inside of the droplet interfacial region and are much more crowded than in a W/O microemulsion formed from the same surfactant system. When we refer to the diameter of a droplet, this is just some average value and different methods of determination will give different values. It is instructive to follow the estimate of water core size given by Overbeek *et al*. [15] for the water (300 mM NaCl)–cyclohexane–sodium dodecyl sulfate (SDS)-co-20% pentanol system. These workers used the

limiting slope of the interfacial tension versus log[surfactant] curve to give the surface excess from the Gibbs adsorption equation (see Chapter 1). Their results gave one SDS molecule for every 0.9 nm^2, so that each SDS molecule would have three pentanol molecules 'associated' with it in the macroscopic interface, which is, of course, flat. Assuming that this molecular area is also that at the curved surface of the water core of the droplet, we can calculate the water droplet surface area, A_D (in m^2), for the number of moles of surfactant in the system, n_{SDS}, as follows:

$$A_D = 9 \times 10^{-19} (n_{SDS} N_A) \approx 5.4 \times 10^5 n_{SDS} \tag{7.15}$$

As the molar volume of water is $1.8 \times 10^{-5}\,kg\,m^{-3}$, the volume of water in the system is

$$V_{H_2O} = 1.8 \times 10^{-5} n_{H_2O} \tag{7.16}$$

Recalling that for a sphere $V/A = \frac{\pi}{6} D^3/\pi D^2 = D/6$, we can use Equations 7.15 and 7.16 to give the approximate diameter (in nanometres) of the water core as

$$D_{H_2O} = \frac{6 V_{H_2O}}{A_{H_2O}} \approx 0.2 \left(\frac{n_{H_2O}}{n_{SDS}} \right) \tag{7.17}$$

We can make a similar estimate for the O/W case, although in this instance the SDS:pentanol ratio is 1:2 [10, 15]. The estimates of droplet size are fairly close to those measured [15].

The measurement of the limiting slope of the γ_{OW} versus $\log(c_{SA})$ curve is not easy as the interfacial tensions are very low. A useful technique to use is the 'spinning drop' method. In this case, the denser phase (usually water) is placed in a capillary tube. The tube is rotated at high speed and a small amount of the less dense phase (usually the oil) is added. The centrifugal forces keep the droplet centred in the tube. These forces also elongate the drop. The axial ratio is measured using a travelling microscope and its value is limited by the value of the interfacial tension and the angular velocity of the capillary. This technique is the only one readily available for the measurement of 'ultra-low' interfacial tensions.

REFERENCES

1. Dickinson, E. (1992) *An Introduction to Food Colloids*, Oxford University Press, Oxford.

2. Walstra, P. (1983) Formation of emulsions, in *Encyclopaedia of Emulsion Technology*, Vol. **II** (ed. P. Becher), Marcel Dekker, New York, pp. 57–77.
3. Saito, H. and Shinoda, K. (1970) *J. Colloid Interface Sci.*, **32**, 647.
4. Shinoda, K. and Kunieda, H. (1973) *J. Colloid Interface Sci.*, **42**, 381.
5. Griffin, W. C. (1949) *J. Soc. Cosmet. Chem.*, **1**, 311.
6. Davies, J. T. and Rideal, E. K. (1963) *Interfacial Phenomena*, Academic Press, London.
7. Israelachvili, J. (1991) *Intermolecular and Surface Forces*, Academic Press, London.
8. Winsor, P. A. (1968) *Chem. Rev.*, **68**, 1.
9. Prince, L. M. (1977) *Microemulsions*, Academic Press, New York.
10. Hunter, R. J. (1989) *Foundations of Colloid Science*, Vol. **II**, Oxford University Press, Oxford.
11. Candau, F., Leong, Y. S. and Fitch, R. M. (1984) *J. Polym. Sci., Polymer Chem. Ed.*, **101**, 167.
12. Jahn, W. and Strey, R. (1988) *J. Phys. Chem.*, **92**, 2294.
13. Overbeek, J. Th. G. (1978) *Discuss. Faraday Soc.*, **65**, 7.
14. Ruckenstein, E. and Krishnan, R. (1980) *J. Colloid Interface Sci.*, **76**, 201.
15. Overbeek, J. Th. G., de Bruyn, P. L. and Verhoeckx, F. (1984) Microemulsions, in *Surfactants* (ed. Th. F. Tadros), Academic Press, London, pp. 111–132.

8

Characterization of Colloidal Particles

8.1 INTRODUCTION

It is a prerequisite to understanding the way in which our colloidal systems behave to be able to know the details of the particles in our dispersions such as the particle size, the size distribution, the shape, the charge and the dimensions of any adsorbed layers. There is a wide range of methods available and we shall review some of them in this chapter. In many situations, we may be restricted by which ones can be applied, and also by only some of these techniques being readily available. However, it is always true that the better our system is characterized, the better our understanding and the easier it is to manipulate the formulation to optimize the material.

In this chapter, we discuss the use of microscopy in terms of size and shape distributions. Following this, we address the scattering of radiation by colloidal particles. This focuses on particle size and structure, although it sets the scene for the discussion of concentrated systems in the next chapter. Characterization of the electrical properties is discussed next, and finally the viscosity of dispersions is addressed, as this is a function of the detailed nature of the particles.

Colloids and Interfaces with Surfactants and Polymers, Second Edition James Goodwin

8.2 PARTICLE SIZE

The first question that we ask when we discuss a colloidal dispersion is, 'What is the particle size?'. Of course, we mean the mean size and we would also like to know something about the size and shape distributions and also the specific surface area, but the latter information is often seen as of less importance than 'the size'. Before methods are discussed. it is important to give some basic definitions.

An example size distribution is shown in Figure 8.1 as a histogram. The number-average diameter, $\bar{D}_n$, is calculated in the same way as we calculated the average molecular weight for polymers in Chapter 2:

$$\bar{D}_n = \frac{\sum n_i D_i}{\sum n_i} = \sum f_i D_i \tag{8.1}$$

where n_i is the number and f_i is the fraction of particles that are found in the size class D_i. We group similar sizes together to plot them as a histogram. Experimentally, of course, this grouping might conveniently correspond to the multiple of resolution of the instrumentation. The standard deviation, σ_s and the coefficient of variation, cv, are now

$$\sigma_s = \left[\sum f_i \left(D_i - \bar{D}_n\right)\right]^{0.5}; \quad \text{cv} = \frac{100\sigma_s}{\bar{D}_n} \tag{8.2}$$

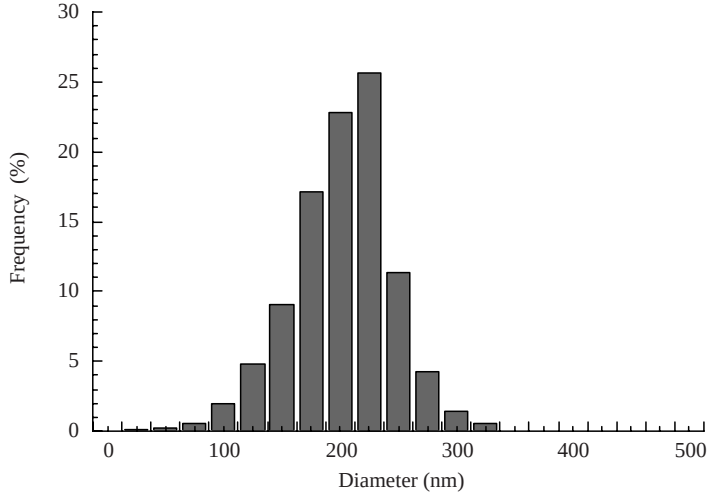


Figure 8.1 Example of a particle size distribution presented as a histogram: number-average diameter = 201 nm; standard deviation = 19 nm; coefficient of variation = 9.4%

We can calculate the various moments of the distribution, M_j, by using the following equation [1]:

$$M_j = \sum f_i D_i^j \tag{8.3}$$

Hence the first moment is the number average given in Equation 8.1 and the weight average is the ratio of the fourth to the third moment:

$$\frac{M_4}{M_3} = \frac{\sum D_i f_i D_i^3}{\sum f_i D_i^3} \quad \text{as} \quad w_i = \frac{\pi}{6} \rho f_i D_i^3 \tag{8.4}$$

where w_i is the mass of particles with a diameter D_i and density ρ. In the example distribution plotted in Figures 8.1 and 8.2, the number-average diameter is 201 nm with a standard deviation of 19 nm, which is equal to a coefficient of variation of 9.4%. The weight-average diameter is 226 nm. Here we can see how the average changes as we use different weighting. We must keep in mind that size determinations that use different physical properties of the particles can correspond to different averages so quite different sizes may be observed, especially if the size distribution is broad.

The shape of distribution is indicative of how the system was produced. The example illustrated above has a skew showing a tail towards smaller sizes. This is typical of the distribution that we find from a dispersion which is prepared by a particle nucleation and growth procedure.

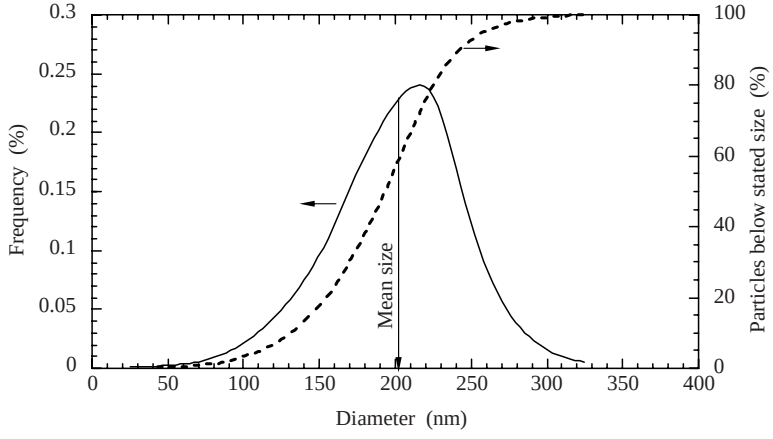


Figure 8.2 Frequency and cumulative distribution curves for a (colloidal) particle: number-average diameter =201 nm; standard deviation = 19 nm; coefficient of variation = 9.4%

The tail indicates that new particles are formed until late in the process. If the system has been prepared by comminution, the tail is usually towards the larger sizes as we progressively reduce the size down to a limiting threshold. Above we have talked about the moment of the distribution about zero with the first moment, M_1, being the number average or mean. We can make use of this concept of the moments of the distribution if we use the mean as the reference point. The skew of the distribution can now be expressed formally as the third moment about the mean:

$$M_{m3} = \sum_i f_i \left(D_i - \bar{D}_n\right)^3 \tag{8.5}$$

where the subscript m indicates the moment about the mean. The skew of the distribution shown in Figure 8.2 is −234, the negative value indicating that the asymmetry is the tail on the left-hand side. A symmetrical distribution has zero skew. The second moment is the 'variance', that is, the square of the standard deviation.

It is conventional to refer to a system being 'monodisperse' if the coefficient of variation on the mean diameter is $\leq$10%. When viewing a size distribution, especially when presented on a logarithmic scale, it is important to keep in mind that a small size fraction can dominate the number and area of the system. In addition, it should be noted that the large size fraction of the distribution contains most of the mass of the disperse phase.

There is a wide range of techniques available for the determination of particle size. A summary of the more common methods is given in Figure 8.3. Some of the methods give the full distribution, whereas others provide just a mean value. The average is not always the number average and the measure of the distribution width is often determined by an algorithm which is embedded in the instrument software, the details of which are not available to the user.

8.3 MICROSCOPY

8.3.1 Optical Microscopy

The optical microscope is a valuable tool in any colloid laboratory. With modern optics, we can see much of the colloidal size range, although we are still restricted to the larger end of the range for size data. The state

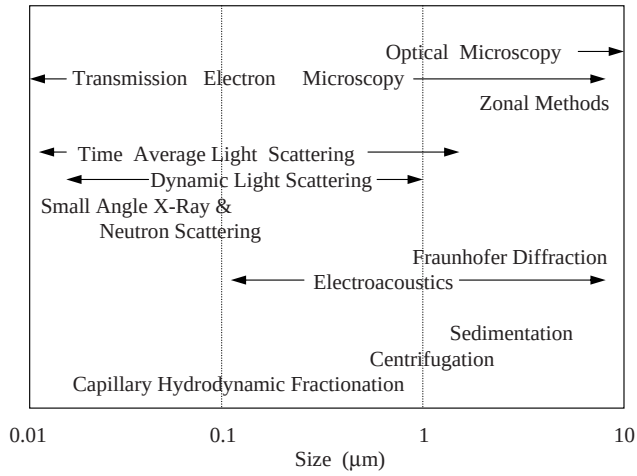


Figure 8.3 Summary of techniques available for the measurement of particle size

of the dispersion is immediately apparent as it is observed in the 'unperturbed' state. Aerosols are an exception here as the motion is so rapid that high-speed photography would be required. With particles dispersed in a liquid, we can see immediately if the system is well dispersed, coagulated or weakly flocculated. In the last case, particles can be seen to move together for some time and then to separate. The Brownian motion of particles dispersed in any liquid presents problems with resolution if an accurate size is required. Dried samples may solve this problem, but often the study of the particles in the wet state is important. A good photographic recording system is always required to optimize the method whether normal transmission, dark field or fluorescence microscopy is employed.

Once an image has been obtained, it has to be analysed manually or usually by a computerized image analysis system. If the particles are spherical, the problem of particle size is straightforward – we simply define the diameter as the size. However, for non-spherical particles the problem is not so easy, as illustrated in Figure 8.4. The particle is sketched as a platelet which would be similar to a kaolinite clay particle. The clay crystal has surface topographical detail which is lost in the projected image. The various dimensions are also defined in the figure (the dimensions shown in Figure 8.4f and g would be obtained by techniques other that microscopy, of course, but are included here to illustrate the difficulties often encountered when 'sizes' are compared for

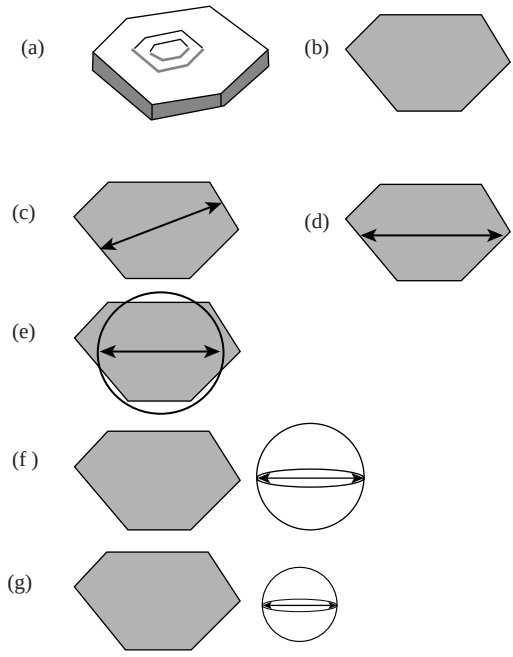


Figure 8.4 Illustration of the determination/definition of particle size for a non-spherical particle, in this case a platelet: (a) three-dimensional 'image', showing the surface topography; (b) projected 'image', which loses the surface topography; (c) perpendicular distance between tangents of opposite sides – Feret's diameter, d_F; (d) length bisecting the particle – Martin's diameter, d_M; (e) diameter of a circle of equal projected area, d_a; (f) diameter of a sphere with the same surface area (equivalent spherical area diameter), d_s; (g) diameter of a sphere of the same volume (equivalent spherical volume diameter), d_v

the same material. Calibration requires an image of a stage micrometer, the best precision being ~3%.

8.3.2 Optical Resolution

Although we can use lenses with higher magnification to increase the apparent size of the objects that we are studying, the resolution of the objects is a function of both the wavelength of light that we are using and the amount of light that is collected from the specimen. The point resolution is the minimum distance between two objects that can be distinguished as separated. Modern optics minimize losses in the optical

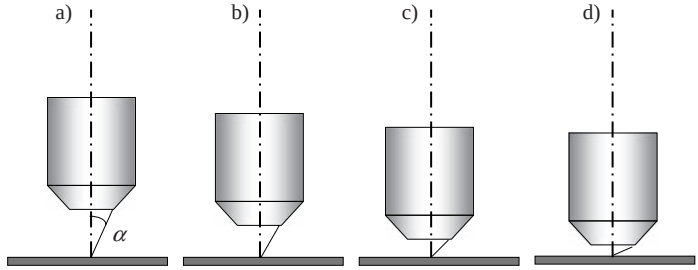


Figure 8.5 The numerical aperture, *NA*, of a dry objective related to the semi-cone angle of the collected light. (a) magnification = 10×, $\alpha = 15^\circ$, $NA = 0.25$; (b) magnification = 20×, $\alpha = 20^\circ$, $NA = 0.34$; (c) magnification = 40×, $\alpha = 45^\circ$, $NA = 0.77$; (d) magnification = 100×, $\alpha = 60^\circ$, $NA = 0.86$

train but we need to maximize the light collected by the objective lens. Light coming from an object in the focal plane goes through the coverslip, then through air until it reaches the objective lens. The maximum possible collection angle in dry air is 90° either side of the normal to the lens surface. This objective would have a numerical aperture, *NA*, of 1. The focal length of the lens, f, and the radius of the objective limit the collection angle, α_c. Figure 8.5 illustrates the relationship between *NA* and the focal length. The numerical aperture is calculated from

$$NA = n_r \sin \alpha_c \tag{8.6}$$

where n_r is the refractive index of the medium between the coverslip and the objective.

In practice, light goes through the sample media, a coverslip and then the air and is collected over all angles less than the *critical angle*, α_{crit}, at the coverslip glass–air interface. By using an oil with the same refractive index as the glass ($n_r \rightarrow 1.5$), more light is collected. Of course, there is some refraction at the sample–coverslip interface, so the $NA \leq 1.5$.

In order to understand the resolution, we need to examine what happens when light is diffracted by a small object of diameter d_a (i.e. from a point source) in its path. The diffracted light is spread radially over a small forward angle, θ, giving an image formed at an image plane at a distance x from the aperture. The intensity falls rapidly with angle in an oscillating manner (the radial intensity is described by a first-order Bessel function [2]), with the second peak being <2% of that of the main peak. The geometry is shown in Figure 8.6 along with the decrease in

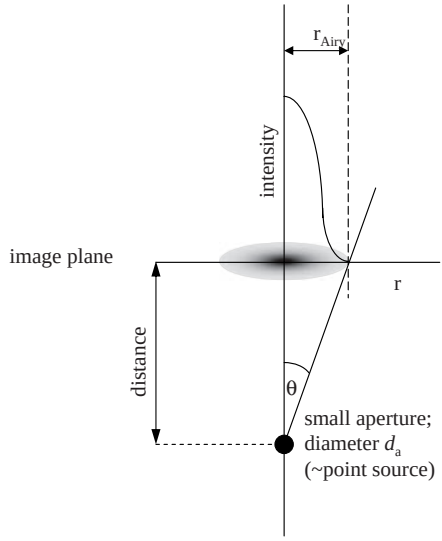


Figure 8.6 The radial intensity of light diffracted from a small (point-like) object. Only the first peak is shown. The radius of this peak defines the bright central disc of the diffraction pattern which is known as the Airy disc [2]

intensity of the zero-order diffraction peak. If the radius of the image of the diffraction rings at the objective image plane is r_D, then for small angles

$$r_D = x \tan\theta \simeq x \sin\theta \tag{8.7}$$

and the intensity, I, is

$$I = f\left(\pi \frac{d_a \sin\theta}{\lambda}\right) = f(\pi m) \tag{8.8}$$

Table 8.1 gives the positions of the minima and maxima of the diffraction peaks.

Figure 8.6 illustrates the illuminated disc at the image plane of the objective lens formed by the main diffraction peak. This disc is called the 'Airy disc' and the radius is

$$r_{Airy} = 1.22\lambda \left(\frac{\text{focal length}}{\text{lens diameter}}\right) = 1.22\lambda \left(\frac{1}{2NA}\right) \tag{8.9}$$

Table 8.1 The positions of the minima and maxima of the diffraction peaks from a point aperture

Peak Order	Minimum	Maximum	Relative Intensity
0	–	0	1
1	1.22π	1.64π	0.175
2	2.23π	2.68π	0.0042
3	3.24π	3.09π	0.00078

The factor 1.22 is calculated from the position of the first diffraction minimum as this defines the radius of the Airy disc that we see. The resolution of our microscope is the minimum distance that we can detect between two points, that is, small objects. The 'Rayleigh criterion' is this minimum distance and is given by the diameter of the Airy disc, that is, when the centres of the two diffraction peaks are $2r_{\text{Airy}}$ apart, the peak positions are still discernible as there is a small decrease in intensity between them. The decrease is only ~26% and with slightly more overlap it will be markedly smaller. Therefore, Equation 8.9 gives us the resolution of our microscope and we can rewrite this in terms of the minimum resolvable distance, d_{min}:

$$d_{\text{min}} = \frac{0.61\lambda}{NA} \tag{8.10}$$

When we close the condenser aperture down, NA_{con} can become less than that of the objective lens and we must take the average value as in the equation

$$d_{\text{min}} = \frac{1.22\lambda}{NA + NA_{\text{con}}} \tag{8.11}$$

In the special case where the condenser aperture is closed down, we must substitute the sum of the numerical aperture of the objective and that of the condenser for the value of the objective alone in Equation 8.10 [2]. Figure 8.7 illustrates how the resolution increases with numerical aperture for dry lenses using a wavelength of 550 nm, which is in the middle of the visible region.

Other factors that vary with the magnifying power of a microscope are the depth of field and the working distance. The depth of field is an important factor both when looking for fine details in a larger structure and for the quantification of information (in particle counting, for example).

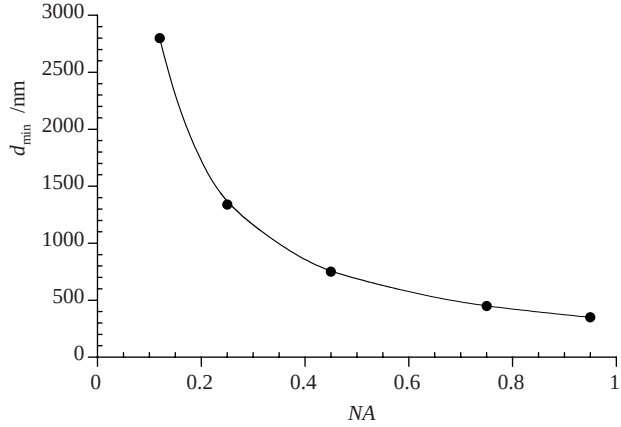


Figure 8.7 The variation of the optical resolution as a function of the numerical aperture of the objective lens

The depth of field, *DoF*, is the thickness of the specimen in which all the objects are in focus simultaneously and is given by [2]

$$DoF = \frac{n_1 \lambda}{(NA)^2} \tag{8.12}$$

For dry objectives, $n_r = 1$, it is a multiple of the wavelength of light and decreases rapidly with increase in the numerical aperture of the objective. The working distance of the lens decreases with increase in the magnification of the lens, as illustrated pictorially in Figure 8.5 and graphically in Figure 8.8. The small values of the working distance of the

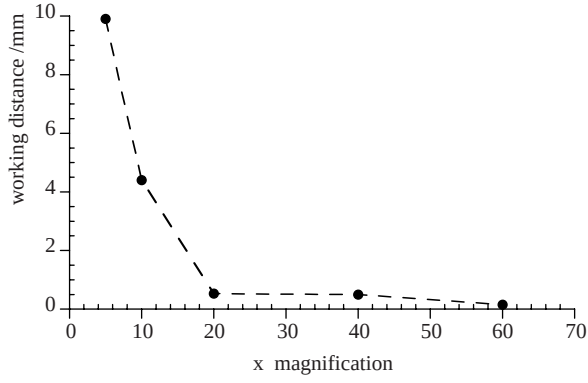
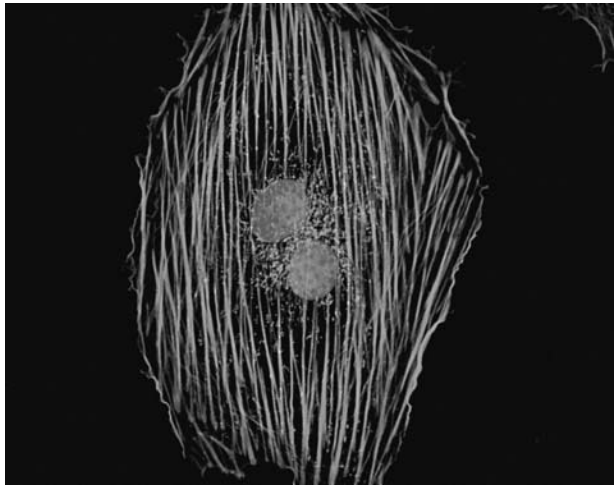


Figure 8.8 The working distance of a dry objective lens as a function of the magnification of the lens

higher magnification objectives do not represent a problem in microscopy if slides with coverslips are being used. However, the use of special optical cells often requires long working distance objectives.

8.3.3 Fluorescence Microscopy

This is a branch of microscopy that is widely used in the life sciences and it has become an extremely powerful tool for studying cells and tissues. A fluorescent molecule or nanoparticle absorbs light at one frequency and then emits it at a lower frequency. For fluorescence to occur with molecules, light at a specific narrow band of wavelengths is required. The electrons are raised to an excited state and, after a set amount of time – the *fluorescence lifetime* – they drop back to a lower level by emitting light at a longer wavelength as some of the energy has been lost from the excited state. The displacement between the peak absorption and the peak emission is called the *Stokes shift* and Figure 8.9 is a schematic of the absorption and emission peaks. The fine structure on the curves indicates the energy levels in the ground and excited states. Note that the curves are mirror images of each other. This shift in wavelength means that the experiments can be made much more sensitive by using filter sets that only pass the emitted wavelengths, so by labelling the features that we wish to study with a fluorophore, we can cut out all the extraneous background and concentrate on the features that we have labelled. The Stokes



A Sika deer skin cell, photo by J. Kilgore, Copyright of Molecular Probes at Life Technologies, Eugene, OR 97402, USA.

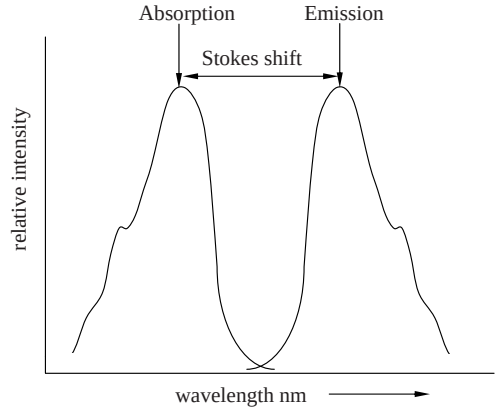


Figure 8.9 A schematic diagram of the absorption and emission peaks of a fluorophore showing the Stokes shift

shift is a function of the structure of the molecule and can vary widely. For example, it is ~20 nm for fluorescein and ~200 nm for porphyrin.

8.3.3.1 Fluorescence

Some background on fluorophores is necessary if we are to optimize the labels for our experimental programme. Most fluorophores are rigid molecules containing conjugated double bonds in addition to some unconjugated double bonds and functional groups. The details of the molecular structure define the possible electronic energy levels of the molecule and hence the absorption spectrum. Figure 8.10 illustrate s schematically the energy levels in a *Jablonski* diagram. Once in the excited state, an electron can lose energy and it comes to rest in the lowest excited state prior to falling to one of the ground states. This decrease in energy occurs without the radiation of a photon, is termed *internal conversion* and the energy is lost as heat. This is shown as the 'wavy' arrow in Figure 8.10. The energy lost in the transition from the excited singlet state to the ground state is emitted as a photon with the wavelength of the emission peak. There is also a possibility that an excited electron can move to the triplet state (a different spin state). This is termed *intersystem crossing* and is important in terms of *photobleaching*. Photobleaching occurs when molecules have electrons in the excited triplet state and they may interact chemically with other molecules such as oxygen. The change in molecular structure causes the fluorescence activity to be lost permanently.

In a fluorescence spectrometer, we measure the emitted fluorescence intensity with the detector perpendicular to the incident beam. The

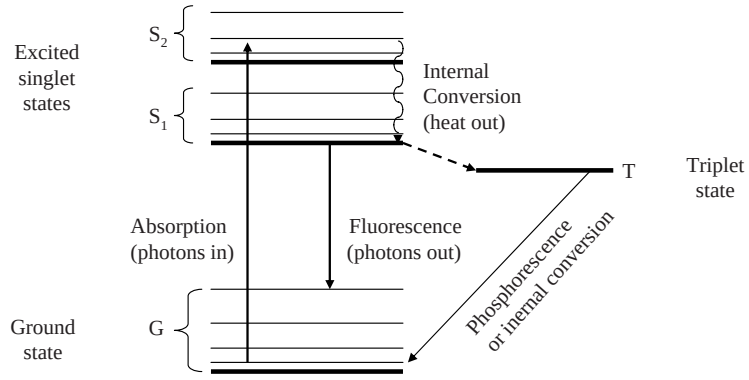


Figure 8.10 Jablonski diagram showing the electronic transitions from the ground state, G, to the excited singlet states, S_1 and S_2. The wavy arrow indicates a non-radiative decay (internal conversion)

Beer–Lambert law expresses the relative intensity of the transmitted light to the molar concentration, $[c]$, the molar extinction coefficient, ε_m, and the pathlength, l:

$$\frac{I}{I_0} = 10^{-\varepsilon_m[c]l} \tag{8.13}$$

and the absorption is

$$A = \varepsilon_m[c]l \tag{8.14}$$

Now, the fluorescence intensity, F, is related to the absorption and the incident intensity, I_0, as

$$F = I_0\varepsilon_m[c]l\,Q \tag{8.15}$$

where Q is the *quantum yield* of the fluorophore. Q is just the ratio of the number of photons emitted to those absorbed. The relative quantum yield is a quantity that is often measured by using a Rhodamine solution as a reference.

The *fluorescence lifetime* is the characteristic time that an electron remains in the excited state. The decay is a first-order rate process and so, if we excited a group of fluorophores with a light pulse, the fluorescent intensity would decrease exponentially with time:

$$I(t) = I_{t=0}\exp\left(\frac{t}{\tau_0}\right) \tag{8.16}$$

The rate constant is

$$k_f = \tau_0^{-1} \tag{8.17}$$

and $\tau_0 \approx 10^{-9}$ s. However, there are other non-radiative decay possibilities that can occur, such as internal conversion. Therefore, when we consider the situation where we are illuminating a sample, we must add all the contributions to the decay and the measured lifetime, τ_f, becomes

$$\tau_f^{-1} = \sum k \tag{8.18}$$

In this steady illumination condition, we can use Equations 8.17 and 8.18 to give the quantum yield, as if all the photons absorbed decayed by fluorescence emission, the characteristic time would be τ_0 but, as the other processes contribute to the emission, we measure τ_f. Therefore,

$$Q = \frac{\tau_f}{\tau_0} = \frac{k_f}{\sum k} \tag{8.19}$$

As τ_0 can be calculated, the fluorescence lifetime is a measure of the quantum yield.

Quenching of the fluorescence occurs when non-radiative energy transfer occurs due to interactions with other molecules. A potential quencher moiety may diffuse close enough to an excited molecule for energy transfer to occur, thereby removing the opportunity for radiative energy loss. The close proximity may be the result of intermolecular or intramolecular motion as conformational changes occur. These conformational changes may also be the result of changes in the chemical environment so that changes in fluorescence can be used as a quantification of those changes. The close proximity of a quencher species and a fluorophore can be the result of molecular complex formation. This interaction is more specific than the types of interaction that occur with solvent molecules with dipole–dipole interactions, for example. A molecule in an excited state has a greater dipole moment and so if it is dissolved in a polar solvent, the dipolar interactions result in a decrease in the energy of the excited state and a red shift of the fluorescence peak is observed. The binding of ionic species to the fluorophore can change the emission so that it is possible to use fluorescence as a local indicator for pH or calcium ion concentration. Although useful on a macroscopic level, the application to colloidal systems through modern microscopy is particularly exciting.

Prior to discussing the design of microscopes and experiments that exploit fluorescence, we need to consider the phenomenon of *fluorescence resonance energy transfer* (FRET). This occurs when two fluorophores with different spectral characteristics come into close proximity. More explicitly:

1. the separation between the two fluorophores must be $< 10\,\text{nm}$;
2. the emission peak of one fluorophore must have significant spectral overlap with adsorption peak of the other.

The first condition means that we have a distance calibration which is below the resolution limits of our optical microscope, which can be useful in indicating when units come into close approach as when we are studying the interactions between components within cells.

It is important to understand that the process involves the non-radiative transfer of energy from the excited fluorophore to the other fluorophore – it is not the reabsorption of emitted radiation. In practical terms, this means that we observe a large Stokes shift as the fluorescent emission is now at the wavelength of the second molecule which would not normally be excited by the wavelengths passed by the experimental filter set. We define the energy transfer efficiency as

$$E_{\mathrm{T}} = \frac{\tau_{\mathrm{fA}}}{\tau_{\mathrm{f}}} \tag{8.20}$$

where $\tau_{f\mathrm{A}}$ is the fluorescence lifetime of the donor fluorophore in the presence of the acceptor fluorophore. The rate of energy transfer, K_{T}, is defined in terms of the separation of the two fluorophores, r, and a critical distance, the Förster distance, R_0, where the probability of transfer is equal to that of de-excitation of the donor molecule in the absence of the acceptor, thus [3]

$$K_{\mathrm{T}}\tau_{\mathrm{f}} = \left(\frac{r}{R_0}\right)^6 \tag{8.21}$$

Now the Förster distance is a function of the refractive index of the medium, n_1, and the quantum yield of the donor [3]:

$$R_0 = \left[5.83 \times 10^{-25} n_1^{-4} QJ\,(\lambda)\right]^{\frac{1}{6}} \tag{8.22}$$

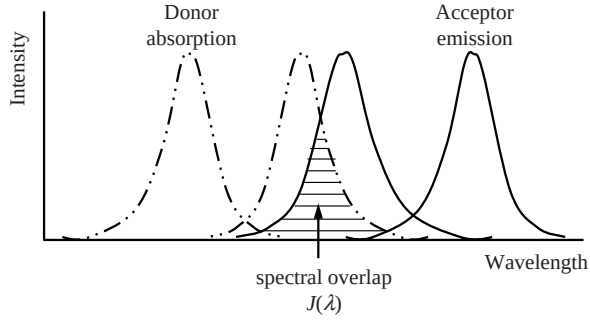


Figure 8.11 Fluorescence resonance energy transfer (FRET), where the non-radiative energy transfer by dipole–dipole interactions occurs due to the close proximity of two fluorophores

where $J(\lambda)$ is the area of spectral overlap between the normal donor emission and the normal acceptor absorbance. This is illustrated in Figure 8.11.

The energy transfer efficiency and the Förster distance can be used to estimate the separation of the donor and acceptor moieties [3]:

$$r = R_0 \left(\frac{\tau_f}{\tau_{fA}} - 1 \right)^6 \tag{8.23}$$

8.3.3.2 Mode of illumination

The normal transmissive illumination mode can be used for fluorescence microscopy, although this demands very high-quality filter sets as a bright field condenser brings a great deal of light into the objective. Epi-illumination is the more usually preferred mode. In this method, the illumination is brought into the microscope body between the eyepiece and the objective and is directed on to the specimen through the objective by reflection from a dichroic mirror. The dichroic mirror is a long-pass filter which reflects shorter wavelengths towards the specimen and allows longer wavelengths to pass through, that is, the fluorescence from the specimen passes through towards the eyepiece. The dichroic mirrors are constructed from multiple thin layers of dielectric material and can reflect and pass multiple band pairs.

A xenon light source has a flat spectral output and hence is a popular choice. Filter sets for the common fluorophore wavelengths are used as the dichroic mirrors do not have a sharp enough cut-off to render these unnecessary for the excitation and emission discrimination. The filters

are accommodated in a block or wheel so that they can be changed rapidly and digital cameras used for recording data. Short exposure times are often required if photobleaching is to be minimized. Specimens labelled with fluorescent nanocrystals are less subject to photobleaching and can be brighter, although nanocrystals require their own filter sets and disappointing results will be obtained if the filter sets used are those designed for the common fluorescent molecules.

8.3.3.3 Imaging

Fluorescent probes are commonly used for staining organelles, parts of cells containing particular species or particular molecular sequences in DNA or RNA. Some probes bind selectively to the target moieties. For example, 4′,6-diamidino-2-phenylindole (DAPI) complexes strongly with DNA. Ethidium bromide is an alternative DNA stain. Some stains are bound electrostatically by the membrane potential. Rhodamine 123 binds to strongly to functioning mitochondria in live cells by this mechanism. (The mitochondria are the organelles which have important functions such as energy production by the oxidation of lipids and oxidative phosphorylation.) When the cells are fixed, however, the membrane potential is lost and the Rhodamine stain can be washed away. MitoTracker Red is an alternative stain that is retained.

Multi-colour staining is frequently used to identify or map a variety of components in the cell simultaneously. One of the problems is that permeabilization of the cell membranes is required to allow access to all the cell components and nuclear staining is often the most demanding as the selected probe has to gain entrance not only to the cell but also to the nucleus. With modern digital photography, it is a simple matter to photograph the same selected area using different filter sets and build the compound colour image. A wide range of bright probes are available from Life Technologies [4] as their Molecular Probes™ brand.

Immunofluorescence is a useful technique for labelling moieties in cells. In this case, the fluorescent probe is covalently conjugated to an antibody which binds to the target. Antibody–antigen interactions are highly specific; however, the fluorophore density can be low so that the background fluorescence from autofluorescence from fixatives such as glutaraldehyde can be a problem. Ethanol, acetone and formaldehyde do not autofluoresce and so can alleviate the problem. However, the signal can be enhanced using a secondary fluorescent antibody complex that will bind to the first. To illustrate this type of amplification strategy for the study of mouse cells, for example, we could use a an

Alexa Fluor 488 probe conjugated to a rabbit anti-mouse IgG antibody (or a Fab_2 fragment) to give the first complex. This is followed with an Alexa Fluor 488 probe conjugated to a goat anti-rabbit IgG to form a second complex, thus building up more fluorophores at the target and thus amplifying the signal.

Fluorescence microscopy is widely used in a biocolloid context but has wider applications to the general colloid and surface area of study. For example, the structure of monolayers of octadecylamine has been studied on a Langmuir trough using an epifluorescent microscope [5]. The structures of the liquid condensed and liquid expanded regions were imaged at a variety of temperatures and substrate types.

Fluorescence in situ hybridization (FISH) uses a fluorophore covalently bound to a sequence of DNA. This fluorescent DNA section is then incubated with the double-stranded DNA in a chromosome so that hybridization can occur, so that the sequences become labelled and their location mapped. This can be a powerful tool for diagnosing genetic abnormalities.

8.3.3.4 Confocal laser scanning microscopy (CLSM)

This is a powerful optical method for studying three-dimensional objects. The structure of fluorescently labelled components of a specimen as thick as 50 μm can be imaged by this technique. Not all specimens have to be fluorescent. The surface topography of a reflective specimen can also be imaged successfully. However, biology is the field where most use is made of the instrumentation.

The x–y plane is the plane of the microscope slide and the technique scans parallel layers stacked in the z-direction. A topographical or three-dimensional image is then constructed by the computer software and can be studied in addition to x–y, x–z or y–z slices. Of course, the data are recorded continuously and so temporal changes can be studied. In addition, it is possible to use multiple colour tags of different features, thus providing a five-dimensional data set, namely x, y, z, t and colour. Additional features are (a) the possibility of using the colour intensity quantitatively, so bringing assays, such as of local pH value, into the frame, and (b) the ability to vary the magnification electronically so that one objective lens can be used with a range of magnification values.

A disadvantage of the method is the tendency for faster photobleaching of organic fluorophores as areas are repeatedly scanned by an intense laser beam. However, if Q-Dots^{TM} are the selected fluorophores, then

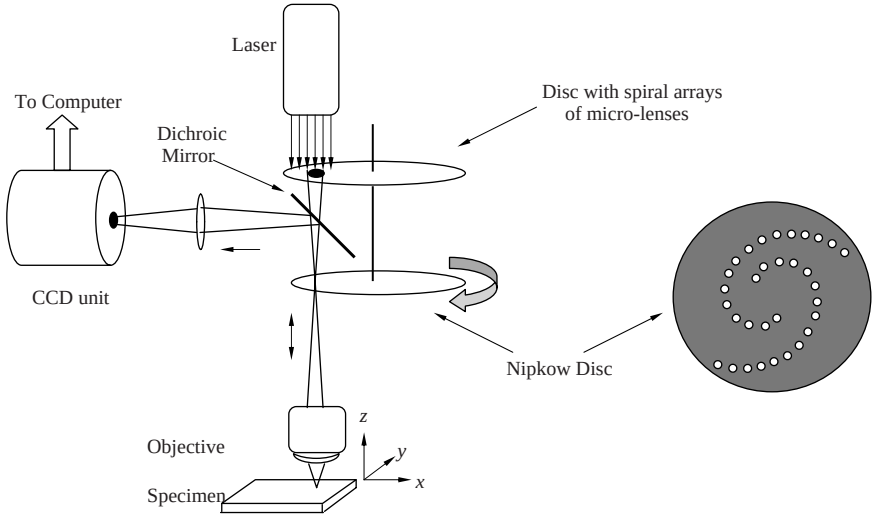


Figure 8.12 Schematic arrangement of a laser scanning confocal microscope showing the Nipkow disc with spiral arrays of pinholes coupled in tandem with a corresponding array of small lenses. The laser is focused on the pinhole, on the specimen and also on the CCD camera, resulting in pixels with brightness corresponding to the fluorescence intensity of each scanned part of the specimen

photobleaching is no longer a problem as these materials are much more stable (see Chapter 1). An additional advantage gained with Q-Dots^{TM} is the use of a single excitation wavelength for multi-colour read-outs. The size of the nanocrystals is significantly larger than that of organic molecular fluorophores and so there is a place for both types in our toolkit.

The construction of a CLSM is illustrated in Figure 8.12. The rotating disc with the multi-spiral pinhole array (the Nipkow disc) results in a spiral raster pattern on the camera. The coupled array of micro-lenses and the objective lens ensure that the laser is confocal with both the Nipkow disc and the specimen. The use of the dichroic mirror enables the advantages of fluorescence to be exploited. The technique is widely applied in determining structures in cell biology but also has a place in the general colloid and polymer field. Some of the 'early adopters' of the method studied structures in colloidal systems which were difficult to observe with other methods. Van Blaaderen *et al.* [6] synthesized 400 nm silica particles with a fluorescent core and studied the gelation of dispersions with $\varphi \approx 0.15$ as a function of temperature close to a glass surface. Verhaegh and van Blaaderen [7] followed the crystallization of 200 nm silica spheres in both hydrophilic and hydrophobic environments. For

the latter experiments, the silica particles were surface modified with octadecanol to provide a steric barrier to aggregation. The fluorophores, Rhodamine isothiocyanate (RITC) and fluorescein isocyanate (FITC), were in the core of the particles. The resolution was sufficient to resolve the liquid phase and the crystalline phase boundary and follow the advance of the boundary at a rate of 4 $\mu m\, s^{-1}$. Li *et al.* [8] studied copolymer film formation from a mix of polymer colloid particles; 1 μm particles of polystyrene and fluorescently tagged poly(methyl methacrylate) were mixed in a 10:1 ratio, dried and melt pressed into a film at 110 °C. The subsequent films were annealed at 180 °C and the aggregation of particles followed. Food research involves the study of heterogeneous colloidal systems, often with both hydrophobic and hydrophilic phases in the presence of polymers. CLSM can be used to study the phases and indicate the location of components such as proteins, starch and lipids stained with fluorescein, Texas Red and Nile Blue, respectively [9]. The improvements in computing power and the memory capacity of our computers have made the technique an important tool in our laboratories.

8.3.4 Electron Microscopy

Transmission electron microscopy (TEM) is the preferred technique for size analysis if optimal precision is required. A sample of the dispersion is placed on a mesh grid coated with a thin polymer or carbon film. When it has dried, the grid is placed in the microscope column, which is then evacuated. The beam of electrons produces an image which is focused on to a fluorescent screen. The principle is exactly analogous to a 'normal' optical microscope but in this case the lenses are electromagnets. The wavelength of the electron beam is controlled by the accelerating voltage used in the microscope, 10 or 20 kV being typical values, so that a resolution of <1 nm is available. The limit, however, is not inherent in the instrument but more often is governed by both the electron density of the particle and its thermal stability. Particles of high atomic number elements scatter electrons more strongly than low atomic number elements such as carbon. The heating effect of the beam is greatest in the highly focused intense beam used at the highest magnification and this can cause particles to shrink or decompose. The lowest magnification is limited by 'pincushion' distortion of the image which is readily observed on the screen. The reliable range for particle size analysis covers particles from $\sim$10 nm to 10 μm. The magnification has to be calibrated with each sample and the standards used are crystal spacing for the higher magnifications and carbon replicas of gelatine casts taken from diffraction

gratings for the lower magnifications. As in optical microscopy, a precision of 3% is the best that can usually be achieved.

Scanning electron microscopy (SEM) is frequently available from the same instrument. In this mode, the electron beam is focused down to a spot of ~5 nm and scanned across the specimen. The latter is mounted on an aluminium stub and usually coated with a conductor such as a thin film of gold. This prevents electrostatic charge building up, which interferes with the image. The scanned beam produces low-energy secondary electrons from the sample surface which are collected and displayed on a TV monitor. The image that we see gives us a three-dimensional perspective of the sample surface which can be immensely valuable and is always of interest. However, in terms of size analysis, the 'foreshortening' results in a wider distribution of measured sizes than is actually the case. The resolution is less than with a TEM and is limited by the size of the focused and scanned spot. If the conducting coating is absent, the energy distribution of the secondary electrons can be analysed and the composition of the surface layers can be determined. This technique is termed *electron probe microanalysis*.

In general, with electron microscopy we are limited to dry specimens as a high vacuum is required in order to prevent scattering of the electron beam by gas molecules. Microscopes are available in which the sample chamber is isolated from the rest of the microscope except by a small aperture. The evaporation rate is a slowed by the aperture and with a sufficiently powerful vacuum pump system a well-defined beam can be achieved. This enables wet samples to be imaged. However, the system is evaporating and therefore is rapidly changing, which limits the utility of the technique.

8.3.5 Atomic Force Microscopy

Atomic force microscopy (AFM) is also a scanning technique. In this technique, a fine stylus mounted on a delicate cantilever spring is brought into close proximity to the surface being studied. The surface is displaced backwards and forwards until the selected area has been examined. The displacement of the stylus is monitored using a laser beam set up to give a large magnification of the displacement by an optical lever and an image is built up on a computer monitor. A group of piezoelectric crystals are used to provide the fine three-dimensional displacement of the specimen. The optical lever and the spring constant control the resolution available and this is similar in magnitude to that with SEM. It is relatively easy to

use the technique with wet surfaces without the complications of electron microscopy.

With colloidal systems, the most important application of the AFM has been in the study of colloidal particle interactions. A polystyrene latex particle is a convenient model, spherical, colloidal particle and, by attaching a latex particle to the stylus with cyanoacrylate adhesive, the displacement of the stylus as it is brought into close proximity to another particle is monitored. From a knowledge of the spring constant and the displacement, a force–distance curve can be constructed. The limitation is that the particles must be large enough to be attached to the stylus without the material of the stylus playing a part in the interaction. Particles larger than 10^3 nm have been used successfully with this technique.

8.4 ZONAL METHODS

In these methods, a dilute colloidal dispersion is caused to flow through a well-defined zone and changes in either the electrical properties or optical properties of the zone are monitored. The magnitude of the change is related to the size of the particle and the number is given by a direct count. A major problem is that we must take care to avoid the coincidence that two or more particles are in the sensing zone simultaneously. If this occurs, they will be read as a single particle. The statistical correction to the number should be no more than a few percent of the total count. To ensure this, the concentration of the dispersion must be very low. The number of particles per unit volume in a dispersion is directly proportional to the solids concentration and inversely proportional to the cube of the particle diameter. Hence the smaller the size, the greater is the dilution that has to be made. This can represent a problem with colloidal dispersions of particles made up of soluble and even sparingly soluble materials. The sensitivity of the methods is such that they work best with dispersions with particles with a diameter >1 μm.

8.4.1 Electrical Sensing

This was first commercialized with the production of the Coulter Counter, which was originally designed simply for counting human red blood cells. It was very quickly developed to enable size distributions to be determined. The principle of this technique is that a dilute dispersion is drawn into a tube through a small cylindrical orifice, with the latter manufactured from sapphire. An electrode is placed on either side of

the orifice so that it can be operated as a small conductivity cell. As a particle passes through the orifice, the conductivity changes, and as the aperture current is set, the magnitude of the voltage pulse is proportional to the volume of the particle in the sensing zone. The pulse counter is set to progressively lower threshold values during repeated counts so that a cumulative distribution is obtained of the number of particles with a volume greater than the threshold value. The data are then converted to the diameter of the equivalent spherical particle and the numbers adjusted for the coincidence of two particles being within the zonal volume simultaneously and for the background levels of particles in the continuous phase. Filtered media must be used to keep the background down to low levels.

For this conductivity method to work, the final dilution has to be into an electrolyte with a high conductivity. The high conductivity results in easily measured changes in voltage, whereas if a low-conductivity medium is used, only small voltage changes occur and resolution is lost. With aqueous systems, a concentration between 1 and 5 M sodium chloride is required. Non-aqueous systems can be used if a suitable organic electrolyte can be found which is compatible with the dispersion. The very large dilution means that coagulation rates are slow and so if measurements are made immediately, aggregation does not affect the data. Otherwise, a suitable non-ionic stabilizer for the particles is required. The instrument calculations assume that the particles are non-conducting. If this is not the case, the data must be corrected to move the distribution to larger size values.

The range of the instrument is a function of the aperture size. If the particle diameter is $<5\%$ of that of the orifice, poor discrimination is obtained. If the particle size is $>40\%$ of the orifice diameter, blockages will occur because when two or three particles arrive simultaneously at the orifice, they can prevent each other from entering. Smaller orifices making the blocking problem more severe but enable smaller particles to be counted. The usual range of orifice diameters used is from 500 down to 30 μm. With the smallest orifices, high aperture currents are needed to produce large enough voltage pulses. The problem that occurs under these conditions is that of heating of the solution in the aperture and, although high electrolyte concentrations can be employed, the noise created by the rapid heating of the fluid in the sensing zone is a limiting factor.

A final point that the operator must bear in mind is that the system is stirred and that the sensing zone is not limited to the actual orifice. The sensing zone extends out a small distance from the ends of the orifice

so that particles passing through these regions will also be counted. The coincidence correction is based on random placement of particles in a given volume and a 'cross-flow' through either end of the sensing volume is not allowed for. The stirring then has to be as slow as possible but sufficient to prevent either sedimentation or creaming.

8.4.2 Optical Sensing

In this case, we are not referring to equipment that carries out angular light scattering from single particles. Such equipment is available, but it should be considered with light scattering in general. Here, we are looking at the change in transmission of light across a defined zone as particles move through that zone. The extinction is proportional to the total amount of light scattered in addition to that absorbed.

A rectangular cross-sectional flow chamber is used with a well-collimated light source (usually a laser) on one side of the chamber and a photodiode detection system on the other. Apertures can define the optical area very well and so there is not a problem at the edges of the zone as we have with the electrical sensing system. Laminar flow must be used, however, as turbulence must be avoided to prevent local recirculation in the eddies. Counting and pulse height analysis are similar to those used in the electrical zonal system, plus the same problem of a coincidence correction is present. The sensitivity is only in part controlled by the incident intensity as the major factor is the cross-sectional area of the particle relative to the illuminated area. Hence the particle sizes of the dispersion have to be $>1\ \mu m$ if the system is to be analysed.

8.5 SCATTERING METHODS

Colloidal particles scatter electromagnetic radiation and the size and shapes of the particles can be calculated by analysing the intensity of the scattered radiation measured as a function of angle around the scattering volume. The intensity is a function of the particle size and the wavelength of the radiation, and also the intensity of the latter and the ability of the former to scatter the radiation. The upper end of the wavelength scale is provided with light with lasers giving plane polarized light in the wavelength range 250–700 nm. For the lower end of the scales, high-energy X-rays from a synchrotron source and low-energy neutrons from a reactor with a zone cooled with liquid deuterium are employed. The

synchrotron gives radiation with an intensity about 100 times brighter than conventional laboratory sources and the wavelength can be varied over a range from approximately 0.05 to 1 nm. The neutrons, tapped from a high-flux reactor using the cold deuterium moderator to reduce the frequency, have a wavelength range of 0.1–1 nm. A mechanical collimator is then used to select a narrow range from this. An alternative source is to make use of a circular accelerator with a target which produces a large pulse of neutrons when hit. This is known as a 'spallation' source. Pulses are continuously produced with a range of wavelengths in each pulse. The scattering of a particular wavelength is isolated using its 'time-of-flight', which is easy to monitor as each pulse provides a start time as a reference. The light-scattering photometer is a bench-top piece of laboratory equipment. The X-ray and neutron spectrometers are shared systems in central locations and, of course, much more expensive. They are therefore only occasionally used for size determination alone.

The scattering of light and X-rays occurs from the interaction of the radiation with the electrons of the atoms making up the particles being studied. The neutrons, on the other hand, interact with the nuclei of the atoms. This leads to an extremely important property in that there is a very marked difference in the scattering of neutrons from hydrogen, which just contains a proton as the nucleus, compared with deuterium, with a neutron and a proton making up the nucleus. Thus, deuterated materials can be introduced into particles, polymers, surfactants or solvents to change the contrast and so hide or bring out structural features of the particles. This is analogous to changing the refractive index in a light-scattering experiment, something which is very difficult to do with simple hydrocarbon-based materials.

We are familiar with the interaction of light with matter in that this is used routinely to provide physico-chemical information. It can be reflected or refracted by macroscopic objects. It can also be diffracted by edges of components which are large compared with the wavelength and we make use of this *Fresnel diffraction* in measuring the size of particles at the upper end of the colloid size scale. Light may be absorbed and the energy dissipated as heat or re-emitted at a lower frequency (*fluorescence*). The discussion in this section is centred around the interaction and re-emission of light at the same wavelength. The energy is directly proportional to the frequency of the wave, ω rad s^{-1} (via the Planck constant). There being no dissipation, the scattered wave has the same frequency as the incident wave. Figure 8.13 illustrates the geometry for a simple scattering experiment. Here, the incident wave is scattered through an angle θ and, as there is no energy change, the magnitude of

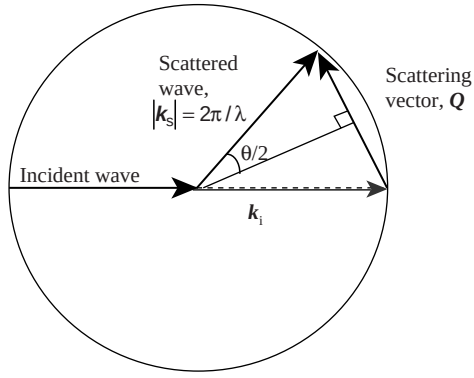


Figure 8.13 Illustration of the geometry employed in a simple scattering experiment (plan view)

the scattered wave is the same as that of the incident wave, as follows:

$$|\boldsymbol{k}_s| = |\boldsymbol{k}_i| = \frac{2\pi}{\lambda} \tag{8.24}$$

where λ is the wavelength of the radiation. The scattering vector, $\boldsymbol{Q}$, is given by the following relationship:

$$\boldsymbol{k}_s = \boldsymbol{k}_i + \boldsymbol{Q} \tag{8.25}$$

From the construction in Figure 8.13, we can see that the magnitude of the scattering vector is as follows:

$$|\boldsymbol{Q}| = Q = \frac{4\pi}{\lambda} \sin\left(\frac{\theta}{2}\right) \tag{8.26}$$

The scattering vector has units of reciprocal distance and we should note that the distance that we probe in a scattering experiment is about $2\pi/Q$. Measurements at different values of Q may be made by either measuring the scattered radiation at different angles at a fixed wavelength or by varying the wavelength at a fixed scattering angle. Figure 8.14 shows the scattering angle plotted against the dimension probed for radiation with wavelengths of 0.35 and 350 nm. The angle used for investigating colloidal particles with neutrons or X-rays is usually $<10^\circ$. Hence we refer to *small-angle neutron scattering* (SANS) and *small-angle X-ray*

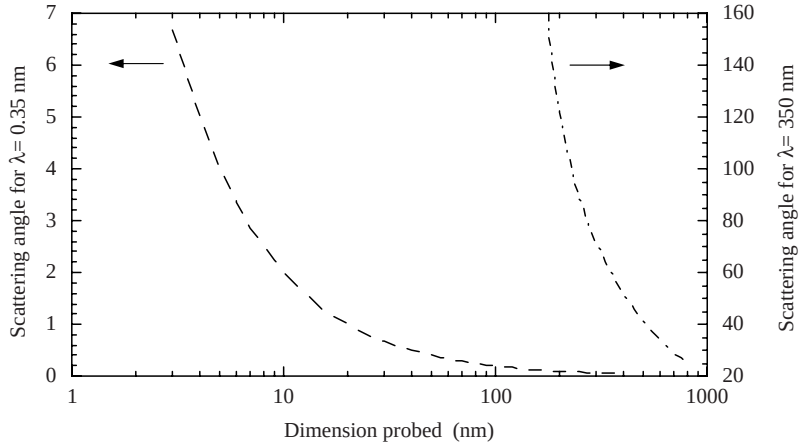


Figure 8.14 The scattering angle, θ, as a function of the dimension probed (distance, $2\pi/Q$) for neutrons and light

scattering (SAXS). With light the range of angle normally used varies from ~20 to 160°, although smaller angles can be probed successfully.

8.6 ANALYSIS OF SCATTERED RADIATION

We can only discuss an outline of this large and important topic here, but there are some excellent texts available [1, 10–13] that cover the area in detail. Light scattering from colloidal systems has a long history. The complexity of the analysis, however, increases as the particle radius approaches the wavelength of the radiation and we will start with the most straightforward situation.

8.6.1 Rayleigh Scattering

In light-scattering experiments, we normally use monochromatic, plane-polarized radiation. When the particle radius is <5% of the wavelength of the light being used, a colloidal particle can be adequately treated as a point in the electric field of the light ray. The oscillating electric field causes the outer electrons of the atoms making up the particle to oscillate at the same frequency. We can treat the particle as single oscillating dipole. This oscillating dipole radiates light of the same frequency in all directions, where the intensity is a function of the angle to the incident

direction and plane of polarization. The first case to consider is that where we are measuring the intensity as a function of angle in the horizontal plane and the electric field of the light is polarized normal to this plane, that is, it is vertically polarized. The intensity of the scattered light, $I(\theta)$, relative to the incident intensity, $I(0)$, is given by the following equation [10]:

$$\left[\frac{I(\theta)}{I(0)}\right]_{\mathrm{V}} = \frac{16\pi^4}{r^2\lambda^4}\left(\frac{\alpha}{4\pi\varepsilon_0}\right)^2 \tag{8.27}$$

where the subscript V indicates that the radiation is vertically polarized and that the scattering angle is measured in the horizontal plane. The intensity is inversely proportional to the square of distance to the detector, r. The wavelength of the radiation is λ and ε_0 is the permittivity of free space; α is the polarizability of the particle and is a function of the refractive index of the particle, n_2, relative to that of the suspending medium, n_1, and the volume of the particle, v_{p}:

$$\alpha_{21} = 3\varepsilon_0\left(\frac{n_{21}^2 - 1}{n_{21}^2 + 2}\right)v_{\mathrm{p}} \tag{8.28}$$

where $n_{21} = n_2/n_1$. Clearly, as Equation 8.28 indicates, the particle will not scatter light if it has the same refractive index as the medium. Substituting Equation 8.28 into Equation 8.27 and also writing the volume in terms of the particle diameter, D_{p}, we have

$$\left[\frac{I(\theta)}{I(0)}\right]_{\mathrm{V}} = \frac{\pi^4}{4r^2}\frac{D_{\mathrm{p}}^6}{\lambda^4}\left(\frac{n_{21}^2 - 1}{n_{21}^2 + 2}\right)^2 \tag{8.29}$$

This equation indicates that we will see no angular dependence of the scattered intensity for vertically polarized light.

When the plane of the polarization is horizontal, the angular dependence will change. We will not see the radiation from the particle when we observe the particle along the polarization axis. In this case, the relative intensity is given by

$$\left[\frac{I(\theta)}{I(0)}\right]_{\mathrm{H}} = \frac{\pi^4}{4r^2}\frac{D_{\mathrm{p}}^6}{\lambda^4}\left(\frac{n_{21}^2 - 1}{n_{21}^2 + 2}\right)^2 \cos^2\theta \tag{8.30}$$

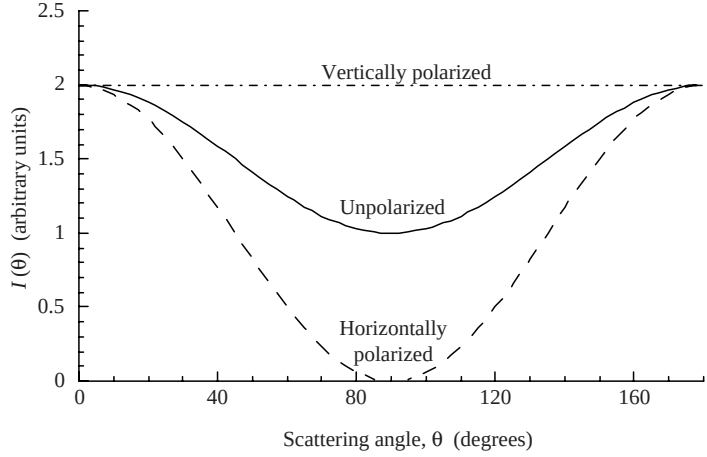


Figure 8.15 Variation of the angular intensity as a function of the scattering angle for a 'Rayleigh scatterer'

If the radiation is unpolarized, the incident intensity is equally divided between the vertical and horizontal components and we will see the relative scattered intensity as the sum of the two:

$$\left[\frac{I(\theta)}{I(0)}\right] = \frac{1}{r^2}\left[\frac{\pi^4 D_p^6}{8\lambda^4}\left(\frac{n_{21}^2 - 1}{n_{21}^2 + 2}\right)^2 (1 + \cos^2\theta)\right] \tag{8.31}$$

Figure 8.15 shows the variation in angular intensity that we would see for the scattering from a small particle using unpolarized light and also horizontally or vertically polarized light. In Equation 8.31, r is an instrument constant and the term in square brackets on the right-hand side is termed the *Rayleigh ratio*, $R(\theta)$. When we have light scattered from a dilute dispersion, each particle contributes and so the intensity as a function of angle is

$$I(\theta) = \frac{I(0)}{r^2} R(\theta)\rho \tag{8.32}$$

where ρ is the number density of particles in the dispersion in the scattering volume. This assumes that the separation between particles is large and so the light scattered by each particle is not subsequently scattered a second or third time by neighbouring particles, that is, there is no *multiple scattering*. This is easily checked by showing that there is a

linear dependence on concentration. Note that Equation 8.31 is very sensitive to the particle size as it contains the particle diameter to the sixth power. Hence, if there is a distribution of particle sizes, the larger fraction contributes much more strongly to the scattered intensity than the smaller sizes and the average diameter calculated from the scattering will weighted to the 'larger end' of the size distribution.

8.6.2 Rayleigh–Gans–Debye Scattering

As the particle diameter becomes a significant fraction of the wavelength of the radiation, we can no longer treat the particles as 'point scatterers'. Radiation scattered from different parts of a particle will have different distances to travel to the detector and there will be interference between the waves, the extent of which will be a function of the angle. The problem is outlined in Figure 8.16, in which the particle diameter is $\sim 1.25\lambda$.

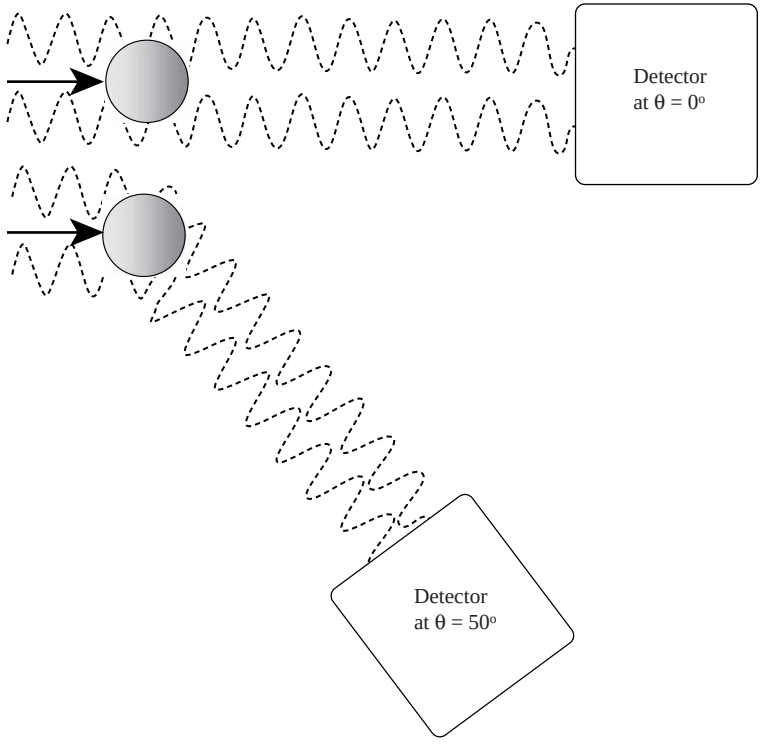


Figure 8.16 Light scattered at 0 and 50°, showing the angular variation of the interference between the rays emanating from different parts of a particle with a diameter $D_p \approx 1.25\lambda$

The intensity of scattered radiation is modified from that calculated with Equation 8.32 by the *form factor* $P(\theta)$ as follows:

$$I(\theta)_{\mathrm{RGD}} = I(\theta)_{\mathrm{R}} P(\theta) \tag{8.33}$$

where the subscripts RGD and R indicate the intensities from Rayleigh–Gans–Debye and Rayleigh calculations, respectively. The theory is applicable to particles where $(n_2 - n_1) D_p / 2\lambda < 1$.

For small deviations from the Rayleigh condition, the form factor can be approximated for any shape particle [12] by the following:

$$P(\theta) = 1 - \frac{(QR_{\mathrm{G}})^2}{3} + \cdots \tag{8.34}$$

where R_{G} is the radius of gyration of the particle and so Equation 8.34 can be used to calculated the angular variation for rods or discs, in addition to spheres. For example, the radius of gyration of a spherical particle with a homogeneous density profile is given by

$$R_{\mathrm{G}} = \left(\frac{3}{5}\right)^{0.5} \frac{D_{\mathrm{p}}}{2} = 0.39 D_{\mathrm{p}} \tag{8.35}$$

If we measure the intensity at two angles, we can use Equations 8.33 and 8.35 to determine the particle size. Angles of 45 and 135° are usually chosen for this and the ratio of the two intensities is known as the *dissymmetry ratio*. Now, for our homogeneous sphere we can write the dissymmetry ratio, after substituting for Q from Equation 8.27, as follows:

$$\frac{I(45)}{I(135)} \approx \frac{1 - 1.17\left(D_{\mathrm{p}}/\lambda\right)^2}{1 - 6.83\left(D_{\mathrm{p}}/\lambda\right)^2} \tag{8.36}$$

Therefore, we can determine the diameter simply from measuring the dissymmetry ratio.

Alternatively, we can use the value of $I(90)$ as a reference intensity (the $\cos\theta$ term in Equation 8.31 becomes zero), and so we can write

$$I(\theta) \approx I(90)\left(1 - \cos^2\theta\right)\left[1 - (QR_{\mathrm{G}})^2/3\right] \tag{8.37}$$

Equation 8.37 gives us the radius of gyration of our particle. For homogeneous particles, the radii of gyration for various shapes are given by

the following:

$$\begin{array}{ll} \text{spherical ring of radius } a\text{:} & R_G = a \\ \text{sphere of radius } a\text{:} & R_G = a\sqrt{3/5} \\ \text{thin disc of radius } a\text{:} & R_G = a/\sqrt{2} \\ \text{thin rod of length } L\text{:} & R_G = L/\sqrt{12} \end{array} \tag{8.38}$$

and so we can determine the particle dimension from the ratio of the intensities.

As the particle size increases further from the Rayleigh region, the form factor can no longer be adequately described using only the first term in the series expansion (Equation 8.34). For homogeneous spheres of radius a, we may write the form factor as follows [12]:

$$P(\theta) = \left\{ \frac{3\left[\sin(Qa) - Qa\cos(Qa)\right]}{(Qa)^3} \right\}^2 \tag{8.39}$$

Figure 8.17 shows the variation of $P(\theta)$ as a function of Q calculated for two particle sizes when using a wavelength of 600 nm. There is a marked structure in the angular variation of the form factor as the radius approaches the wavelength of the radiation. As the particle size distribution broadens, the sharp troughs become shallower and are smoothed out for a sample with a broad size distribution.

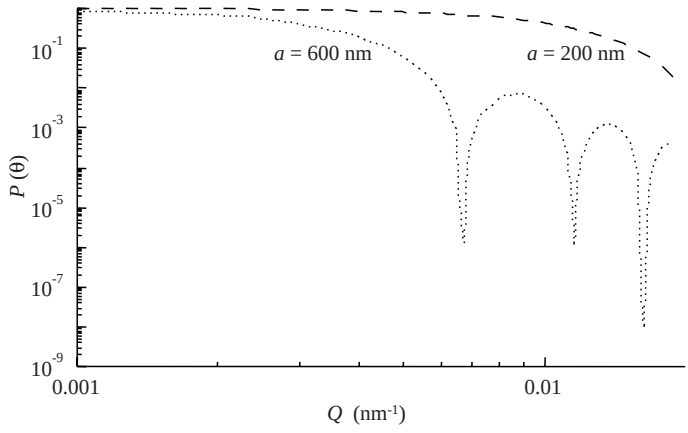


Figure 8.17 Variation of the particle form factor, $P(\theta)$, as a function of Q for two different particle sizes

For large particles, $(n_2 - n_1)a/\lambda \approx 1$, the theoretical analysis, known as the Mie theory [10], is much more complex and is limited to spherical particles. Computer programs are available to match the structure of the angular intensity scan from the theory with that observed experimentally [10, 14]. The problem is not just computational, however, but is also one of measuring absolute intensities accurately so that great care is required to ensure good data. This is always true for light scattering as small amounts of airborne dust can easily get into the dispersions. As most dust particles are in the upper end of the colloid size domain, the scattering from such contaminants can result in significant errors.

8.6.3 X-rays and Neutrons

The above discussion is written in terms of the refractive index of the disperse and continuous phases. The analysis is the same for both X-rays and neutrons except that the scattering length density of the particles is used instead of the refractive index. For a particle, the scattering length density is calculated by summing the scattering lengths, b_i, of all the atoms in the particle. As we noted above, scattering of X-rays occurs from the electrons of the atom and is therefore the product of the atomic number, z_i, of the atom and the scattering length of an electron (which is 2.8×10^{-15} m). Neutrons, however, are scattered from the nucleus and the scattering length does not vary regularly with the atomic number of the atom. The values are tabulated, of course, so this does not represent a problem. We can calculate the scattering length densities for neutrons and X-rays of our particles, ρ_{SN} and ρ_{SX}, respectively, as

$$\begin{aligned} \rho_{SN} &= N_m \sum_i b_{Ni} \\ \rho_{SX} &= N_m \sum_i b_{Xi} = 2.8 \times 10^{-15} N_m \sum_i z_i \end{aligned} \tag{8.40}$$

where N_m is the number of molecules per unit volume made up of atoms of type i in the particle, namely

$$N_m = \frac{\rho_m N_A}{M_m} \tag{8.41}$$

where ρ_m and M_m are the density and molar mass of the molecules, respectively, and N_A is Avogadro's number. Similar calculations must be

Table 8.2 Scattering lengths for neutrons and X-rays of some typical atoms

Atomic species	b_{SN} (10^{-15} m)	b_{SX} (10^{-15} m)
H	−3.74	2.85
D	6.67	2.85
C	6.65	17.1
O	5.81	22.8
Na	3.36	31.3
Si	4.15	39.0
Cl	9.58	48.4
Cd^{2+}	3.7	131.1

done for the continuous phase, of course. Values of the scattering length of some atoms are given in Table 8.2. We can express the refractive index in terms of the scattering length density and the wavelength as follows [15]:

$$n = 1 - \rho_{SN}\lambda^2/\pi \tag{8.42}$$

and we would not observe scattering from the particle if its refractive index matches that of the medium. This ability to *contrast match* can often be exploited.

By using Equations 8.40 and 8.41 with the values taken from Table 8.2, the scattering length density for water is as follows:

$$\frac{10^3 \times 6.022 \times 10^{23}}{0.018}[5.81 - (2 \times 3.74)] \times 10^{-15} = -0.56 \times 10^{14}\text{m}^{-2}$$

whereas that for deuterium oxide is 6.34×10^{14} m^{-2}. Many hydrocarbon polymers fall in between these values so that polymer particles or coatings can be visualized separately if combinations of hydrogenated and deuterated materials are used. This technique is known as *contrast matching*. As an example that may be used to illustrate how we could apply this technique, consider a small polystyrene particle (say $r \approx 25$ nm) coated with a monolayer of dodecanoic acid as a stabilizer. In the experiment to study the coating, we could use d_{23}-decanoic acid. Now, the scattering length density for h_8-polystyrene is 1.41×10^{-14} m^{-2} and that for the deuterated dodecanoic acid is 5.3×10^{-14} m^{-2}. In a mixture of approximately 25% D_2O–75% H_2O we will have a scattering length density match with the core particle and will 'see' the layer as a hollow spherical shell. On the

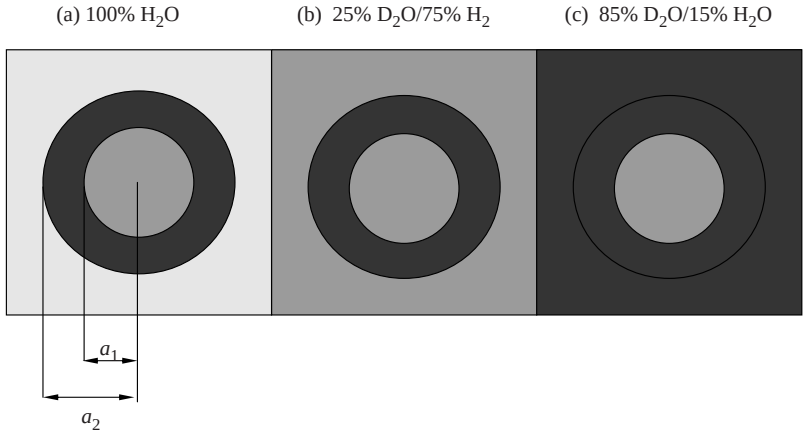


Figure 8.18 Illustration of the 'contrast matching' of a polystyrene particle with a deuterated dodecanoic acid layer using different H_2O–D_2O mixtures. (a) Both the particle and the layer are 'visible'; (b) just the dodecanoic acid layer is 'visible' as the particle and suspending medium are matched; (c) just the core particle is 'visible' as the adsorbed layer and the suspending medium are matched

other hand, when we use a mixture of 85% D_2O–15% H_2O we will no longer be able to 'see' the surfactant shell with the neutron beam and will just scatter from the core polystyrene particle. In 100% water, of course, we will scatter from both the core and the shell. Figure 8.18 illustrates the effect.

It is convenient to express the scattering of neutrons from a dispersion of particles in the form

$$I(\mathrm{Q}) = A\rho \left(\frac{4\pi a_\mathrm{p}^3}{3} \right)^2 (\rho_\mathrm{p} - \rho_\mathrm{m})^2 P(\underline{Q}) \tag{8.43}$$

where A is the instrument constant that includes the wavelength, the incident intensity and the distance to the detector. (Neutron detectors are made up of two-dimensional arrays of 1 cm^2 measuring elements.) The terms ρ_p and ρ_m refer to the scattering length densities of the particle and medium, respectively, a_p is the radius of the particle and ρ is the number density of particles in the dispersion. The equivalent expression for concentric spheres, that is, a dispersion of particles with attached

layers, is [16]

$$\frac{I(Q)}{A\rho\left(\frac{4\pi}{3}\right)^2} = \left(\begin{array}{l}\left(\rho_{\mathrm{p}}-\rho_{\mathrm{m}}\right)\left\{3a_1^3\left[\dfrac{\sin\left(Qa_1\right)-Qa_1\cos\left(Qa_1\right)}{\left(Qa_1\right)^3}\right]\right\} \\ +\left(\rho_{\mathrm{L}}-\rho_{\mathrm{m}}\right)\left\{3a_2^3\left[\dfrac{\sin\left(Qa_2\right)-Qa_2\cos\left(Qa_2\right)}{\left(Qa_2\right)^3}\right]\right. \\ \left.-3a_1^3\left[\dfrac{\sin(Qa)_1-Qa_1\cos\left(Qa_1\right)}{\left(Qa_1\right)^3}\right]\right\}\end{array}\right)^2 \tag{8.44}$$

The first term on the right-hand side of Equation 8.44 represents the scattering from the core particle and the second term is that from the shell with a scattering length density of ρ_{L}. Clearly, as the scattering length density of the medium is adjusted to equal that of the layer or the particle, we will be able to determine the dimension of the particle or the layer, respectively.

8.6.4 The Guinier Approximation

When $Qa \ll 1$, Equation 8.34 provides a reasonable approximation for $P(Q)$, so we have

$$I(Q) = A\rho v_{\mathrm{p}}^2\left(\rho_{\mathrm{p}}-\rho_{\mathrm{m}}\right)^2\left(1-\frac{(QR_{\mathrm{G}})^2}{3}+\cdots\right) \tag{8.45}$$

Writing this as the exponential form instead of the series:

$$I(Q) = A\rho v_{\mathrm{p}}^2\left(\rho_{\mathrm{p}}-\rho_{\mathrm{m}}\right)^2\exp\left[-\frac{(QR_{\mathrm{G}})^2}{3}\right] \tag{8.46}$$

which in logarithmic form gives

$$\ln\, I(Q) = \ln\, I\,(0) - \frac{(QR_{\mathrm{G}})^2}{3} \tag{8.47}$$

Therefore, from a plot of $\ln I(Q)$ versus Q^2 we obtain the radius of gyration of a particle of arbitrary shape from the initial slope as $-R_{\mathrm{G}}^2/3$ and the intercept as $I(0) = A\rho v_{\mathrm{p}}^2(\rho_{\mathrm{p}}-\rho_{\mathrm{m}})^2$.

8.6.5 Porod's Law

This describes the scattering response at high values of Q. As the angles and wavelengths become large, the intensity of scattered radiation varies as Q^{-4}. The limiting equation is

$$I(Q) \approx 2\pi S \left(\rho_{\mathrm{p}} - \rho_{\mathrm{m}}\right)^2 Q^{-4} \tag{8.48}$$

where S is the surface area of the material. It is useful to think in terms of the *fractal dimension* of the surface, d_{sf}. Then the proportionality is

$$I(Q) \propto Q^{-(6-d_{\mathrm{sf}})} \tag{8.49}$$

When we think of a characteristic distance on the surface, r, the area is proportional to r^2, and so the fractal dimension is 2. However, when we have a rough surface, there is a three-dimensional character so that the area is more closely proportional to r^3 and then $d_{\mathrm{sf}} \to 3$ and therefore

$$\begin{aligned} &I(Q) \propto Q^{-4}\text{: smooth surface} \\ &I(Q) \propto Q^{-3}\text{: rough porous surface} \end{aligned} \tag{8.50}$$

8.7 NEUTRON REFLECTION

Macroscopic surfaces can be investigated using a small-angle neutron reflection experiment. The intensity of the neutron beam is measured as a function of the incident angle, θ, and of the wavelength. The reflectivity is dependent on the properties of the surface, such as the thickness of any surface layer and its scattering length density. The reflectivity $R(Q)$ is determined from the variation across the interfacial region, that is, in the z-direction:

$$R(Q) = \frac{16\pi^2}{Q^4} \left| \int \frac{\partial \rho(z)}{\partial z} \exp\left(i\,Qz\right) \mathrm{d}z \right|^2 \tag{8.51}$$

where we see the Q^{-4} decay that we expect from the flat surface. The reflectivity is a complex quantity (note the $i\,Qz$ exponential term), as we would expect for any waveform.

The technique has been used effectively to study monolayers of surfactants on water. The water layer can be rendered 'invisible' by using a D_2O:H_2O ratio of 0.08:0.92 as the scattering length density of the mix is

now zero. The reflection is then just from the monolayer. The scattering length density of the monolayer, ρ_F, is

$$\rho_F = \frac{\sum_i b_{Ni}}{A_m l_m} \tag{8.52}$$

where A_m is the area per molecule in the layer and l_m is the length of the molecule in the layer (layer thickness). (Note that the latter will not necessarily be the length of a stretched surfactant – it is just the thickness of the layer.) We can treat the scattering length density as uniform across the surfactant monolayer so that the reflectivity is now

$$R(Q) \approx \frac{16\pi^2}{Q^4}\rho_F^2 \times 2\left[1 - \cos\left(Qz\right)\right] \tag{8.53}$$

A plot of $R(Q)Q^4$ versus Q varies sinusoidally with a total amplitude of $32\pi^2\rho_F{}^2$ with the peak at π/l_m. Hence the length of the surfactant in the monolayer can be determined from the value of Q at the first peak and the area per molecule from the amplitude of this peak via Equation 8.53.

8.8 DYNAMIC LIGHT SCATTERING

The time-average scattering that we have discussed so far is *elastic* scattering in that the frequency of the scattered radiation is the same as that of the incident radiation. The particles in the suspension are moving with Brownian motion and we can expect a broadening of the wavelength due to a Doppler shift. In principle, this shift would yield the diffusion coefficient and thus give a measure of the particle size. However, this shift is small and difficult to measure.

The technique that we use to follow the motion is known as *photon correlation spectroscopy* [17], the name of which is indicative of the experimental methodology. Lasers produce coherent light and the equipment is focused to give a very small scattering volume. The particles in the dilute dispersion scatter the light and, because the light is coherent, the phase relationships are maintained. This means that the dispersion acts like a three-dimensional diffracting array. The structure is random (as the dispersion is dilute and therefore the particles are non-interacting by definition) and we see a random diffraction pattern. As

the particles are in motion, the diffraction pattern also moves. This is termed a 'speckle pattern'. From the detector (photomultiplier), we look at the small area defined by the scattering volume and the bright spots or speckles move in and out of our vision at a rate that is dependent on the Brownian motion. The photomultiplier is connected to a photon correlator, and the intensity is measured as a function of time. The total time can be set and also the intensity (number of photons) is measured over a series of short time intervals until the total time is reached. The chosen time interval is termed the 'correlation delay time', τ_c. Remember that the intensity is rapidly fluctuating up and down as the speckles move in and out of the field of view. Hence, at two times close together there is a strong correlation in the intensities, but this will not be the case for times which are widely separated. The instrument constructs the correlation function, $g(\tau_c)$, by comparing the product of the number of photons counted initially with those counted in the correlation delay time intervals measured over longer and longer times with the square of the long-time value, that is, the average value. We can express this mathematically as follows:

$$g^2(Q, \tau) = \frac{\langle I(0) I(\tau) \rangle}{\langle I \rangle^2} \tag{8.54}$$

The correlation function decays exponentially, as follows [17]:

$$g(\tau_c) = \exp\left(-\tau_c D_s Q^2\right) \tag{8.55}$$

where D_s is the self-diffusion coefficient of the particle. Note that $D_s Q^2$ is just the time taken for a particle to diffuse through a distance of Q^{-1}. The data are plotted as the natural logarithm versus Q^2, which should yield a straight line, and then the diffusion coefficient is found from the slope. The Stokes–Einstein relation is used to give the hydrodynamic radius of the particle:

$$a_H = \frac{k_B T}{6\pi \eta_0} \tag{8.56}$$

where η_0 is the viscosity of the continuous phase. Of course, if the dispersion consists of particles of monodisperse hard spheres, the hydrodynamic radius should correspond to that determined by microscopy. If the particles are non-spherical or are covered with a stabilizer layer

which has a thickness that is a non-negligible fraction of the radius, the hydrodynamic radius will be larger than that from microscopy.

As the size distribution becomes broader, the range of diffusion times increases in direct proportion. The correlation function then becomes the sum of each of the exponential decaying terms, weighted by the probability of each time occurring. The experimental plot has to be analysed with this in mind as all dispersions have some degree of polydispersity. The problem is that we cannot obtain a unique solution when we try to invert the summation. (Mathematically it is termed an 'ill-conditioned' problem.) The method used by each instrument manufacturer is to set up a model which is calibrated against particles of known size. This usually works well for mixtures of monodisperse particles of well-separated sizes, but may not be so reliable with a single but broad distribution. In addition, as the intensity of scattered light is being measured, the mean will be weighted to the larger sizes.

8.9 CHARACTERIZATION OF THE ELECTRICAL PROPERTIES OF PARTICLES

In this section, we will concentrate on the properties of the surfaces of colloidal systems rather than bulk properties such as the dielectric behaviour. The surface charge density of particles can be determined by titration. The larger the particle size, the more difficult this becomes as the total titratable charge is inversely proportional to the specific surface area. Electrokinetic techniques are used to determine the ζ-potential. The larger part of our discussion will centre around the latter, as this is an important factor for determining the interactions between particles.

8.9.1 Surface Charge by Titration

This technique is appropriate for particles with fixed ionogenic groups on the surface of the particles. A good example of such a surface is a polymer latex particle with charged polymer groups. These can be the terminating groups on the ends of the polymer chains or from a charged copolymer. A good example of the latter is the latexes which have a considerable fraction of acrylic acid copolymerized in the particles. These latexes can be used as modifiers in Portland cement formulations where they improve texture, strength and waterproofing properties.

As it is the strongly attached or covalently bound charge that we are measuring, physically adsorbed material such as weakly adsorbed

surfactants have to be removed. In addition, all the counter-ions to the surface groups must be converted to ones which can be readily titrated, such as protons or hydroxyl ions. Free material in solution must also be removed. Treatment with a mixed-bed ion-exchange resin which has its components in the H^+ and the OH^- forms can achieve all of these objectives, leaving acid groups in the hydrogen form and all salts, acids or bases converted to water.

The dispersion is then titrated with acid or base as appropriate. The 'equivalence point' is determined either conductometrically or potentiometrically. The experimental difficulty is that dilute acids or bases may have to be used, for example at concentrations of 10^{-2} M or less, and we have to be careful to exclude carbon dioxide. Conductometric data are easier to obtain and give the total charge with the equivalence points being shown by clear changes of slope. The latter reflect how strong or weak the acids or bases are. For example, strong acids give a marked negative initial slope as mobile protons are replace by lower mobility sodium ions. Weak acids have a slight positive initial slope as the acid groups in the hydrogen form are poorly dissociated so only a few protons are contributing to the measured conductivity, but all of the sodium ions that have been used to replace them are ionized.

Potentiometric data are less easy to analyse but gives additional information in terms of the pK_a values of the groups. We should keep in mind with particle surfaces that the ionization of one group is affected by the neighbouring groups, that is, the surface has a polyelectrolyte character and so can have a wide range of pK_a values rather than a single value. Reactive surface groups can be quantified and reactions monitored by titration using ion-selective electrodes other than pH electrodes. For example, a chloride-selective electrode can be used to monitor the interaction of proteins with polymer particles with chloromethyl groups at the surface. Chloride ions are produced as the amine groups on the proteins react with the chloromethyl groups to link the proteins covalently by the formation of an amide bond.

8.9.2 Electrokinetic Methods

All of these methods rely on the movement of the suspending medium past the charged interface. As such, they can in some cases be applied to macroscopic surfaces in addition to colloidal particles. The property that is determined is the ζ-potential and not the surface potential. This is because we are making measurements from the relative motion of the

Table 8.3 Types of electrokinetic effects

Technique	Conditions	Procedure
Streaming potential	Static interface	Move liquid; measure potential
Electro-osmosis	Static interface	Apply potential; measure liquid motion
Sedimentation potential	Moving interface	Particles sediment; measure potential
Electrophoresis	Moving interface	Apply potential; measure particle motion
Primary electroviscous effect	Moving interface and liquid	Measure suspension viscosity
Ultrasonic vibration potential	Moving interface and liquid	Apply ultrasound; measure AC potential

electrolyte solution and the surface. It is assumed that the counter-ions that are strongly bound in the Stern layer are static on the surface and that the relative motion starts between this layer and the hydrated ions plus water molecules next to the Stern layer. This gave rise to the concept of a 'shear plane'. This is a convenient simplification for our calculations, but we must be aware that it is no better than the modelling assumption that we normally use of a planar particle surface with the charge smeared out uniformly over the surface. Hence, although the models are satisfactory for distances that are many molecular diameters away from the surface, the precision becomes lower at very close distances. Hence we are left with a potential that we can measure reasonably accurately at a distance from the surface that will be around 0.5–1 nm, but the value of which we do not know with precision. The important point, however, is that it represents a good description of the potential that a test probe will experience as it is brought from a large distance towards the surface. The Stern potential can be much lower and even of the reverse sign to the surface potential due to the strong adsorption of multivalent counter-ions or surfactant molecules. There are a number of methods available to determine the ζ-potential of surfaces. A list is given in Table 8.3 and we will briefly discuss each of these.

Throughout the discussion, we are going to assume that the liquid is Newtonian, that is, the shear stress, σ, is directly proportional to the shear rate, $\dot{\gamma}$, produced in the liquid and the proportionality constant is the coefficient of viscosity, η. The definitions of these terms are illustrated in Figure 8.19.

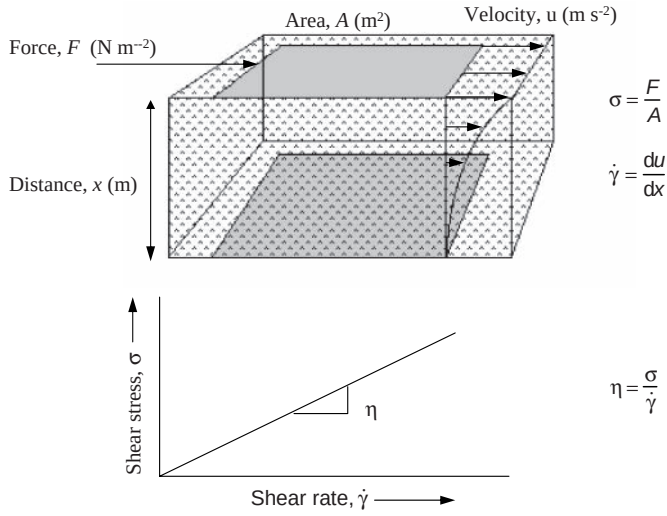


Figure 8.19 Definitions of the shear stress, σ, the shear rate, $\dot{\gamma}$ and the viscosity, η, for a Newtonian liquid. Note that if x is sufficiently small, then the velocity gradient will be linear and there will be a constant shear rate between the surfaces of area A. This is the case for most rotational viscometers. Note: σ is the commonly used symbol for shear stress in most rheological texts

In general, the fluid velocity gradient is not linear over large distances and it is important to know how this varies with distance as liquid flows in, for example, a capillary tube which can be used to determine the ζ-potential of macroscopic surfaces. The geometry that we will use for flow through a capillary is shown in Figure 8.20.

For steady flow through the tube, the applied force is equal to the viscous drag force:

$$\Delta P \pi r^2 = -\frac{\mathrm{d}v(r)}{\mathrm{d}r}\eta \times 2\pi r L \tag{8.57}$$

The boundary conditions for the integration of Equation 8.57 are as follows:

$$\begin{gathered} v = 0 \text{ at } r = a \\ v = v_0 \text{ at } r = 0 \\ \int_{v(r)}^{0} -\mathrm{d}v(r) = \frac{\Delta P}{2L\eta}\int_{r}^{a} r\,\mathrm{d}r \end{gathered} \tag{8.58}$$

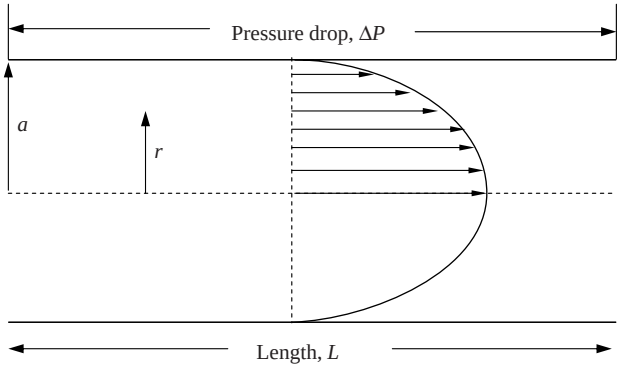


Figure 8.20 Geometry of the flow through a capillary tube

This gives us the well-known parabolic velocity profile for liquid in laminar flow through a tube:

$$v(r) = \frac{\Delta P}{4L\eta}\left(a^2 - r^2\right) \tag{8.59}$$

We can use Equation 8.59 to calculate the volume flowing through any element and by multiplying by the area of that element, $2\pi r\mathrm{d}r$, and then integrating across the radius to give the Poiseuille equation for the volume flow rate, Q, through a capillary tube, as follows:

$$Q(r) = v(r) \times 2\pi r\mathrm{d}r \tag{8.60}$$

so that

$$Q = \frac{\Delta P\pi}{2L\eta}\int_0^a \left(ra^2 - r^3\right)\mathrm{d}r = \frac{\Delta P\pi a^4}{8L\eta} \tag{8.61}$$

8.9.3 Streaming Potential

The simplest form of this experiment is to use a capillary tube with an electrode compartment at each end, flow electrolyte through the tube and measure the potential developed between the electrodes. This is illustrated in Figure 8.21. A high-impedance voltmeter is required as the resistance between the electrodes is high. To obtain reproducible

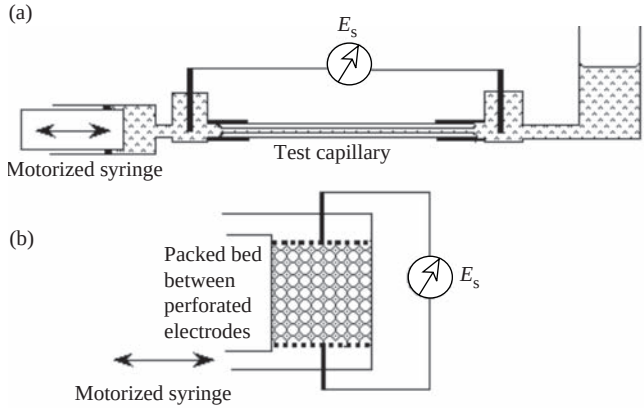


Figure 8.21 Schematics of an apparatus used to measure the streaming potential: (a) test capillary arrangement; (b) cell containing a packed bed of particles or fibres

measurements, care has to be taken to avoid the development of turbulent flow due to too high a flow rate.

In the diffuse layer, there is a higher concentration of counter-ions than co-ions. This, of course, balances the net charge at the Stern plane. The reason why a potential develops is that the flow of the liquid moves the ions in the diffuse layer along the tube. There is therefore a net charge flux through the tube and a 'back-current' is produced which is equal and opposite to this. Far from the wall, the counter- and co-ion fluxes are equal and so only the diffuse layer contributes. Ohm's law is used to calculate the current at the electrodes from the potential developed between the electrodes and the resistance of the contents of the tube.

Our problem then is to calculate the number of ions per unit volume at any position in the diffuse layer, the space charge density $\rho_c(r)$ and their velocities. Integration across the tube then gives the total current and hence the *streaming potential*, E_s.

The first step is to write the back-current in terms of the ion flux:

$$i = -2\pi \int_0^a \rho_c(r) v(r) r \, dr \tag{8.62}$$

Although it is convenient to take the integration out to the tube wall and ignore the thickness of the Stern plane, this, of course, does not imply that we are looking at the surface and not the Stern plane. We can relate the space charge density to the curvature of the diffuse layer potential

using Poison's equation. In addition, it is acceptable to use the planar form of the equation provided that κa is very large, that is, we are not using microscopic pores at low electrolyte concentrations:

$$\frac{\mathrm{d}^2\psi(x)}{\mathrm{d}x^2} = \frac{-\rho_c(x)}{\varepsilon} \tag{8.63}$$

where ε is the permittivity of the solution and x is the distance from the charged surface. It is therefore convenient to recast Equation 8.62 in terms of distance from the wall of the tube, x, using $r = (a - x)$:

$$i = -\frac{\pi \Delta P}{2L\eta} \int_0^a \left(2ax - x^2\right)(a - x)\,\rho_c(x)\mathrm{d}x \tag{8.64}$$

Now, as $a \gg x$, we only require the leading term from the expansion of the terms in parentheses in this equation and hence we have

$$i = \frac{\pi a^2 \varepsilon \Delta P}{L\eta} \int_0^a x \frac{d^2\psi}{\mathrm{d}x^2}\mathrm{d}x \tag{8.65}$$

The boundary conditions for Equation 8.66 are as follows:

$$\psi(x) = \zeta \text{ at } x = 0$$
$$\frac{\mathrm{d}\psi(x)}{\mathrm{d}x} = 0,\ \psi(x) = 0 \text{ at } x = a$$

Equation 8.65 is integrated by parts and with these boundary conditions gives the current as follows:

$$i = \frac{\pi a^2 \varepsilon \Delta P}{L\eta}\zeta \tag{8.66}$$

Noting that the resistance of the tube can be written in terms of molar concentration of the electrolyte, c, and the equivalent conductance, Λ:

$$R = \frac{L}{c\Lambda\pi a^2} \tag{8.67}$$

By using Ohm's law, Equation 8.66 can be written in terms of the streaming potential, E_s, with Equation 8.67. It is also useful to use

Equation 8.61 to give the result in terms of the volume flow rate as

$$E_s = \frac{Q}{c\Lambda}\frac{8L\varepsilon}{\pi a^2}\zeta \tag{8.68}$$

Hence we can vary Q and plot E_s as a function of Q. This should give a straight line with a slope that is directly proportional to the ζ-potential. The proportionality constant is $8L\varepsilon/c\Lambda\pi a^2$.

We have to be careful about the development of a *standing potential* between the electrodes, indicating slight surface differences and polarization of the electrode surfaces. In order to minimize these effects, the electrodes should be 'shorted' when not being used for measurement in addition to the flow direction being reversed and the measurement repeated. We will have a deviation from a straight line if the flow rate is such that it will result in turbulence, or if a packed bed of fibres or coarse particle is used, the pressure drop is high enough to change the packing and hence the path through the bed.

The capillary (see Figure 8.21b) can be constructed of the material in which we are interested or we can coat the internal surface with this material. This may be achieved by physical adsorption of materials such as proteins or surfactants, by chemical grafting of material or even by casting the capillaries from solutions or melts. The cell shown in Figure 8.21b can contain fibres or coarse particles. If these are fine and low electrolyte concentrations are employed, we should be aware that the conductivity that we might measure would become dominated by the contribution from transport in the diffuse layers, that is, surface conductance, and the calculated ζ-potential would be lower than if the effect is allowed for.

8.9.4 Electro-osmosis

Electro-osmotic flow can be observed when an electrical potential is applied across a capillary or a porous plug and it is due to the motion of ions moving in the potential gradient. Of course, in the bulk there is no net flow as the mass transports due to anions and cations balance each other. This is not the case when one of the types of charge has a fixed location but the other not, as is the case with a surface where the diffuse layer ions are able to move freely in a direction parallel to the surface. We will consider the case of a capillary with a charged wall to illustrate the effect. The electrical field strength is the applied potential per unit

length of the capillary, E. Consider a small element, dx thick and of unit area. The force on this element due to the applied field on the solvated ions is then

$$\text{electrical force} = E\rho_c(x)\mathrm{d}x \tag{8.69}$$

This force produces the motion of the element and, when the flow is steady, it is balanced by the viscous drag force. This drag on the surface of the fluid element is due to the relative velocity change across the element and the stress is determined from the rate of change of the stress across the element multiplied by the distance over which it is changing, as follows:

$$\text{viscous force} = \eta\frac{\mathrm{d}^2 v(x)}{\mathrm{d}x^2}\mathrm{d}x \tag{8.70}$$

Equating the forces for the steady flow condition and substituting for the space charge density using the planar form of Poisson's equation (Equation 8.63) and then integrating twice with respect to x to gives the fluid velocity a long distance away from the wall, yields

$$-E\varepsilon\int\frac{\mathrm{d}^2\psi(x)}{\mathrm{d}x^2}\mathrm{d}x = \eta\int\frac{\mathrm{d}^2 v(x)}{\mathrm{d}x^2}\mathrm{d}x \tag{8.71}$$

The boundary conditions for the equation are as follows:

$$\begin{aligned} &x = 0, \quad \psi = \zeta, \quad v(x) = 0; \\ &x = a, \quad \frac{\mathrm{d}\psi}{\mathrm{d}x} = 0, \quad \psi = 0 \text{ and } \frac{\mathrm{d}v(x)}{\mathrm{d}x} = 0 \end{aligned} \tag{8.72}$$

and these give the velocity at a distance x away from the wall as

$$v(x) = \frac{E\varepsilon}{\eta}\left[\zeta - \psi(x)\right] \tag{8.73}$$

When $x \gg 1/\kappa$, the velocity no longer changes with distance and we have 'plug flow' of the liquid through the tube, that is, the velocity profile is flat over most of the tube radius with the change in velocity occurring in the region close to the wall. This is a situation which appears similar to 'wall slip' as the bulk of the liquid moves unsheared through the capillary, although it arises from a different origin. The maximum velocity, which

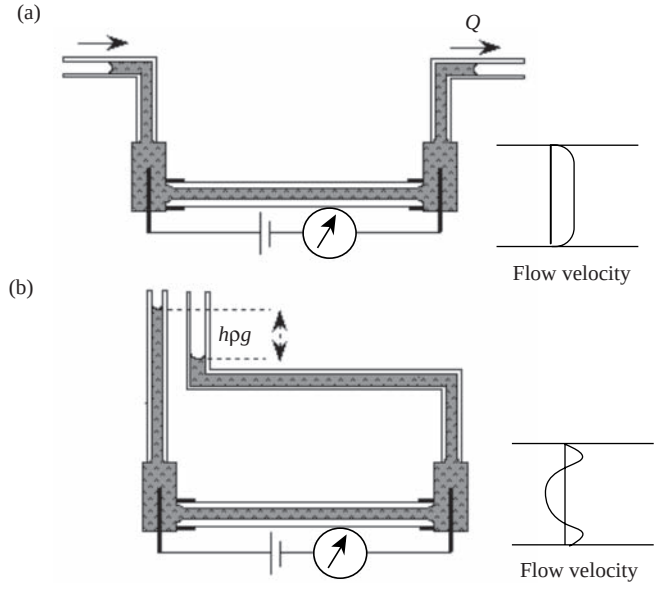


Figure 8.22 Schematics of two experimental arrangements used for the measurement of electro-osmotic flow in a capillary. In the sealed system (b), the liquid is forced to recirculate by moving back along the tube centre

we may call the electro-osmotic velocity, v_{eo}, from Equation 8.73, is given by the following:

$$v_{eo} = \frac{E\varepsilon}{\eta}\zeta \tag{8.74}$$

Two schematic arrangements whereby the electro-osmotic flow in a capillary tube of radius a can be measured are illustrated in Figure 8.22. A packed bed of coarse particles or fibres could be substituted for the capillary in each case. The bed would need to be held in place by perforated electrodes. Figure 8.22a is the simplest arrangement. Here, the volumetric flow rate, Q, is measured and this is equal to the electro-osmotic velocity multiplied by the cross-sectional area of the capillary. In addition, Ohm's law can be used to cast the relationship in terms of the measured current and electrolyte concentration:

$$EL = i\frac{L}{c\Lambda\pi a^2} \tag{8.75}$$

The flow rate is then given by

$$Q = i\frac{\varepsilon}{c\Lambda\eta}\zeta \tag{8.76}$$

and a graph of the flow rate versus the current yields the ζ-potential from the slope.

In Figure 8.22b, the exit tubes are arranged vertically and the result of the applied field is to produce flow resulting in a difference in height of the liquid in the two limbs. Under steady conditions, the liquid appears static with a difference in the head. This does not mean that the flow near the capillary walls has ceased. What is happening is that the liquid is forced to recirculate back down the centre of the tube. The flow profiles of both arrangements are sketched in Figure 8.22. Making use of the Poiseuille equation (Equation 8.61), we can equate the two flow rates and then the electro-osmotic pressure becomes

$$h\rho g = i\frac{8L\varepsilon}{\pi a^4 c\Lambda}\zeta \tag{8.77}$$

A plot of the pressure versus the current then enables the ζ-potential to be determined.

As with streaming potential, this technique may be used for material which can be made in the form of capillaries, to study the adsorption of surfactants or proteins on those surfaces and also for packed beds of particles and fibres.

8.9.5 Sedimentation Potential

This is also known as the *Dorn effect*. As large colloidal particles sediment in an electrolyte, there is relative motion between the particles and the fluid in a similar manner to fluid flowing through a packed bed in a streaming potential experiment. A potential is developed between the top and bottom of the sedimenting region. The velocity of the motion is governed by Stokes law for very dilute systems and is slowed by the 'back-flow' of liquid as the volume fraction of the suspension exceeds 0.1. The experimental difficulties in using this effect to estimate the ζ-potential then are first that the potential is small as the velocity is small, and second that in order to vary the velocity, a centrifuge would be

required and this makes the measurement of the potential rather difficult. Hence this is not a methodology that is of more than passing interest.

8.9.6 Particle Electrophoresis

This is the 'inverse' of the sedimentation potential as, in this case, we apply an electric field and measure the velocity of the particles. This is the most widely used technique for the determination of the ζ-potential of colloidal particles. It is also the most widely studied of all of the electrokinetic techniques with a wide variety of cell types being used. Traditionally, cylindrical cells were used with an ultra-microscope arrangement to view the particles. The ultra-microscope illuminates the particles at 90° to the viewing direction so that the particles are observed as bright spots due to the light scattered from them. The motion was then measured as a function of the applied field. Currently, laser illumination is employed and the motion is measured using a variant of photon correlation spectroscopy. The results are then given as the electrophoretic mobility, u_e, which is the velocity of the particle per unit applied field strength:

$$u_e = \frac{v_e L}{E} \quad \text{(in units of m}^2\ \text{s}^{-1}\ \text{V}^{-1}\text{)} \tag{8.78}$$

Large particles will sediment and this vertical contribution to their velocity can be ignored as long as the electric field is applied and the motion is measured in a horizontal direction. Brownian motion, however, is random and not unidirectional and so cannot be eliminated from the measurement. This effect is superimposed on the measured particle motion and is responsible for a spread of velocities with even the most uniform dispersions. For example, if we have a particle with a diameter of 0.5 μm and $\zeta \sim 70$ mV, we would expect a 5% spread in the measured electrophoretic mobility values just from the Brownian motion alone. If a cylindrical cell is used for the measurements, electro-osmotic flow will occur due to charges on the wall of the cell. If a sealed system is used, measurements can be made at the point of the recirculating electrolyte where there is no net motion of the fluid. There is a cylindrical surface inside the cell where the motion of the liquid in one direction is just counterbalanced by the motion in the other direction and this level is known as the *stationary level* (see Figure 8.22b). The net velocity is

calculated from summing the electro-osmotic velocity and the velocity from the Poiseuille flow in the reverse direction. The null position is at a distance of $a/\sqrt{2}$ from the centre of the tube. The equipment must be carefully adjusted to make measurements at this position. However, the measurements can also be made between electrodes in a an optical cuvette where there is no cell surface between the electrodes. This completely avoids the problems stemming from the use of capillary cells and the cells can be placed in dynamic light-scattering equipment with very little modification.

8.9.6.1 The mobility at $\kappa a > 100$

With this condition, the diffuse layer is small compared with the radius of the particles. This enables us to treat the diffuse layer as planar relative to the particle surface and both the lines of force from the electric field and the fluid flow lines are then taken as being parallel to the particle surface. Hence we have the condition for a large particle at a moderate to high electrolyte level. This is the case which is nearly always assumed in the software that is supplied with most pieces of equipment when the ζ-potential is printed out as opposed to the mobility.

Figure 8.23 illustrates the model that is employed for the calculation. The particle is moving at a velocity v_e through the electrolyte solution due to the application of an electric field and hence there is a velocity

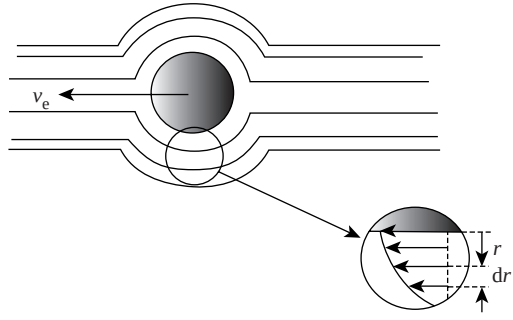


Figure 8.23 Model for the electrophoretic mobility used to derive the Smoluchowski equation (Equation 8.84) with the simplified flow lines around a particle moving with a velocity v_e due to an applied field. The particle is assumed to be non-conducting and the electrical field is also distorted in order to be parallel to the surface of the sphere

gradient close to the particle surface and we have an analogous situation to that which we discussed with electro-osmosis. When we consider a small element of unit area at a distance r from the surface, the change in the viscous drag on this element is the result of the change in the velocity gradient across it. So, by integrating with respect to r, we obtain the viscous stress at the surface of the particle:

$$-\eta \int \frac{\mathrm{d}\dot{\gamma}}{\mathrm{d}r}\mathrm{d}r = -\eta \int \frac{\mathrm{d}^2 v(r)}{\mathrm{d}r^2}\mathrm{d}r \tag{8.79}$$

The negative sign just indicates that the velocity decreases as we go further away from the particle surface. Now, the electrical force is the charge multiplied by the field strength, and the total charge on the particle, σ_{T}, is the charge density at the shear plane multiplied by the area of the particle:

$$E\sigma_{\mathrm{T}} = -E \times 4\pi a^2 \varepsilon \left.\frac{\mathrm{d}\psi(r)}{\mathrm{d}r}\right|_{r=0} \tag{8.80}$$

The planar form of Poisson's equation (Equation 8.63) can be used as the diffuse layer dimension is so much smaller than the radius of the particle. When there is steady motion, we equate the forces. To do this, we must multiply the viscous stress by the surface area of the particle:

$$E \times 4\pi a^2 \varepsilon \left.\frac{\mathrm{d}\psi(r)}{\mathrm{d}r}\right|_{r=0} = 4\pi a^2 \eta \frac{\mathrm{d}v(r)}{\mathrm{d}r} \tag{8.81}$$

In order to calculate the velocity, we must integrate with respect to r. giving

$$E\varepsilon\zeta = \eta v_{\mathrm{e}} \tag{8.82}$$

by using the boundary conditions $r = 0$, $\psi(r) = \zeta$, $v = v_{\mathrm{e}}$. In terms of the mobility, we have

$$u_{\mathrm{e}} = \frac{\varepsilon\zeta}{\eta} \tag{8.83}$$

This is the *Smoluchowski equation* for the electrophoretic mobility of colloidal particles. Note that there is no size term in the equation as the surface area terms cancel out.

8.9.6.2 The mobility at $\kappa a \ll 1$

In this limit, the particle is assumed to act as a point charge so that the lines of force of the electric field are unperturbed by the non-conducting volume of the particle. Again, as we have steady motion, we balance the forces on the particle. In this model, the drag on the particle is the sum of two terms. The first of these, the viscous drag, is taken as the velocity multiplied by the Stokes drag factor, $v_e \times 6\pi\eta a$. The second force slowing the particle motion is the electrical force on the ions of the diffuse layer as they will try to migrate in the opposite direction to the particle as their charge is of opposite sign. Hence, in balancing the forces, we can write

$$\sigma_T E = 6\pi\eta a v_e + E f\,[\rho_c(r)] \tag{8.84}$$

where r is the distance from the particle centre. The electrophoretic velocity is now

$$v_e = -\frac{E \times 4\pi a^2 \varepsilon}{6\pi\eta a} \left.\frac{d\psi(r)}{dr}\right|_{r=a} - \text{correction term} \tag{8.85}$$

The correction term is due to the motion created by the movement of the diffuse layer ions due to the applied electric field. It is straightforward to calculate the force due to the field on a thin spherical shell in the diffuse layer as

$$4\pi r^2 E\rho_c(r)dr = 4\pi r^2 E\varepsilon \frac{d^2\psi(r)}{dr^2} \tag{8.86}$$

This force would induce a motion so that there would be a relative velocity, $dv(r)$, between layers which would produce the drag force between adjacent layers, so approximating each layer to a continuous shell and balancing the drag the result is:

$$6\pi\eta r\,dv(r) = 4\pi r^2 E\varepsilon \frac{d^2\psi(r)}{dr^2} \tag{8.87}$$

Integrating with the boundary condition that at large distances the potential gradient is zero and rearranging to give the liquid velocity at a

long distance from the particle, v_l, yields

$$v_l = \frac{2E\varepsilon}{3\eta} \int \frac{d^2\psi(r)}{dr^2} dr \tag{8.88}$$

We can now substitute Equation 8.88 into Equation 8.85 and with integration by parts with the boundary condition that at $r = a$, $\psi(r) = \zeta$:

$$v_e = \frac{2\varepsilon E\zeta}{3\eta} \tag{8.89}$$

and expressing this in terms of the electrophoretic mobility gives

$$u_e = \frac{2\varepsilon\zeta}{3\eta} \tag{8.90}$$

This limiting equation is the *Hückel equation*. The mobility that it yields is less than that we would expect for the Smoluchowski equation. However, both are limiting approximations and we require a better estimate for many colloidal dispersion when we have $1 < \kappa a < 100$. Over most of the size range we cannot treat the particle as a point charge in a uniform field but we have to allow for the distortion of the electric field around the particle. This means that the field is variable with position around the particle and only when the particle is very large can the field be considered to have lines of force parallel to the surface at all points (as in the Smoluchowski model). We need a distortion which is a function of radius. The local field acts on the diffuse layer ions and alters the retardation correction to the particle velocity [1, 18]:

$$u_e = \frac{2\zeta\varepsilon}{3\eta} [1 + Kf(\kappa a)] \tag{8.91}$$

where K is calculated from the conductivity of the particles and the medium, and for non-conducting particles $K = 1/2$. This is the *Henry equation*. The function of κa can be written as a power series. For example for $\kappa a > 5$:

$$\frac{f(\kappa a)}{2} = \frac{1}{2} - \frac{3}{\kappa a} + \frac{25}{(\kappa a)^2} - \frac{220}{(\kappa a)^3} + \ldots \tag{8.92}$$

This function provides a smooth transition from the Hückel result to the Smoluchowski result as κa increases to values of ~100.

However, there also needs to be a correction for the distortion of the diffuse layer from spherical symmetry. A full analysis requires numerical computation as there is no simple analytical result. Some of the results of such calculations from the literature were published in tabular form by Ottewill and Shaw [19] and these show that for a 1:1 electrolyte the Henry equation (Equation 8.91) provides an adequate description, although this is not the case with multivalent species such as 2:2 electrolytes, for example, where there can be considerable deviations in the intermediate κa regime and then a numerical computation is required [20].

8.9.7 Electro-acoustics

We have already seen how a DC potential produces the motion of an electrolyte past a surface. Now, if an AC potential is applied to a dispersion, we can expect the particles to follow the oscillating field, as will the diffuse layer distortion. When the frequency is increased to levels in excess of ~1 MHz, the particle inertia can become large enough to prevent significant particle motion although the ions can still respond, producing a pressure wave which can be monitored with an ultrasonic transducer. Alternatively, an ultrasonic vibration can be applied and the relative motion of the diffuse layer produces an AC potential (cf. the streaming potential). When the frequencies exceed ~20 MHz, the diffuse layer can no longer respond, the signal is lost and the particles appear to be uncharged. The process has been analysed in detail [21] and is used to determine a dynamic mobility, u_d. In addition, the frequency variation may be used to determine the mean particle diameter.

The current commercially available equipment applies an ultrasonic vibration and measures the resultant AC potential. The particle inertia is both a function of the particle size and the difference in the density, $\Delta\rho$, between the particles and the medium. The technique operates most efficiently with dense particles. In addition, the amplitude of the signal is proportional to the volume fraction of the dispersion as the higher the volume fraction, the more oscillating dipoles there are and so the stronger the signal. We may use this to advantage as there are no optical requirements as we had in electrophoretic measurements and work at moderate volume fractions when the dispersions are opaque. The

dynamic mobility is given by

$$u_{\mathrm{d}} = \frac{\varepsilon\zeta}{\eta} G(\varphi, \Delta\rho) \tag{8.93}$$

where G is a correction factor that is a function of the dispersion concentration, and also the particle and medium densities. The current state of the theory assumes that the diffuse layers around each particle do not interact with that of the neighbouring particles. This means, however, that we are limited to dilute dispersions and we must check that the mobility is linear with concentration. For large particles and thin diffuse layers, we could still work with a high solids content. However, with small particles and dilute electrolyte concentrations, we can often be limited to $\varphi < 0.1$.

8.10 VISCOSITIES OF DILUTE DISPERSIONS

Measurement of the viscosity of a dispersion is a useful method of characterizing the particles in that dispersion. The viscosity is, of course, sensitive to the concentration of particles but, in addition, it is a function of particle charge and shape, and also the dimensions of any adsorbed layers. Capillary viscometry can be precise and is most suitable for use with dilute dispersions. This is because, as we have noted above, the fluid velocity has a parabolic profile so when we calculate the shear rate from $\mathrm{d}v(r)/\mathrm{d}r$ we find that there is a linear change of shear rate from zero at the tube centre to a maximum value at the wall. For example, a wall shear rate in excess of $10^3\ \mathrm{s}^{-1}$ is common. As a result, we should avoid systems which are non-Newtonian; that is, there is not a linear dependence of shear stress on the shear rate. However, this is not normally a problem with dilute dispersions.

8.10.1 Dependence on Volume Fraction

Particles in a flowing fluid produce a dilation of the field as the fluid flows around them. Figure 8.24 illustrates schematically the flow around single particles and also how pairs of colliding particles interact. Particles move at the velocity of the streamline in line with the particle centre. The fluid flowing past the upper surface of particles in this figure is moving more rapidly than the particles and conversely that at the lower surface is moving more slowly. This produces a rotation of the particles and the

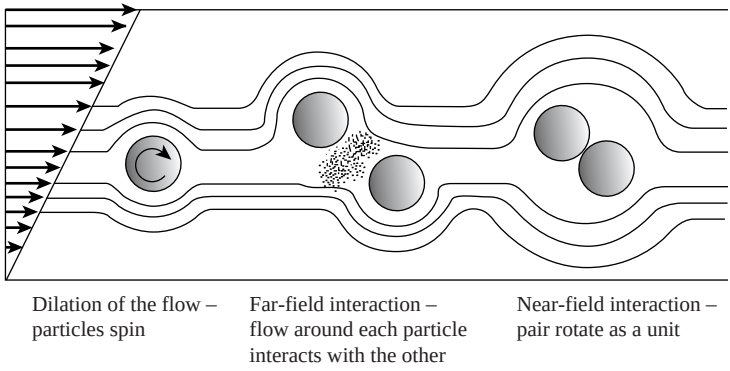


Figure 8.24 Schematic illustration of the interactions of particles in shear flow

rotation rate is equal to half the shear rate, $\dot{\gamma}/2$. This is the *vorticity* of the shear field.

Figure 8.24 also shows the manner in which colliding pairs of particles interact with the flow field. The 'near-field' interaction is how we first think of a collision, in that the particles come together and rotate as a single unit. The 'far-field' interaction, however, must also be included. In this case, the particle centres are not on a collision trajectory but are close enough that the dilation of the flow around one particle is disturbed by that around the other. These interactions with the shear field enhance the energy dissipation rate and we measure this as an increase in the viscosity. The hydrodynamics of the interactions are well established and we may write the viscosity of dispersions of hard spheres in shear flow as follows:

$$\eta(0) = \eta_0 \left[1 + 2.5\varphi + 6.2\varphi^2 + \mathrm{O}\left(\varphi^3\right)\right] \tag{8.94}$$

where φ is the volume fraction of the dispersion and η_0 is the viscosity of the continuous phase. The viscosity of the dispersion is $\eta(0)$ and we are using the '(0)' to indicate the low shear limiting behaviour so that the spatial arrangement of the particles is not perturbed by the shear rate. This condition means that the diffusional time for a particle, τ, is less than the characteristic shear time, so that $\tau\dot{\gamma} < 1$. The first coefficient in Equation 8.94 was derived by Einstein [22] and the second by Batchelor [23]. We do not have a rigorous hydrodynamic derivation of the coefficient for the three-body collision and the equation is therefore limited to dispersions where $\varphi < 0.1$.

Many flows have an extensional component and some can be mainly extensional with only a small shear component. When a fluid is sprayed, there is a large extensional contribution as the fluid is forced through a small nozzle. This is also the case in blade and roller coating. For extensional flow, the appropriate relationship is [23]

$$\eta(0) = \eta_0 \left(1 + 2.5\varphi + 7.6\varphi^2 + \mathrm{O}(\varphi^3)\right) \tag{8.95}$$

This indicates that we can expect differences in the viscous behaviour of dispersions of moderate to high concentration in different types of flow as then the two- and higher-body interactions dominate the response. It is also important to note that it is *never* possible to achieve a steady extensional flow and we must be aware that the residence time in the extension regime should also be considered. Reviews of dispersion viscosity can be found in the books by Hunter [1] and Goodwin and Hughes [24].

Every deviation from the particles being hard spheres results in changes in the Einstein and Batchelor coefficients. Hence the dispersion viscosity may be used to give information on adsorbed layers, particle charge, particle shape and the fluid nature of the particles. The experimental problem is that we must work with dilute dispersions in order to have rigorous hydrodynamic descriptions, which means that the variation in the viscosity can be fairly small. We can rewrite Equation 8.94 in the general form

$$\frac{\eta(0)}{\eta_0} = 1 + [\eta]\,\varphi + k_{\mathrm{H}}\,[\eta]^2\,\varphi^2 + \ldots \tag{8.96}$$

where $[\eta]$ is the intrinsic viscosity and k_{H} is the *Huggins coefficient* [24]. It is usual to express this in linear form, as follows:

$$\eta_{\mathrm{red}} = \left[\frac{(\eta(0)/\eta_0) - 1}{\varphi}\right] = [\eta] + k_{\mathrm{H}}[\eta]^2\varphi \tag{8.97}$$

A plot of the reduced viscosity, η_{red}, versus concentration then yields the intrinsic viscosity and deviations from intrinsic viscosity values of '2.5' indicate deviations from hard-sphere behaviour. However, even if that difference is 50%, this will only change the viscosity of a dispersion at a volume fraction of 0.05 by ~6%. Now, if we are using a capillary viscometer where we are measuring an efflux time of ~200 s, then 6% would mean a change of 12 s and we would like to obtain good precision

so that would mean that subsequent times should agree within <0.5 s at worst. This is a demanding experiment.

8.10.2 Adsorbed Layers

Stabilizer layers are often present on colloidal particles. These may be short-chain surfactants or polymers. In order to optimize the stabilizing properties, the surface concentrations are maximized and hence the layers cannot be regarded as 'free draining'. This means that the fluid must flow around the layer and not through it. With polymer layers, this will be an approximation as there can be some fluid motion within the layer, although much restricted. At the outer periphery, where the population is the longer tails, significant flow will occur. Hence the layer thickness must be termed a *hydrodynamic thickness* and this may be an underestimate of the particle–particle interaction thickness. The way in which the adsorbed layer is included is by using an *effective volume fraction*, φ':

$$\varphi' = \varphi \left(\frac{a + \delta}{a} \right)^3 \tag{8.98}$$

in Equation 8.94, where δ is the adsorbed layer thickness. In terms of the intrinsic viscosity, we could write

$$[\eta] = 2.5 \left(\frac{a + \delta}{a} \right)^3 \tag{8.99}$$

An adsorbed layer thickness of 10% of the particle radius will increase the intrinsic viscosity by 33%. As most stabilizer layers have $\delta < 10$ nm, this method for determining the thickness is limited to particles with a radius less than ~150 nm if the results are going to have reasonable precision.

8.10.3 Fluid Droplets

When the disperse phase is a fluid, the forces exerted by the shear field can deform the particle. The force maxima are at 45° to the flow direction and so the droplets become elongated as the shear force varies from a compression to a tension as we scan the angle to the shear direction from 0° through 45° to 135°. In addition to the continuous deformation

of the droplet, the field vorticity results in the outer surface moving continuously around the droplet, resulting in some circulation within. Taylor [25] carried out the analysis for a droplet where the interface allows the ready transfer of energy across it, that is, it is thin and not rigid. In this case, the circulation of the fluid inside the drop resulted in a decrease in the rate of energy dissipation compared with that for rigid particles and with the viscosity of the internal phase being η_i we have the intrinsic viscosity as

$$[\eta] = 2.5\left(\frac{\eta_i + 0.4\eta_0}{\eta_i + \eta_0}\right) \tag{8.100}$$

For this equation, we can consider the following limiting conditions:
$\eta_i = \infty,\ [\eta] = 2.5;\ \ \eta_i = \eta_0,\ [\eta] = 1.75;\ \ \eta_i = 0,\ [\eta] = 1.$

In many emulsions, the viscosities of the two phases are similar and so the middle condition will be reflected. The last condition is relevant to gas bubbles and is what we would find from measuring the viscosity of low-quality foams, that is dilute foams. In the case of a particle with a thick, low-density polymer layer, we can expect some analogous behaviour in that some distortion of the layer could occur and some motion of the liquid within the layer would decrease the overall rate of energy dissipation compared with a rigid particle of the same size as the particle plus polymer layer and the intrinsic viscosity would be less than 2.5. Interpreting the viscosity then to give the layer thickness would clearly lead to an underestimate. However, the detailed hydrodynamics would be rather different so the analogy should not be taken as far as attempting to estimate a 'layer viscosity'.

8.10.4 Electroviscous Effects

There are a number of electroviscous effects and these have been reviewed extensively by van de Ven [26]. The first one to be identified, and hence known as the *primary electroviscous effect*, was the effect of the distortion of the diffuse layer caused by the flow of liquid past the particle. The similarity to electrophoresis and electroacoustics immediately comes to mind. The result for the intrinsic viscosity for large κa and small potentials is

$$[\eta] = 2.5 + \frac{6\,(\varepsilon\zeta)^2}{K\eta_0 a^2} \tag{8.101}$$

where K is the specific conductance of the electrolyte solution. This equation was derived by von Smoluchowski and has the same restrictions as his equation for the electrophoretic mobility (Equation 8.83). As the value of κa falls to below 100, we should have a variation with κa and, although there are analyses available which have analytical approximations, the solution for any potential and any κa value must be carried out by numerical calculation [27]. However, under most conditions the increase in the value of the intrinsic viscosity above the value of 2.5 is small and requires great precision in the viscosity measurement if an estimate of the ζ-potential is required. The exception is for small particles in dilute electrolyte, and then the numerical calculation is required.

The interaction between to charged particles in flow leads to a change in the second coefficient in Equation 8.95. This effect is the *secondary electroviscous effect*. The most complete treatment was given by Russel *et al.* [28], which gives

$$\frac{\eta(\dot{\gamma})}{\eta_0} = 1 + [\eta] + \left[2.5 + \frac{3}{40}\left(\frac{r_0}{a}\right)^5\right]\varphi^2 \qquad (8.102)$$

where r_0 is the centre-to-centre distance between the particles as they collide. It is calculated from the balance between shear forces pushing the particles together and the electrostatic repulsion from the particle surfaces. Thus r_0 will vary from $\sim 2a$ at very high shear to much larger values at low shear. Of course, at very low shear, Brownian forces are still operating and they will control the closes distance of approach. In many stable colloidal systems, we find that at a surface separation of $\sim 5/\kappa$ there is $\sim 1 k_B T$ of repulsive energy and so as a first estimate we should have $r_0 \approx 2a + 5/\kappa$ as a low shear limit. Note that we now have a shear dependence of the viscosity, that is, the dispersion is non-Newtonian and this means that we will no longer have a simple interpretation of the flow times as the dispersion viscosity varies across the capillary due to the large variation of shear rate with radius.

Polyelectrolytes are effective stabilizers of colloidal dispersions as they provide an 'electrosteric barrier' to aggregation. The charged groups along the chains of the polyelectrolytes repel each other and causes the chain to be in an expanded conformation. Added electrolyte and, if the charged groups are weak acid groups, changes in pH cause changes in the conformation. This is the *tertiary electroviscous effect* as the dimensions of the polymer effect the value of the viscosity. When the polymer is attached to a particle surface, the adsorbed layer can be

expanded or contracted by altering the chemical environment – thus δ in Equation 8.99 will change.

8.10.5 Particle Shape

The rotation of anisometric particles in a shear field is periodic and the orbit depends on the initial orientation. The hydrodynamic analysis is available for prolate and oblate ellipsoids [29] but not for particles of arbitrary shape. In order to introduce the problem, we shall consider the limiting rotational orbits of a rod in a shear field. When the major axis of the rod is aligned with the flow direction, the rod will rotate 'end-over-end' with a constantly changing angular velocity dependent on the angle of the axis to the shear plane. The rod effectively 'flips' over. This orientation is one that corresponds to a high energy dissipation rate compared with a sphere of the same volume. However, the rod could be aligned with its major axis in the shear plane but perpendicular to the flow direction. In this orientation, it simply 'rolls' and this corresponds to a lower rate of energy dissipation than for a sphere of similar volume. Rotary Brownian motion will tend to randomize the orientations and so we have to average over all possible configurations. If the Brownian motion is weak compared with the shear field, there are solutions available in the literature [29]. The calculations indicate that axial ratios >5 are required before the increase in the intrinsic viscosity becomes large enough to be measured with precision. However, when the Brownian motion is strong compared with that from the shear field, the effect is stronger as the time aligned with the flow is reduced. In addition, we find that *shear thinning* occurs when the rotational rates are of the same order. Thus, for an axial ratio of 5, the intrinsic viscosity increases by 30% at high shear rates whereas the increase is 100% at low shear rates [29].

8.11 SEDIMENTATION OF DISPERSIONS

If particles in a colloidal dispersion are of a different density to the continuous phase, gravity will cause them to move in a uniform direction. If the density of the particles is greater than that of the medium, the net motion is towards the bottom of the container and is referred to as *sedimentation*. Alternatively, if the density of the particles is less than that of the medium, the motion is towards the upper surface and is termed

creaming and is frequently observed with emulsions and dispersions of gas bubbles. The rate at which the particles move is a characteristic of the particle density, size and shape. Measurement of the sedimentation velocity was frequently used to estimate the size of particles in the upper portion of the colloid size range. However, at 1*g* it is slow and has largely been superseded by more rapid, mainly optical techniques. The sedimentation properties though are still of great importance.

We can start the discussion of the problem with a spherical particle suspended in a 'simple' or Newtonian liquid. The gravitational body force on the particle causes it to move down and the motion is resisted by the Stokes drag force. When the motion, v_0, is steady, these two forces are in equilibrium:

$$\frac{4\pi a^3 \Delta\rho g}{3} = v_0 6\pi \eta_0 a \tag{8.103}$$

which can be rearranged to give Stokes law:

$$v_0 = \frac{2a^2 \Delta\rho g}{9\eta_0} \tag{8.104}$$

8.11.1 The Disc Centrifuge

This technique allows us to measure the particle size distribution of a dispersion by following the sedimentation rate. The use of high values of *g* makes the technique rapid and convenient to use. The equipment consists of a narrow, hollow disc, about 20 cm in diameter and with transparent sides. In operation, the disc is set spinning and filled with liquid medium through an opening on the axis of the disc. A thin layer of dilute dispersion is added to the surface and the progress of the particles is tracked optically by light absorption. Assuming that the particles are of uniform shape and density, the larger particles track through the light beam first and the smaller particles last. The optical density plotted against inverse time reflects the fraction of particles as a function of size. The *g* force increases with radius and is not constant within the fluid. In order to reduce the acceleration of the particles as they move radially outwards within the disc, a density gradient is produced as the medium is initially added to the disc, that is, the body

force is

$$\mathbf{F} = f\left(r\frac{\mathrm{d}\rho_0}{\mathrm{d}R}, \Omega^2 r\right) \tag{8.105}$$

where ρ_0 is the density of the medium, Ω is the angular velocity of the disc and r is the radial distance from the centre of the disc. This permits a better resolution of the particle size distribution. Although a stable gradient can be set up initially, the quality is gradually compromised as large numbers of samples are analysed.

The instrument is calibrated using dispersions of monodisperse particles. It provides a rapid method of sizing for a wide range of particle sizes. However, the data are not reliable if the dispersions are very heterogeneous with respect to particle density or shape, such as can be the case for systems that have been flocculated and subsequently redispersed mechanically.

8.11.2 Batch Sedimentation

When we place a well-mixed colloidal dispersion in a closed container, such as a graduated cylinder or a can, the particles sediment at a rate which is a function of the particle diameter, the density difference between the two phases, the viscosity of the continuous phase and the volume fraction of the dispersion. Consider a fairly dilute dispersion ($\varphi \approx 0.1$) of monodisperse hard spheres. After a period of time, we can identify four distinct regions. At the top of the container a clear region devoid of particles has developed. Below this there is a free-settling region where the volume fraction is $\varphi \approx 0.1$ and the particles are moving at a constant velocity. A sediment can be observed at the bottom of the container and as the system is a dispersion of hard spheres, the volume fraction of the sediment is $\sim$0.64 – the value for dense random packing. Just above the sediment there is a transition zone were the concentration gradient is governed by the laws of diffusion; see the analyses of Russel *et al.* [28] and Kynch [30]. We can define a Péclet number for the sedimentation in this region by equating the sedimentation rate and the Brownian diffusion rate away from the concentrated sediment as

$$Pe_{\mathrm{sed}} = \frac{4\pi a^4 \Delta\rho g}{3k_{\mathrm{B}}T} \tag{8.106}$$

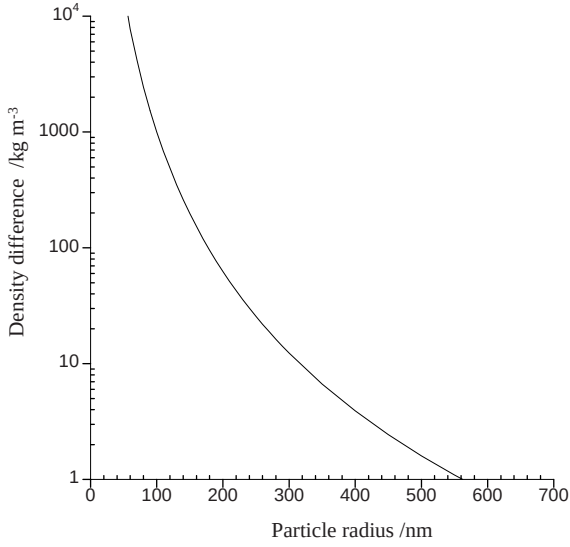


Figure 8.25 The density difference between particles and medium at which a sediment will form for particles of different radii with a Péclet number = 10^3 using Equation (8.106)

For a clear sediment to form, a value of $Pe_{\text{sed}} \approx 10^3$ is required, as shown by Figure 8.25, where the this value of the Péclet number is used with Equation 8.106 to calculate the density difference required for a sediment to form as a function of particle radius. The result indicates that a polystyrene latex dispersion will show a sediment over time if the particle diameter is $>0.5\,\mu$m and that a titanium dioxide pigment will form a sediment for particle diameters $>0.15\,\mu$m.

The sedimentation rate in the free-settling region is modified from the Stokes law value due to three factors. The mass transfer of material to the bottom of the container is balanced by the backflow of continuous phase. This means that the particles are moving down through liquid flowing in the opposite direction, thus hindering the relative settling rate by a factor of -5.5φ. The near-field hydrodynamic interactions between neighbouring particles further slows the relative rate by a factor of -1.55φ and finally there is a small enhancement from the pressure gradient of $+0.5\varphi$. The net result gives the sedimentation rate as [31]

$$v = v_0\left[1 - 6.55\varphi + O\left(\varphi^2\right)\right] \tag{8.107}$$

Implicit in the derivation of Equation 8.107 is that the spatial distribution of the particles was random. When the pair potential of the particles takes the systems away from the hard-sphere limit, we see changes to the volume fraction dependence. With a long-range repulsion, the closest distance of approach increases and reduces the near-field interactions but the backflow increases the retardation [31–33] so that the coefficient is now >6.55. Secondary minimum attraction, however, also perturbs the random distribution and increases the near-field interaction but the clusters have a lower contribution from the backflow so that the coefficient is <6.55.

8.11.3 Sedimentation in Polydisperse Sols

We can write the relative rate of sedimentation as

$$\frac{v}{v_0} = 1 + K\varphi \tag{8.108}$$

where $K = -6.55$ for monodisperse uncharged spheres. Batchelor and Wen [32] laid the theoretical groundwork for the sedimentation of spherical particles with discrete particle size and density differences. Their result gives the sedimentation rate for spheres of type i at a constant volume fraction as

$$\frac{v_i}{v_{0i}} = 1 + \varphi \sum_{j=1}^{m} K_{ij} \frac{\varphi_j}{\varphi} + O(\varphi) \tag{8.109}$$

where

$$\begin{aligned} K_{ij} &= f(\lambda_s, \gamma_s) \\ \lambda_s &= \frac{a_j}{a_i}; \quad \gamma_s = \frac{\Delta\rho_j}{\Delta\rho_i} \end{aligned} \tag{8.110}$$

The simplest example is the case of large spheres (i) settling in a dispersion of small (j) spheres [31–33]. In this case, $\lambda_s < 1$ and only the viscous drag and the backflow are important, so that

$$K_{ij} = -2.5 - \gamma_s + O(\lambda_s) \tag{8.111}$$

Table 8.4 Estimated sedimentation coefficient for 248 nm PVC particles (density 1380 kg m^{-3}) in a mixture with 29 nm PS particles (density 1.054 kg m^{-3})

φ_i/φ_j	$K = \sum K_{ij}\frac{\varphi_j}{\varphi}$	K
2:1	$-\left(\frac{6.55 + 2 \times 2.69}{3}\right)$	−3.98
1:1	$-\left(\frac{6.55 + 2.69}{2}\right)$	−4.62
1:2	$-\left(\frac{2 \times 6.55 + 2.69}{3}\right)$	−5.26

Equation 8.112 indicates that the larger spheres are settling through an 'effective medium' with the viscosity of the dispersion of small spheres [$\eta = \eta_0(1 + 2.5\varphi_j)$] and with a density modified by $\gamma_s\varphi$. The sedimentation coefficient of the smaller sphere is given by [31]

$$K_{ji} = -\frac{1}{\gamma_s}\left(1 + \frac{3}{\lambda_s} + \frac{1}{\lambda_s^2}\right) + O(\lambda_s) \tag{8.112}$$

Partridge [34] studied the sedimentation of poly(vinyl chloride) (PVC) particles ($a_i = 248$ nm) in mixtures with small polystyrene (PS) particles ($a_j = 29$ nm). The experiments were carried out using 10^{-2} mol dm^{-3} sodium chloride in order to minimize any charge effects and the value of $\gamma_s = 0.148$. $K_{11} = -6.55$ (PVC particles interacting with each other); $K_{22} \to -\infty$ (the ps particles do not sediment at this size); and $K_{12} = 2.69$ from Equation 8.111 but with a small addition for the primary electroviscous effect using the calculations of Watterson and White [27]; $K_{21} \approx -10^3$ (Equation 8.108 in addition to the small size minimizing the sedimentation). Table 8.4 shows the estimated sedimentation coefficients for the PVC particles in the mixtures and the estimates were in reasonably good agreement with the experiments [34].

Russel *et al.* [20] explored the variation of the value of K_{ij} with variation of γ_s and λ_s. They concluded that neutrally buoyant spheres ($\gamma_s = 0$) simply increase the viscosity so $K_{ij} = -2.5$ (Equation 8.111) and that when the particle radii are equal:

$$K_{ij} = -2.52 - 0.13\gamma_s \tag{8.113a}$$

except that as

$$\gamma_s \to 1;\ K_{ij} \to -5.62 \tag{8.113b}$$

There are many opportunities to explore the theoretical predictions for model size distributions and compare those predictions with experiment.

REFERENCES

1. Hunter, R. J. (1987) *Foundations of Colloid Science*, Vol. **I**, Oxford University Press, Oxford.
2. Murphy, D. B. (2001) *Fundamentals of Light Microscopy and Electronic Imaging*, Wiley-Liss, New York.
3. Herman, B. (1998) *Fluorescence Microscopy*, 2nd edn, Springer, New York.
4. Life Technologies (2008) *The Handbook – A Guide to Fluorescent Probes and Labeling Technologies*, available at http://probes.invitrogen.com/handbook/; last accessed 1 December 2008.
5. Stine, K. J. and Stratman, D. T. (1992) *Langmuir*, **8**, 2509.
6. van Blaaderen, A., Imhof, A., Hage, W. and Vrij, A. (1992) *Langmuir*, **8**, 1414.
7. Verhaegh, N. A. M. and van Blaaderen, A. (1994) *Langmuir*, **10**, 1427.
8. Li, L., Sosnowski, S., Kumacheva, E. and Winnik, M. (1996) *Langmuir*, **12**, 2141.
9. Vodovotz, Y., Vittadini, E., Coupland, J., McClements, D. J. and Chinachoti, P. (1996) *Food Technol.*, **50**, 74.
10. Kerker, M. (1969) *The Scattering of Light and Other Electromagnetic Radiation*, Academic Press, New York.
11. van de Hulst, H. C. (1955) *Light Scattering by Small Particles*, Wiley-Interscience, New York.
12. Guinier, A. and Fournet, G. (1955) *Small Angle Scattering of X-rays*, Wiley-Interscience, New York.
13. Lindner, P. and Zemb, Th. (eds) (1991) *Neutron, X-ray and Light Scattering*, North-Holland, Amsterdam.
14. Rowell, R. L., Farinato, R. S., Parsons, J. W., Ford, J. R., Langley, K. H., Stone, J. R., Marshall, T. R., Parmenter, C. S., Seaver, M. and Bradford, E. B. (1979) *J Colloid Interface Sci.*, **69**, 590.
15. Hunter, R. J. (1989) *Foundations of Colloid Science*, Vol. II, Oxford University Press, Oxford.
16. Ottewill, R. H. (1982) Small angle neutron scattering, in *Colloidal Dispersions* (ed. J. W. Goodwin), Royal Society of Chemistry, London, pp. 143–164.
17. Pusey, P. N. (1982) Light scattering, in *Colloidal Dispersions* (ed. J. W. Goodwin), Royal Society of Chemistry, London, pp. 129–142.
18. Hunter, R. J. (1981) *The Zeta Potential in Colloid Science*, Academic Press, London.
19. Ottewill, R. H. and Shaw, J. N. (1972) *J. Electroanal. Chem.*, **37**, 133.
20. O'Brien, R. W. and White, L. R. (1978) *J. Chem. Soc., Faraday Trans. 2*, **74**, 1607.
21. O'Brien, R. W. (1988) *J. Fluid Mech.*, **190**, 71.
22. Einstein, A. (1911) *Ann. Phys.*, **34**, 591.

23. Batchelor, G. K. (1977) *J. Fluid Mech.*, **83**, 97.
24. Goodwin, J. W. and Hughes, R. H. (2000) *Rheology for Chemists – An Introduction*, Royal Society of Chemistry, Cambridge.
25. Taylor, G. I. (1934) *Proc R. Soc. London, Ser. A*, **146**, 501.
26. van de Ven, T. (1989) *Colloidal Hydrodynamics*, Academic Press, New York.
27. Watterson, I. G. and White, L. R. (1981) *J. Chem. Soc., Faraday Trans. 2*, **77**, 1115.
28. Russel, W. B., Saville, D. A. and Schowalter, W. R. (1989) *Colloidal Dispersions*, Cambridge University Press, Cambridge.
29. Leal, L. G. and Hinch, E. J. (1971) *J. Fluid Mech.*, **46**, 685.
30. Kynch, G. J. (1952) *Trans. Faraday Soc.*, **48**, 166.
31. Batchelor, G. K. (1972) *J. Fluid Mech.*, **52**, 245.
32. Batchelor, G. K. and Wen, C. S. (1982) *J. Fluid Mech.*, **124**, 495.
33. Reed, C. C. and Anderson, J. L. (1980) *AIChE J.*, **26**, 816.
34. Partridge, S. J. (1982) A study of sedimentation in mono- and polydisperse systems, MSc Thesis, University of Bristol.

9

Concentrated Dispersions

9.1 INTRODUCTION

Much of this volume has dealt with dilute colloidal dispersions where we have considered the properties of single particles or two particles interacting with each other. However, as we increase the concentration of a dispersion, multi-body interactions become increasingly important and we have a condensed phase. This occurs when the interparticle forces produce a structure which is space-filling. The forces may be strongly attractive, as occurs with clays used for the manufacture of ceramics, but structures may also be due to weakly attractive forces or indeed simply to interparticle repulsion. We saw early in the volume how concentrated aggregates of surfactant molecules produce three-dimensional structures and these are just examples of colloidal condensed phases.

The macroscopic properties of the structures formed by concentrated colloidal dispersions take us into the behaviour of thickened liquids and gels. Such systems are becoming increasingly referred to as *complex fluids*, *soft matter* or *soft solids*. There are two important questions that have to be addressed with such materials. The first of these is, 'what is the structure?'. Intimately connected with this is how it may be determined. The second question is, 'what are the physical properties?'. This question primarily concerns the rheological or handling properties of the dispersion. The first question can be successfully addressed for model colloidal dispersions such as monodisperse spheres, while in answering the second we can always make measurements to record the behaviour, which, although often of critical importance to commercially important systems,

Colloids and Interfaces with Surfactants and Polymers, Second Edition James Goodwin

may be difficult to predict for any but the more simple systems. However, the work on model systems has greatly increased our understanding and is an important guide to the interpretation of more complicated systems.

9.2 THE STRUCTURE OF CONCENTRATED DISPERSIONS

The scattering of radiation as a means of providing information on particle size and so on was described in Chapter 8. The scattering of X-rays by organized molecular structures as opposed to individual molecules is a concept that is generally familiar. This is also the case for the scattering of X-rays, neutrons and light from concentrated colloids. The angular intensity of scattered radiation is the product of that scattered from the individual particles and that from the periodic structure in the dispersion. Hence for a dispersion of monodisperse spheres we may write

$$I(Q) = \left(\rho_{\mathrm{p}} - \rho_{\mathrm{m}}\right)^2 N_{\mathrm{p}} V_{\mathrm{p}}^2 P(Q) S(Q) \tag{9.1}$$

To obtain $S(Q)$ from the intensity data, we measure the intensity of a dilute dispersion where $S(Q) = 1$ and then divide the $I(Q)$ which we obtained for the concentrated system with the values for the dilute one scaled by the ratio of the concentrations. The value of $S(Q)$ varies with Q in a periodic manner when Q^{-1} is of a similar magnitude to the periodicity in the structure of the dispersion. We are used to observing Bragg peaks from the scattering from a highly ordered structure such as we find with a crystal. We have an analogous situation with concentrated dispersions, although the interparticle forces often allow much more motion and the peaks tend to be broader and may be similar to the peaks that we find from molecular liquids. The structure is dynamic (from the Brownian motion) and we are observing an average over all the structures.

The problem then is to go from the structure factor to a description of the spatial arrangement of the particles. In principle, we should be able to carry out a Fourier transformation of the measured structure factor to calculate the pair distribution function, which gives us the local density of particles with reference to a central particle, that is, it is simply the probability of finding a particle at any distance r from the centre of a reference particle. This inversion may not be straightforward [1] and is not

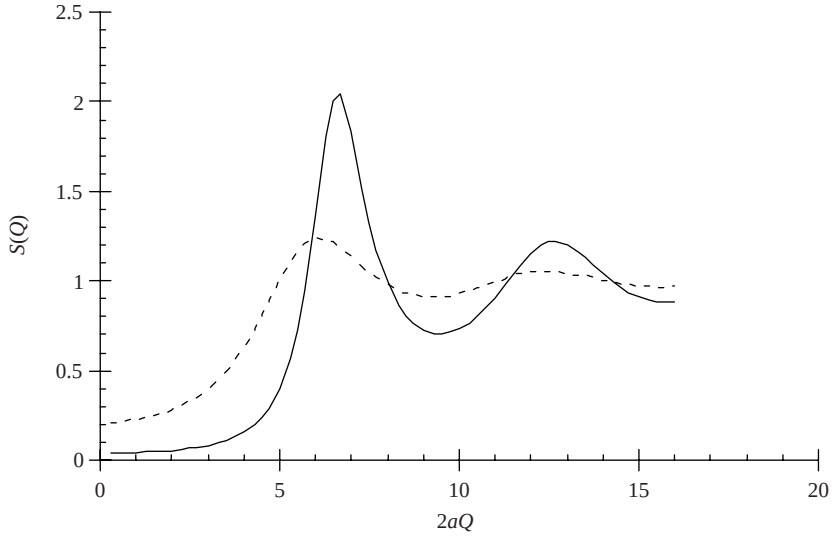


Figure 9.1 The structure factor calculated for dispersions of hard spheres at volume fractions of 0.2 (—) and 0.4 (- - -). Note that as $Q \to 0$, $S(Q)$ reduces at higher concentrations, indicating that the compressibility is lower, that is, the modulus is higher

the usual route. It is more common either to calculate a pair distribution function and then use that to determine an $S(Q)$ for comparison, or to use a computer simulation. Figure 9.1 shows structure factors calculated [1, 2] for 'hard-sphere' fluids. This figure serves to illustrate general points. The first is that as systems become more concentrated, the periodic structure becomes increasingly well defined. The second point is related to the behaviour at low Q values. Note from Figure 9.1 how $S(0)$ is greater for the lower concentration system. $S(0)$ is equal to the osmotic compressibility of the system [1, 3] and it should be no surprise that a more concentrated system is more difficult to compress – in other words, it is less *compliant*. The relationship between the pair distribution, $g(r)$, function and the structure factor is [1, 3]

$$S(Q) = 1 + \frac{4\pi\rho_0}{Q} \int_0^\infty r\,[g(r) - 1] \sin(Qr)\,\mathrm{d}r \qquad (9.2a)$$

where $\rho_0 (= N_p / V)$ is the average number density of particles in the dispersion. Now, the Fourier transform of this equation gives the pair

distribution function, $g(r)$ [1, 3]:

$$g(r) = 1 + \frac{1}{2\pi^2 r\rho_0} \int_0^\infty [S(Q) - 1] Q \sin(Qr)\, \mathrm{d}Q \qquad (9.2b)$$

where the limits of $g(r)$ are $r \to 0$, $g(r) \to 0$; $r \to \infty$, $g(r) \to 1$.

The shape of the pair distribution function is dependent on the details of the interactions between the particles and is related to the potential of mean force [1]:

$$g(r) = \exp\left[-\frac{\Phi(r)}{k_B T}\right] \qquad (9.3)$$

The potential of mean force is the reversible work done in bringing two particles together from an infinite separation. Of course, in the dilute limit this is just the colloid pair potential, but in a concentrated system there are all of the interactions with other particles along the way. Hence the potential of mean force has an oscillatory character and this shows in the structure of $g(r)$.

The structures illustrated in Figure 9.2 are for two common concentrated colloidal situations. The structure for the weakly attractive system (Figure 9.2a) shows the particles in close contact but, although the packing is dense, it is not highly ordered. This would be typical of, say, a depletion-flocculated system or a sterically stabilized system where there is no long-range repulsion. The pair distribution function, $g(r)$, is also shown and the shape is typical for a system which is weakly attractive with a very sharp peak corresponding to the first nearest-neighbour shell at $r/2a = 1$. Figure 9.2b presents the situation for a system with long-range repulsion. The separation between the particles would correspond to a volume concentration of ~25%. Here, the first nearest-neighbour peak corresponds to $\sim r/2a = 1.2$. There is slightly more short-range order here, but with such a large surface–surface separation (about $0.4a$) the Brownian motion results in a rapid fall in the structure relative to the central particle. As the concentration is increased with this sort of system, colloidal crystals can be formed. However, the usual structures are nearer to a *colloidal glass*. This is because a considerable time is required for the reorganization to occur to produce well-defined crystals [usually with face-centred cubic (fcc) symmetry]. In addition, the material is readily deformed and the structure reduced in size.

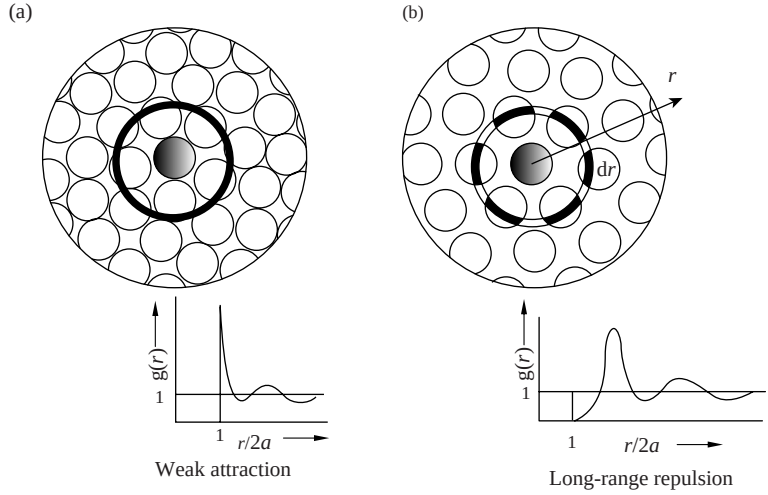


Figure 9.2 Illustrations of the structures in concentrated dispersions of particles with (a) weak attraction and (b) long-range repulsion between particles. The corresponding pair distribution functions are also shown

This means that we have a very important observation which actually applies to both systems. The structures that we are nearly always working with are non-equilibrium. At first sight, this would seem to be a major problem when we are characterizing the macroscopic behaviour. This is not the case, however, as we shall see as we can obtain good predictions of the behaviour in the laboratory by just considering the nearest-neighbour interactions to be the dominant term. However, we must be wary. For example, we may have a weakly flocculated concentrated dispersion which appears to be space-filling – in other words, we do not see sedimentation occurring. However, we may see that many hours later, the structure suddenly compacts and we observe a rapid sedimentation. This is simply the structure locally densifying to reduce the local energy. This can only occur at the expense of supports elsewhere, and at some point the 'overburden' becomes critical as the structure becomes progressively weaker in some regions concomitant with densification in other regions. We then observe sudden collapse. This represents an important problem in designing non-sedimenting dispersions which can be subjected to long-term storage.

There is one other structure which we should consider at this point. This is the type of structure that is formed by particles with a strong interparticle attraction, that is, in coagulated dispersions. If we carry

out computer simulations or careful experiments in the laboratory, these diffusion-limited aggregates have a fractal structure. In a concentrated dispersion, the growth of the fractal aggregates means that they interpenetrate. In practical terms, the complexity involved in attempting to describe such situations is rarely worth tackling as the systems that we use are all heavily shear-processed and this dominates the structure. For example, we may coagulate a monodisperse concentrated dispersion, but when sheared we find that we have a system of spherical aggregates in which the particles are in dense random packing [4]. These systems show a marked change, which is irreversible, when first subjected to shear, but subsequently change very little and so the initial state is of minor importance to that produced by shear processing.

Returning to the pair distribution function, we have seen how we have a periodic function which gives the concentration of particles at a distance r from the central particle relative to the global average number density. Hence, we can define the number of particles in a shell dr thick from the volume of the shell, the global average density and $g(r)$, as follows:

$$\text{number} = \rho_0 \times 4\pi r^2 g(r)\mathrm{d}r \tag{9.4}$$

Hence, if we integrate this expression over a distance corresponding to the first nearest-neighbour shell, we have the coordination number for the structure, z, as follows:

$$z = 4\pi\rho_0 \int_{2a}^{r_{\min}} r^2 g(r)\mathrm{d}r \tag{9.5}$$

Of course, the number density may be written in terms of the volume fraction of the dispersion and the particle volume, giving

$$z = \frac{3\varphi}{a^3} \int_{2a}^{r_{\min}} r^2 g(r)\mathrm{d}r \tag{9.6}$$

The results of such a calculation are shown in Figure 9.3 for the two situations illustrated in Figure 9.2, namely a weakly flocculated system and a strong repulsive system with long-range interactions. The coordination number for the weakly attractive system [5] corresponded to a synthetic polymer latex in 'high salt' but sterically stabilized with a

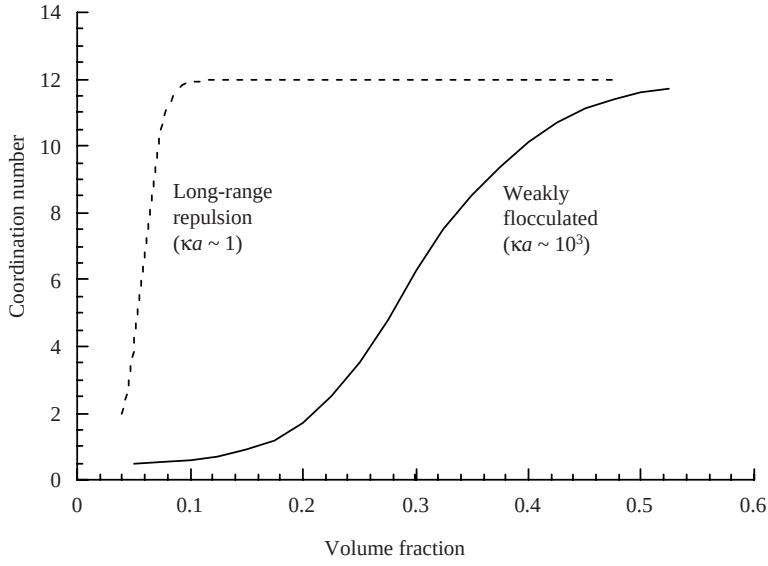


Figure 9.3 The coordination numbers calculated for the structures formed from both weakly flocculated particles and those with long-range repulsion between particles

non-ionic surfactant. The pair potential had an attractive minimum of $7k_BT$ at ~10 nm from the surface. The shape of the curve is sigmoidal, and as the volume fraction is increased we see that the coordination number increases most rapidly at $\varphi \approx 0.3$. What is happening in this structure is that the coordination number is increasing while the particles are at close separation, that is, in the potential minimum. On a practical application note, it is common to encourage the separation of particles by causing some aggregation and then centrifuging or using a filter press. When the coordination number increases rapidly, the structure becomes very strong and 'dewatering' becomes difficult. We noted that as $Q \to 0$, $S(Q)$ gives the osmotic compressibility. This is the excess osmotic pressure, that is, that due to the particle–particle interactions. Hence from Equation 9.2a we can see immediately that we expect the compressibility to be reduced as $g(r)$ increases. Figure 9.3 also shows the equivalent result for a system with long-range repulsion between the particles. The $\kappa a \approx 1$ condition means that the particles become highly ordered at a low volume fraction. When such a system is concentrated, the structure (fcc) is the same short-range structure throughout most of the concentration range but the separation is changing; which is qualitatively completely different from the weakly attractive system.

A consequence of the structure that we see developing as we concentrate dispersions is that *the compressibility of the structure* is reduced. This means that the structure resists changes in volume as we might expect to see in a centrifugation or filtration experiment. We are talking about the bulk elastic modulus of the structure (not including that of the suspension medium), so that we are observing solid properties with the concentrated dispersion. This resistance to shape change of the structure also shows in shear or extension experiments and the resistance is a function of $g(r)$ and the rate of change of the interparticle force with changes in separation. However, we also find that under high stress, or sometimes under low stress applied for long times, the systems flow. This area of study is *rheology* and is a key feature of most of the concentrated dispersions that we use every day.

9.3 RHEOLOGY

9.3.1 Definitions

When we study the deformation and flow of concentrated systems, we must define the forces and deformations carefully. The stress is defined as the force per unit area over which it is applied and so has dimensions of pascals (Pa) (the SI unit of pressure and stress). The strain is the deformation relative to the original dimension and so is dimensionless. An arbitrary applied stress that results in a deformation has to be described in all three dimensions. There are shear stresses in addition to stresses normal to the reference planes in the material. To describe this, we need to resort to tensor algebra and this has operational rules for the manipulation of the equations which may not be familiar to many of those interested in colloids. The complexity is avoided in the laboratory by carefully controlling the way in which the materials deform or the way in which the stress is applied. Thus, we can limit our mathematical manipulations to simple linear algebra and some complex number algebra in the contents of the present chapter. The symbol commonly used for stress in much of the rheological texts is σ and that used for the strain is γ, and we will use these here, despite the risk of confusion due to their use to denote charge and surface tension. Moreover, it will implied that they are for shear stresses and strains, as shown in Figure 9.4.

In addition to stress and strain, the other parameter that we need to define is the timescale of the experiment. It is important to relate this to a characteristic timescale of the material behaviour. This is the type of thing

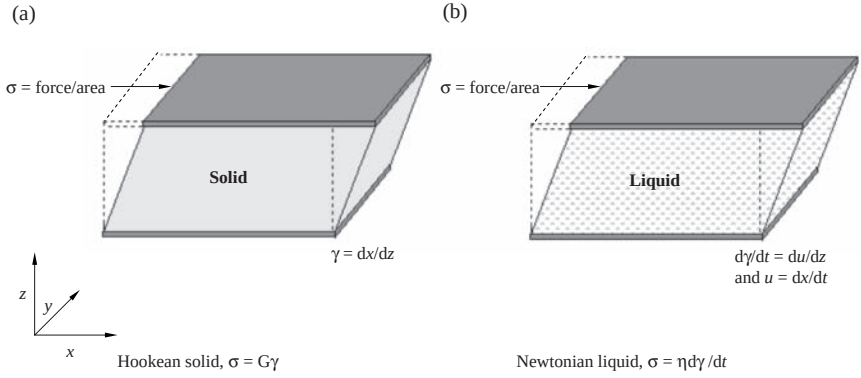


Figure 9.4 Definitions of the shear stress, σ, the shear strain, γ, and the rate of shear, $\dot{\gamma}$, for (a) a Hookean solid and (b) a Newtonian liquid

that we do automatically when, for example, we carry out an experiment were a chemical change is taking place in a wide range of situations such as working with radioisotopes or simply observing an expiry date on a commercial product, such as a foodstuff, in our everyday lives. In a rheological experiment, we are considering movement – the timescale that is relevant here is that of the diffusional process. It is the slowest process which is rate determining and with a condensed phase we need to define a long-time self-diffusion time. The Einstein–Smoluchowski equation gives the average time for a particle to diffuse through one particle radius, and by using the Stokes–Einstein equation, which defines the diffusion coefficient, we have the characteristic time, τ, as follows:

$$\tau = \frac{a^2}{D} = \frac{6\pi\eta(0)a^3}{k_B T} \tag{9.7}$$

The important point to note in Equation 9.7 is that the limiting viscosity is that of the colloidal dispersion under low-shear-rate conditions and not that of the continuous phase, η_0. This is the *effective medium* concept, which gives a first-order estimate of all the multi-body hydrodynamic interactions. If the experimental time is t_{ex}, we can define the dimensionless group known as the *Deborah number* as follows:

$$D_e = \frac{\tau}{t_{ex}} \tag{9.8}$$

If a material is only deformed by an applied stress for such a short time that the particles do not have time to diffuse very much, then when the stress is removed, the original shape is regained and the deformation is termed *elastic*. The reason is that as the structure formed by the particles is deformed, work is done and this is stored in that structure. Removal of the stress allows the structure to move back to a lower energy configuration. The situation where the experiment continues for a time which is much longer than the time for particle diffusion results in a permanent or *viscous* deformation of the material as the structure moves towards a low-energy configuration in the new shape. We say that the stress has relaxed and the characteristic time, τ, is thus the *stress relaxation time*. There are three ranges of Deborah number that can be identified, as follows:

$D_e \ll 1$	$D_e \sim O(1)$	$D_e \gg 1$
liquid-like	viscoelastic	solid-like

and it is immediately clear that there is progression from the behaviour of liquids through to that of solids. This appears to contradict our normal experience where the liquid–solid transition is very sharp, but this only means that the relaxation time changes dramatically with a very small change in temperature, concentration or pressure (stress). With concentrated colloids, surfactants and polymer solutions, the change in relaxation time is not so sudden as the structural components can move over larger distances and we can often make use of the broader viscoelastic range to produce the correct handling properties for colloidal products.

It is important to note that experimental timescales are not arbitrarily chosen by laboratory instrumentation. Such instruments are built with timescales that either simulate a process that we require numerical data for, such as the brushing of a paint film or the sedimentation of a colloidal dispersion on storage, or to quantify an observation that we have made. This last aspect takes the problem to our in-built 'bio-timescale'. We are aware of changes taking place over times ranging approximately from 10^{-3} to 10^{3} s. Hence, if the relaxation time lies within the range 10^{-4} s $< \tau < 10^{4}$ s, viscoelastic responses will be observed as we handle the material.

Many materials are produced as colloidal dispersions because of their liquid-like behaviour under at least certain conditions. Often, this is under high stresses and/or high strains and the property that is required is the 'correct' viscosity. In practice, the last stage in a formulation is

often the adjustment of the concentration by the addition of solvent to a particular viscosity value. The characteristic timescale for a continuous shear process is just the reciprocal of the shear rate, $\dot{\gamma}^{-1} = \mathrm{d}t/\mathrm{d}\gamma$, which is the time for unit strain to occur. As this is the characteristic experimental timescale, it immediately suggests the dimensionless group which is known as the *Péclet number*, P_e, which for a concentrated colloid is given by

$$P_e = \tau\dot{\gamma} = \frac{6\pi\eta(0)a^3\dot{\gamma}}{k_B T} \tag{9.9}$$

Note that again we are using the limiting suspension viscosity value at low shear rate as the viscosity of the effective medium. We should explore this in more detail. At low Péclet numbers, the relaxation time is short compared with the characteristic shear time and the structure is relatively unperturbed by the shear action. However, at high Péclet numbers, the shear rate is dominant and the structure is controlled by the hydrodynamic forces as the particles have insufficient time to diffuse to allow for the shape change of the structure. This means that at $P_e \approx 1$ the structure should be intermediate. We will explore how this should effect the viscosity in the next section while we note at this stage that the viscosity must be a function of the structure.

9.3.2 The Viscosities of Concentrated Dispersions

In Chapter 8, the viscosity of a dilute dispersion was presented and the equation describing the variation with volume fraction was based on rigorous hydrodynamic analysis and in the dilute limit the result due to Einstein was as follows:

$$\eta = \eta_0 (1 + [\eta]\varphi + \cdots) \tag{9.10}$$

where the intrinsic viscosity $[\eta] = 2.5$ for hard spheres at a volume fraction φ. As rigorous hydrodynamics cannot take the result much above $\varphi \approx 0.05$, the effective medium approach due to Krieger and Dougherty [6] will be used. The most straightforward analysis was presented by Ball and Richmond [7] and is duplicated here as it emphasizes some important points. The following equation gives us the rate of increase in

viscosity with volume fraction:

$$\frac{\mathrm{d}\eta}{\mathrm{d}\varphi} = [\eta]\eta_0 \tag{9.11}$$

Now, when we consider the small replacement of a volume of the continuous phase by some additional particles, we may expect a similar rate of change in viscosity and so

$$\mathrm{d}\eta = [\eta]\eta(\varphi)\mathrm{d}\varphi \tag{9.12}$$

Here, we are using $\eta(\varphi)$ as the viscosity of the system that the newly added particles will 'experience'. As the new particles have to have been added to a constant-volume system (by replacing medium with more particles), the change in concentration must be corrected to the available volume:

$$\frac{\mathrm{d}\eta}{\eta(\varphi)} = [\eta]\frac{\mathrm{d}\varphi}{\left(1 - \dfrac{\varphi}{\varphi_\mathrm{m}}\right)} \tag{9.13}$$

Here, the maximum concentration at which flow can occur is φ_m and so $\varphi/\varphi_\mathrm{m}$ is the excluded volume of the dispersion, that is, $(1 - \varphi/\varphi_m)$ is the liquid volume that could be replaced by more particles. Integration through the volume fraction range, with the boundary condition that as $\varphi \to 0$, $\eta(\varphi) \to \eta_0$, gives the *Krieger–Dougherty equation* for the volume-fraction dependence of the viscosity as

$$\frac{\eta(\varphi)}{\eta_0} = \left(1 - \frac{\varphi}{\varphi_\mathrm{m}}\right)^{-[\eta]\varphi_\mathrm{m}} \tag{9.14}$$

For a monodisperse system of hard spheres, $[\eta] = 2.5$. The value of φ_m varies [5] from the value of the liquid–solid transition for hard spheres under quiescent (i.e. 0.495 for freezing and 0.54 for melting) to 0.605, corresponding to flow of hexagonally packed layers at high shear rates. Hence we can expect the viscosity to be a function of shear rate. These values give for uniform hard spheres the limiting viscosity at high and low shear rates as

$$\frac{\eta(0)}{\eta_0} = \left(1 - \frac{\varphi}{0.5}\right)^{-1.25}; \quad \frac{\eta(\infty)}{\eta_0}\left(1 - \frac{\varphi}{0.605}\right)^{-0.51} \tag{9.15}$$

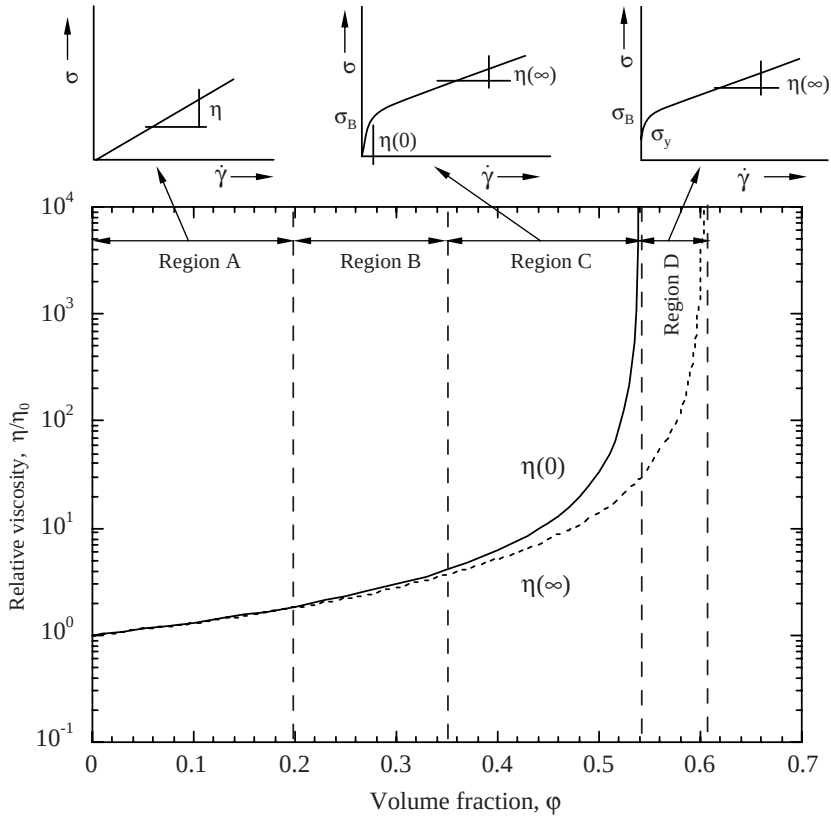


Figure 9.5 The relative viscosity as a function of volume fraction in the high- and low-shear limits. The dispersion is Newtonian in Region A, 'pseudoplastic' in Region C and viscoelastic ('plastic') in Region D (some shear-thinning can be observed in Region B)

We have used $\varphi_m = 0.5$ as the maximum volume fraction in the lower shear limit but we should recognize that the liquid and solid phases can coexist between volume fractions of 0.495 and 0.54, and this will allow some flow to occur between these limits. The shear-thinning behaviour is illustrated in Figure 9.5, in which the high and low shear limits of the viscosity are plotted as a function of the volume fraction of the dispersion.

There are four regions shown in Figure 9.5. In Region A, the dispersions behave as simple Newtonian fluids with no discernible shear dependence, whereas in Region B some shear thinning may be observed. Throughout Region C, viscoelastic liquid-like behaviour can be found with both the high-shear-limiting viscosity, $\eta(\infty)$, and low-shear-Newtonian limit, $\eta(0)$, being accessible. The concentrated dispersion in

Table 9.1 The characteristic parameters for a viscoelastic colloid, where $G(\infty)$ is the high-frequency limiting shear modulus, σ_y the static yield value, σ_B the dynamic or Bingham yield value, ω the frequency of an applied small-amplitude shear oscillation, $\dot{\gamma}$ the shear rate and τ the characteristic time

Region	Linear responses – small stresses and/or strains[a]		Non-linear responses – large stresses or strains[b]
Liquid-like Region C	$G(\infty)$ (Pa)	$\eta(0)$ (Pa s)	$\eta(0)$; $\dot{\gamma}^{-1} \gg \tau$ (s)
	$\omega^{-1} \gg \tau$ (s)		$\eta(\infty)$; $\dot{\gamma}^{-1} \ll \tau$ (s)
Solid-like Region D	$G(\infty)$ (Pa)	σ_y (Pa)	$\eta(\infty)$; $\dot{\gamma}^{-1} \ll \tau$ (s)
	$\omega^{-1} \ll \tau$ (s)		σ_B

[a] For example, oscillations.
[b] For example, continuous shear.

this region can be characterized as a weak gel as significant elastic behaviour can be observed. We can also call a colloid in this region a *complex fluid*. In Region D, the behaviour is that of a viscoelastic solid so that $\eta(0)$ is no longer accessible but there is a yield stress and when this is exceeded, the structure will melt and flow occurs. This is the *soft solid* stage. At high shear stresses, a value of $\eta(\infty)$ can be measured. At the upper limit of this region, the material will fracture under a sufficiently high stress and we will not observe flow. Table 9.1 summarizes the measurable parameters which are characteristic of the material. Figure 9.5 also shows the type of shear stress–shear rate curves that are measured, so that in Region A the linear response of a Newtonian fluid is shown, Region C has typical *pseudoplastic* behaviour and Region D is a *plastic* material. The static and dynamic yield stresses are shown where appropriate.

The shear-thinning response of the pseudoplastic dispersion is a smooth change from the low-shear Newtonian behaviour to the high-shear plateau. The curve is well described by the *Krieger equation* [8], which is written in terms of the reduced stress, σ_r:

$$\eta(\sigma_r) = \eta(\infty) + \frac{\eta(0) - \eta(\infty)}{1 + (b\sigma_r)^n} \tag{9.16}$$

The reduced stress, which was derived from dimensional analysis, is directly related to the *Péclet number*:

$$\sigma_r = \frac{a^3\sigma}{k_B T} = \frac{P_e}{6\pi} \tag{9.17}$$

For a monodisperse system $n = 1$ in Equation 9.16 and it exceeds this value for polydisperse systems. Using a value of the exponent $n > 1$, the viscosity shear-thins over a wider range of stresses than the monodisperse system. A value of b for a particular system is obtained from the mid-point of the curve. Recalling that P_e is the ratio of the convective to thermal timescales for the particle motion, means that midway between the Brownian motion-dominated structure (the zero-shear plateau) and the shear-dominated structure (the high-shear plateau) $b\sigma_r = 1$, so that $b = 1/\sigma_{rc}$. The latter parameter, the critical reduced stress, is the value of the reduced stress at the mid-point of the viscosity curve. It is instructive to think of this in terms of the characteristic time τ for the relaxation of the structure. This concept suggests that a spectral range of relaxation times should be included. For example, as the size distribution broadens, the range of relaxation times also broadens because the diffusive motion is a function of particle size and local concentration (and also pair potential for particles other than hard spheres). As a starting point, we can sum the contributions of each component to the stress weighted for its probability, ρ_i, and then we may re-write Equation 9.16 as

$$\frac{\eta(\sigma_r) - \eta(\infty)}{\eta(0) - \eta(\infty)} = \sum_i \left(\rho_i \frac{1}{1 + \sigma_{rci}} \right) \tag{9.18}$$

where σ_{rci} is critical stress for the ith component. Figure 9.6 illustrates the response for a system with a single relaxation time and how an example spectrum of times broadens the response from about three or four orders of magnitude to at least six. We should keep in mind that for a monodisperse system, we expect a spectrum of relaxation times from the fluctuations of the local concentration due to the collective diffusion of the particles [5]. We will see later how this concept of using a spectrum of relaxation times fits naturally with the practice adopted in analysing linear viscoelastic data. This is most appropriate when used with polymer solutions, where we can usually identify the low-shear viscosity with the low-frequency dynamic viscosity (the *Cox–Mertz rule*). With particulate systems, this correspondence is not always good but it is an important characterization nevertheless.

As the colloidal forces between the particles becomes significant, the stress required to move them relative to one another increases. This means that the total stress for the flowing system is the sum of the hydrodynamic and the colloidal terms. Note that for a simple hard-sphere system, the colloidal term is the Brownian term. At very low shear rates,

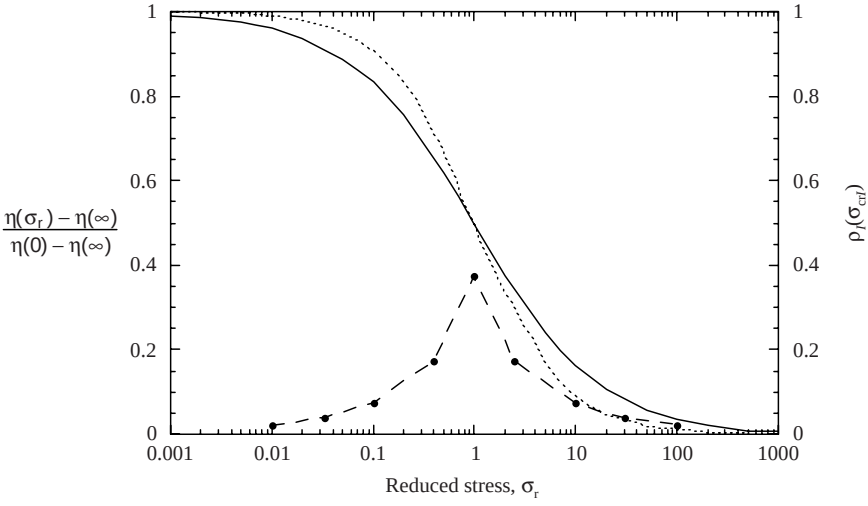


Figure 9.6 The effect of a range of relaxation times on the shear-thinning response of a dispersion: (- - -) single relaxation time; (—) spectral range; (- • -) spectrum

the hydrodynamic term becomes much smaller than the colloidal contribution. The result is that the zero-shear viscosity is controlled by the colloidal forces. There are statistical mechanical models available [9] to calculate the viscosity which integrate the contributions from the colloidal forces over the structure. The latter is given by the pair distribution function, which is distorted by the shear field. For the case of charge-stabilized dispersions, we may treat the particles as 'effective' hard spheres where the 'effective' radius is given in the low-shear limit as the collision radius of the particles during a Brownian motion encounter. This may be carried out by equating the electrostatic repulsive force with the thermal force. Russel *et al.* [10] used this idea to derive the closest distance between particle centres, which is, of course, the value of the 'effective' hard sphere diameter, $r_0(0)$:

$$r_0(0) \approx \kappa^{-1} \ln \left\{ \frac{\alpha}{\ln\left[\alpha / \ln\left(\alpha / \ln \cdots\right)\right]} \right\} \qquad (9.19)$$

$$\text{where } \alpha = \frac{4\pi\varepsilon\kappa\,(a\zeta)^2 \exp(2\kappa a)}{k_B T}$$

The consequence of this 'effective' radius is just an increase in the excluded volume of the particle and so modifies the value of the maximum volume fraction at which the viscosity diverges. Hence the low-shear

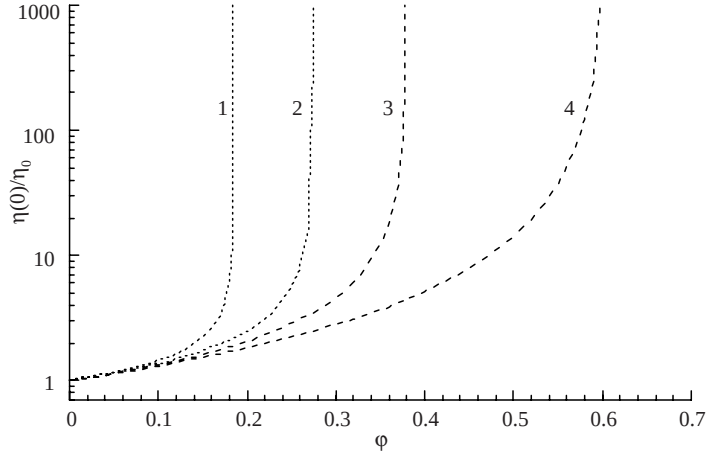


Figure 9.7 Zero-shear limiting viscosity as a function of particle radius: $\zeta = -50\,\text{mV}$; $\sigma_\delta = 1\,\text{mC}\,\text{cm}^{-2}$: (1) $a = 50\,\text{nm}$; (2) $a = 100\,\text{nm}$; (3) $a = 250\,\text{nm}$; (4) high-shear limit

limiting viscosity for charge-stabilized dispersions now becomes

$$\eta(0) = \eta_0 \left(1 - \frac{\varphi}{\varphi'_{\text{m}}}\right)^{\frac{5\varphi'_{\text{m}}}{2}}$$
$$\text{where } \varphi'_{\text{m}} = 0.495 \left[\frac{2a}{r_0(0)}\right]^3 \tag{9.20}$$

Figure 9.7 shows how sensitive the low shear viscosity is to particle size for small charged particles at moderate electrolyte concentrations. Here, the calculation used values of $\zeta = 50\,\text{mV}$ with a charge density of the Stern layer of $\sigma_\delta = 1\,\mu\text{C}\,\text{cm}^{-2}$, with the latter quantity being used to calculate the value for $\kappa(\varphi)$ as it gives the number of counter-ions in the diffuse layer. Equation 9.19 gave a value of $r_0(0) \approx 2a + 5/\kappa$, which serves to illustrate the importance of the slow decay of the pair potential with distance and indicates the origin of the large excluded volume of small charged particles. It should also be noted that the very steep rise in the viscosity as the liquid–solid transition is approached will be less sharp in practice due to the coexistence of the two phases in the volume fraction range 0.5–0.54.

Similar results are found with particles with attractive forces dominating the long-range part of the pair potential, resulting in a weakly flocculated system. The modelling requires a different approach and the particle size dependence is different. In most practical cases, colloidal

formulations have many components, so it is not always clear how the behaviour relates to the theoretical calculations which are frequently used for model dispersions of monodisperse spheres. However, there are some common origins so that we may formulate some general observations.

9.3.3 Viscosity: Summary

1. There are three forces which govern the stress that we observe when we measure the viscosity as a function of shear rate, namely hydrodynamic forces, Brownian (thermal) forces and colloidal forces arising from the form of the pair potential.
2. In the high-shear limit, the dispersion structure and measured stress are dominated by the hydrodynamic forces.
3. The high-shear limit of φ_m is a function of the nature of the disperse phase. Polydispersity increases the value [5], as do fluid particles due to particle deformation, although this does not occur until high volume fractions are reached as the surface forces are strong for particles with colloidal dimensions. Particle anisometry increases the value.
4. It should be noted that in many applications, much higher shear rates occur than can be accessed using most laboratory rheometers, and also there is frequently an extensional component present. With particulate systems, this does not represent a problem as the shear thinning occurs over a relatively narrow range of shear rates or applied stress so the high-shear limit can be reached or at least approached. However, this is not the case with solution polymers, which have a much broader response due first, to the broad molecular weight distributions, which are much broader than the particle size distribution produced with most particulates. Second, the relaxation mode of each chain varies from timescales ranging from the centre-of-mass diffusional time of the whole molecule down to that of a single segment.
5. In the low-shear limit, the dispersion structure and excluded volume of the particles are governed by the balance of the Brownian and colloidal forces. Measurements under low-shear conditions are therefore particularly useful in identifying changes in colloidal interactions as formulation conditions are varied.
6. The prediction of the shear-thinning behaviour is usually not possible as the complexity of most practical systems rules this out and we have to rely on experimental measurements.

7. Many formulations are produced in the concentration regime between where the low-shear viscosity increases rapidly and the high-shear limit. In this region, it is usually the viscoelasticity that is most important, as this dominates our perception of the 'handling properties' of the dispersion, whereas the extremes of behaviour are the most important parameters for the applications. For example, brushing, spreading and spraying are all high-shear processes with shear rates $\dot{\gamma} \geq 10^3\,\mathrm{s}^{-1}$, whereas sedimentation is a 'long time' low-stress process where $\dot{\gamma} \to 0$.

9.4 LINEAR VISCOELASTICITY OF COLLOIDAL DISPERSIONS

The discussion of viscoelasticity in this text is restricted to linear responses, where the viscosity is Newtonian and the elasticity obeys *Hooke's law*, and so the limiting constitutive equations for shear are

$$\sigma = \eta\dot{\gamma}; \quad \sigma = G\gamma \tag{9.21}$$

This means that in the laboratory we have to restrict our measurements to low stresses and strains, but moreover, we must check that this is the case if we are to use the analysis of the data to give characteristic parameters of the dispersion rheology. Hence we should determine that if the stress or strain is changed, the corresponding strain or stress changes in proportion.

9.4.1 Constitutive Equations

When we make measurements in the laboratory, it is convenient to fit the data to a curve so that we may summarize the data in terms of as few characteristic parameters as give an adequate description of our observations. This serves two purposes. It enables us to make a quantitative comparison of different materials, which could be new batches of the same formulation or different formulations. Thus, the subjectivity is minimized. In addition, we may be able to interpret the characteristic parameters in terms of the interactions between the microstructural components. Even if the latter becomes difficult, the former is an imperative.

When we devise experiments to determine the viscoelastic response of a material, we need to vary stresses or strains over a range of timescale to explore the response over a range of Deborah numbers. For example, we can apply a step stress/strain and follow the response over time, or we may oscillate the material over a range of frequency. The measured data are in terms of the elastic modulus if we measure the stress as a function of applied strain or the compliance if the strain is recorded as a function of the applied stress. The problem that we then have is to fit the experimental curves. As an aid to this end, we can invoke the response of mechanical analogues to help us to derive suitable equations. Note that it is also possible to use the response of electrical circuits for the same purpose. The algebra is similar in both cases. The utility of the constitutive equations is primarily to describe the correct response of the material over the range studied and, if possible, to yield the limiting behaviour.

The mechanical analogues use Hookean springs for the elastic behaviour and Newtonian dashpots or dampers to give the viscous response. These can be coupled in many combinations to give mechanical responses similar to the experimental curves that we are attempting to simulate. A few of these are shown in Figure 9.8 with the corresponding constitutive equations relating the stress and strain. Which ones we make use of depends on the form of the experiment that we carry out. The appropriate equation is then used to calculate the form of the modulus or compliance of *the model* for the appropriate form of the temporal dependence of the applied strain or stress. Although it would be possible to construct very complicated analogues, they would be of little utility as we lose uniqueness of the fit if we have many 'characteristic parameters'. We will therefore restrict our use to a very few of the possible models, keeping in mind that that the object is to compare and record the responses of materials with as few characteristic parameters as possible. A quantitative microstructural interpretation of these may occasionally be possible for the simplest systems, but that is not the primary objective of the process.

9.5 PHENOMENOLOGY

This section will describe the various experiments that are carried out to characterize linear viscoelasticity. The first method described – the application of an oscillating deformation – is one frequently used and derivation of the responses from a mechanical analogue is illustrated in

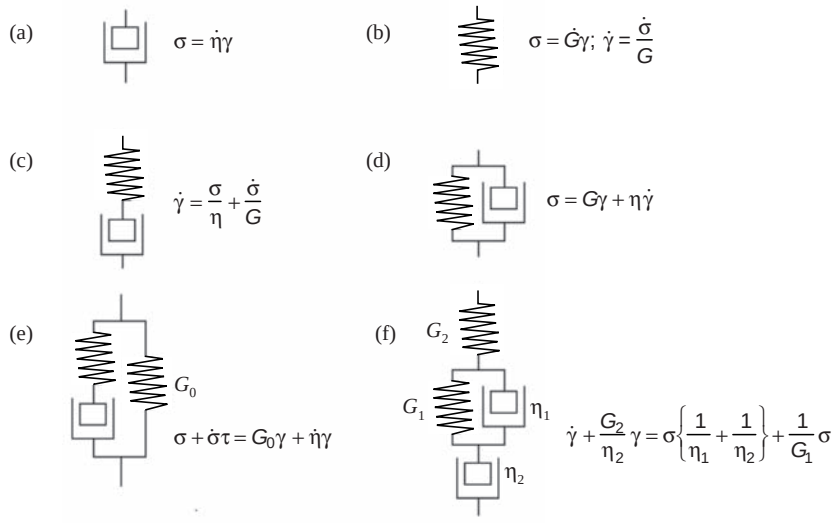


Figure 9.8 Mechanical models of linear viscoelastic materials with the corresponding equations. The characteristic time for all models is $\tau = \eta/G$

some of these as examples. The results of the algebraic analyses for the other types of experiment are simply stated.

9.5.1 Oscillating Strain

In this experiment, the concentrated colloid is usually subjected to a sinusoidally oscillating strain and the resulting stress is measured at a variety of frequencies. During the experiment the rheometer records three parameters:

- the peak strain that is applied
- the resulting peak stress
- the difference in phase between the strain and stress wave forms at each frequency.

Figure 9.9 illustrates the form of the stress that we would measure for a material subjected to an oscillating strain, which we will assume to have a frequency of 1 rad s^{-1}. In this example, the maximum strain $\gamma_0 = 0.05$ and the maximum stress σ_0 is 25 Pa. Now, the stress per unit strain is

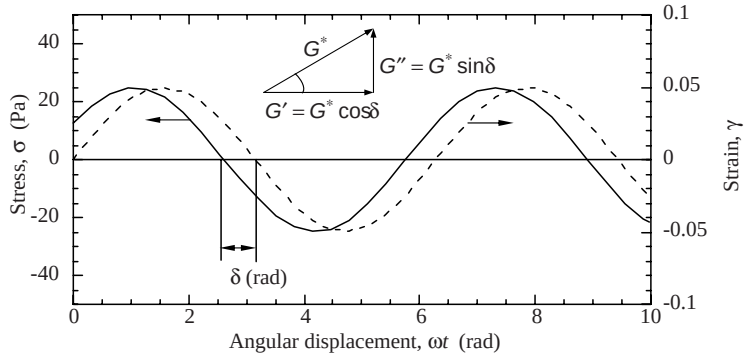


Figure 9.9 Forced oscillation experiment for a material with $G^* = 0.5$ kPa at a strain of 5%: $\delta = 30^\circ$; G′= 0.433 kPa; G″= 0.25 kPa

the modulus at an oscillation frequency ω, as follows:

$$G^*(\omega) = \frac{\sigma_0}{\gamma_0} \tag{9.22}$$

which for the example data in Figure 9.9 gives a value of 0.5 kPa for the *complex modulus*, G^*. The complex modulus has contributions from both a storage (elastic) and a loss (viscous dissipation) term. If the material were a simple Hookean solid, the position of the maximum stress would coincide with that of the maximum strain. On the other hand, if the experiment was carried out on a Newtonian liquid, the maximum stress would be coincident with the maximum rate of strain. The phase shift, δ, enables us to assign values to the storage and loss contributions to the complex modulus, as follows:

$$\begin{aligned} G^*(\omega) &= G'(\omega) + i\,G''(\omega) \\ G'(\omega) &= G^*(\omega)\cos\delta \\ G''(\omega) &= G^*(\omega)\sin\delta \end{aligned} \tag{9.23}$$

Here we are using the normal complex number algebraic notation where $i^2 = -1$ and

$$G'(\omega) + i\,G''(\omega) = G^*(\omega)\,(\cos\omega t + i\sin\omega t) = G^*(\omega)\mathrm{e}^{i\omega t} \tag{9.24}$$

Alternatively, the complex viscosity can be defined as follows:

$$\eta^*(\omega) = \frac{G^*(\omega)}{\omega} = \eta'(\omega) + i\eta(\omega) \tag{9.25}$$

and the dynamic viscosity is the real or in phase part of the complex viscosity:

$$\eta'(\omega) = \frac{G''(\omega)}{\omega} \tag{9.26}$$

Returning to the curves in Figure 9.9, the values for the storage and loss moduli are $G'(\omega) = 0.433\,\text{kPa}$ and $G''(\omega) = 0.25\,\text{kPa}$ at $1\,\text{rad}\,\text{s}^{-1}$, respectively. Experimentally, we determine the moduli over as wide a range of frequency as is practicable, but the problem is then to condense these data to as few characteristic material constants as will provide an accurate fit to such data.

To illustrate the process further, we can take the constitutive equation for a *Maxwell model*, which has a single spring in series with a dashpot, as shown in Figure 9.8c. Hence for an oscillating strain we have

$$\dot{\gamma}^* = \frac{\dot{\sigma}^*}{G} + \frac{\sigma^*}{\eta} \tag{9.27}$$

and the complex stresses and strains are written in terms of the peak stresses and strains that are measured:

$$\begin{aligned} &\gamma^* = \gamma_0 \exp(i\omega t); \quad \dot{\gamma}^* = i\omega \exp(i\omega t) = i\omega\gamma^*; \\ &\sigma^* = \sigma_0 \exp[i(\omega t + \delta)]; \quad \dot{\sigma}^* = i\omega\sigma_0 \exp[i(\omega t + \delta)] = i\omega\sigma^* \end{aligned} \tag{9.28}$$

Substitution of these into Equation 9.27 and rearranging gives

$$\frac{i\omega\gamma^*}{i\omega\sigma^*} = \frac{1}{G^*(\omega)} = \frac{1}{G} + \frac{1}{i\omega\eta} \tag{9.29}$$

Noting that the characteristic or time of the material is $\tau_\rho = \eta\,(\text{Pa}\,\text{s})/G\,(\text{Pa})$, we can then re-arrange Equation 9.29 to give

$$G^*(\omega) = G\left(\frac{i\omega\tau_\text{r}}{1 + i\omega\tau_\text{r}}\right) \tag{9.30}$$

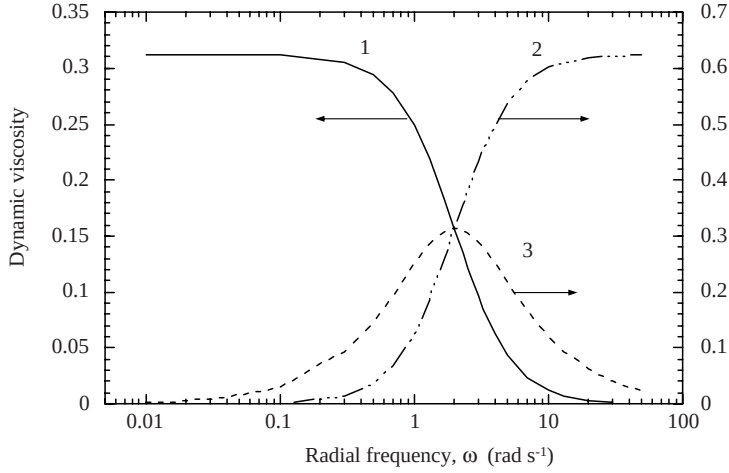


Figure 9.10 Frequency response for a Maxwell liquid with $\eta(0) = 0.32$ Pa s, $G(\infty) = 0.63$ Pa and $\tau_M = 0.5$ s: (1) dynamic viscosity; (2) storage modulus; (3) loss modulus

We can separate the complex modulus into the storage and loss moduli, and multiplying Equation 9.30 throughout by $(1 - i\omega\tau_r)$ gives

$$G'(\omega) + iG''(\omega) = G\frac{(\omega\tau_r)^2}{1 + (\omega\tau_r)^2} + iG\frac{\omega\tau_r}{1 + (\omega\tau_r)^2} \tag{9.31}$$

If for our example we assume the relaxation time to be 0.5 s, the value of G from the data given can be readily calculated:

$$G''(1) = G\left(\frac{0.5}{1 + 0.5^2}\right)$$

which gives $G = 0.625$ kPa and $\eta = 0.5 \times 0.625 = 0.313$ kPa s.

The frequency dependence of the storage and loss moduli storage is as illustrated in Figure 9.10. Features of the curves that are typical of a Maxwell model are that:

- The storage modulus reaches a plateau at high frequency, $G'(\omega) \rightarrow G(\infty)_{\omega\rightarrow\infty}$, noting that $G(\infty) = G$ in the model.
- $G''(\omega) > G'(\omega)$ when $\omega < 1/\tau_r$ but $G''(\omega) < G'(\omega)$ when $\omega > 1/\tau_r$.
- $G''(\omega)_{\omega=1/\tau_r} = 0.5G(\infty)$.
- At low frequencies, the dynamic viscosity reaches a plateau and so $\eta'(\omega)_{\omega\rightarrow 0} \rightarrow \eta(0)$, noting that $\eta(0) = \eta$ in the model.

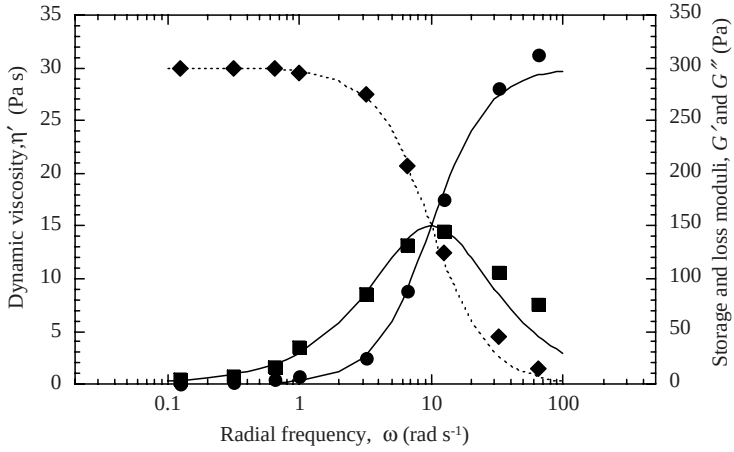


Figure 9.11 Oscillation response for a sample of shower gel. The symbols represent experimental data, with the curves calculated for a Maxwell fluid with $\eta' = 30\,\text{Pa s}$ and $\tau = 0.1\,\text{s}$

At this point it, is useful to consider some experimental data obtained using a colloidal product. The material used was a shower gel. The formulation consisted of a concentrated surfactant system at a high enough concentration that a condensed phase of 'worm-like' micelles can form. The rheology of the formulation had been 'fine-tuned' with a small amount of polymer. The experimental data obtained from this system using an applied strain of 10% are plotted in Figure 9.11. The data are shown as symbols with the curves being calculated for a Maxwell model. The relaxation time was 0.1 s and $\eta(0)$ was 30 Pa s. Both were easily determined from the curves and these values gave $G(\infty) = \eta(0)/\tau_r = 300\,\text{Pa}$. Note that the fit becomes poor at frequencies above 10 Hz. This is caused by the applied frequency approaching a resonance frequency of the measuring unit. However, it is evident from the form of the rest of the plot that the shower gel is a viscoelastic fluid and a Maxwell model provides an adequate description of the material with a well-defined characteristic time and a clear limiting viscosity at low frequencies. Hence these are the only parameters which we need to record to characterize the material and permit a batch-to-batch comparison.

The next example is a household cleaner used for kitchens. Like the shower gel, it contained a large amount of surfactant in order to remove greasy deposits. However, it also contained an abrasive powder to aid the cleaning action. This abrasive was calcium carbonate and the particle

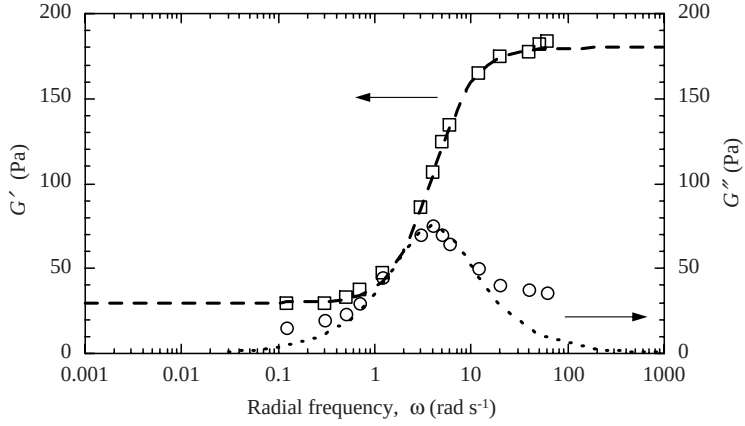


Figure 9.12 Oscillation response for a sample of household cleaner: □, storage moduli; ○, loss moduli. The curves are plotted using the mathematical model shown in Figure 9.8e, using the following parameters: $G(0) = 30\,\text{Pa}$; $\tau_1 = 0.25\,\text{s}$; $G_1 = 150\,\text{Pa}$ [giving $G(\infty) = 180\,\text{Pa}$]

size was in the upper part of the colloidal range; that is, 1 μm < mean diameter <10 μm. If the particle size is too large the material feels 'gritty' and not creamy; if too small, the abrasive action is sacrificed. Now, the density difference between calcium carbonate and the surfactant concentrate is $\sim 2000\,\text{kg m}^{-3}$. If this was the final formulation, the particles would sediment slowly as the system would be a viscoelastic fluid like the shower gel. However, poly(acrylic acid) was also included as a 'rheological modifier'. Divalent Ca^{2+} binds strongly to $-COO^-$ groups and cross-links the polymer. This cross-linking makes such a system a viscoelastic solid and sedimentation is prevented as the limiting viscosity at low stresses $\eta(0) \to \infty$. Careful choice of the polymer concentration limits the cross-link density and the structure is readily *strain melted* (i.e. broken). The experimental data are shown in Figure 9.12, with the curves representing the values calculated for the mechanical analogue in Figure 9.8e, that is, a Maxwell model in parallel with a single spring providing the solid-like response at long times. It is important to recognize that the presence of a $G(0)$ value with the curves of the storage and loss moduli diverging is indicative of solid-like behaviour and clear control of sedimentation. The curves need three parameters and the values used were $G(\infty) = 180\,\text{Pa}$, $G(0) = 30\,\text{Pa}$, $\tau_1 = 0.25\,\text{s}$; $\eta(0) = \infty$, of course. In terms of the model shown in Figure 9.8e, this gives $G_0 = 30\,\text{Pa}$, $G_1 = (180 - 30) = 150\,\text{Pa}$ and $\eta_1 = (150 \times 0.25) = 37.5\,\text{Pa s}$. The fit is moderately reasonable in terms of the storage modulus. However, the

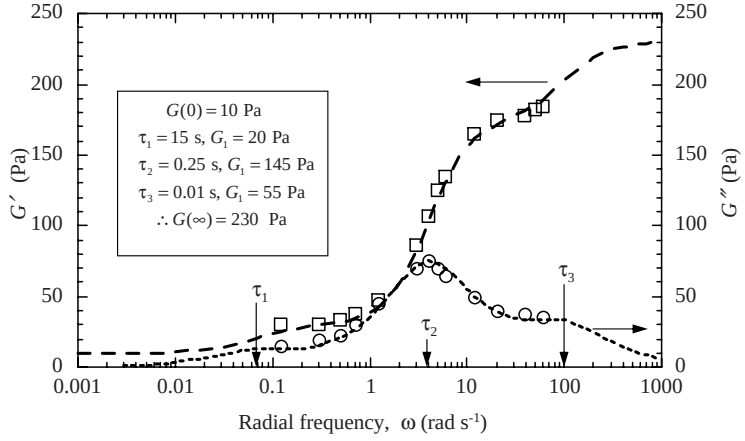


Figure 9.13 Corresponding data to those shown in Figure 9.12 for a sample of household cleaner: □, storage moduli; ○, loss moduli. The curves are plotted in this case by using a four-element model consisting of three Maxwell elements with a spring in parallel (e.g. Figure 9.14b with $n = 3$), giving a much improved fit to the experimental data

loss modulus is poorly described at frequencies both higher and lower than the characteristic frequency. This broadening of the response is indicative of there being more than one relaxation process present – not a surprising result for a system consisting of 'worm-like' micelles, cross-linked polymer chains and large particles of colloidal size.

We obtain a much better fit to the experimental data when we use a more complicated model. A good fit was obtained using a four-element model consisting of three Maxwell models and a spring in parallel. The values for each of the components are given in Figure 9.13. Note that the main peak is barely changed and that a smaller $G(0)$ is predicted, but it is still of the same order of magnitude as in the simpler model. For routine quality control purposes, the data obtained from the simpler approach illustrated in Figure 9.12 would often be adequate. This may not always be the case, however, and in terms of product development it would be important to determine, numerically, how the values shown in Figure 9.13 vary with component changes. Of course, the data range shown in the figure should be broadened as much as possible, but the concept introduced here is an important one.

We are seeing a simple illustration of a spectral response, that is, there are a series of processes occurring which all contribute to the overall response. Mathematically such a series is described as the *generalized*

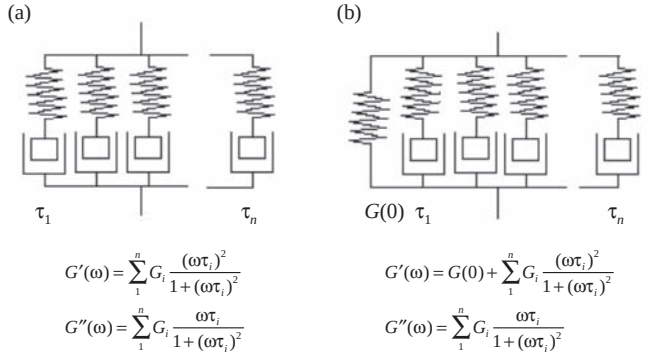


Figure 9.14 Illustration of generalized Maxwell models for (a) a viscoelastic fluid and (b) a viscoelastic solid

Maxwell model (see Figure 9.14), which is represented as follows:

$$\begin{aligned} G'(\omega) &= G(0) + \sum_1^n G(\infty)_n \frac{(\omega\tau_n)^2}{1+(\omega\tau_n)^2} \\ G''(\omega) &= \sum_1^n G(\infty)_n \frac{\omega\tau_n}{1+(\omega\tau_n)^2} \end{aligned} \tag{9.32}$$

It is very important to recognize that this is qualitatively different from the viscous response that we discussed earlier for plastic materials in Section 9.3.2. In that case, we were looking at the shear melting of the structure as the rates of deformation became too rapid for parts of the structure to accommodate the change. This was a non-linear measurement. In the case we are now discussing, the measurement is linear and the structure is not significantly modified by the experiment, that is, this is a non-destructive experiment, whereas the former was a destructive experiment. We may be able to fit our data using Equation 9.32, but we should also recognize that the way in which that equation is written implies that each relaxation process has equal weighting. An alternative way of expressing the spectral response is through an integral equation instead of a summation, as follows:

$$\begin{aligned} G'(\omega) &= G(0) + \int_{-\infty}^{+\infty} H \frac{(\omega\tau)^2}{1+(\omega\tau)^2} \mathrm{d}\ln\tau \\ G''(\omega) &= \int_{-\infty}^{+\infty} H \frac{\omega\tau}{1+(\omega\tau)^2} \mathrm{d}\ln\tau \end{aligned} \tag{9.33}$$

The H in Equation 9.33 is known as the *relaxation spectrum* and is equivalent to each of the $G(\infty)$ values at each time multiplied by the probability of that process. Note that here the temporal behaviour is expressed on a logarithmic scale as often we observe very broad responses. Of course, if our material is a viscoelastic liquid, then $G(0) = 0$ in Equations 9.32 and 9.33. Hence for a viscoelastic liquid such as the shower gel example in Figure 9.11, we can write the 'generalized Maxwell model' equations as follows:

$$\begin{aligned} G'(\omega) &= \sum_{1}^{n} G(\infty)_n \frac{(\omega\tau_n)^2}{1+(\omega\tau_n)^2} \\ G''(\omega) &= \sum_{1}^{n} G(\infty)_n \frac{\omega\tau_n}{1+(\omega\tau_n)^2} \end{aligned} \tag{9.34}$$

which in integral form are

$$\begin{aligned} G'(\omega) &= \int_{-\infty}^{+\infty} H \frac{(\omega\tau)^2}{1+(\omega\tau)^2} \mathrm{d}\ln\tau \\ G''(\omega) &= \int_{-\infty}^{+\infty} H \frac{\omega\tau}{1+(\omega\tau)^2} \mathrm{d}\ln\tau \end{aligned} \tag{9.35}$$

At this point, we can see the similarity with the characterization of materials by spectroscopic methods in the microwave, infrared, and so on. However, we rarely see sharp, well-separated peaks. Hence, although our relaxation spectrum is really a summary of our material behaviour, recording the maximum peak height(s) and the frequency at which it (they) occur does not provide enough information. This is because in nearly all cases we are looking at multiple process which are located close in time and we should note from Figure 9.10 how a single process has a curve which has a 'half-width–half-height' of ~1.5 orders of magnitude in frequency.

9.5.2 The Limiting Behaviour at the Extremes of Frequency

To estimate the limiting response at high frequency, that is, as $\omega \to \infty$, it should be noted that for a *viscoelastic liquid*

$$G^*(\omega)_{\omega\to\infty} \to G(\infty) = \sum_{1}^{n} G(\infty)_n \text{ or } = \int_{-\infty}^{+\infty} H \mathrm{d}\ln\tau \tag{9.36}$$

$$\text{as } \frac{(\omega\tau)^2}{1+(\omega\tau)^2} \to 1 \text{ and } \frac{(\omega\tau)}{1+(\omega\tau)^2} \to 0.$$

The low-frequency limit is also of interest, and in this case it is expressed in terms of the limiting value of the complex viscosity or the zero shear viscosity:

$$\eta^*(\omega)_{\omega\to 0} = \frac{G^*(\omega)_{\omega\to 0}}{\omega} \to \eta(0) = \sum_1^n G(\infty)_n \tau_n \text{ or } = \int_{-\infty}^{+\infty} H\tau \mathrm{d} \ln \tau \tag{9.37}$$

$$\text{as} \frac{\omega\tau^2}{1+(\omega\tau)^2} \to 0 \text{ and } \frac{\tau}{1+(\omega\tau)^2} \to \tau.$$

The modelling of the zero-shear viscosity has already been discussed for some colloidal dispersions and there are models in the literature for polymers in solution [11]. The modelling of the high-frequency modulus will be discussed later in this chapter.

Carrying out the same exercise for a *viscoelastic solid* yields

$$G(\infty) = G(0) + = \sum_1^n G(\infty)_n \text{ or } = G(0) + \int_{-\infty}^{+\infty} H \mathrm{d} \ln \tau \tag{9.38}$$

The low-frequency limit is

$$\frac{G^*(\omega)_{\omega\to 0}}{\omega} \to \eta(0) = \frac{G(0)}{\omega \to 0} + \sum_1^n G(\infty)_n \tau_n \to \infty \text{ or}$$
$$= \frac{G(0)}{\omega \to 0} + \int_{-\infty}^{+\infty} H\tau \mathrm{d} \ln \tau \to \infty \tag{9.39}$$

which is of course the correct result for a *solid*! As an illustrative exercise, we will take the spectra used to fit the data shown in Figure 9.13 with and without the $G(0)$ term to show the frequency dependence of the complex viscosity. It is clear from Figure 9.15 that without the $G(0)$ value in the four-component spectrum, the viscosity reaches a plateau

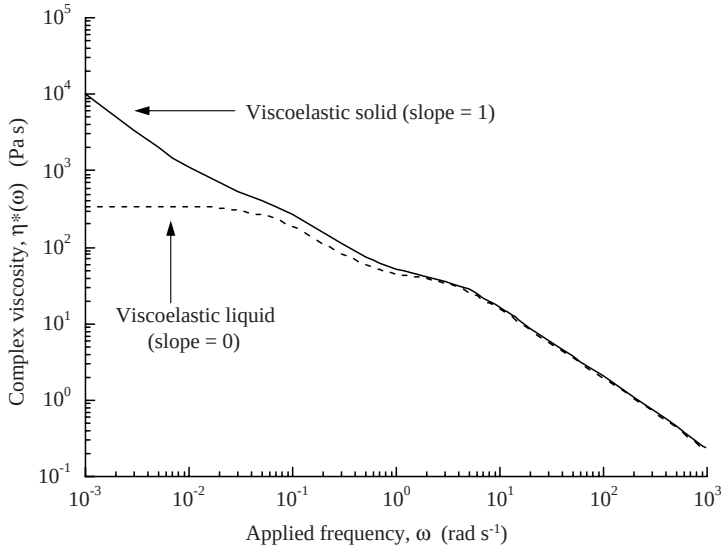


Figure 9.15 The complex viscosity calculated for the four-component viscoelastic spectrum used in Figure 9.13 and for a viscoelastic liquid from the same spectrum with $G(0) = 0$

value of $\sim$350 Pa s at frequencies below 10^{-2} rad s^{-1}. When there is a solid-like component, the viscosity increases steadily towards infinity with unit slope.

It is appropriate to consider the implications in terms the product behaviour. When such a formulation has been prepared to be non-sedimenting, the solid-like response would appear to be desirable. However, the plateau value of the low shear viscosity of the fluid-like system is sufficiently high that, for abrasive particles such as calcium carbonate, if the radius of the largest particles or aggregates is $<$2–3 μm the sedimentation rate is so slow that it would take $\sim$3 months for a 1 mm layer of clear fluid to form. This may be acceptable if the length of time spent by the formulation in storage is usually less than this.

9.5.3 Stress Relaxation or the Step Strain Experiment

This experiment is the mechanical analogy of the temperature jump and the pressure jump experiments used to study chemical equilibria. In this case, we apply a small shear strain to the material as rapidly as possible and then follow the resulting stress as a function of time.

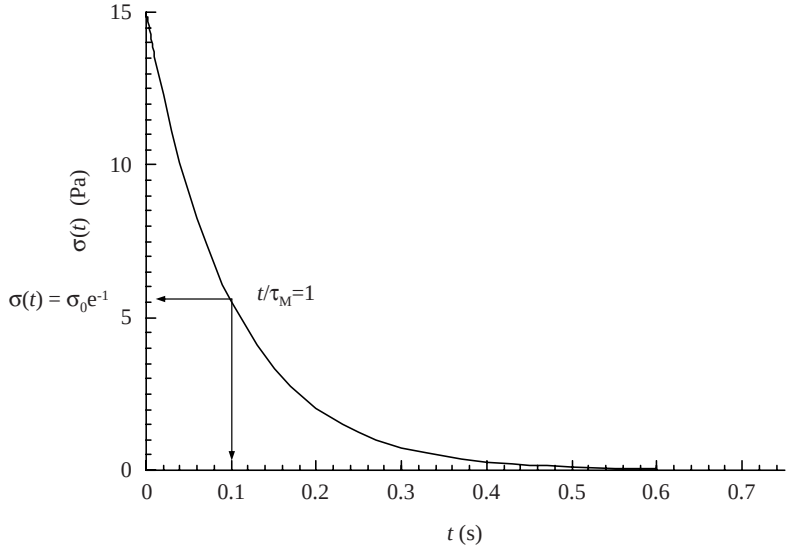


Figure 9.16 Stress relaxation for a linear viscoelastic 'Maxwell 1' fluid with a characteristic time of 1 s after being subjected to a step strain of 5%

A suitable starting point to describe the behaviour is to consider the response of a single Maxwell model, that is, a linear viscoelastic fluid with a single relaxation time. This is then the analogue of a first-order chemical reaction approach to equilibrium and we follow the stress from its initial value, σ_0, to zero over time with the constant strain of γ_0. A curve such as that shown in Figure 9.16 will be obtained.

We can express the rate of change of stress at any time as

$$\frac{\mathrm{d}\sigma(t)}{\mathrm{d}t} = -k\sigma(t) \tag{9.40}$$

where k is the first-order rate constant for this relaxation and thus has units of s^{-1}. It is the reciprocal of the characteristic or Maxwell relaxation time, τ_M. We can write Equation 9.40 in terms of the stress relaxation modulus, $G(t)$, by dividing through by the value of the constant applied strain:

$$\frac{\mathrm{d}G(t)}{G(t)} = -k\mathrm{d}t \tag{9.41}$$

This equation is then integrated to give the value of the relaxation modulus at time t, as follows:

$$\int \frac{\mathrm{d}G(t)}{G(t)} = -\frac{1}{\tau_{\mathrm{M}}} \int \mathrm{d}t \tag{9.42}$$

The limits for the integration are $t = 0$ and t, so we have

$$\ln G(t) - \ln G(t \to 0) = -\frac{t}{\tau_{\mathrm{M}}} \tag{9.43}$$

that is

$$G(t) = G(t \to 0) \exp\left(-\frac{t}{\tau_{\mathrm{M}}}\right) \tag{9.44}$$

This can be generalized to include as many processes as are required to describe the curves:

$$G(t) = \sum_{i=1}^{i=n} G_i \exp(-t/\tau_i) \qquad \text{(fluid)} \tag{9.45a}$$

$$G(t) = G(0) + \sum_{i=1}^{i=n} G_i \exp(-t/\tau_i) \qquad \text{(solid)} \tag{9.45b}$$

Figure 9.17 shows the plots for the Maxwell liquid shown in Figure 9.16 (corresponding to data found for the shower gel shown in Figure 9.11) and the four-component model used to describe the solid-like behaviour of the domestic cleaner in the oscillation experiment plotted in Figure 9.13. The integral forms of the stress relaxation equations for the generalized Maxwell models are follows:

$$G(t) = \int_{-\infty}^{+\infty} H\left(\ln t\right) \exp\left(-t/\tau\right) \mathrm{d} \ln \tau \qquad \text{(fluid)} \tag{9.46a}$$

$$G(t) = G(0) + \int_{-\infty}^{+\infty} H\left(\ln t\right) \exp\left(-t/\tau\right) \mathrm{d} \ln \tau \qquad \text{(solid)} \tag{9.46b}$$

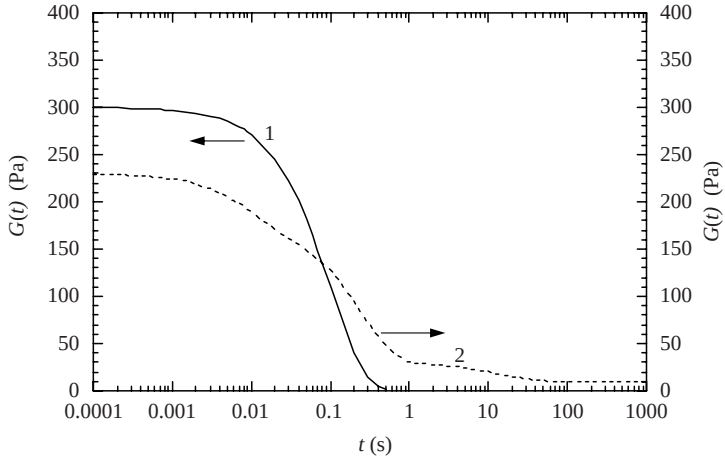


Figure 9.17 Stress relaxation modulus as a function of time for (1) a Maxwell liquid with $G = 300\,\text{Pa}$ and $\tau = 0.1\,\text{s}$ and (2) a four-component solid with $G(0) = 10\,\text{Pa}$, $G_1 = 20\,\text{Pa}$, $G_2 = 145\,\text{Pa}$, $G_3 = 55\,\text{Pa}$, with $\tau_1 = 15\,\text{s}$, $\tau_2 = 0.25\,\text{s}$, $\tau_3 = 0.01\,\text{s}$

The stress relaxation experiment has some strong points in its favour. For example, it is much more rapid to carry out than applying a series of discrete oscillation frequencies, especially if very low frequencies (long timescales) are important. However, there are some points that we need to keep in mind. The first is that rapid responses can be lost during the period taken to apply the step strain. Second, large strains are always applied during the loading of the instrument with the sample and these may not be applied evenly. It is therefore useful to apply an oscillating strain with an initially large but decreasing amplitude to remove any residual directional stress components and then allow sufficient recovery time for structural 'rebuild' prior to starting the experiment. This is also good practice with the oscillation experiment. Third, we should keep in mind that the solution to a sum of exponential functions is difficult mathematically. It is known as an 'ill-conditioned' problem and will not give a unique solution. The simplest approximation to estimating the spectrum from the relaxation curve [11] is built into some equipment software and approximates the relaxation spectrum to the slope of the relaxation modulus curve, as follows:

$$H(\ln t) \approx \frac{-\mathrm{d}G(t)}{\mathrm{d}\ln t} \tag{9.47}$$

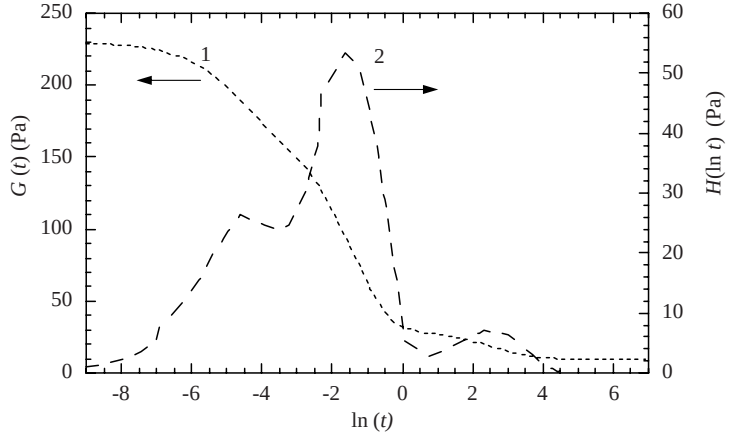


Figure 9.18 The stress relaxation modulus (1) and the relaxation spectrum (2) as a function of time: $\tau_1 = 15$ s, $\tau_2 = 0.25$ s, $\tau_3 = 0.01$ s

The stress relaxation modulus calculated for the four-component model that is plotted in Figure 9.17 is replotted in Figure 9.18 as a function of ln t, along with the relaxation spectrum calculated from Equation 9.47. Peaks corresponding to the three relaxation times are clearly visible. The area under the spectrum, when added to the $G(0)$ value of 10 Pa, is equal to the value of $G(t \to 0) = G(\infty) = 230$ Pa (see Equation 9.38). The individual relaxation times are resolved but only the value of G_1 can be readily estimated just by approximating the peak to a triangular shape of height 10 Pa and a width of 4, giving an area of 20 Pa. Estimation of the values of G_2 and G_3 would require a model of the peak shapes as the two peaks overlap each other.

9.5.4 Creep Compliance

In this experiment, a step stress is applied and the compliance, $J(t)$ (the measured strain per unit applied stress), is recorded as a function of time. With the modern range of controlled stress instruments, it is easier to apply the stress in a very short time than the strain in the step strain experiment. Models such as those illustrated in Figure 9.8d and f are typical of those used to describe the response. Figure 9.8d shows a model of a solid with a retarded response – this is known as a *Kelvin* or *Voight* model. Springs or dashpot dampers can be added to provide an instantaneous elastic response or a fluid response, respectively. The

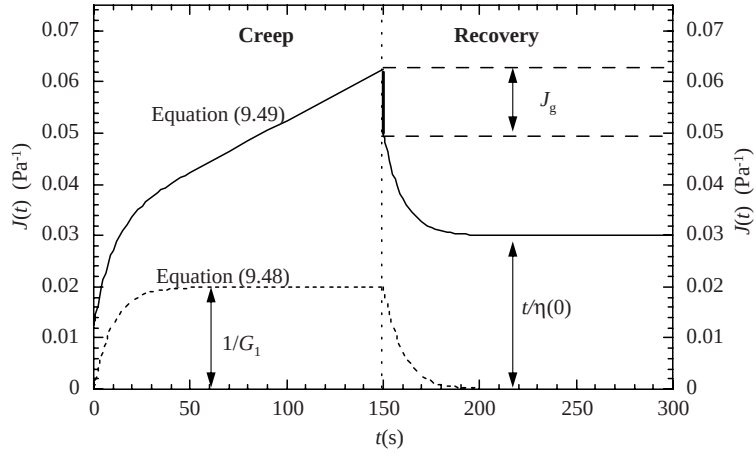


Figure 9.19 Creep and recovery curves for the Kelvin model and the Burger body calculated by using Equations (9.48) and (9.49): $G_1 = 50$ Pa; $\tau_K = 10$ s; $G_2 = 1/J_g = 80$ Pa; $\eta_2 = \eta(0) = 5$ kPa s

Burger body shown in Figure 9.8f has both. The compliances of the individual elements in series are simply added to give the compliance of the whole model, as follows [12]:

$$J(t) = \frac{\gamma(t)}{\sigma_0} = \frac{1}{G}\left[1 - \exp\left(-\frac{t}{\tau_K}\right)\right] \qquad \text{Kelvin model (solid)} \quad (9.48)$$

$$J(t) = \frac{1}{G_2} + \frac{1}{G_1}\left[1 - \exp\left(\frac{t}{\tau_K}\right)\right] + \frac{t}{\eta_2} \qquad \text{Burger body (liquid)} \quad (9.49)$$

where $\tau_K = \eta_1/G_1$ and is known as the *Kelvin retardation time*, Note that here η_2 is the zero shear viscosity, $\eta(0)$, and G_2 is equal to $G(\infty)$ of the material described by the model.

In addition, it should be noted that $1/G(\infty) = J_g$, the *glassy compliance* of the material. Curves calculated from Equations 9.48 and 9.49 are plotted in Figure 9.19 for both the creep and recovery response after the application of a step stress of 1 Pa. Note that only the elastic components give recovery – the viscous compliance is permanent.

The limiting slope of the creep curve gives the zero shear viscosity [i.e. $\eta(0) = \eta_2$ in Equation 9.49], but there is often an experimental problem associated with its determination from this part of the data. We have to be sure that we are still working at small enough strains that linear

behaviour is occurring and fairly large strains may be used before the experimental curve appears to have become linear. It is straightforward to see if too large a strain has occurred as the elastic recovery should be the same as the initial elastic strain. Therefore, although it is tempting to leave the stress applied to the sample until the compliance response becomes linear, the likely result will be a very small elastic recovery. The inverse of the slope then is not the zero shear viscosity because the structure has been markedly perturbed by the strain – here, strain melting has occurred. The solution is to apply the stress for a short enough time to ensure that only small strains are induced and then to analyse the recovery curve.

The response described by the constitutive equation for the Burger body (Equation 9.49) has three elements. It may be generalized by adding a range of retardation times, either as a summation or in an integral form, as follows:

$$J(t) = J_g + \sum_1^n \left[1 - \exp\left(-\frac{t}{\tau_i}\right)\right] \qquad \text{(solid)} \tag{9.50a}$$

$$J(t) = J_g + \sum_1^n \left[1 - \exp\left(-\frac{t}{\tau_i}\right)\right] + tJ_v \qquad \text{(liquid)} \tag{9.50b}$$

where the steady state or viscous compliance is $J_v = \eta(0)$. The integral forms of these equations are as follows:

$$J(t) = J_g + \int_{-\infty}^{+\infty} L(\ln t)\left[1 - \exp\left(-\frac{t}{\tau}\right)\right] \mathrm{d}\ln\tau \qquad \text{(solid)} \tag{9.51a}$$

$$J(t) = J_g + tJ_v + \int_{-\infty}^{+\infty} L(\ln t)\left[1 - \exp\left(-\frac{t}{\tau}\right)\right] \mathrm{d}\ln\tau \qquad \text{(liquid)} \tag{9.51b}$$

where $L(\ln t)$ is the retardation spectrum of the material. This may also be determined from oscillating stress experiments carried out at different frequencies. It should be noted that for a linear viscoelastic material that the complex compliance is

$$J^* = G^{*-1} \tag{9.52}$$

The dynamic compliances are

$$J'(\omega) = \int_{-\infty}^{+\infty} L(\ln t) \frac{1}{1+(\omega\tau)^2} \mathrm{d} \ln \tau \tag{9.53a}$$

$$J''(\omega) = \int_{-\infty}^{+\infty} L(\ln t) \frac{\omega\tau}{1+(\omega\tau)^2} \mathrm{d} \ln \tau \tag{9.53b}$$

The manufacturers of controlled stress rheometers frequently output the data as dynamic moduli using the following relationships:

$$G'(\omega) = \frac{J'(\omega)}{[J'(\omega)]^2 + [J''(\omega)]^2} \tag{9.54a}$$

$$G''(\omega) = \frac{J''(\omega)}{[J'(\omega)]^2 + [J''(\omega)]^2} \tag{9.54b}$$

However, it is important to note that the relationships expressed by Equations 9.54a and 9.54b are only true as long as the material is within the linear range of behaviour. As with all the experiments described in this section, the first stage in the analysis of a new material should be to determine its linear range of response.

9.5.5 The Limiting Behaviour at Long or Short Times

The limiting responses at long or short times can be important in the end uses of colloidal materials and are often important as characterization parameters. For example, Equations 9.36 and 9.38 give the high-frequency limit to the storage modulus in terms of the integral of the whole relaxation spectrum for a liquid-like system and a solid-like system, respectively, whereas Equation 9.37 gives the zero shear viscosity for the liquid-like system derived from the spectrum. The challenge is to predict these values from our knowledge of the colloidal particles. The zero-shear viscosity of colloidally stable and concentrated dispersions of charged spherical particles has already been discussed. The problem of the viscosity of long rods has much in common with that of semi-dilute

polymer solutions [13] utilizing reptation dynamics. Some of the models have also been reviewed by Goodwin and Hughes [5].

Modelling of the high-frequency elastic response has not been discussed so far, although in some cases it is relatively straightforward to carry out [5]. It is useful to think in terms of the family of material properties consisting of the internal energy density, the excess osmotic pressure (which is just the contribution due to the interactions between the particles) and the high-frequency shear modulus of the dispersion [5]:

$$\frac{\bar{E}a^3}{k_B T} = \frac{9\varphi}{8\pi} + \frac{3}{2}\varphi \int_0^\infty r^2 g(r) \frac{V(r)}{k_B T} \mathrm{d}r \tag{9.55a}$$

$$\frac{\Pi a^3}{k_B T} = \frac{3\varphi}{4\pi} - \frac{3\varphi^2}{8\pi a^3} \int_0^\infty r^3 g(r) \frac{\mathrm{d}}{\mathrm{d}r}\left[\frac{V(r)}{k_B T}\right] \mathrm{d}r \tag{9.55b}$$

$$\frac{G(\infty)a^3}{k_B T} = \frac{3\varphi^2}{40\pi a^3} \int_0^\infty g(r) \frac{\mathrm{d}}{\mathrm{d}r}\left\{r^4 \frac{\mathrm{d}}{\mathrm{d}r}\left[\frac{V(r)}{k_B T}\right]\right\} \mathrm{d}r \tag{9.55c}$$

In each of these equations, we are integrating the interaction over dispersion structure, whether the interaction energy, the interparticle force or the 'colloidal spring constant'. The particle size and volume fraction are readily determined experimentally and we can calculate the pair potential. In most colloidal systems, the first nearest-neighbour interactions dominate the response and so the use of the pair potential is often a reasonable approximation. The problem is then the pair-distribution function. In principle, this may be determined from the inversion of the structure factor obtained from scattering, but it may also be modelled using the statistical mechanical techniques for molecular liquids. Comparison of the values calculated from Equation 9.55c with those obtained from experiments using model of spherical particles [14], with a short-range attractive minimum in the pair potential, show excellent agreement. These dispersions contained particles with a weak attractive minimum between the particles. It is interesting that the longer the range of the interaction, the lower is the volume fraction at which elastic responses appear. Fluid-like systems can be found with the longer range attractions, whereas rigid solid phases are produced when the interaction

is of very short range compared with the particle radius. It should also be noted that we are dealing here with non-equilibrium structures. Just the process of mixing systems or placing them in a rheometer cell for measurement tends to homogenize the structures, which is one of the reasons why methods developed for the liquid-state calculations can be so successful for systems that are clearly solids in that yield-behaviour $G(0)$ values can be seen.

The situation can be simplified when there is a long-range repulsion between the particles with a narrow size distribution. In this case, local structure becomes ordered, which simplifies the structure factor. The separation is the same for all the nearest neighbours and the structure has face-centred cubic (fcc) symmetry. In some cases, long-range order can be produced, that is, we have a colloidal crystal. However, it is much more usual that the order is short range and we should think in terms of a 'colloidal glass' (see Chapter 8). The regularity of the spatial aspects of the structure enable Equations 9.55a–c to be written more simply for a spherical cell with a coordination number z and mean centre-to-centre particle separation as

$$\frac{\bar{E}a^3}{k_\mathrm{B}T} = \frac{9\varphi}{8\pi} + \frac{3\varphi z}{8\pi}\frac{V(R)}{k_\mathrm{B}T} \quad \text{with} \quad R = 2a\left(\frac{\varphi_\mathrm{m}}{\varphi}\right)^{\frac{1}{3}} \tag{9.56a}$$

$$\frac{\Pi a^3}{k_\mathrm{B}T} = \frac{3\varphi}{4\pi} - \frac{\varphi z}{8\pi}\frac{\mathrm{d}}{\mathrm{d}R}\left[\frac{V(R)}{k_\mathrm{B}T}\right] \tag{9.56b}$$

$$\frac{G(\infty)a^3}{k_\mathrm{B}T} = \frac{\varphi z R^2}{40\pi}\frac{\mathrm{d}^2}{\mathrm{d}R^2}\left[\frac{V(R)}{k_\mathrm{B}T}\right] \tag{9.56c}$$

A similar result to Equation 9.56c can be derived using a 'zero-Kelvin lattice model' [5] (zero-Kelvin so there is no diffusive motion and therefore no relaxation). With a slightly different spatial averaging process from the spherical cell surface giving Equation 9.56c the result was

$$\frac{G(\infty)a^3}{k_\mathrm{B}T} = \frac{3\varphi z R^2}{256\pi}\frac{\mathrm{d}^2}{\mathrm{d}R^2}\left[\frac{V(R)}{k_\mathrm{B}T}\right] \tag{9.56d}$$

The effectiveness of Equation 9.56d in accurately describing the elastic response has been demonstrated using a range of charge stabilized monodisperse particles. Again, it should be noted that the structures are non-equilibrium due to preparation and handling and are similar to

colloidal glasses. The equilibrium structure would be in the form of colloidal crystals and these can be observed with small particles just above the phase transition. At higher concentrations, the equilibrium structure is rarely if ever attained. The forces are so weak that the long-range structure is readily broken on very gentle stirring although the short-range structure is maintained.

Polymers are often used to produce weak gels so that sedimentation is controlled at the same time as high-shear viscosity is maintained at a chosen level. With polymer gels, we are not usually concerned with the high-frequency limit of the storage modulus but the plateau or network modulus. This is observed over a long range of timescales [11, 13]. The faster processes which result in the $G(\infty)$ are characterized by the relaxation times associated with the motion of one or two repeat units on the chain. The much slower cooperative motion of reptation governs the other end of the timescale observed for simple homopolymers. With cross-linked polymers, the lifetime of the links is much longer than that of an entanglement. These cross-links can be due to covalent bonds, ion–chain interactions, particle–chain interactions or hydrophobe self-assembly. The simplest treatment is to model the system as a swollen elastomeric network so that the network modulus, G_N, is simply proportional to the number of links contributing to the network, N_L:

$$G_N = N_L k_B T \tag{9.57}$$

There are models available to calculate N_L for the various types of polymer network [5]. It should be noted that N_L is the number density of *effective* links and so requires some statistical modelling in most cases.

9.5.6 Time–Temperature Superposition

The stress relaxation process takes place via the diffusion of the microstructural components making up the material we are studying. With polymer systems, the relaxation occurs by multiple diffusional modes. The slowest mode, the centre-of-mass diffusion of the molecules, does not lead to structural relaxation as their motion does not enable the molecule to 'wriggle' past its neighbours. The next slowest mode is the two halves of the molecule moving relative to each other although still connected. We can divide the molecule into smaller and smaller sections so that each is moving relative to its neighbours without

losing connectivity. These are known as Rouse modes [5], with the zero mode being the centre-of-mass diffusion mode. The limiting mode corresponds to the motion of the building blocks of the chain. This means that a very wide range of diffusional times are contributing to the relaxation of a high polymer. Twelve orders of magnitude are not uncommon.

Viscosity is a function of temperature and, for simple liquids, the change of viscosity with temperature is well described by the Arrhenius relationship [5]:

$$\eta_T = A \exp\left(\frac{E}{RT}\right) \tag{9.58}$$

where E is the activation energy for flow in the system and R is the gas constant. The diffusional relaxation time is directly proportional to the 'viscosity' of the system that we are studying. By viscosity we mean both the hydrodynamic and thermodynamic contributions to the motion of the microstructure. We can anticipate the same strong dependence on temperature with the complex viscosity and related viscoelastic functions.

Figure 9.20 shows the experimental data for the frequency dependence of both the storage and loss moduli of 50 Pen bitumen that is used as part of a road-surfacing procedure. The moduli move to higher frequencies with increasing temperature. Note that at frequencies in excess of 30 rad s^{-1} the data become unreliable as the resonance frequency of the instrument as set up is reached. We can define the characteristic frequency as the point at which the storage and loss moduli are equal and this also defines a characteristic time $\tau_c = {\omega_c}^{-1}$. The dependence of the characteristic frequency on absolute temperature is shown in Figure 9.21 and a good fit to the data is given by

$$\log \omega_c = 0.125T - 35.34 \tag{9.59}$$

and in terms of a shift in characteristic frequency with temperature, a_T, we have

$$\log a_T = \log\left(\frac{\omega_{c2}}{\omega_{c1}}\right) = 0.125\,(T_2 - T_1) \tag{9.60}$$

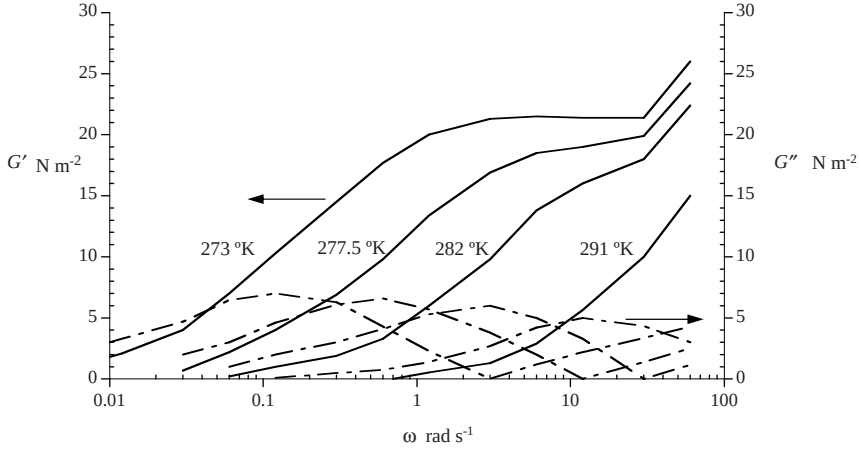


Figure 9.20 Storage and loss moduli for a 50 Pen bitumen measure as a function of temperature. (—) G'; (– · –) G''

For polymer systems that have a glass transition temperature, T_g, the WLF equation [11, 15] provides a better description:

$$\log a_T^{-1} = -\frac{17.4\,(T - T_g)}{51.6 + (T - T_g)} \tag{9.61}$$

where we have used the glass transition temperature as our start point.

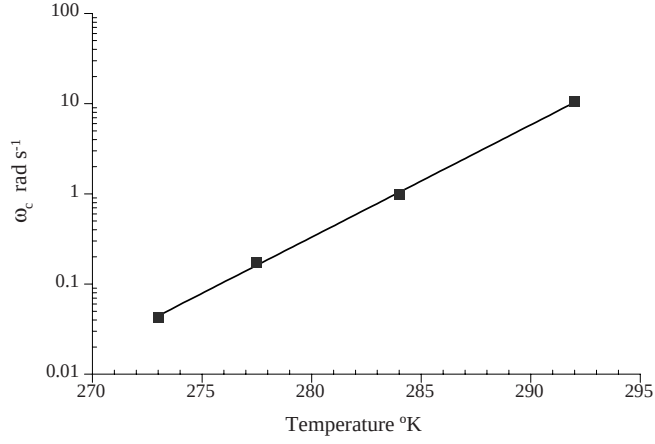


Figure 9.21 Logarithm of the characteristic frequency as a function of absolute temperature for a 50 Pen bitumen

By applying Equation 9.59 to the data obtained for the bitumen sample at 0 °C to estimate the behaviour at 18 °C we can estimate the response of the material at the higher temperature to a higher frequency than was obtainable with our instrument, that is, to $\omega > 30\,\mathrm{rad\,s^{-1}}$. The result is shown in Figure 9.22 and the limited data obtained at 18 °C are also plotted for comparison. An estimate of the response at 18 °C up to a frequency of $10^4\,\mathrm{rad\,s^{-1}}$ is obtained. The storage modulus overlap indicates that the estimate is reasonable; however, the fit of the loss modulus is less satisfactory as the experimental data indicate a ~28% decrease in $G''(\omega)$ as the temperature increases from 0 to 18 °C. The following observations are important:

- We can see from the higher temperatures that the system is a viscoelastic fluid as the loss modulus diverges from the storage which goes to zero. We could not say this from just the low-temperature experiment.
- There is clearly a more rapid relaxation process than the one for which we can obtain a clear picture. Lower temperatures would have been required to reveal the detail of that one. The crossover of the storage and loss moduli occurs just downfield of the peak loss

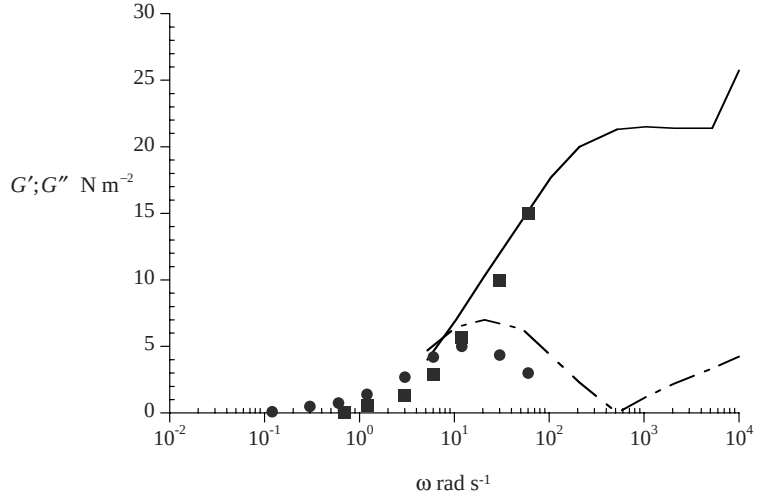


Figure 9.22 Storage and loss modulus of a 50 Pen bitumen at 291 K. (—) G' calculated from Equation 9.59, (■) experimental data; (– · –) G'' calculated from Equation 9.59, (•) experimental data

and at ~25% of the plateau of the storage modulus, indicating that we have captured a large proportion of the behaviour.
- We are observing the slowest relaxation process which will dominate the limiting viscous response.

Bitumen is a complex colloidal material that is made up of asphaltenes, which are very large, insoluble stacks of aromatic molecules peptized by aromatic resins in a hydrocarbon matrix [16]. The latter are referred to as maltenes. Bitumens of the 50 Pen type contain ~15% asphaltenes with ~20% of the stabilizing resins. This composition results in a partially peptized asphaltene dispersion so that an interconnected network is formed and is known as a 'gel' bitumen. The bitumen is hot rolled with aggregate on to the road surface and 'cures' with time, losing lower molecular weight hydrocarbons in the process. At temperatures >45 °C, wheel tracks can begin to form and hence the rheological properties over the normal climatic temperature range are extremely important. With a polymer network, we expect the plateau value of the storage modulus to increase with temperature as in Equation 9.57. With a partially peptized colloidal network, the extent of peptization and hence the cross-link density could vary and result in a subtle variation in moduli with temperature, even though the shift in relaxation time with temperature is reasonably well predicted by the time–temperature superposition principle.

9.5.7 Processing Effects

The key factors that we need to aid our understanding of the behaviour of concentrated dispersions are the interactions between particles and the structure of the system. It is usually possible to calculate the pair potential even if the potential of mean force may be more difficult to estimate. The structure is the major problem. Whenever we measure our materials, they will have been subjected to large deformations and the effects of this are frequently important. For example, we can set up a model experiment to study diffusion-limited aggregation and start to observe fractal clusters being formed. As soon as the system is transferred to a container or an instrument cell, it is subjected to large deformations and we will then be studying a *processed* system. This may work in our favour in many instances but it may also lead to long-term changes on storage. To clarify these points, we will consider one or two examples.

In some instances, concentrated systems which are coagulated are used. Clay slurries for slip-casting, or at higher concentrations for ceramic materials, are an example. If a dispersion is caused to coagulate and then pumped or spread, the shear processing will do two things: it will tend to densify the local structure and then break this up into flow units made up of many particles. Spherical polymer particles coagulated with electrolyte and sheared in a controlled fashion produce monodisperse dense spherical clusters [5]. An example of these is illustrated as Figure (d) on page (365). The interacting units then become large clusters so any model based on the summation of all the interactions between all the single particles in the system would fail as a large overestimate of the resistance to continuous motion.

The models used above for the behaviour of concentrated dispersions of uniform-sized particles, which are structured due to either strong repulsive interactions or weak attractions, are based on a structural uniformity that is based on short-range structures. In both cases the lowest energy state that we can visualize would be large crystals rather than the glassy or liquid-like structures that we use. These models work in a satisfactory manner because the intensive mixing that is a part of getting the systems into containers or measuring equipment produces these metastable structures. The relaxation times are often very long – days, weeks or even years. If the weakly attractive system is explored a little further, we can sometimes see unexpected phenomena such as delayed settling. What occurs is that the shearing forces are large enough to pull the individual particles apart and distribute them evenly. On the cessation of shearing, the particles form a uniform liquid-like structure with a coordination number that is a function of the volume fraction. The motion of the individual particles allows slow local densification to occur, for example, 12 nearest neighbours is a lower energy state than eight. The local densification leads to other regions with fewer particles. As these weaknesses grows, a point can be reached where the structure collapses and the liquid is expelled from the voids. The collapse can be rapid once the system reaches a critical state, although it can take a long time to reach that condition.

Mixtures formed with different particle sizes and/or densities can show structural changes as the result of low shear forces. For example, the structure of some systems may become sufficiently broken down, that is, the viscosity has been lowered, by the low frequency, although large-amplitude deformations produced when transporting bulk shipments may mean that large or dense particles can separate out. Simple jar tests in the laboratory may not show this.

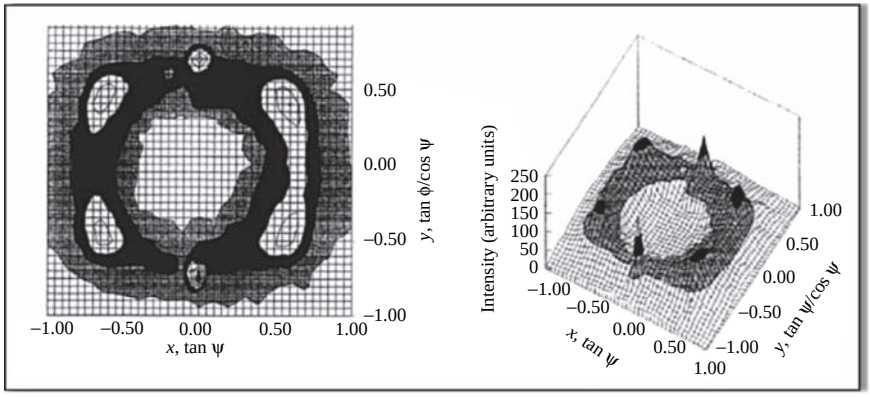


Figure 9.23 Hexagonal spot patterns obtained from diffraction in an optical rheometer for a sample under oscillatory shear [15]. Reprinted with permission from H. Sakabe, the Structure and Rheology of Strongly Interacting Suspensions, PhD Thesis, University of Bristol, Bristol, UK Copyright (1995) University of Bristol

Another interesting example of the effect of low shear is the effect on concentrated systems formed from stable particles. Oscillation can enhance the structuring rate and large amplitudes are more effective than small. Figure 9.23 shows the optical diffraction pattern produced in an optical rheometer [17]. Systems of bimodal particles can be induced to crystallize by slow rotation of the container holding them if the particle size ratio and concentrations are carefully chosen [18].

9.6 SEDIMENTATION IN CONCENTRATED DISPERSIONS

Batch sedimentation in dilute systems is discussed in Chapter 8. We now need to extend this to a broader range of conditions that apply to many systems of practical importance. Indeed, the control of the sedimentation or creaming behaviour is a common problem with formulations of many colloidal systems. The formulation scientist often adds material to 'thicken' the product. Soluble polymers or clays are common choices and low levels of additives are needed to minimize changes to the final application of the product and also to keep costs as low as possible. At first sight, thickening the dispersion is a simple concept, but the results are often more complex due to the interaction between the large number of components. In the first two systems discussed below, the Péclet number

is large so that diffusional contributions can be neglected in the analysis without introducing significant error. However, as attractive forces between particles increase sufficiently for significant three-dimensional structures to be formed, diffusion again becomes important in allowing temporal changes in the structures formed.

9.6.1 Thickened Dispersions

A common response to the problem is to add a 'rheology modifier' or 'thickener'. Soluble polymers, dispersions of fine particles such as montmorillonite clays or mixtures of both are commonly chosen additives. At moderate concentrations, these additives produce a medium which is pseudoplastic, that is, there is a limiting viscosity at low stresses, the zero shear viscosity, $\eta(0)$. Now the gravitational force on a particle is

$$F_g = \frac{\Delta\rho 4\pi a^3 g}{3} \tag{9.62}$$

The shear stress at the surface is

$$\sigma_s = \frac{\Delta\rho a g}{3} \tag{9.63}$$

As an example, a mineral particle of radius 1 μm and with a density difference of $2 \times 10^3\,\mathrm{kg\,m^{-3}}$ would exert a stress of $\sim 6 \times 10^{-3}$ Pa. The critical stress at which many thickened media shear thin is typically two orders of magnitude higher and so the zero shear viscosity is the appropriate term to use in the Stokes law settling calculation (see Equation 8.105):

$$v_{s0} = \frac{2a^2\Delta\rho g}{9\eta(0)} \tag{9.64}$$

To put this in practical terms, we can consider a titanium dioxide particle in an aqueous latex paint formulation thickened with an ethyl hydroxyethylcellulose such as Modocoll Ex 600 at ~1.4%, giving the medium a viscosity of ~1 Pa s. The sedimentation rate of the pigment would be ~1 cm per year and there would be a significant bed on the bottom of the can. To prevent the sedimentation completely, the system must have

a static yield value and not be pseudoplastic. Surfactants are used to provide a well-dispersed pigment and polymer latex particles. The polymers are effectively non-absorbing and produce a depletion flocculated dispersion, thus providing additional structure. At the solids concentration used in a latex paint, 1–2% of polymer is sufficient to produce a weak attraction between particles, giving a thixotropic gel with a static yield value which prevents sedimentation of the pigment. Dilution will change the concentrations of both particles and polymers and sedimentation will occur as the structuring properties are changed dramatically.

9.6.2 Sedimentation in Stable Dispersions

The concentration dependence of the sedimentation rate of dispersions at low concentrations is given by Equation 8.108 in the general form

$$\frac{v_s}{v_{s0}} = 1 + K\varphi + O(\varphi^2) \tag{9.65}$$

K can be obtained by measurement using dilute dispersions, so for systems behaving like hard spheres, $K = -6.55$. With systems with repulsion between the particles or with some degree of polydispersity, the value of K would be significantly lower. Russel *et al.* [10] obtained a reasonably good fit to published data with the equation

$$\frac{v_s}{v_{s0}} = (1 - \varphi)^{-K} \tag{9.66}$$

Equation 9.65 has the same form as Equation 9.10 and we may use the same 'effective medium' approach as was used to derive Equation 9.14 from Equation 9.10 to give

$$\frac{v_s}{v_{s0}} = \left(1 - \frac{\varphi}{\varphi_m}\right)^{-K\varphi_m} \tag{9.67}$$

With $K = -6.55$ and $\varphi_m = 0.64$ we have the expression for dispersions of hard spheres (0.64 is the concentration appropriate for dense random packing).

Equation 9.67 was derived by Buscall *et al.* [19] and applied to charged polymer particles and Figure 9.24 shows a plot of the pair potential of

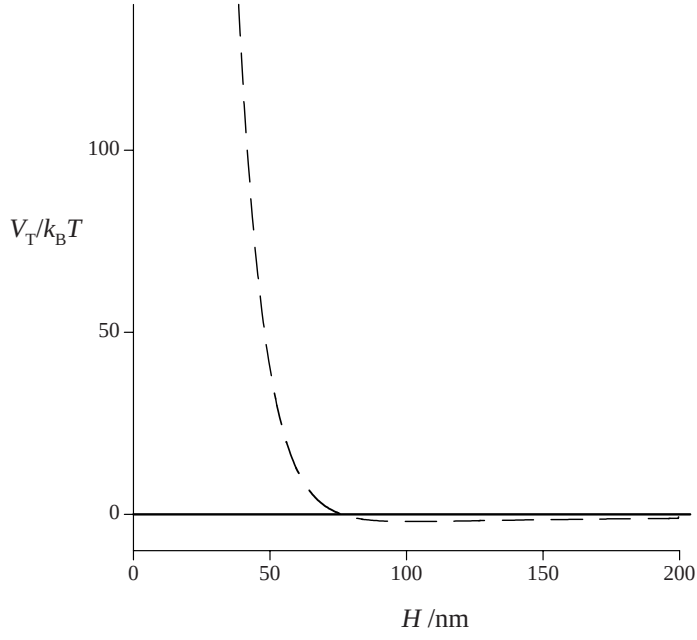


Figure 9.24 Part of the pair potential for polystyrene particles with $a = 1550\,\text{nm}$, with a ζ-potential of $-70\,\text{mV}$ and [NaCl] = 1 mM. There is a strong repulsion as the particles approach but with a shallow secondary minimum attraction of $2k_BT$ at 100 nm

the polystyrene particles used. A good fit to their data can be obtained with

$$\frac{v_s}{v_{s0}} = \left(1 - \frac{\varphi}{0.74}\right)^{(5.4\times 0.74)} \tag{9.68}$$

Figure 9.25 shows plots of Equations 9.66–9.68. Here the value of $K = -5.4$ in Equation 9.68 was the value that provided a good fit to the data at the dilute part of the concentration range, reflecting the change from a random distribution of particles caused by the repulsive pair potential but with a shallow long-range attraction modifying the near-field hydrodynamic contribution to the relative sedimentation rate. In this case $\varphi_m = 0.74$, the value for fcc packing which can be formed by uniform spheres with a repulsive potential. With $\kappa a \approx 150$, the interparticle separation in the sediment would provide only a very small correction to 0.74.

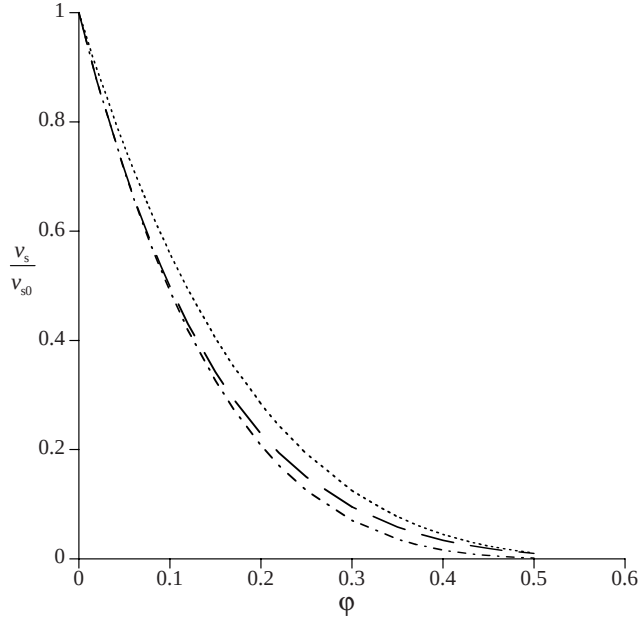


Figure 9.25 Relative sedimentation velocity as a function of volume fraction. (——) Equation 9.66, $K = -6.55$; (-·-·-) Equation 9.67 with $K = -6.55$, $\varphi_m = 0.64$; (····) Equation 9.68 with $K = -5.4$, $\varphi_m = 0.74$

9.6.3 Strongly Aggregated Systems

Coagula in a strongly aggregated system of particles will settle until a percolating network is formed. The kaolinite dispersions studied by Michaels and Bolger [20] are an excellent example of this type of system. A little consolidation at the base of the bed may occur due to the relative mass of the network. The effects of shear processing are important and well-defined compact aggregates can be formed, and it is these aggregates which become the sedimenting units.

Systems prepared at volume fractions in excess of the percolation threshold do not sediment due to the strong interparticle forces. When strongly aggregated particulate systems are prepared at volume fractions in excess of this threshold, the structures are dominated by the shear processing that has occurred in the preparation and dense, strong, non-sedimenting structures are formed. It should be noted that the structures are usually non-equilibrium, but due to the strong forces, the timescale of changes towards equilibrium are extremely long and any changes that do occur normally strengthen the structures – an effect that is readily

observed when using modelling clay, which becomes more malleable as it is worked, only to increase in rigidity on standing.

9.6.4 Weakly Aggregated Dispersions

This is a very important class of colloidal materials and many products fall into this category as we endeavour to formulate them to be non-sedimenting while being readily stress melted so that their application is easily achieved. We noted in Section 9.6.2 that a shallow secondary minimum can slow the sedimentation rate by enhancing the near-field hydrodynamic effects due to the perturbation from a random distribution of particles. In the case illustrated in Figure 9.24, the maximum well depth was only $2k_BT$. Deeper wells can be obtained with higher electrolyte concentration, lower ζ-potentials or denser particles. It should also be noted that the width of the well is several tens of nanometres. We also discussed in Section 9.6.1 that frequently when soluble polymeric thickeners are added to a dispersion, a synergism is created by the polymer causing the particles to be depletion flocculated – another form of weak aggregation. In this case, the width of the potential well scales with the radius of gyration of the polymer additive. The probability of a particle escaping from a shallow potential well of depth V_{sm} is proportional to $\exp(V_{sm}/k_BT)$. If the well is $\sim 3k_BT$, 5% of the particles have sufficient energy to escape. However, the residence time is also a function of the width of the well [21] as the particle has to diffuse up the potential gradient until $exp(V_{sm}/k_BT) \ll 1$. In Section 9.2, we saw how the structure evolves with volume fraction for weakly aggregated systems. An example was illustrated in Figure 9.3. The energy required to escape from a particle's neighbours increases with the coordination number.

Buscall *et al.* [22] have brought together and analysed the data from several systems which had attractive pair potentials of -5 to $-15k_BT$. All the systems formed space filling networks at moderate concentrations and initially formed weak gels that appeared to be non-sedimenting. These gels showed the phenomenon of *delayed sedimentation*, where after a period of time, which could vary from minutes to months, the gels rapidly collapsed to a denser sediment. The delay time increased with both the initial particle number density and the well depth. The mechanism is as follows. The weak gels are formulated and the 'uniform' but non-equilibrium structure is space filling and is strong enough to support itself. Over time, the particles diffuse into lower energy configurations, that is, denser regions with higher coordination numbers.

Concomitant with the local densification is the formation of weaker regions with lower particle number densities. The process continues until the weaknesses combine so that the gel strength becomes critical and collapse ensues. As the dense regions move down, dispersion medium moves up through the fault lines, often creating clearly visible tracks.

The parameter that we measure is a *delay time*, τ_d. For the change to be observed, there has to have been a large number of individual particle 'diffusional events' occurring whereby particles are changing their neighbours. Following Buscall *et al.* [22], we can write that

$$\tau_d \propto \tau_{esc} \tag{9.69}$$

where τ_{esc} is the time that it takes for a particle to escape from the first nearest-neighbour shell, that is, the long-time self-diffusion time. We can obtain a simple approximation for this if we initially consider a pair of particles with an attractive pair potential in the form of a square well of depth V_{min} and width Δ. The diffusional time for a particle to move a distance Δ is

$$\tau_\Delta = \frac{6\pi\eta_0 a\Delta^2}{k_B T} \tag{9.70}$$

The escape frequency will be reduced by the Boltzmann factor:

$$\tau_\Delta = \frac{6\pi\eta_0 a\Delta^2}{k_B T \exp\left(\dfrac{-V_{min}}{k_B T}\right)} \tag{9.71}$$

The total interaction of a reference particle in the structure can be taken as the well depth multiplied by the coordination number of the particles in the structure, z. This is a reasonable first approximation to the potential of mean force given the scaling of the range of the attraction with the particle radius for nearly all experimentally accessible systems. The coordination number in the structure is a function of the volume fraction as expressed by Equation 9.6 and we may rewrite Equation 9.71 as

$$\tau_d \propto \tau_{esc} = \frac{6\pi\eta(0) a\Delta^2}{k_B T \exp\left(\dfrac{-zV_{min}}{k_B T}\right)} \tag{9.72}$$

The change from η_0 to $\eta(0)$ is a crude approximation to add a contribution to the slowing of the diffusion due to hydrodynamics and it should be the value for hard spheres (Equation 9.15) so that it does not include a contribution to the colloid interaction. The value of $V_{\min}$ may be calculated from the sum of the components of the pair potential given in Chapters 3 and 4. The width of the well, Δ, is the radius of gyration for a depletion flocculation system but for a secondary minimum aggregation we need to define a cut-off. A reasonable approximation would be to calculate the distance, H, when the binding energy

$$z V_{\min} \leq k_B T \tag{9.73}$$

Even though we may have done a high level of shear processing to produce a 'uniform structure', the local density (coordination number) will change above and below the initial mean value as the changes progress. The result is that a distribution of coordination numbers will occur as the structural changes progress. As the binding energy increases, the initial shear processing may not achieve a uniform distribution at the outset. The result is that the general case would be a spectrum of binding energies so that the single exponential in Equation 9.72 would have to be replaced by the sum of exponentials weighted with a probability reflecting their contribution [22]:

$$\tau_d \propto \langle \tau_{esc} \rangle = \frac{6\pi\eta(0) a \Delta^2}{k_B T \left[\sum_i w\ (z_i) \left(\frac{-z_i V_{\min}}{k_B T} \right) \right]} \tag{9.74}$$

Now, if the binding energy is low, the initial processing will produce a uniform structure as the shear forces are large in comparison with the interparticle forces, but also and importantly the particles can diffuse in and out of local structures, thereby maintaining greater uniformity. This means that we can anticipate the limiting response at low binding energies to be described by the single exponential dependence given in Equation 9.68. With stronger binding energies, the structural variation could increase with the denser regions becoming more like particle 'sinks' with less two-way traffic, which may account for the shorter delay times than would be expected from a simple exponential dependence on binding energy. From an analysis of a wide range of data, Buscall *et al.* [22] found the following functional dependence of the normalized relative

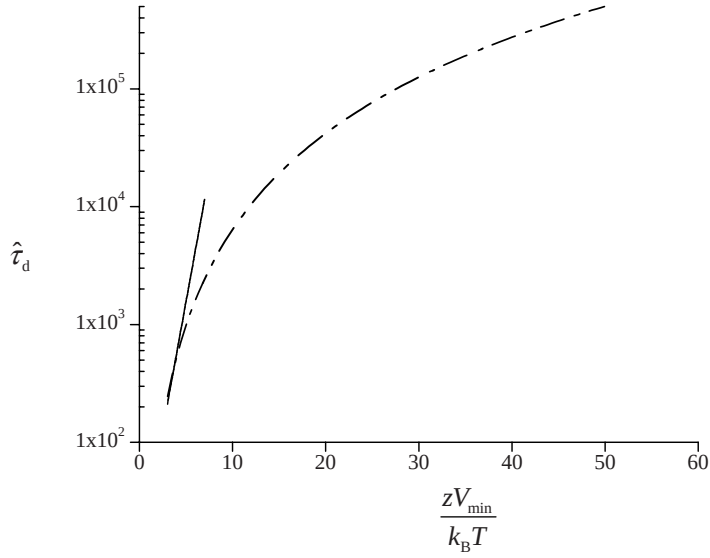


Figure 9.26 Normalized relative delay time, $\hat{\tau}_d$, plotted as a function of the relative binding energy, $-zV_{min}/k_B T$. (—) $\hat{\tau} \propto \exp(-zV_{min}/k_B T)$; (–·–·–) $\hat{\tau}_d \propto (-zV_{min}/k_B T)^{2.7}$

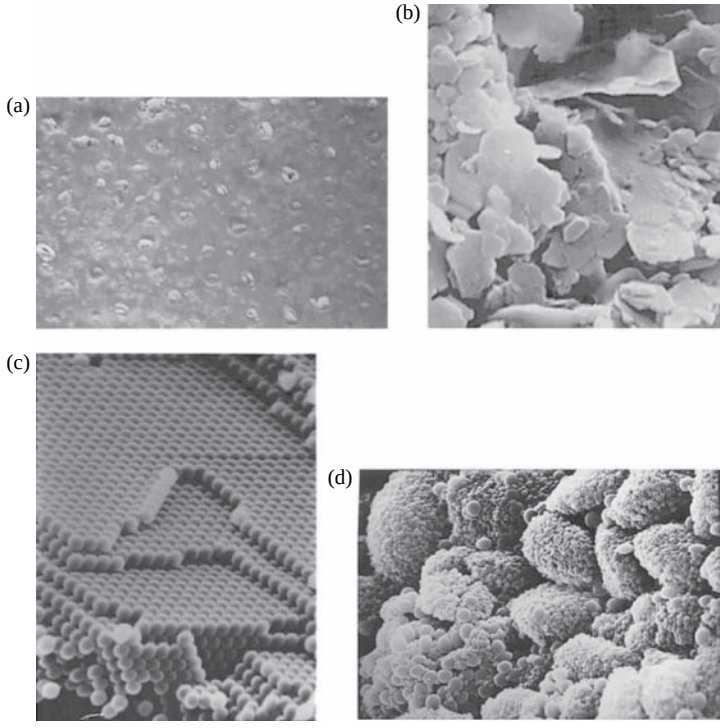
delay time, $\hat{\tau}_d$, on the initial binding energy as

$$\hat{\tau}_d \propto \exp\left(\frac{zV_{min}}{k_B T}\right) \quad \text{for} \quad \left(\frac{zV_{min}}{k_B T}\right) < 5$$
$$\hat{\tau}_d \propto \left(\frac{zV_{min}}{k_B T}\right)^{2.7} \quad \text{for} \quad \left(\frac{zV_{min}}{k_B T}\right) > 5 \tag{9.75}$$

These functions are plotted in Figure 9.26.The highest value of the binding energy analysed by Buscall *et al.* [22] was ~60, so the transition from weakly to strongly aggregated systems was not highlighted. The practical problem in the product formulation laboratory is that of increasing the binding energy by manipulation of the chemical environment in order to take the delay time to such a high value that sedimentation will not occur during the shelf-life of the product. The challenge then becomes to limit the stress-melting threshold of the structure to low values so that the use of the product is not adversely affected.

REFERENCES

1. Hunter, R. J. (1989) *Foundations of Colloid Science*, Vol. **II**, Oxford University Press, Oxford.
2. Baxter, R. J. (1968) *Aust. J. Phys.*, **21**, 563.
3. Ottewill, R. H. (1982) Concentrated dispersions, in *Colloidal Dispersions* (ed. J. W. Goodwin), Royal Society of Chemistry, London, pp. 197–217.
4. Goodwin, J. W. and Mercer-Chalmers, J. (1998) Flow induced aggregation of colloidal particles, in *Modern Aspects of Colloidal Dispersions* (eds R. H. Ottewill and A. Rennie), Kluwer, Dordrecht, pp. 61–75.
5. Goodwin, J. W. and Hughes, R. W. (2000) *Rheology for Chemists, an Introduction*, Royal Society of Chemistry, Cambridge.
6. Krieger, I. M. and Dougherty, T. J. (1959) *Trans. Soc. Rheol.*, **3**, 137.
7. Ball, R. and Richmond, P. (1980) *Phys. Chem. Liquids*, **8**, 99.
8. Woods, M. E. and Krieger, I. M. (1970) *J. Colloid Interface Sci.*, **34**, 417.
9. Ohtsuki, T. (1981) *Physica A*, **108**, 441.
10. Russel, W. B., Saville, D. A. and Schowalter, W. R. (1989) *Colloidal Dispersions*, Cambridge University Press, Cambridge.
11. Ferry, J. D. (1980) *Viscoelastic Properties of Polymers*, 3rd edn, Wiley-Interscience, New York.
12. Tanner, R. I. (1985) *Engineering Rheology*, Oxford University Press, Oxford.
13. Doi, M. and Edwards, S. F. (1986) *The Theory of Polymer Dynamics*, Oxford University Press, Oxford.
14. Goodwin, J. W., Hughes, R. W., Partridge, S. J. and Zukoski, C. F. (1986) *J. Chem. Phys.*, **85**, 559.
15. Williams, M. L., Landel, R. F. and Ferry, J. D. (1955) *J. Am. Chem. Soc.*, **77**, 3701.
16. Read, J. and Whiteoak, D., (2003) *The Shell Bitumen Handbook*, 5th edn, Thomas Telford, London.
17. Sakabe, H. (1995) The structure and rheology of strongly interacting suspensions, PhD Thesis, University of Bristol.
18. Johnson, G. D. W., Ottewill, R. H. and Rennie, A. (1998) Characterization of particle packing, in *Modern Aspects of Colloidal Dispersions* (eds R. H. Ottewill and A. Rennie), Kluwer, Dordrecht, pp. 89–100.
19. Buscall, R., Goodwin, J. W., Ottewill, R. H. and Tadros, Th. F. (1982) *J. Colloid Interface Sci.*, **85**, 78.
20. Michaels, A. S. and Bolger, J. C. (1962) *Ind. Eng. Chem. Fundam.*, **1** (3), 24.
21. van Biejeren, H. (2003) *J. Stat. Phys.*, **110**, 1397.
22. Buscall, R., Choudhury, T. H., Faers, M. A., Goodwin, J. W., Luckham, P. A. and Partridge, S. J. (2009) *Soft Matter*, **5**, 1345.



Scanning electron micrographs of some colloidal particles:

(a) A film of printing ink showing well-dispersed titanium dioxide pigment (3400×).
(b) Sodium kaolinite (china clay) particles in an open 'card-house' aggregate (9100×).
(c) Dried polystyrene latex showing how the order in a concentrated dispersion is maintained on drying (9100×).
(d) Coagulated bimodal latex mixture after shear processing showing densely packed, uniform aggregates (700×). Reproduced with permission from 'Modern Aspects of Colloidal Dispersions' edited by Ottewill and Rennie. Copyright (1998) Kluwer Academic Publishers.

Index

Note: page numbers in italics refer to figures or tables

Colloids and Interfaces with Surfactants and Polymers, Second Edition James Goodwin

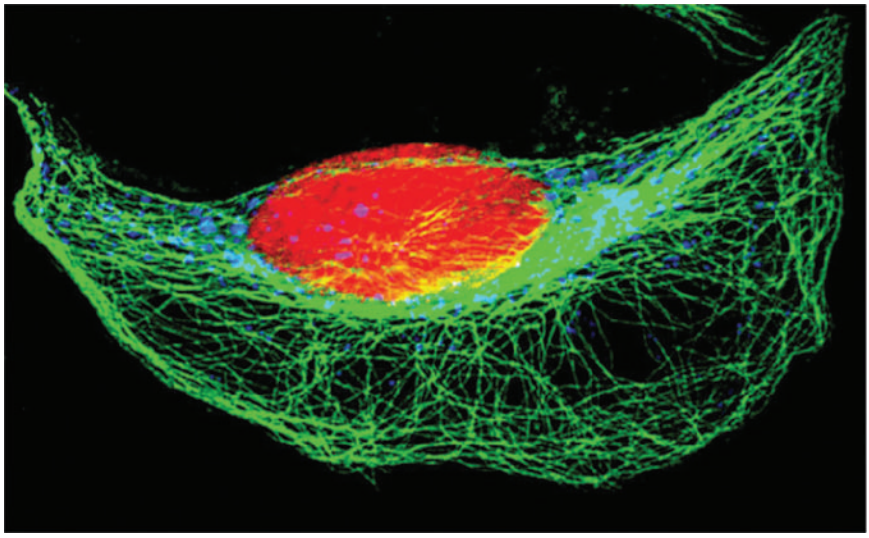


Plate 1 Three-color staining of HeLa cells using fluorescent Qdot® nanocrystal conjugates. The intracellular structures in fixed HeLa cells were visualized using a red-fluorescent Qdot® 655 $F(ab')_2$ goat anti-mouse IgG (Q11021MP, Q11022MP)(nuclei), a yellow-fluorescent Qdot® 585 $F(ab')_2$ goat anti-rabbit IgG (Q11411MP)(Golgi), and a green-fluorescent Qdot® 525 streptavidin conjugate (Q10141MP)(microtubules). Reprinted with permission Life Technologies, Eugene, Oregon, USA.