Physical Forces Organizing Biomolecules

III. Electrodynamic (van der Waals) forces

Say "van der Waals force" to most people and they think "van der Waals gas" (if they think anything at all). When a gas becomes dense enough, it is no longer ideal. Interactions between molecules can no longer be ignored. It is the attraction between molecules in a gas that most people think of as the van der Waals force, a rather weak charge-fluctuation force that dies away as the sixth power of molecular separation. So weak that one can think of it as very small even compared to thermal energy kT. So sparsely distributed that one can safely compute its effect as the sum of interactions between the gas molecules considered two at a time, any two molecules feeling each other without significant influence of other near-by molecules.

In fact, charge fluctuation forces are far more powerful within and between condensed phases than they are in gasses. The electric fields that billow out from charges moving about all through the solid or liquid act on many other atoms or molecules at the same time. In no way can one think of the electromagnetic charge fluctuation forces in liquids and solids as one does about forces in a gas.

Rather than think of a solid or liquid as an especially dense gas, the modern and practical way to look at van der Waals forces is in terms of the electromagnetic properties of the condensed materials themselves—in particular their dielectric response and their electromagnetic absorption spectrum.   Why? Because *the frequencies at which charges spontaneously fluctuate are the same as those at which they can resonate to absorb electromagnetic energy from applied electromagnetic waves.*   This is the substance of the "fluctuation dissipation theorem" which states explicitly that the spectrum (frequency distribution) over which charges in material fluctuate is the same as the spectrum of their ability to dissipate (absorb) electromagnetic waves passing across them. [See the Appendix on dielectric susceptibility]

Charge fluctuation force computation is, at least in principle, essentially a matter of converting observed absorption spectra. This possibility holds for *all* forms of matter -- solids and liquids as well as gasses. The absorption spectrum measured on a liquid or solid automatically includes all the interactions and couplings between its constituent atoms or molecules. No need to construct a connection from observing the properties of the disconnected atoms or molecules in a gas of the same material.

Between 1894 when it was first realized that this connection between absorption spectra and charge fluctuation forces must exist, to 1952 when the connection was actually made, to the early 1970's when people finally began to convert measured spectra to predicted forces, people still thought in terms of van der Waals forces between condensed materials as though they were the same as in gasses.[1]

---

[1] The man who made the early connection between spectra and forces was P. N. Lebedev whose early thoughts in his doctoral dissertation were summarized in a 1967 article by his stepson, another outstanding scientist, B. V. Derjaguin.

In this manual, heuristic derivations of the modern forms of the van der Waals force are given in later sections. For the present we'll stick to words, pictures and leading-term formulae.

The Simplest Case

Begin with the interaction between two plane parallel surfaces, separated by a distance *l* across a medium "m". Each of the bodies is semi-infinite, filling the space to the left or the right of the surface.

$$\varepsilon_A \qquad \varepsilon_m \qquad \varepsilon_B$$

$$< l >$$
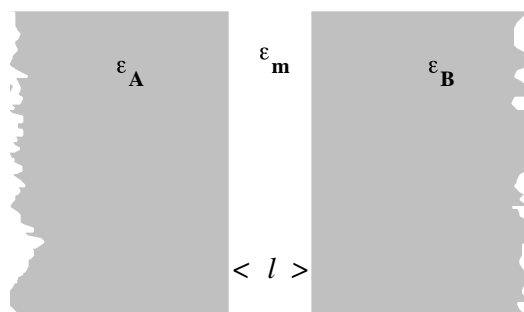
While each of these bodies is electrically neutral, it is composed of moving charges so that at any given instant there is can be a net positive or negative charge at any given location, an instantaneous configuration of charges throughout the space occupied by the bodies and a corresponding electromagnetic field throughout those bodies *and* the space around them. When "A" and "B" are far apart the dance of their charges and field will depend only on their own material properties and that of the surrounding space. When they come to a finite separation, the fields emanating from each will act on the other. If the dance of the charges and fields continued as if the other body were not nearby, the average effect on the energy of interaction would be zero. But ..

the fields coming out of each body (and out of the medium in between) will distort the dance in such a way to favor the probability of configurations that have a lower electrical and magnetic energy. (Moving charges create magnetic fields, which should not be forgotten even though here they are usually not as important as electric fields.) It is this change in probabilities of lower energy charge configurations, this mutual perturbation of the dance of the charges, that creates the charge-fluctuation (or van der Waals or electrodynamic) force.

What kinds of charges are important in these fluctuations? In practice all kinds. Electrons moving about atoms. Vibrating and rotating dipoles. Mobile ions in solution and mobile electrons

in metals. Every charge movement that can respond to an applied electric field is a charge movement that can create transient electric fields.

You can think in terms of wildly dancing charge configurations or think of the ever-changing electric fields set up by these moving charges or think of the spectrum of collective waves sent between the interacting bodies at different separations. The closer the separation, the stronger the coupling between fluctuations and the stronger the electrical signals between "A" and "B".

The ease of time-varying charge displacement, measured as the time-dependent polarizability, is expressed by the dielectric susceptibility function, ε. This susceptibility is function of frequency; it measures the ability of a material to react to an electric field at each frequency.

The electrodynamic work, or free energy, $G_{AmB}(l)$ to bring bodies A and B to separation l from an infinite separation in medium m depends on there being a *difference* in the dielectric susceptibilities ($\varepsilon - \varepsilon_m$) and ($\varepsilon - \varepsilon_m$) of each of the bodies and the medium. (If the ε's were the same for two adjacent materials, the interface between materials would be electromagnetically invisible. No electromagnetic interface, no "separation" l!) [2]

There is a major price for the right to use whole-material dielectric susceptibilities. One is assuming that the two bodies A and B are far enough apart that they do not see molecular or atomic features in their respective structures. This is the "macroscopic-continuum" limit where material are treated as macroscopic bodies on the laboratory scale and all polarizability properties

---

[2] I have used dielectric "susceptibility" here as a generalization of dielectric "constant". Why? When dielectric constants were introduced as a way to fudge coulomb's law for charge-charge interactions, it was to recognize that the medium itself responded to the static or glacially changing electric fields set up by the charges. A good way to think about that response is to imagine the material placed between the two plates of a capacitor. (Up to microwave frequencies, a capacitor is actually practical, not just conceptually convenient for measuring dielectric properties. At higher frequencies -- infrared, visible, ultraviolet -- one uses absorption and reflection of light.) One applies a known voltage and measures the amount of charge it takes to reach that voltage. That amount of charge is proportional to the dielectric constant of the material between the plates. The amount of charge put on the plates is equal and opposite to what the material between the plates delivered to its outer boundaries at the plates.
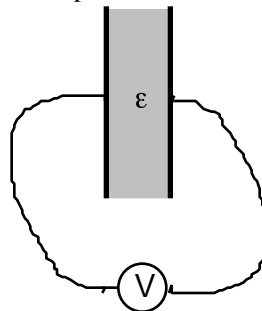
Here, with charge fluctuations and electrodynamics, one needs to know the response of the material at *all* frequencies. It is as though one conceived an ideal capacitor that could apply electric fields at all frequencies and measure a susceptibility of the material at each of those frequencies.

Susceptibilities are elaborated in great detail in Appendix EE (files iv-eps*.doc)
.

are averaged out much as in they average out in a capacitance measurement or in transmission/reflection measurements on macroscopic bodies.

In fact, the price is not too high. Separations l must be much greater than the graininess of the atomic packing. For most materials, separations down to ~20 Angstroms are probably allowed. This limiting distance still allows a very large number of important applications.[3]

Mathematical form of the charge-fluctuation free energy of interaction of material A with B across medium m.

The formula has a lot of pretty parts. Because I want to avoid any fear of using it, I'll write it out in a *nearly* exact version, then take it apart piece by piece.

This simplified version of the interaction formula is for the case where there isn't a big relative difference in susceptibilities , , $_m$ and where there is a negligible difference in the velocity of light in the three media.

For historical reasons, the interaction between planar parallel surfaces is written in terms of a "Hamaker Coefficient"[4] so that the free energy of interaction per unit area $G_{AmB}(l)$ relative to infinite separation is in the form

$$G_{AmB}(l) = -\frac{A_{Am/Bm}}{12\pi\, l^2}$$

The Hamaker coefficient $A_{Am/Bm}$ (in modern language!) is to first approximation

$$A_{Am/Bm} \quad \frac{3kT}{2} \sum_{n=0}' \quad _{Am} \quad _{Bm} \, R_n,$$

is a function of distance l and of each material's dielectric susceptibility. The index n is to designate a sum over frequencies described below.

Except for the fact that the coefficient $A_{Am/Bm}$ itself varies with separation, the immediate appearance of the change in free energy with separation is that it goes as the inverse *square* of the separation l. There is additional distance dependence in the "relativistic screening function" R_n to

---

[3] It is unfortunate that this macroscopic-continuum limitation is sometimes forgotten in over-zealous application. The same limitation also holds in the theory of the electrostatic double layers where one often makes believe that the medium is a featureless continuum. Neglect of structure in double layers is equally risky though and even more common than in the computation of van der Waals forces.

The m-c limitation is sometimes circumvented by creating spatially varying dielectric susceptibilities and solving the charge fluctuation equations with these more detailed structures.

Lest anyone think that the macroscopic-continuum nature of the modern theory renders it inferior to its predecessors, recall that the alternative is to give all the material A, B, m the polarizability properties of a *gas* rather than of a liquid or a solid, an assumption so gross as to kill any hope of further logical thought.

[4] One still hears the archaic designation "Hamaker *Constant*" from the time when people did not recognize that the coefficient could itself vary with separation l. However its dependence on l is usually much weaker than that of the l^2 in the denominator and this coefficient in modern usage can be a very useful measure of the strength of van der Waals forces.

be elaborated below. At short distances $R_n = 1$; and long distances summation with this factor creates another inverse power in separation to make the total energy of interaction go as $1/l^3$.

Form of dependence on ε's:

The susceptibilities $\quad$, $\quad$, $_m$ come in as relative differences, that is, differences over sums in the quantities $\overline{\quad}_{Am}$ and $\overline{\quad}_{Bm}$:

$$\overline{\quad}_{Am} = \frac{\varepsilon_A - \varepsilon_m}{\varepsilon_A + \varepsilon_m} \, ,$$

$$\overline{\quad}_{Bm} = \frac{\varepsilon_B - \varepsilon_m}{\varepsilon_B + \varepsilon_m} \, .$$

Only relative differences in dielectric response show up in the van der Waals interaction.[5]

Frequencies at which ε's are evaluated:

The $\quad$'s are sampled over an infinite series of what are unfortunately called "imaginary" frequencies, a terrible name that for decades has probably stymied people from taking advantage of the modern theory. Why "imaginary"?

Because the dance of the charges, is described in terms of their spontaneous fluctuation from their time-average positions and their gradual, exponential-in-time return to average positions. "Imaginary" frequencies describe this exponentially varying, rather than a sinousoidally varying, (sine wave) process.[6]

The $\quad$'s are evaluated at discrete frequencies $i\,_n$ where

$$_n = \frac{2\ kT}{\hbar}\,n$$

---

[5] Again, if there were no difference in the way charges danced in the different materials, there would be no sense talking about a dielectric "interface" between them. For all electromagnetic intents and purposes they would be the same material; no interface. No energy depends on location of invisible interfaces.

   Subtly, the formula for the $\overline{\quad}$'s suggests that it is the relative difference in susceptibilities that matters. Think of the sum in the denominators ($\quad + \,_m$) and ($\quad + \,_m$),as the average of the two $\quad$'s. Because the dielectric properties of each material are different functions of frequency, the $\overline{\quad}$'s will also depend on frequency.

[6] Using the "imaginary" number

$$i = \sqrt{-1}\,,$$

the exponential becomes

$$e^{\iota\quad} = \cos(\ ) + i\,\sin(\ ).$$

   If $\quad$ increases with time, $\quad = \quad t$, we speak of $\quad$ as the "radial frequency" in radians/second (the same as $2\ f$ (or $2\quad$) where frequency f (or $\quad$) is the normal frequency in cycles/second or Hertz). In this case the sine, cosine and exponential go through the usual "sinusoidal" oscillations.

   But if $\quad$ were itself an imaginary quantity,

$$\quad = i\quad\,,$$

with $\quad$ a positive real number, then the exponential $e^{i\,\omega t}$ would have an argument $i\quad\quad t = i^2\quad t = -\quad t$. The formerly oscillatory $e^{i\quad t}$ now decays exponentially, $^{-\quad t}$.

with a running index of summation n = 0 to infinity (and the prime ' in the summation for $G_{AB}(l)$ meaning that the n = 0 term is to be multiplied by 1/2).

Lots of neat things happening already.

- kT, Boltzmann's constant times absolute temperature, reflects the contribution of thermal agitation to charge fluctuation.
- $2\pi\hbar$ is Planck's constant, reflecting the fact that many charge fluctuations are quantum mechanical in origin. $\hbar \times \xi_n$ is the energy of a photon.[7] The formula for ξ_n tells us that the energy of the interesting photons is to be measured in thermal energy units. The computation of the ε's at these ξ_n is described below.

At room temperature, the coefficient $\dfrac{2\pi\, kT}{\hbar}$ in $\xi_n = \dfrac{2\pi\, kT}{\hbar}n$ is big enough that ξ_1, the first frequency after $\xi_0 = 0$, corresponds to the lifetimes of infra-red vibrations. The sampling over the set of ξ_n then involves one zero-frequency term, then a few terms in the infra-red followed by a very large number of terms in the visible and ultraviolet parts of the spectrum. It is for this reason that most of the time one expects ultraviolet spectral properties to dominate van der Waals interactions although at these frequencies the quantities

$$\overline{\Delta}_{Am} = \frac{\varepsilon_A - \varepsilon_m}{\varepsilon_A + \varepsilon_m} \quad\text{and}\quad \overline{\Delta}_{Bm} = \frac{\varepsilon_B - \varepsilon_m}{\varepsilon_B + \varepsilon_m}$$

are very small.

This expectation is not realized in aqueous systems though where the huge differences in $\varepsilon(0)$ for water and for non-polar materials can make for very big $\overline{\Delta}_{Am}$'s.

Effect of the finite velocity of the electromagnetic signal traveling across the medium m:

The last term in the summation for $A_{Am/Bm}$ is R_n, a dimensionless "relativistic retardation correction term" due to the finite velocity of light. Here is another of those complicated-sounding terms for a simple idea. Think about the dancing charges. Think of the time it takes them to do a particular step. Think of the travel time for the electric pulse created by that dancing charge to go a distance l to another body plus the time it takes for the charges on that other body to respond and send a signal back. Altogether it's a trip 2 l in length; the time it takes is this back-and-forth length divided by the velocity of light in the intervening medium.

The characteristic times are the periods of charge fluctuation or lifetimes $1/\xi_n$ that are from the set of characteristic frequencies ξ_n, the rhythms of a set of simultaneous dances with interaction being the coupling of dances at each independent rhythm ξ_n.

---

[7] The Planck constant h = $2\pi\hbar$ = 6.62517 x 10^{-27} erg seconds = 6.62517 x 10^{-34} joule seconds; $\hbar$ = 1.0544 x 10^{-27} erg seconds = 1.0544 x 10^{-34} joule seconds.

The energy of a photon is Planck's constant h times frequency f (or ν), E = hν or $\hbar\omega$ when $\omega = 2\pi\nu$.

Rearrange the definition of ξ_n and you have

$$\hbar\xi_n = 2\pi\, kT\, n$$

as though the energy of these interesting imaginary-frequency photons were multiples of $2\pi\, kT$.

The travel time of the signal, across the gap l and back again is the length $2l$ divided by the velocity of light $c/\sqrt{\varepsilon_m(i\xi_n)}$ in medium m.[8]. That is,
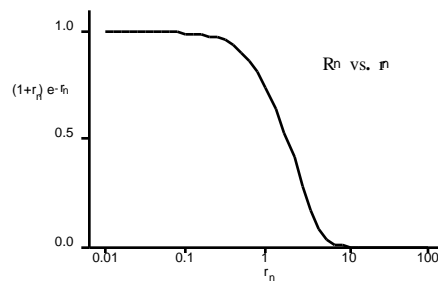
$$\frac{2l}{c/\varepsilon_m^{1/2}}$$

The parameter of interest is a ratio, r_n, which is this travel time relative to the fluctuation lifetime, $1/\xi_n$.

$$r_n \;=\; \frac{2l}{c/\varepsilon_m^{1/2}} \;\bigg/\; \frac{1}{\xi_n}$$

The relativistic screening of a particular frequency fluctuation goes as

$$R_n = (1 + r_n)\, e^{-r_n}$$

which is less than or equal to one. It looks like

When the ratio r_n is much less than 1, that is when the light signal travels across and back much faster than the length of time $1/\xi_n$. that a fluctuation endures, then there is no loss of signal between the two bodies.

$R_n = (1 + r_n)e^{-r_n} \sim 1$ Then the finite velocity of light does not affect the van der Waals interaction from frequency ξ_n.

When $r_n \gg 1$, the time of travel of the signal across the gap l and back is longer than the lifetime of the electromagnetic fluctuation, then the damping term goes to zero almost exponentially

$$(1 + r_n)e^{-r_n} \quad r_n e^{-r_n} \qquad 0$$

so that there is no longer a contribution to van der Waals forces at large distances or from high frequencies (large ξ_n, small period of fluctuation $1/\xi_n$). At these large distances the signal takes too long for the movement of charges to fall into step. The "first" charge has moved too far by the time the signal comes back from the other side.

---

[8]To keep language consistent, I used the square root of dielectric susceptibility here rather than its equivalent, the index of refraction

$$n_{ref} \;=\; \sqrt{\varepsilon_m(i\xi_n)}.$$

When separations are small enough that $r_n = 0$ ($R_n = 1$) for all contributing frequencies, the interaction between A and B across m goes as the inverse square of separation

$$G_{AmB}(l) = -\frac{A_{Am/Bm}}{12\ l^2}$$

with a Hamaker coefficient

$$A_{Am/Bm} \approx \frac{3kT}{2} \sum_{n=0}^{'} \overline{\ \ _{Am}\ \ _{Bm}\ \ },$$

that is, not a function of separation (through R_n).

At larger separations $A_{Hamaker}$ also varies with separation so that the energy does not go as a simple l^2.

How big are these long-range van der Waals forces? When do we think of them as responsible for organizing large molecules or aggregates? One criterion is whether they are large compared to thermal energy kT. When are the forces strong enough to overcome thermal agitation? Several worked examples are given later for computing van der Waals forces with different materials.

A useful rule of thumb is that for non-conducting materials, the sum $\sum_{n=0}^{'} \overline{\ \ _{Am}\ \ _{Bm}\ \ }$ is often of the order of unity. The coefficient $A_{Am/Bm}$ is of the order of kT. Since $G_{AmB}(l) = -\frac{A_{Am/Bm}}{12\pi\ l^2}$ is an energy per unit area, we can expect anti-thermally significant attraction when the area of interaction is large compared with $12\ l^2$. More simply state, let the size of the bodies be small compared to their separation and that separation be small enough that there be no significant retardation screening.

The same reasoning holds for non-planar bodies, spheres or cylinders can be expected to enjoy strong attraction when their separations are small compared with their radii.

Screening of low-frequency fluctuations in ionic solutions.

        As already emphasized, the full computation of charge-fluctuation forces necessarily includes movement by all kinds of charge. Traditionally the most important charges were thought to be the electrons that were the greatest number of charges in any material. While this expectation is usually verified in "dry" systems, it does not necessarily hold for polar liquids such as water wherein there can be charge fluctuations from dipolar vibrations (at infra-red frequencies), from rotations of polar molecules (at microwave frequencies), and from ionic fluctuations (from microwave down to zero frequency). The full theory as developed in modern from implicitly includes all these contributions by via the form of the dielectric susceptibilities that can also include zero-frequency charge movements that correspond to the currents carried by mobile charges be they electrons in a metal or ions in a salt solution[9].

        This last case, salt solution, is of special interest in biological and colloidal systems and merits special attention on its own. Because of their capacity to form diffuse electrostatic double layers, the fluctuations of mobile ions displays a particularly pretty coupling between charge fluctuations and screening of the electric fields that come from those fluctuations. Because

        The first example of this coupling is seen as a "salt-screening" of very low frequency charge fluctuations. Imagine that medium m is a salt solution with a Debye screening constant and screening length $_D = 1/$. A low-frequency electric field emanating from body A will be screened by the salt solution with an attenuation typical of double layers between parallel planar bodies. That is, it will die as e^{-x} as a function of the distance x from the interface where the signal enters medium m from body A. By the time the signal travels the distance l to body B it will be screened to an extent e^{-l}. The response of B back to A will also suffer a screening by a factor e^{-l}.

        The across-and-back screening of charge fluctuations has a form remarkable similar to the relativistic screening due to the finite velocity of light. We will call it R_0 to emphasize that it occurs for $_n = 0$ and write it in the approximate form

$$R_0 = (1 + 2 \quad l)\, e^{-2\ l}.$$

or

$$R_0 = (1 + 2l/\ )\, e^{-2l/} .$$

Recall from the section on electrostatic double layers that the screening length in a .1 Molar solution is ~ 10 Angstroms. The ionic screening of low-frequency fluctuations can be very strong. For a separation $l = 10$ Å (already rather too small for the continuum limit where the van der Waals theory holds) the screening will amount to a factor
$(1 + 2)\, e^{-2} \quad .4$; for $l = 20$ Å, $(1 + 3)\, e^{-3} \quad .2$, an 80% screening of the signal.

        Ionic screening of low-frequency fluctuations is even more potent when the surfaces bear permanent charge and the ~ Molar local ion concentrations have effective   ~ 3-Angstrom screening lengths.

        Other aspects of ionic fluctuations will be deferred to a separate section.

---

[9] Please see section IV-xx with the derivation of the full formula.