

Chapter 6

Electrostatic origin of forces

6.1 Microscopic origin of forces

reading: JacobI, 2.6, 5.1, 5.2, 5.5, 5.7, 8.6

In these lectures, we have to some extent reversed the customary order of topics, in which the microscopic origin of forces are treated first, as a prelude to interactions of larger objects such as polymers and interfaces. Instead, we have focused on role of short-range “steric” repulsions and configurational fluctuations in determining self-assembled structure, and so have finessed questions related to the specific origin of forces. Until this point, it sufficed to describe forces with a χ parameter or a hard-core repulsion. But in the treatment of surfactants, which self-assemble typically in water and are often charged, we are finally obliged to give an account of interactions involving charges.

From the point of view of physical chemistry, there is a great variety of microscopic forces: Coulomb forces between charged species, charge-dipole and dipole-dipole interactions, interactions between charges or dipoles and polarizable species, and finally London or Keesom interactions between thermally or quantum-mechanically fluctuating dipoles. All of these interactions ultimately arise from the Coulomb interaction that acts on electrons and nuclei. In this lecture, we shall review the electrostatic origin and particular properties of this panoply of interactions.

Review of electrostatics.

We begin with a brief review of the equations and main results of electrostatics, the starting point for which is Coulomb’s law, which describes the interaction energy of two point charges Q_1 and Q_2 at separation r in a medium of dielectric constant ϵ :

$$U = \frac{Q_1 Q_2}{\epsilon r} \quad (6.1.1)$$

(Here we adopt the Gaussian system of units.)

Coulomb’s law motivates the definition of electrostatic potential ϕ , which for a point charge is

$$\phi(r) = \frac{Q}{\epsilon r} \quad (6.1.2)$$

and which superposes for a collection of charges. If we have a charge distribution $\rho(r)$, the potential becomes

$$\phi(r) = \int dr' \frac{\rho(r')}{\epsilon |r - r'|} \quad (6.1.3)$$

Correspondingly, the energy of a point charge Q in a given potential ϕ is

$$U = Q\phi \quad (6.1.4)$$

This relation expresses one of the useful consequences of defining the electrostatic potential; namely, once we know $\phi(r)$, we know the potential energy for a charge added to the system.

In turn, to compute the potential from the charge distribution, we may in principle use the integral form of Coulomb's law above. But it is also useful to have a differential equation relating ϕ to the charge density; this is given by the Poisson equation,

$$\nabla^2\phi = -4\pi\rho/\epsilon \quad (6.1.5)$$

The Poisson equation may be verified as equivalent to Coulomb's law by observing that for a point charge at the origin $\rho(r) = Q\delta(r)$, the potential $\phi(r) = Q/\epsilon r$ is indeed a solution.

In terms of the potential ϕ , we define the electric field E as

$$E = -\nabla\phi \quad (6.1.6)$$

In the same way that ϕ is useful because it gives the energy of an added charge in a given arrangement of other charges that give rise to ϕ , E is useful because it gives the force on the added charge,

$$F = -\nabla U = QE \quad (6.1.7)$$

Electric field lines drawn emanating from positive charge have outwardly directed arrowheads, indicating the direction of force on a second positive charge.

In terms of the electric field, the Poisson equation takes the form

$$\nabla \cdot E = 4\pi\rho/\epsilon \quad (6.1.8)$$

Now we recall the divergence theorem from multivariable calculus, which relates the volume integral of the divergence of a vector field to the surface integral of that field:

$$\int dV \nabla \cdot E = \int dS \cdot E \quad (6.1.9)$$

Applied to the Poisson equation for E , we obtain Gauss' law:

$$\int dS \cdot E = 4\pi/\epsilon \int dV \rho = 4\pi Q/\epsilon \quad (6.1.10)$$

in which Q is the total charge enclosed by the surface S .

Gauss' law is very convenient for computing fields in high-symmetry situations. As an example, consider a charge distribution $\rho(r)$, spherically symmetric about the origin. The electric field E in this situation must point radially, and depend only on the distance to the origin. So if we apply Gauss' law to a spherical surface of radius r , the integral $\int dS \cdot E$ is simply $4\pi r^2|E|$, because $|E|$ is constant on the surface. As a result, Gauss' law implies that

$$E(r) = \hat{r}Q(r)/\epsilon r^2 \quad (6.1.11)$$

in which $Q(r)$ is the total charge from $\rho(r)$ contained within the sphere of radius r .

Dipoles

A dipole can be thought of as a pair of cancelling charges of magnitude Q , separated by a vector a . (More realistically, a dipole may be a smear of positive charge displaced by some small vector from a canceling smear of negative charge.) A dipole is characterized by its dipole moment p , given in the idealized case of point charges by

$$p = Qa \quad (6.1.12)$$

The length scale a of a dipole is molecularly small, of order a few Angstroms. The energy of the equal and opposite charges of a dipole in an electrostatic potential would be zero if the potential were uniform, and thus depends on the gradient of the potential. To see this, write

$$U = Q\phi(r + a/2) - Q\phi(r - a/2) \approx Qa \cdot \nabla\phi(r) = -p \cdot E \quad (6.1.13)$$

in which we have expanded $\phi(r \pm a/2)$ to first order about the center position r (or equivalently, recognized the right-hand side above as a finite-difference approximation to the derivative of ϕ in the direction a).

Likewise, the electrostatic potential of a dipole can be obtained as the derivative of the point-charge potential, this time with respect to the location of charge. We may think of the dipole potential as nearly canceling contributions from the nearby equal and opposite charges of the dipole:

$$\begin{aligned} \phi(r; p = Qa) &= -a \cdot \nabla \frac{Q}{r} \\ &= -Qa \cdot \nabla \left(\frac{1}{(r^2)^{1/2}} \right) \\ &= -Qa \cdot \left((-1/2) \frac{2r}{(r^2)^{3/2}} \right) \\ &= (p \cdot r)/r^3 \end{aligned} \quad (6.1.14)$$

(In the above, we have indicated how to conveniently take gradients of functions of r , which is to write them as functions of r^2 , use the chain rule, and note that $\nabla_i r^2 = 2r_i$.)

The interaction energy between a charge Q and a dipole p is simply $Q\phi(r; p)$, which evidently falls off as $1/r^2$ rather than $1/r$, on account of the one derivative having been taken, or equivalently on the near-cancellation of the interaction energy of the two ends of the dipole with a distant point charge.

The interaction energy between two dipoles follows from same trick of taking derivatives. This time we start with the dipole potential $\phi(r; p)$; the two ends of the second dipole are located at $r \pm a'/2$, which leads to a derivative with respect to the point at which the potential is measured:

$$\begin{aligned} U(r; p; p' = Qa') &= Qa' \cdot \nabla\phi(r; p) \\ &= p'_i p_j \nabla_i \left(\frac{r_j}{r^3} \right) \\ &= p'_i p_j \left(\frac{\delta_{ij}}{r^3} - \frac{3r_i r_j}{r^5} \right) \\ &= \frac{p \cdot p' - 3(p \cdot n)(p' \cdot n)}{r^3} \end{aligned} \quad (6.1.15)$$

(In the above, n is the unit vector in the direction of the vector r .) The resulting interaction falls off yet faster than the charge-dipole interaction, this time as $1/r^3$, again because of the additional

gradient taken, or equivalently the additional near-cancellation of the interaction of the two ends of the second dipole with the potential from the first dipole.

In the above, we once again indicate how to manage the bookkeeping in taking gradients of functions of the scalar distance r and the vector r_i . Here, we find it useful to use indices on the various vectors to keep track of which is dotted into which. We write $\nabla_i r_j = \delta_{ij}$, where δ_{ij} is the Kronecker delta symbol, equal to unity if $i = j$ and zero otherwise. Also, we adopt the summation convention that repeated indices are summed over (i.e., the corresponding vectors are dotted together).

Note the dependence on the orientation of the dipoles: to minimize the energy, p and p' are both parallel to the vector between the two. If the dipoles are constrained to be perpendicular to their separation (as for vertically oriented molecules at an interface), the energy is minimized if the dipoles point antiparallel. If they are constrained instead to be parallel (for example, because they are surfactants, obliged to be oriented all the same way with respect to the interface), the resulting repulsive interaction builds up over long distances in such a monolayer, and leads ultimately to the presence of stripes or other domains of finite width.

6.2 van der Waals forces

In liquids of low dielectric constant, free charges are typically not thermodynamically stable, because the binding energy of charges of opposite sign far exceeds kT , so that translational entropy is never enough to stabilize unbound charges. Thus in such liquids, forces of electrostatic origin are limited to those involving dipoles. In this lecture, we consider the consequences of dipole-dipole interactions, particularly with regard to the cumulative effects on mesoscopic objects (particles or interfaces) of many such pairs of interacting dipoles.

Consider a liquid in which molecules have permanent dipoles. If interactions between these dipoles are weak compared to kT , and the dipoles are free to rotate (as they are in a liquid), then we expect for the most part the dipoles in such a liquid will be randomly oriented. As we can see from the orientational average of the dipole-dipole interaction, the average interaction energy of a pair of randomly oriented dipoles is zero.

However, if we consider the field of a given dipole, oriented in a particular direction at a given moment, it will tend to bias the orientations of neighboring dipoles, leading to a small net polarization of those nearby dipoles. Then, the given dipole will have a nonvanishing attractive interaction with the slightly oriented neighboring dipoles. This interaction, which we shall describe in more detail below, is called a Keesom interaction.

In some liquids, the molecules have no permanent dipoles. However, all molecules have quantum-mechanical dipole fluctuations, which arise (crudely speaking) from the different positions in which the electrons may be found within the electron “clouds” (wavefunctions) surrounding the nuclei in the molecule. These dipole fluctuations do not depend on thermal excitations; nonetheless, a mechanism very like the Keesom interaction results, in which a momentary dipole fluctuation on a given molecule biases the fluctuations on a nearby molecule, inducing a weak net oriented dipole on the neighbor, with which it then has an attractive interaction, called a London or dispersive interaction.

Together, Keesom and London interactions, which both vary as $1/r^6$ (except at sufficiently large distances that so-called “retardation” effects are important), constitute the ubiquitous van der Waals interaction — ubiquitous, because all materials are polarizable, and so all have some van der Waals attractions. Because these interactions are always attractive, they accumulate, such that even weak interactions on a per-molecule basis can lead to significant forces between macroscopic objects (particles, surfaces).

To give an account of the microscopic origin of van der Waals interactions, we need a description of polarizability, of both classical and quantum-mechanical origin. The latter is for the most part beyond the scope of this course, but we present in the next section a crude argument that is sufficient at least to give the idea, and a rough order of magnitude of the effect.

Polarizability

When an electric field E is imposed on an atom or molecule, it develops an induced dipole moment p in response. We write

$$p = \alpha E \tag{6.2.1}$$

in which α is the molecular polarizability. If the molecule has no permanent dipole, the applied electric field induces a dipole by distorting the electron clouds in the molecule from their unperturbed arrangements.

Here we give a very crude picture of this process, for a single atom with one valence electron. Because of quantum mechanics, this electron may be thought of as a smear of charge centered on

the nucleus, with some characteristic radius a_0 . The nucleus (screened by the more tightly bound core electrons) has an effective charge e for the present case of a single valence electron.

If we displace the nucleus with respect to the center of the cloud, there must be an increase in energy, and hence a restoring force, originating in the electrostatic attraction of the electron cloud and the nucleus. By “enlightened dimensional analysis”, this must take the form

$$U(r) \sim \frac{e^2 r^2}{2a_0^3} \quad (6.2.2)$$

We can argue for this form in several ways: the attractive energy is electrostatic in origin (hence proportional to e^2) and harmonic (hence proportional to r^2); the only remaining length in the problem is a_0 , which must appear in the denominator as a_0^3 to give an energy. (Or, we consider the change in electrostatic energy from rigidly displacing the cloud by a small distance, assuming some convenient form for the charge distribution, and using Gauss’ law to find the force on the nucleus.)

Now we apply an electric field E to the atom, which adds a term to the energy:

$$U = \frac{e^2 r^2}{2a_0^3} - er \cdot E \quad (6.2.3)$$

Minimizing with respect to r , we have $r = a_0^3/eE$, whereupon the induced dipole $p = er$ is $p = a_0^3 E$. Thus we find the polarizability α scale according to

$$\alpha \sim a_0^3 \quad (6.2.4)$$

which in retrospect, we could have identified by dimensional analysis, since E goes as charge over length squared, and p goes as charge times length.

Note that the induced displacement r is typically a very small fraction of a_0 , since $r/a_0 = E/(e/a_0^2)$ is the ratio of the applied field to the internal field of the atom. The electrostatic energy in an atom, scaling as e^2/a_0^2 is of order electron volts, which means the internal electric fields are of order volts per Angstrom (a_0 being of order a fraction of an Angstrom); whereas, typical external applied fields may be of order volts per centimeter. Even fields from permanent dipoles on nearby molecules will be smaller than internal molecular fields, roughly by the ratio of $(a_0/r)^3$, where r is the distance to the neighboring molecule. So the perturbation of the atom by the inducing field is always small.

In the case where a molecule has a permanent dipole, randomly oriented by thermal fluctuations, an applied electric field will bias the random orientations, leading to a net average induced dipole. We use equilibrium statistical mechanics to calculate the induced dipole:

$$\langle p \cdot z \rangle = \frac{\langle p \cos \theta e^{\beta p E \cos \theta} \rangle}{\langle e^{\beta p E \cos \theta} \rangle} \approx \beta p^2 E \langle \cos^2 \theta \rangle = (1/3) \beta p^2 E \quad (6.2.5)$$

in which we have approximated $pE \ll kT$, which for any externally applied fields is always well satisfied. The polarizability is thus

$$\alpha = (1/3) \beta p^2 \quad (6.2.6)$$

Charge - induced dipole interaction

Now we revisit the calculation of electrostatic interactions, to consider the case of charges or dipoles interacting with induced dipoles. Suppose we have a charge at the origin, and a polarizable molecule at r . Field from the charge at the site of the polarizable molecule is

$$E(r) = \frac{Qn}{r^2} \quad (6.2.7)$$

which induces a dipole of strength $p = \alpha E$.

One might think then that the interaction between the charge and the polarizable molecule would be simply

$$U = -p \cdot E = -\frac{\alpha Q^2}{r^4} \quad (6.2.8)$$

However, this is not quite right.

One way to see that there is a problem, is to compare the work done in bringing in the polarizable molecule from infinity to r , to the work done in bringing a permanent dipole in from infinity to r . Far away, the induced dipole is vanishingly small; it grows larger as the polarizable molecule approaches. In contrast, of course the permanent dipole has the same magnitude over the entire process. So clearly, the work done in the two cases is different. In fact, carrying out this work integral is one way to find the right answer, which turns out to be that the interaction energy is one-half of the permanent dipole result.

A more direct way to see this is to construct a simple model for polarizability, as follows. The polarizable molecule has an internal energy U_{int} , which we write as

$$U_{int} = (1/2)Ka^2 \quad (6.2.9)$$

for some “spring constant” K , which imposes an energetic cost for polarizing the molecule. (This is essentially what we did to give a crude estimate for the quantum-mechanical polarizability of a single atom.) The polarization is then determined by minimizing the sum of U_{int} and the energy of the induced dipole in an applied field. This leads to

$$0 = \partial_a (U_{int} - Qa \cdot E) = Ka - QE \quad (6.2.10)$$

This result allows us to identify the spring constant K as satisfying $Q^2/K = \alpha$, or $K = Q^2/\alpha$.

The corresponding total energy of the induced dipole in the field is

$$U_{int} - p \cdot E = -(1/2)p_{ind} \cdot E = -(1/2)\alpha E^2 \quad (6.2.11)$$

The internal “formation energy” of the dipole turns out to be one half of the energy of the dipole in the field. This is simply the result of assuming that the energy is quadratic in the dipole strength, which itself just corresponds to a Taylor expansion of the energy to lowest order in p . (The energy, being a scalar with no preferred direction for p , must start at order p^2 .)

Keesom and London interactions

We find an analogous result for the energy of a dipole interacting with a polarizable molecule. The dipole potential was

$$\phi(r) = \frac{p \cdot r}{r^3} \quad (6.2.12)$$

We take the gradient to find the electric field,

$$\begin{aligned} E_i &= -\nabla_i \phi = -p_j \nabla_i \left(\frac{r_j}{r^3} \right) \\ &= -p_j \left(\frac{\delta_{ij}}{r^3} - \frac{3r_i r_j}{r^5} \right) \\ &= \frac{3(p \cdot n)n_i - p_i}{r^3} \end{aligned} \quad (6.2.13)$$

Now the same argument as for the charge - induced dipole interaction leads to the conclusion that the interaction energy between the dipole and induced dipole is

$$\begin{aligned}
 U &= -(1/2)p_{ind} \cdot E = -(1/2)\alpha E^2 \\
 &= -(1/2)\alpha [3(p \cdot n)n - p]^2 / r^6 \\
 &= -(1/2)\alpha \frac{3(p \cdot n)^2 + p^2}{r^6}
 \end{aligned} \tag{6.2.14}$$

If the dipole itself is randomly oriented (because it is thermally randomized), then we average the above result over orientations of p , writing

$$\begin{aligned}
 \langle (p \cdot n)^2 \rangle &= n_i n_j \langle p_i p_j \rangle \\
 &= n_i n_j (p^2 \delta_{ij}) / 3 = p^2 / 3
 \end{aligned} \tag{6.2.15}$$

Whereupon, we have

$$\langle U \rangle = -\alpha p^2 / r^6 \tag{6.2.16}$$

In the above, we argue that the average of $p_i p_j$ must be proportional to δ_{ij} , say $\langle p_i p_j \rangle = C \delta_{ij}$, because there is no distinguished vector we can use to “build” the answer. We find the proportionality constant by summing with $i = j$ (“taking the trace” or “contracting the indices”), to find that $p^2 = 3C$ or $C = p^2/3$.

If the polarizability arises from a thermally randomized permanent dipole of strength p' , such that $\alpha = \beta p'^2/3$, we have altogether

$$U = -\frac{\beta p'^2 p^2}{3r^6} \tag{6.2.17}$$

(in which we have silently dropped the angle brackets around the U). This interaction between two thermally randomized permanent dipoles, in which we may regard one dipole as polarized by the instantaneous fluctuating field of the other, is called a Keesom interaction.

The Keesom interaction, which relies on “classical” (non quantum mechanical) fluctuations of permanent dipoles, is strongly analogous to the dispersive or London interaction, which arises as a result of quantum-mechanical fluctuations of polarizable atoms or molecules. We may make a very crude argument as to the magnitude of this interaction, as follows. The electron in its orbit, treated within a Bohr model of the atom, results in a fluctuating dipole, of magnitude $a_0 e$, where a_0 is the Bohr radius.

This polarizes a second atom, with polarizability α scaling as a_0^3 . The interaction between the first dipole and the second polarizable atom scales as

$$U \sim -\alpha p^2 / r^6 \sim -a_0^6 (e^2 / a_0) / r^6 \sim -\alpha^2 \epsilon_0 / r^6 \tag{6.2.18}$$

in which ϵ_0 is of the order of the Bohr energy, which is the characteristic electrostatic energy of an electron attracted to the net charge on the nucleus, at a distance of order the Bohr radius.

van der Waals interactions in different geometries

The van der Waals interaction $u(r)$ between two molecules is sum of the Keesom and London interactions, both of which fall off as $1/r^6$ (neglecting retardation effects):

$$u(r) = -C/r^6 \tag{6.2.19}$$

Because the van der Waals interaction is always attractive between molecules, the effects can accumulate into a sizeable force between mesoscopic objects, such as particles or substrates, even

though the per-molecule interactions are weak. In the present section, we work out the dependence on separation distance for the van der Waals interaction between mesoscopic objects in various geometries.

For this purpose, it is useful to define the Hamaker constant, in terms of vdW pair potential, as

$$A \equiv \pi^2 C c_1 c_2 \quad (6.2.20)$$

in which c_1 and c_2 are the number of molecules per unit volume in material elements 1 and 2, and C is the coefficient in the corresponding molecular van der Waals interaction. Typical values of A are on the order of 10^{-19} – 10^{-20} J for material interacting across vacuum or gas.

We find the interaction energy between two macroscopic objects interacting by van der Waals forces, by summing the interactions of each volume element of the first object with each volume element of the second object. The simplest case is the interaction of two distant spherical particles, i.e., separated by a distance r large compared to their radii. Then all the volume elements in the first sphere are at essentially the same distance from all volume elements in the second sphere, whereupon we have

$$U = \frac{V_1 V_2 A}{\pi^2 r^6} \quad (6.2.21)$$

Next, we consider the interaction of a distant sphere and a half-space. Again, “distant” means the sphere radius is small compared to the closest approach distance r , so that the integration over volume elements in the sphere is trivial, giving a factor V_1 . We have:

$$\begin{aligned} U &= \frac{V_1 A}{\pi^2} 2\pi \int_0^\infty dz \int_0^\infty \rho d\rho ((r+z)^2 + \rho^2)^{-3} \\ &= \frac{V_1 A}{\pi} \int_r^\infty dz' \int_0^\infty dy (z'^2 + y)^{-3} \\ &= \frac{V_1 A}{2\pi} \int_r^\infty dz' z'^{-4} \\ &= \frac{V_1 A}{6\pi r^3} \end{aligned} \quad (6.2.22)$$

Note that the contributions to the integral are dominated by the region of the half-space that is within a factor of two or so of the distance of closest approach. The van der Waals interaction is long range (falls off like a power law), but not so long range as to be dominated somehow by distant points.

Finally, we consider the interaction between two parallel half-spaces across a gap of width h . Here, we compute the interaction energy per unit area, by integrating on one side over a thin “column” of material of cross sectional area dA normal to the interface. Evidently, we have translational invariance in the plane, so each such column has the same interaction energy with the opposite half-space.

$$\begin{aligned} U/\text{area} &= \frac{A}{\pi^2} 2\pi \int_0^\infty dz_2 \int_0^\infty dz_1 \int_0^\infty \rho d\rho ((z_1 + z_2 + h)^2 + \rho^2)^{-3} \\ &= \frac{A}{2\pi} \int_0^\infty dz_2 \int_{h+z_2}^\infty dz'_1 z_1'^{-4} \\ &= \frac{A}{6\pi} \int_0^\infty dz_2 (z_2 + h)^{-3} \\ &= \frac{A}{12\pi h^2} \end{aligned} \quad (6.2.23)$$

Again, this integral is dominated by the bits of the half-spaces that are within a factor of two or so of the distance of closest approach h .

Computing Hamaker constants for real materials

The arguments of the previous section, although sufficient to explain the origin and rough magnitudes of the Keesom and London interactions, are not useful for calculating their precise magnitudes. For this, more detailed theory is required, beyond the scope of this book. However, a very useful approximate formula for computing Hamaker constants is given by Israelachvili [JI 11.4] on the basis of the full Lifshitz theory:

$$A = \frac{3kT}{4} \left(\frac{\epsilon_1 - \epsilon_2}{\epsilon_1 + \epsilon_2} \right)^2 + \frac{3h\nu_e}{16\sqrt{2}} \frac{(n_1^2 - n_2^2)^2}{(n_1^2 + n_2^2)^{3/2}} \quad (6.2.24)$$

for material 1 interacting across a region filled with material 2. To evaluate A , values of the dielectric constant ϵ and index of refraction n are needed for materials 1 and 2. In addition, the parameter ν_e is a typical optical excitation frequency for material 1; fortunately, for a broad class of materials these fall into a relatively narrow range, $(3 - 5) \times 10^{15} \text{sec}^{-1}$, so that a typical value may be used in absence of particular values.

Sometimes, values of dielectric constant or refractive index are available at one temperature but not the desired temperature, or for a related chemical substance but not the desired substance. Because the physical origin of both dielectric constant and refractive index in nonpolar materials is molecular polarizability, which does not depend on temperature, the temperature-dependence of these properties arises essentially just from density variations. This fact can also be used to adapt ϵ or n values for homologous molecular liquids (for example, normal alkanes of various length), which arguably have similar molecular polarizabilities, but different densities. Likewise, one can approximately relate ϵ and n for liquid and solid nonpolar materials, in which the chief difference between the two is the density change on freezing.

Using the fact that molecular polarizability is approximately constant, while molecular concentrations may vary, one can derive the Clausius-Mossotti relation

$$\frac{\epsilon - 1}{\epsilon + 2} \propto c \quad (6.2.25)$$

and the analogous Lorenz-Lorentz equation

$$\frac{n^2 - 1}{n^2 + 2} \propto c \quad (6.2.26)$$

in which c is the concentration of molecular species. These relations can be used as described above to infer values of ϵ or n at different molecular concentrations (hence, at different temperatures, or in liquids versus solids) if values at a given concentration are already known.

These relations may only be used with caution if the dielectric response contains contributions from fluctuating permanent dipoles, since temperature changes can affect the polarizability of such materials (drastically, if one considers transforming such a liquid into a solid or glass).

Interfacial tension

There is a direct relation between the attractive interactions of two half-spaces of material 1 across a gap filled with a second material 2, and the interfacial tension γ_{12} associated with the interface between materials 1 and 2. Recalling our result for the van der Waals interaction between two

parallel half-spaces, we see that if we close the gap from infinity down to zero, the attractive interaction diverges. We need a short-distance cutoff, presumably of atomic dimensions, to forestall an infinite result.

As the two surfaces come into contact from a large distance, we have removed two 1–2 interfaces from the system. It makes sense to identify the interfacial tension with the work done in separating the two half-spaces to infinity from intimate contact at the cutoff distance. Thus we may write

$$2\gamma_{12} = \frac{A}{12\pi h_0^2} \quad (6.2.27)$$

in which h_0 is the short-distance cutoff, and $A = A_{12}$ the Hamaker constant appropriate for material 1 to interact with itself across an expanse of material 2.

One might expect that the best value of cutoff distance, for which experimental surface tension values agrees with the above equation, would depend on the materials involved. However, it turns out that a single value of h_0 , of order 1.6\AA , gives a remarkably good estimate of the interfacial tension of a wide variety of nonpolar liquids. Presumably, this happy coincidence obtains because the hard-core dimensions in the “thinnest direction” of many common molecules are quite similar.

Jacob Israelachvili’s book, from which this discussion is adapted, presents a table of estimated Hamaker constants and interfacial tensions for a broad range of liquids, for which agreement with experiment is uniformly good to within 10–20 percent. This result illustrates that the interfacial tension for common liquids and nonpolar solids, for which the cohesive energy arises from attractive van der Waals interactions, is simply a result of the lost attractive interactions when interfaces are introduced.

Derjaguin approximation and surface forces apparatus

Consider a sphere of radius R closely approaching the flat surface bounding a half-space of material. We wish to compute the interaction between the sphere and the halfspace. Suppose we know the interaction locally between different portions of the surface of the sphere and the halfspace. This would be true for dispersive interactions, for example, because we would argue that different portions of the “gap region” look very much like portions of two half-spaces at the local separation. This is a reasonable approximation because as we recall, the interactions in such a case are dominated by the bits of material within a finite factor of the distance of closest approach. There are other circumstances in which the interaction between the two surfaces may also be thought of as locally a function of the gap distance h — for example, if brushes were grafted on the surfaces, and compressed when the sphere and plane were brought together.

In such a case, it make sense to write the interaction energy as an integral of the local interaction energy per unit area, over the area of the contact:

$$W = 2\pi \int_0^\infty U(h(\rho))\rho d\rho \quad (6.2.28)$$

To proceed further, we need to describe how the gap $h(\rho)$ depends on the in-plane radial distance ρ . This we obtain from writing

$$\rho = R \sin \theta \quad (6.2.29)$$

and

$$\begin{aligned} z &= h + R(1 - \cos \theta) \\ &= h + R(1 - \sqrt{1 - (\rho/R)^2}) \\ &\approx h + R(1 - (1 - (\rho/R)^2/2)) \\ &= h + \rho^2/(2R) \end{aligned} \quad (6.2.30)$$

Using this expansion (a parabolic approximation to the shape of the sphere near the contact point), we have

$$\begin{aligned} W &\approx 2\pi \int_0^\infty U(h + \rho^2/(2R))\rho d\rho \\ &= 2\pi R \int_h^\infty U(z)dz \end{aligned} \quad (6.2.31)$$

Remarkably, this implies that the force on the sphere, $F = -\partial W/\partial h$, is directly proportional to the interaction potential between the sphere and the surface:

$$F = -\partial W/\partial h = 2\pi R U(h) \quad (6.2.32)$$

The above discussion plays a key role in the design of a device for making macroscopic measurements of forces between surfaces, known as the surface forces apparatus (SFA). In the SFA, forces are measured between the surfaces of two glass cylinders crossed at a 90 degree angle, to which atomically smooth mica sheets have been glued. (Mica is used because it cleaves into atomically smooth planes, eliminating the effect of roughness in the gap between the cylinders. Crossed cylinders are used rather than e.g. two spheres or a sphere and a plate, because thin mica sheets can be bent in one direction to wrap a cylinder but not a sphere.)

The force is measured by means of a very sensitive mechanical spring, and the separation between the cylinders by means of laser interferometry (observation of the circular interference fringes resulting from the narrow gap between the cylinders at close approach). The relative position of the cylinders is controlled (up to the deflection of the spring) by a combination of fine screws and/or piezoelectric transducers.

It turns out that the same geometry results from for two crossed cylinders at 90 degrees in close approach, as for the case of a sphere in close approach to a plane treated above in our discussion of the Derjaguin approximation. Thus, the force measured in the SFA is directly proportional to the interaction potential per unit area.

6.3 Charged surfaces in solution

In this lecture, we consider the behavior of charged particles or interfaces aqueous systems, in which we expect the counterions to the surface charges to be at least partly solubilized and mobile. In addition to the counterions, aqueous solutions may contain added electrolyte (dissolved salt). We are interested in the electrostatic interactions between such charged particles, which we expect to be repulsive for like-charged particles, resulting in a means by which colloidal suspensions can be stabilized against flocculation due to van der Waals attractions.

To find the interactions between the particles, we need to know not only the areal density of the charges bound to the surface, but also the density of mobile ions (counterions, in the case with no added salt). This leads to a complicated problem, because the mobile ions have two competing tendencies: energetically, they tend to be where the electrostatic potential is favorable; while entropically, they tend to be everywhere, to maximize their translational entropy. Because the electrostatic potential depends on where the charges are, we have a self-consistent problem: to know where the charges are, we need to know what the potential is, and vice versa.

To solve this problem, we must compute at the same time the potential and the ion concentrations. The governing equation is the Poisson-Boltzmann (PB) equation, which combines the Poisson equation that tells how the potential arises from the charge distribution with the Boltzmann factor that tells how likely it is to find charges in a given value of the potential. The PB equation is nonlinear, because the Boltzmann factor depends nonlinearly on the potential. As a result, it can be solved analytically only in special cases (one-dimensional planar geometries).

If we assume the electrostatic potential energy differences are weak compared to kT , we can linearize the PB equation, resulting in the more tractable Debye-Huckel equation. The DH equation can be very effective in describing even highly charged surfaces (which would naively give rise to large surface fields and ultimately large potential energy differences) if we take account of how counterions reaccumulate or “condense” near such surfaces, resulting in a much lower renormalized value of the surface charge density. In fact, we may state a rule of thumb, that the renormalized surface charge in such cases is reduced until the electrostatic energy for additional charges to condense at the surface is of order kT , and thus comparable to the entropy gain of unbound counterions.

Poisson-Boltzmann equation

Combine two eminently sensible (approximate) statements: 1) the concentration of charged species is proportional to the Boltzmann factor for that species in the electrostatic potential:

$$\rho_{\pm} \propto e^{\mp\beta e\phi} \quad (6.3.1)$$

2) the electrostatic potential is given by the Poisson equation, i.e., is generated by the distribution of charged species:

$$\nabla^2\phi = -4\pi e(\rho_+ - \rho_-)/\epsilon \quad (6.3.2)$$

In the above, we write ρ_{\pm} for the concentration (rather than the charge density) of charged species, presumed here to be of charge $\pm e$.

Combining these two equations gives the Poisson-Boltzmann equation,

$$\nabla^2\phi = -\frac{4\pi e}{\epsilon} \left[\rho_{0+} e^{-\beta e\phi} - \rho_{0-} e^{\beta e\phi} \right] \quad (6.3.3)$$

Here $\rho_{0\pm}$ are “reference values” of the concentrations of positive and negative ions present at zero potential. (If instead the total number of ions is known, $\rho_{0\pm}$ must be determined by normalizing the integral of the respective ion concentrations.)

The Poisson-Boltzmann equation is so sensible, we should point out what it neglects; namely, fluctuations in the charge concentration. The Poisson equation is exact; but the “Boltzmann” half of the story, $\rho \propto \exp -\beta e\phi$, is only true in an average sense,

$$\langle \rho(r) \rangle \propto \langle \exp^{-\beta e\phi(r)} \rangle \quad (6.3.4)$$

in which the right-hand average is over all possible configurations of the charges, with $\phi(r)$ the corresponding exact potential for each configuration. That is, the potential $\phi(r)$ fluctuates because the charge density fluctuates, and this is neglected in the PB equation.

If the potential is weak, such that we may expand the exponentials to lowest order, we arrive at the Debye-Huckel approximation,

$$\begin{aligned} \nabla^2 \phi &\approx \kappa^2 \phi \\ \kappa^2 &= \frac{4\pi\beta e}{\epsilon} (\rho_{0+} + \rho_{0-}) \end{aligned} \quad (6.3.5)$$

Here κ^{-1} is the Debye screening length, which evidently determines the length scale over which the potential ϕ will vary in response to imposed boundary values.

In the presence of monovalent salt, the reference concentrations both become $c/2$, whereupon the Debye length satisfies

$$\begin{aligned} \kappa^2 &= 4\pi l c \\ l &= \beta e^2 / \epsilon \end{aligned} \quad (6.3.6)$$

Here l is the Bjerrum length, the separation at which the interaction energy of two point charges in a medium of dielectric constant ϵ is kT . For water, with $\epsilon = 80$, l is about 0.7nm. Evidently the Debye screening length depends on salt concentration; note that both cations and anions contribute equally to the screening length. In a 1mM solution of monovalent salt (NaCl, for example), the Debye length is about 10nm.

We may ask, how many ions are there in a cube of linear dimension equal to the screening length? For a mean-field approximation such as Poisson-Boltzmann (or its linearization, DH) to work, we need to have many ions present per screening volume. Define $c = 1/d^3$, so that the volume per ion is d^3 ; then, using the definition of the Debye length (writing $l_D = 1/\kappa$), we have

$$l_D^3 / d^3 \sim (d/l)^{3/2} \quad (6.3.7)$$

The left hand side above is the ratio of the screening volume to the volume per ion, or the number of ions per screening volume; we see this is large as long as the mean distance between ions d is large compared to the Bjerrum length. This places an upper limit on the ion concentrations for which the mean-field theory is valid, such that the interaction energy between typical nearest-neighbor ions is weak compared to kT .

Screened point charge

Consider a point charge at the origin in a medium described by the Debye-Huckel equation. It turns out that the solution of the DH equation for a point charge is given by

$$\phi(r) = \frac{e}{\epsilon r} e^{-\kappa r} \quad (6.3.8)$$

That this is a solution of the DH equation can be verified by direct substitution. To see that it is the solution corresponding to a point charge at the origin, note that for distances r much smaller

than the screening length, $\phi(r)$ is approximately $e/\epsilon r$, the unscreened potential of a point charge. (This form of potential is called a Yukawa potential.)

We may generalize this solution to a charged sphere of radius b and charge Ze , as follows. We again guess that the potential must be proportional to the Yukawa potential, which is the only spherically symmetric solution to the DH equation that dies off at infinity:

$$\phi(r) = C \frac{Ze}{\epsilon r} e^{-\kappa r} \quad (6.3.9)$$

To find the proportionality constant C , we need to impose a boundary condition on the particle surface, such that the electric field there is that of an as-yet unscreened charged sphere. To do this, we apply Gauss' law on the surface of the sphere:

$$\begin{aligned} \int \nabla \cdot E dV &= \int E \cdot dS = 4\pi b^2 |E| \\ &= 4\pi/\epsilon \int \rho dV = 4\pi eZ/\epsilon \end{aligned} \quad (6.3.10)$$

which implies

$$E = \frac{eZ}{\epsilon b^2} \quad (6.3.11)$$

We can explicitly calculate the electric field from the expression for $\phi(r)$,

$$E = -\nabla\phi = -CZe \left(\frac{-\kappa n}{r} - \frac{n}{r^2} \right) e^{-\kappa r} = CZe(1 + \kappa r) \frac{e^{-\kappa r}}{r^2} \quad (6.3.12)$$

Evaluating E on the sphere surface and equating to the Gauss' law result, we have

$$C = \frac{e^{\kappa b}}{1 + \kappa b} \quad (6.3.13)$$

so that the potential for the charge sphere is

$$\phi(r) = Ze \frac{e^{-\kappa(r-b)}}{\epsilon r(1 + \kappa b)} \quad (6.3.14)$$

Renormalized charge Z^*

There is a problem with the result of the previous section; if the charge Z is too large, the energy to place a counterion on the surface of the particle far exceeds kT . We can estimate this energy from the above, as

$$U = e\phi(b) \approx \frac{Ze^2}{\epsilon b^2 \kappa} \quad (6.3.15)$$

where we have assumed that the particle radius is much larger than the screening length, as is usually the case.

Rewrite the above in terms of the surface charge density σ , such that $Z = 4\pi b^2 \sigma$, as

$$U = \frac{4\pi\sigma e^2}{\epsilon \kappa} = \frac{4\pi\sigma lT}{\kappa} \quad (6.3.16)$$

Evidently, if the surface charge density exceeds one charge per $l\kappa^{-1}$ (which for our typical values of $l = 1\text{nm}$ and $\kappa^{-1} = 10\text{nm}$ is about one charge per 1000\AA^2), then the energy for a counterion to "condense" onto the surface will exceed kT .

Self-consistent estimate of Z^*

If the charge on the sphere is too large, then some of the counterions will condense into a relatively thin layer near the surface of the sphere, resulting in an effective charge Z^* . Here we make a self-consistent scaling argument as to the magnitude of Z^* . The concentration of counterions at the surface is enhanced by a factor of $e^{\beta e \phi(b)}$ relative to the background concentration c far from the sphere. The potential on the surface as a result of the concentration is $\phi(b)$ scaling as $Z^* e / (\epsilon \kappa b^2)$. Then concentration is enhanced over a layer of thickness the order of the Debye screening length, or a volume of order $b^2 \kappa^{-1}$.

The enhanced concentration is thus $c e^{\beta e \phi(b)}$, with a total number of counterions in the enhanced layer of order $c e^{\beta e \phi(b)} b^2 \kappa^{-1}$. This enhanced number is itself of order Z , anticipating that Z^* is much smaller than Z , so that the counterions in the thin layer nearly cancel the “bare” charge of the sphere. So we have

$$c e^{\beta e \phi(b)} b^2 \kappa^{-1} \sim Z \quad (6.3.17)$$

Solving for Z^* (which enters through $\phi(b)$), we have after some arithmetic

$$Z^* \sim \frac{\kappa b^2}{l} \log \left(\frac{Z \kappa}{b^2 c} \right) \quad (6.3.18)$$

Equivalently, we can write the result for the reduced surface charge density σ^* as

$$\sigma^* \sim \frac{\kappa}{l} \log \left(\frac{\sigma \kappa}{c} \right) \quad (6.3.19)$$

So the reduced surface charge density σ^* depends only weakly on the bare value σ . Typical values will tend to be in the neighborhood of κ/l , which for 1mM salt solution in water we estimated as about one charge per 1000 \AA^2 . For a particle of 0.1μ diameter, this gives a reduced charge Z^* of about 3000.

6.4 Poisson-Boltzmann equation

In this lecture, we focus on the full nonlinear Poisson-Boltzmann equation in simple geometries in which it can be solved analytically, despite its nonlinearity. By this method, we can study counterion condensation in a strong potential without invoking the self-consistent scaling argument of the previous section.

We begin with a single charged planar surface with no added salt. The question here to be answered is how far the counterions manage to escape from the charged surface to which they are attracted. The corresponding solution of PB equation gives rise to a new length scale, the Gouy-Chapman length λ , that gives the characteristic thickness of the so-called electric “double layer” of surface charge and counterion cloud. The Gouy-Chapman length can be regarded as a self-consistently determined value of the Debye screening length κ^{-1} , determined such that 1) the concentration of counterions within the double layer gives a screening length κ^{-1} of order λ , and 2) all the counterions necessary for electrical neutrality are contained within the double layer.

Next, we present the PB solution for two parallel charged surfaces separated by a thickness h of water without added salt. In this case, the important question in addition to the distribution of counterions is the magnitude of the repulsive force per unit area between the two surfaces. It turns out that this force per unit area, called the disjoining pressure, may be regarded as arising entirely from the osmotic pressure of the counterions within the gap. Thus, having found the counterion distribution for this problem, we can determine the disjoining pressure between charged surfaces.

Single surface

The first of the PB geometries we treat is a single planar surface with a uniform density σ of negative charges, bounding a semi-infinite region of space in which the only ions are the positive counterions for the surface. (That is, there is no added salt.) In this case the PB equation reduces to

$$\nabla^2\phi = -4\pi e\rho_0\epsilon e^{-\beta e\phi} \quad (6.4.1)$$

in which we have dropped the + subscript on the counterions, since we only have one species here.

To neaten the notation, we define

$$\begin{aligned} \psi &= \beta e\phi \\ \lambda^{-2} &= 8\pi\beta e^2\rho_0/\epsilon = 8\pi l\rho_0 \end{aligned} \quad (6.4.2)$$

We choose the zero of potential so that $\phi = 0$ on the surface ($z = 0$), which implies that ρ_0 is the counterion concentration at the surface (since the right-hand side of the PB equation is just $-4\pi\rho$). Later, we shall relate ρ_0 to the surface charge density, by imposing overall charge neutrality of the system.

In terms of these definitions, the PB equation takes the form

$$\psi'' = -(1/2)\lambda^{-2}e^{-\psi} \quad (6.4.3)$$

To solve this equation, we first multiply by $2\psi'$ on both sides, which leads to both left and right sides integrable (in mechanics, this would be called an “energy integral”):

$$\partial_z(\psi'^2) = \lambda^{-2}\partial_z e^{-\psi} \quad (6.4.4)$$

We integrate from infinity to some z value; we expect ψ' must vanish at infinity (which implies for consistency that ψ must become infinitely large at infinity). This leads to

$$\psi'^2 = \lambda^{-2}e^{-\psi} \quad (6.4.5)$$

Take the square root, and we have a separable equation, which we rearrange as

$$e^{\psi/2} d\psi = \lambda^{-1} dz \quad (6.4.6)$$

This time, integrate from $z = 0$ to some z , and use the fact that $\psi(0) = 0$, to obtain

$$2(e^{\psi/2} - 1) = \lambda^{-1} z \quad (6.4.7)$$

which we rearrange to obtain finally

$$\psi = 2 \log(1 + z/2\lambda) \quad (6.4.8)$$

The potential indeed goes to infinity, albeit only logarithmically, as z becomes large, consistent with our assumption above. The electric field is of course the derivative of the potential, and vanishes as $1/z$ for large z , which gives a small but nonzero force to distant charges that stray far from the surface, directing them back towards the double layer. Thus we may say that a slowly diverging potential is in fact necessary to keep the ions more or less confined.

The corresponding counterion density is

$$\rho = \rho_0 e^{-\psi} = \rho_0 / (1 + z/2\lambda)^2 \quad (6.4.9)$$

We integrate this over all positive z to find the total number of counterions per unit area, which must be equal to σ :

$$\sigma = \rho_0 \int_0^\infty dz (1 + z/2\lambda)^{-2} = 2\rho_0 \lambda \quad (6.4.10)$$

which implies $\rho_0 = \sigma/(2\lambda)$. Combining this with the definition of λ , we find

$$\lambda^{-1} = 4\pi\sigma l \quad (6.4.11)$$

Equivalently, we can use Gauss' law to determine the normalization of the counterion density, by constructing a ‘‘Gaussian pillbox’’ surface with faces parallel to the surface. This leads to

$$\begin{aligned} E &= -\phi'(0) = 1/(\beta e \lambda) \\ &= 4\pi e \sigma / \epsilon \end{aligned} \quad (6.4.12)$$

where the last equality follows from application of Gauss' law to the ‘‘pillbox’’ surface. Combining the equalities above leads to the same expression for λ as we found by charge neutrality.

Gouy-Chapman length

The length scale λ is called the Gouy-Chapman length; we note that the counterion density is roughly speaking confined to a layer of order λ thick near the interface. The corresponding potential energy for a charge e near the interface goes as

$$e\phi = T\psi = 2T \log(1 + z/2\lambda) \quad (6.4.13)$$

We see that the potential energy of a counterion varies by about kT in a distance of order λ .

The length scale λ may be regarded as the ‘‘self consistent length scale for counterion confinement’’, or equivalently the ‘‘self-consistently determined screening length’’. We can (after the fact) make a scaling argument determining λ as follows: the formula for the Debye screening length can be written (with λ as the screening length) as

$$\lambda^{-2} = 4\pi\rho_0 l / e \quad (6.4.14)$$

in which the counterions, at concentration ρ_0 near the surface, are providing the screening. But we assume as well that most of the counterions are found within a screening length of the surface, so that the counterion concentration in this region consistent with overall neutrality is

$$\rho_0 = e\sigma/\lambda \quad (6.4.15)$$

Combining these two expressions leads to the Gouy-Chapman length.

PB equation for two charged surfaces

Now we modify the problem, so that we have two identical, parallel charged surfaces facing each other across a finite expanse of water, of thickness $2h$. The PB equation is the same, but the boundary conditions are different.

We pick up the calculation at the energy integral,

$$\partial_z(\psi'^2) = \lambda^{-2}\partial_z e^{-\psi} \quad (6.4.16)$$

Now, however, we integrate from the midplane between the two surfaces, which we label as $z = 0$. On the midplane, the derivative ψ' must vanish by symmetry. In addition, we choose the zero of potential so that $\psi(0) = 0$ at the midplane. Integrating, we have

$$\psi'^2 = \lambda^{-2}(e^{-\psi} - 1) \quad (6.4.17)$$

Because we assume as before a negative surface charge density, we expect that the potential at the surfaces is negative, so the sign of ψ for positive z is negative. Thus we take the negative square root, and rearrange to obtain

$$\frac{e^{\psi/2} d\psi}{\sqrt{1 - e^\psi}} = -\frac{dz}{\lambda} \quad (6.4.18)$$

This we may integrate, again from $z = 0$, to obtain

$$\begin{aligned} z/\lambda &= -\int_0^\psi d\psi e^{\psi/2}(1 - e^\psi)^{-1/2} \\ &= -2\int_1^{\exp(\psi/2)} dy y(1 - y^2)^{-1/2} \\ &= 2\int_0^{\cos^{-1} \exp(\psi/2)} d\theta \\ &= 2\cos^{-1} \exp(\psi/2) \end{aligned} \quad (6.4.19)$$

in which we have made the substitutions $y = \exp \psi/2$ and then $y = \cos \theta$. Finally solving for ψ , we have

$$\psi = 2 \log \cos(z/2\lambda) \quad (6.4.20)$$

The corresponding counterion density is

$$\rho = \rho_0 e^{-\psi} = \frac{\rho_0}{\cos^2(z/2\lambda)} \quad (6.4.21)$$

In the present problem of two interfaces, ρ_0 is the counterion density at the midpoint (because ψ equals 0 there); we determine its value either by enforcing normalization, or using Gauss' law at $z = \pm h$ as before. Following the latter approach, we have

$$E = -\phi'(h) = 1/(\beta e\lambda) \tan(h/2\lambda) = 4\pi e\sigma/\epsilon \quad (6.4.22)$$

Using the above equality, we find

$$\lambda^{-1} = \frac{4\pi\sigma l}{\tan(h/2\lambda)} \quad (6.4.23)$$

or equivalently

$$(h/2\lambda) \tan(h/2\lambda) = 2\pi\sigma hl \quad (6.4.24)$$

which needs to be solved numerically. Once solved, we can use the definition of λ to determine the counterion reference concentration ρ_0 .

Two limiting cases of the equation for λ are easily treated. In the limit of low charge density on the surface, such that $hl\sigma$ is small, we see above that $h/2\lambda$ must likewise be small, which allows us to expand the tan, to obtain $\lambda^{-2} \approx 8\pi\sigma l/h$. Then using the original definition of λ in terms of ρ_0 , we have

$$\rho_0 \approx \sigma/h \quad (6.4.25)$$

In this limit, ψ is nearly constant (because λ is much larger than h , the argument of the cosine varies very little across the intervening fluid), which means the counterion concentration is nearly constant, equal to ρ_0 everywhere. The counterions in this weakly charged case are essentially a uniform ideal gas between the two surfaces.

In the opposite limit of high charge density on the surface, such that σhl is very large, we must have the argument of the tan nearly equal to $\pi/2$ (so that the diverging tangent function can make the left side of the above equation as large as needed); hence $\lambda \approx h/\pi$. Then we have from the definition of λ in terms of ρ_0 , that

$$\rho_0 \approx \frac{\pi}{8lh^2} \quad (6.4.26)$$

If we compare this midplane counterion concentration to the uniformly distributed result above, we see that the concentration in the present highly charged limit is smaller by a factor of $\pi/(8lh\sigma)$, which is small (as we have assumed $lh\sigma$ is large in this limit). Hence the midplane counterion concentration is much reduced compared to uniform concentration, because most of the counterions stay close to the interfaces. We could interpret this result in terms of an effective surface charge density σ^* , chosen so that the uniformly distributed counterion concentration equalled the actual midplane counterion concentration. This prescription implies

$$\sigma^* = \frac{\pi}{8lh} \quad (6.4.27)$$

Note that the effective surface charge density by this argument is independent of the actual charge density (which, for this argument to work, must be sufficiently high that $lh\sigma$ is large).

We can check the limit of a single interface, by taking the limit of large h . This is rather tricky; it turns out we need to take $h/2\lambda$ to be $\pi/4$ plus a multiple of 2π , so that the tan is unity; then we can expand the function ψ around $z = \pm h$, as

$$\begin{aligned} \cos(z/2\lambda) &= \cos((\delta z - h)/2\lambda) \\ &\approx \cos(h/2\lambda) + \sin(h/2\lambda)\delta z/2\lambda + \dots \\ &\approx 1 + \delta z/2\lambda \end{aligned} \quad (6.4.28)$$

which recovers the single-interface result.

Disjoining pressure

Two opposing charged interfaces in close proximity lead to an enhanced concentration of counterions in the intervening liquid, higher than would be the case for an isolated interface. This leads to a repulsion between the interfaces. There are several equivalent ways to compute this repulsion; for example, we could evaluate the free energy of the entire configuration as a function of the distance $2h$ between the two interfaces.

However, the simplest route to computing the repulsive force per unit area is to observe that on the midplane, the electric field is zero by symmetry (since the two interfaces are identical, the field could not point uniquely towards either one). Hence, if we ask how the left side of the system can exert a force on the right side of the system, the only force exerted across the interface is due to the osmotic pressure of the counterions. The counterions are an ideal gas, so the osmotic pressure at the midplane is $\Pi = kT\rho_0$, which must then be the force per unit area exerted by one interface on the other.

Once we compute (numerically) λ above, we then have the force per unit area. In particular, for the limiting behaviors determined above for the case of highly charged surfaces,

$$\Pi = kT\rho_0 \approx \frac{\pi kT}{8lh^2} \quad (6.4.29)$$

6.5 Surface interactions with added salt

In this lecture, we extend our treatment of charged surfaces in solution to handle the practically important case of electrolytes (salt solutions). In practical situations, even very modest amounts of added salt result in the screening length being determined entirely by the salt concentration and not by the intrinsic counterions associated with charged surfaces. Furthermore, by controlling the salt concentration and thus the screening length, we can control the strength of repulsive interactions between like-charged surfaces. By this means, colloidal suspensions stabilized against flocculation by electrostatic repulsion may be destabilized, by adding sufficient salt.

We begin by solving the problem of two parallel surfaces separated by a region of electrolyte in the Debye-Huckel approximation. This solution is reasonable even for strongly charged surfaces if we take proper account of counterion condensation in renormalizing the surface charge density to a lower effective value, at which the surface potential is of order kT . From this solution, we can compute the disjoining pressure acting between the surfaces, by generalizing our earlier argument that the disjoining pressure results only from the osmotic pressure of the counterions at the midplane. Then, we treat the same problem within the full nonlinear Poisson-Boltzmann equation, which automatically takes account of counterion condensation.

Finally, we investigate the total interaction potential between slabs of material with both disjoining pressure due to the electrostatic double layer, and van der Waals attraction. This combined interaction, called the DLVO potential, allows us to describe the change in the interactions of such surfaces with added salt. At low salt, the disjoining pressure is dominant and the interaction is strongly repulsive; as salt is added, an attractive “secondary” minimum appears at intermediate separations, at which the longer range van der Waals attraction wins out over the exponentially decaying double layer repulsion. At still higher salt concentrations, the repulsive electrostatic barrier at short distances diminishes to the point that the attractive “primary” minimum, due to the singularly strong van der Waals attraction in close contact, becomes accessible.

Debye-Huckel for charged surfaces

Here we solve the DH equation,

$$\nabla^2\phi = \kappa^2\phi \quad (6.5.1)$$

for the case of parallel surfaces. For a single surface at $z = 0$, the solution is evidently of the form $\phi = Ce^{-\kappa z}$; the coefficient C is determined by the boundary condition. Application of Gauss' law gives $E = 4\pi\sigma/\epsilon$ on the surface, which leads to

$$\phi = \frac{4\pi\sigma e}{\epsilon\kappa}e^{-\kappa z} \quad (6.5.2)$$

For two surfaces separated by a distance D , we build a solution satisfying the boundary conditions on the two surfaces at $z = \pm D/2$ from exponential solutions $e^{\pm\kappa z}$. By symmetry, we must have

$$\phi = C(e^{\kappa z} + e^{-\kappa z}) = 2C \cosh \kappa z \quad (6.5.3)$$

To determine C , enforce the boundary conditions on each surface. For this purpose, compute the electric field,

$$E = -\partial_z\phi = -2\kappa C \sinh \kappa z \quad (6.5.4)$$

(In the above, we have taken σ to be the areal number density of charges.) By applying Gauss' law at each interface, we must have

$$E(z = \pm D/2) = \mp \frac{4\pi e\sigma}{\epsilon} \quad (6.5.5)$$

Applying this result, we find the value of C , and thus the potential, as

$$\phi = \frac{4\pi e\sigma}{\epsilon\kappa} \frac{\cosh \kappa z}{\sinh(\kappa D/2)} \quad (6.5.6)$$

The corresponding number density of countercharges is

$$\rho = -\frac{\epsilon}{4\pi e} \nabla^2\phi = -\sigma\kappa \frac{\cosh \kappa z}{\sinh(\kappa D/2)} \quad (6.5.7)$$

which has the expected sign opposite to that of the surface charge density. On the midplane, for surfaces far apart compared to the Debye length, we have

$$\rho(z = 0) = -2\sigma\kappa e^{-\kappa D/2} \quad (6.5.8)$$

Disjoining pressure, DH with salt

We argued previously that the disjoining pressure is the osmotic pressure of the counterions. With added salt, this statement is still true, if we generalize it to refer to the excess osmotic pressure of the counterions above that of the unperturbed, reference electrolyte solution.

Thus we require the individual concentrations of the positive and negative salt ions, whereas at present we have only the net number density of countercharges. At the lowest order in the Debye-Huckel expansion, the screening countercharge in a salt solution would be made up of equal and opposite perturbations in the number density of positive and negative ions, each symmetrically driven towards or away from a given region.

Thus to the lowest order the total concentration of counterions is unaffected by the surfaces. To obtain the separate perturbed concentrations of countercharges of each sign, and thus compute

the disjoining pressure, we must expand to the next higher order, making use of the Boltzmann factors

$$\begin{aligned}\delta\rho_{\pm} &= \rho_0(e^{\mp\beta e\phi} - 1) \\ &= \approx \rho_0(\mp\beta e\phi + (\beta e\phi)^2/2 + \dots)\end{aligned}\quad (6.5.9)$$

Thus, the sum of the perturbed concentrations is

$$\delta\rho_+ + \delta\rho_- \approx \rho_0(\beta e\phi)^2 \quad (6.5.10)$$

Using our DH solution for ϕ , we have for separations D large compared to the Debye length the final result for the disjoining pressure

$$\Pi \approx 8\pi kT\sigma^2 l e^{-\kappa D} \quad (6.5.11)$$

Grahame equation

A useful general result can be derived, relating the surface charge density to the excess concentration of countercharges relative to the midplane (or relative to infinity, if only a single charged surface is considered). We begin with the countercharge number density, written as

$$\rho = \sum_i \rho_{0i} e^{-\beta z_i e \phi} \quad (6.5.12)$$

in which z_i is the magnitude of the charge of the i th species (± 1 , for monovalent ions). Note that

Taking the derivative with respect to z of the above, we have

$$\frac{\partial\rho}{\partial z} = -\beta e \sum_i z_i \rho_{0i} e^{-\beta z_i e \phi} \frac{\partial\phi}{\partial z} \quad (6.5.13)$$

In the above, we recognize the right-hand side of the PB equation, which allows us to replace as

$$\frac{\partial\rho}{\partial z} = \frac{\beta\epsilon}{4\pi} \frac{\partial^2\phi}{\partial z^2} \frac{\partial\phi}{\partial z} = \frac{\beta\epsilon}{8\pi} \frac{\partial}{\partial z} \left(\frac{\partial\phi}{\partial z} \right)^2 \quad (6.5.14)$$

which can be integrated to give

$$\rho(z) - \rho(z_0) = \frac{\beta\epsilon}{8\pi} \left(\frac{\partial\phi}{\partial z} \right)^2 \quad (6.5.15)$$

where z_0 is the midplane, or anywhere the field vanishes.

Finally, we can evaluate the above on the surface, and replace the electric field using Gauss' law, to obtain the Grahame equation,

$$\rho_s - rh_0\sigma = 2\pi l\sigma^2 \quad (6.5.16)$$

where as before we have defined σ as the areal number density of surface charges.

Poisson-Boltzmann for a charged surface with salt

The Grahame equation provides a convenient starting point for solving the PB equation for the case of a single interface with added monovalent salt. We take the reference point z_0 to be at infinity. Writing the ion concentrations as

$$\rho_{\pm}(z) = \rho_0 e^{\mp\beta e\phi} \quad (6.5.17)$$

in which ρ_0 is the reference concentration of each species (where the potential vanishes). We have

$$\frac{\beta\epsilon}{8\pi} \left(\frac{\partial\phi}{\partial z} \right)^2 = \rho_0 (e^{\beta e\phi} - e^{-\beta e\phi} - 2) = 4\rho_0 \sinh^2(\beta e\phi/2) \quad (6.5.18)$$

Taking the square root, and writing the resulting equation in terms of $\psi = \beta e\phi$, we have

$$\frac{\partial\psi}{\partial z} = -2\kappa \sinh(\psi/2) \quad (6.5.19)$$

(We expect $\psi = \psi' = 0$ at $z = \infty$, so assuming ψ at the surface is positive, the derivative ψ' must be negative — hence we have taken the negative square root above.)

The resulting equation is separable, using the integral

$$\int \frac{d\eta}{\sinh \eta} = \log \tanh(\eta/2) \quad (6.5.20)$$

Using this, we find

$$\log \left(\frac{\tanh(\psi(z)/4)}{\tanh(\psi(0)/4)} \right) = -\kappa z \quad (6.5.21)$$

Solving for $\psi(z)$, we find after a bit of algebra

$$\psi = 2 \log \left(\frac{1 + \gamma e^{-\kappa z}}{1 - \gamma e^{-\kappa z}} \right) \approx 4\gamma e^{-\kappa z} \quad (6.5.22)$$

in which the final approximation holds for z well beyond the Debye length, and we have defined

$$\gamma = \tanh(\beta e\phi(0)/4) \quad (6.5.23)$$

This result can be used in combination with our perturbative result for the excess ion concentration, to find the disjoining pressure between two charged surfaces at large separations. If the separation is large, the electrostatic potential on the midplane is well approximated by the sum of the potentials for each surface considered in isolation. This gives rise to

$$P \approx 64kT\gamma^2\rho_0 e^{-\kappa D} \quad (6.5.24)$$

for two surfaces separated by a distance D .

Note that the factor γ is somewhat messy to calculate, depending as it does on the value of the potential at the surface. However, we can see that for large surface charges, γ must approach unity, never growing beyond unity no matter how large the surface charge becomes.

Whereas, for weak surface charging, we expect to be able to expand the tanh, whereupon we have

$$\gamma \approx \frac{\beta e\phi(0)}{4} = \frac{\pi l\sigma}{\kappa} \quad (6.5.25)$$

in which we have used the DH result (appropriate for weak charging) to evaluate ϕ on the surface. Then the pressure reduces to

$$P \approx 8\pi l\sigma^2 kT e^{-\kappa D} \quad (6.5.26)$$

which is identical to our previous DH result (as it should be in the weakly charged limit).

DLVO theory

DLVO theory amounts to adding together the double layer repulsive interaction and the van der Waals attractive interaction, and examining the shape of the potential.

Once the shape of the potential is known, we can apply the Derjaguin approximation to compute the interaction force between spherical particles. (Or, as indicated below, we can make a “poor man’s Derjaguin approximation”.)

The disjoining interaction energy per unit area is the integral of the disjoining pressure with respect to separation. Thus the sum of the two interactions in the strongly charged limit ($\gamma \approx 1$) is

$$U = \frac{64\rho_0}{\kappa} e^{-\kappa h} - \frac{\beta A}{12\pi h^2} \quad (6.5.27)$$

with the Debye length given as usual by

$$\kappa^2 = 8\pi\rho_0 l \quad (6.5.28)$$

This interaction is known as the DLVO potential.

Typical values of the parameters are $A = 5 \times 10^{-20}$ J (hence βA is typically about 10), $l = 7 \text{ \AA}$, and ρ_0 ranging anywhere from 1–1000mM (or equivalently, from about 6×10^{-7} – 6×10^{-4} per cubic Angstrom).

Qualitatively, the behavior of the DLVO potential is clear. At extremely small separations, the singular van der Waals attraction dominates the finite disjoining repulsion. This gives rise to an attractive minimum at close contact, known as the primary minimum. Likewise, at separations large compared to the screening length the van der Waals interaction, which only falls off as $1/z^2$, dominates the exponentially decaying repulsive interaction, giving rise to a secondary attractive minimum.

Just outside the primary minimum, the double layer interaction gives a repulsive barrier, the height of which depends on the salt concentration. For small values of salt concentration, the screening length is large, the repulsive barrier is high, and the exponentially decaying disjoining repulsion extends out so far that the secondary minimum is extremely weak, too weak to give rise to any significant binding. Likewise, the large repulsive barrier effectively prevents surfaces from coming into extremely close contact required to access the primary minimum.

As the salt concentration increases, the screening length decreases, the repulsive barrier weakens, and the secondary minimum becomes stronger and moves closer in, eventually becoming strong enough to bind particles or surfaces together against thermal agitation. At still higher salt concentrations, the repulsive barrier weakens sufficiently that surfaces may approach closely enough to access the primary minimum and adhere strongly.

We can create a family of plots the DLVO potential for a set of typical salt concentrations, to see how the depth and location of the secondary minimum vary with salt concentration. Then, for a given particle radius, we can construct a “poor man’s” Derjaguin approximation by arguing that the effective contact area for particles binding with respect to that minimum must scale as hR , where h is the separation of the minimum. This allows us to make a quick graphical estimate of the binding energy for particles of any radius to see what sort of salt concentration will result in binding of the particles with several times kT .