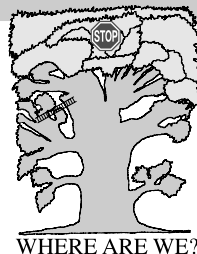


Chapter 13

INTERMOLECULAR INTERACTIONS



Where are we?

We are already in the crown of the TREE.

An example

Why does liquid water exist? Why do molecules stick together at low temperatures? Visibly they attract each other *for some reason*. The interaction is not however very strong since water evaporates when heated (without destroying the water molecules).

What is it all about

THEORY OF INTERMOLECULAR INTERACTIONS

Interaction energy concept (Δ)

p. 684

- Natural division and its gradation
- What is most natural?

Binding energy (Δ)

p. 687

Dissociation energy (Δ)

p. 687

Dissociation barrier (Δ)

p. 687

Supermolecular approach (Δ)

p. 689

- Accuracy should be the same
- Basis set superposition error (BSSE)
...and the remedy
- Good and bad news about the supermolecular method

Perturbational approach ($\diamond\hat{U}\S$)

p. 692

- Intermolecular distance – what does it mean?
- Polarization approximation (two molecules) (Δ)
- Intermolecular interactions: physical interpretation
- Electrostatic energy in the multipole representation and the penetration energy
- Induction energy in the multipole representation
- Dispersion energy in the multipole representation

Symmetry-Adapted Perturbation Theories (SAPT) ($\diamond\hat{U}\S$)

p. 710

- Polarization approximation is illegal
- Constructing a symmetry adapted function
- The perturbation is always large in polarization approximation

- Iterative scheme of the symmetry adapted perturbation theory
- Symmetry forcing
- A link to the variational method – the Heitler–London interaction energy
- When we do not have at our disposal the ideal $\psi_{A,0}$ and $\psi_{B,0}$

Convergence problems (♦)

p. 721

Non-additivity of intermolecular interactions (♦ \bar{U})

p. 726

- Many-body expansion of interaction energy
- Additivity of the electrostatic interaction
- Exchange non-additivity
- Induction energy non-additivity
- Additivity of the second-order dispersion energy
- Non-additivity of the third-order dispersion interaction

ENGINEERING OF INTERMOLECULAR INTERACTIONS

Noble gas interaction

p. 741

Van der Waals surface and radii (Δ)

p. 742

- Pauli hardness and the van der Waals surface
- Quantum chemistry of confined space – the nanovessels

Synthons and supramolecular chemistry (\perp)

p. 744

- Bound or not bound
- Distinguished role of the electrostatic interaction and the valence repulsion
- Hydrogen bond
- Coordination interaction
- Hydrophobic effect
- Molecular recognition – synthons
- “Key-and-lock”, template and “hand-and-glove” synthon interactions

Chapter 8 dealt with the question of why atoms form molecules. Electrons and nuclei attract each other, and this results in almost exact neutralization of matter. Despite this, atoms and molecules interact, because

- two atoms or molecules cannot occupy the same space,
- electrons and nuclei in an atom or molecule may still interact with those in other atoms or molecules.

This chapter will tell us about the very reason for this and will give details of the interaction.

Why is this important?

What is the most important message humanity ever learned about matter? According to Richard Feynman the message would be: “*The world is built of atoms, which repel each other at short distances and attract at longer ones*”. If the intermolecular interactions were suddenly switched off, the world would disintegrate in about a femtosecond, that is in a single period of atomic vibration (the atoms simply would not come back when shifted from their equilibrium positions). Soon after, everything would evaporate and a sphere of gas, the remainder of the Earth, would be held by gravitational forces. Isn’t it enough?

What is needed?

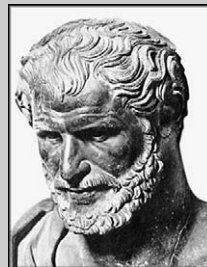
- Perturbation theory (Chapter 5, absolutely).

- Variational method (Chapter 5, recommended).
- Appendix X, p. 1038 (absolutely).
- Many-Body Perturbation Theory (MBPT) (Chapter 10, p. 554, necessary).
- Reduced resolvent (Chapter 10, p. 554, necessary).
- Appendix Y, p. 1050 (recommended).
- Appendix T (mentioned).

Classical works

Such an important subject was recognized very early. The idea that the cohesion of matter stems from the interaction of small indivisible particles (“atoms”) comes from Democritus. ★ An idea similar to that cited by Feynman was first stated clearly by the Croat scientist Rudjer Bosković in “*Theoria Philosophiae naturalis*”, Venice, 1763. ★ Padé approximants were first proposed in the PhD thesis of Henri Padé entitled “*Sur la représentation approchée d’une fonction pour des fractions rationnelles*”, which was published in *Annales des Sciences d’Ecole Normale Supérieure, Suppl. [3]*, 9 (1892) 1. ★ The role of intermolecular interactions was highlighted in the work of Johannes Diderik van der Waals, especially in “*Die Kontinuität des gasförmigen und flüssigen Zustandes*”, Barth, Leipzig (1899, 1900). From that time on, intermolecular interactions are often called van der Waals interactions. ★ The concept of ionic radii was first proposed by Linus Pauling in “*The Sizes of Ions and the Structure of Ionic Crystals*”, *Journal of the American Chemical Society*, 49 (1927) 765. ★ The quantum mechanical explanation of intermolecular forces, including the ubiquitous dispersion interactions, was given by Fritz London in “*Zur Theorie und Systematik der Molekularkräfte*”, *Zeitschrift für Physik*, 63 (1930) 245 and in “*Über einige Eigenschaften und Anwendungen der Molekularkräfte*” from *Zeitschrift für Physikalische Chemie (B)*, 11 (1930) 222. ★ Linus Pauling, invited Baker Lecturer to Cornell University, wrote one of the most seminal books in chemistry “*The Nature of the Chemical Bond*”, Cornell Univ. Press, Ithaca, 1948, where *inter alia* he elucidated the role of hydrogen bonds in forming structures. ★ The hydrophobic effect was first highlighted by Walter Kauzmann in a paper “*Some Factors in the Interpretation of Protein*

Democritus of Abdera (ca. 460 B.C. – ca. 370 B.C.), Greek philosopher, founder of the first atomic theory. According to him, nature represents a constant motion of indivisible and permanent particles (atoms), whose interactions result in various materials. *It turned out after almost 25 centuries that this hypothesis was basically correct!* All the written works of Democritus have been lost, but his



ideas continued to have an important impact on science for centuries.

Rudjer Josip Bosković (1711–1787), a Croat physicist, mathematician, astronomer and philosopher from beautiful Dubrovnik.



Johannes Diderik van der Waals (1837–1923), Dutch physicist, professor at the University of Amsterdam. His research topic was the influence of intermolecular forces on the properties of gases (equation of state of the real gas, 1873) and liquids. In 1910 van der Waals received a Nobel Prize “*for his work on the equation of state for gases and liquids*”.



Denaturation”, in *Advances in Protein Chemistry*, 14 (1959) 1, the effect was further elaborated by George Nemethy, Harold Scheraga, Frank Stillinger and David Chandler among others. ★ Resonance interactions were first described by Robert S. Mulliken in an article “*The Interaction of Differently Excited Like Atoms at Large Distances*”, in *Physical Reviews*, 120 (1960) 1674. ★ Bogumił Jeziorski and Włodzimierz Kołos extended the existing theory of intermolecular forces to intermediate distances (“*On the Symmetry Forcing in the Perturbation Theory of Weak Intermolecular Interactions*”, *International Journal of Quantum Chemistry*, 12 Suppl. 1 (1977) 91).

THEORY OF INTERMOLECULAR INTERACTIONS

There are two principal methods of calculating the intermolecular interactions: the supermolecular method and the perturbational method. Both assume the Born–Oppenheimer approximation.

13.1 INTERACTION ENERGY CONCEPT

The idea of interaction energy is based on the Born–Oppenheimer (clamped nuclei, see eq. (6.4)) approximation. Let us define interaction energy at the configuration \mathbf{R} of the nuclei as

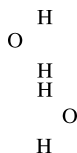
$$E_{\text{int}}(\mathbf{R}) = E_{ABC\dots}(\mathbf{R}) - [E_A(\mathbf{R}) + E_B(\mathbf{R}) + E_C(\mathbf{R}) + \dots], \quad (13.1)$$

where $E_{ABC\dots}(\mathbf{R})$ is the electronic energy (corresponding to $E_0^{(0)}$ from eq. (6.21)) of the total system, and $E_A(\mathbf{R})$, $E_B(\mathbf{R})$, $E_C(\mathbf{R})$, ... are the electronic energies of the interacting subsystems, calculated *at the same positions of the nuclei* as those in the total system.

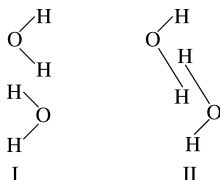
13.1.1 NATURAL DIVISION AND ITS GRADATION

Although the notion of interaction energy is of great practical value, its theoretical meaning is a little bit fuzzy. Right at the beginning we have a question: interaction of *what*? We view the system *as composed of particular subsystems*, that once isolated, then have to be put together.

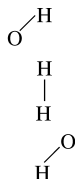
For instance, the supersystem



may be considered as two interacting water molecules, but even then we still have an uncertainty, whether the two molecules correspond to (I) or to (II):



In addition the system might be considered as composed of a hydrogen molecule interacting with two OH radicals:



etc.

Choice of subsystems is of no importance from the point of view of mathematics, but is of crucial importance from the point of view of calculations in theoretical chemistry.

The particular choice of subsystem should depend on the kind of experiment with which we wish to compare our calculations:

- we are interested in the interaction of water molecules when studying water evaporation or freezing;
- we are interested in the interaction of atoms and ions that exist in the system when heating water to 1000 °C.

Let us stress, that in any case when choosing subsystems we are forced to single out *particular atoms* belonging to subsystem¹ A and B. It is not sufficient to define the kind of molecules participating in the interaction, see our examples I and II.

If when dividing a system into n subsystems in two ways (I and II), we obtain $|E_{\text{int}}|_{\text{I}} < |E_{\text{int}}|_{\text{II}}$, division I will be called more *natural* than division II.

natural
subsystems

13.1.2 WHAT IS MOST NATURAL?

Which division is *most* natural? We do not have any experience in answering such questions. What? Why should we have any difficulties? It is sufficient to consider all possible divisions and to choose the one which requires lowest energy. Unfortunately, this is not so obvious. Let us consider two widely separated water molecules (Fig. 13.1.a).

¹This means that the interaction energy idea belongs to classical concepts. In a quantum system, particles of the same kind are indistinguishable. A quantum system does not allow us to separate a part from the system. Despite this, the interaction energy idea is important and useful.

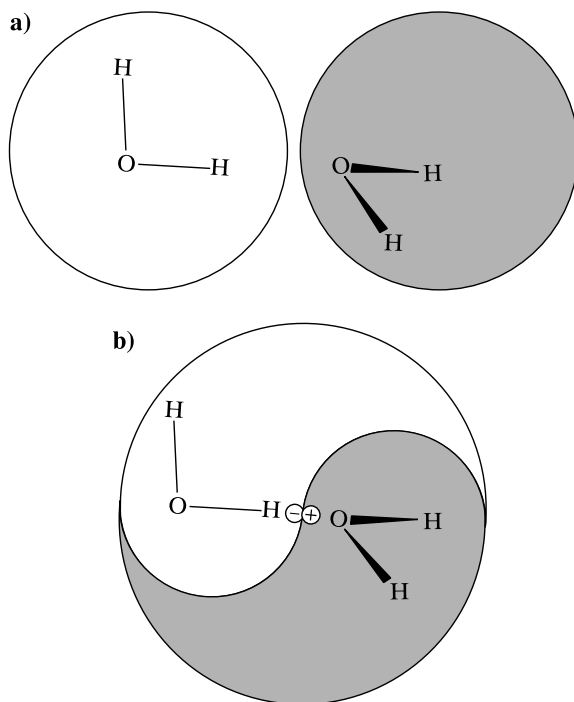


Fig. 13.1. Part-entity relationship. Two distinct ways of dividing the $(\text{H}_2\text{O})_2$ system into subsystems. Division (a) is traditional. The interacting objects are two isolated water molecules and the interaction energy is equal to about -5 kcal/mol (attraction). Division (b) is more subtle – a certain point in space is treated as being composed of two fictitious charges $q > 0$ and $-q$, and one of the charges is ascribed to one, and the other to the second molecule. In this way two *new subsystems* are defined, each of them composed of a water molecule *and* the corresponding point charge. The value of q may be chosen in such a way as to produce the interaction energy of the new subsystems close to 0. Therefore this is a more natural choice of subsystems than the traditional one. The total interaction energy of the two water molecules is now absorbed within the interactions of the fictitious point charges with “their” water molecules. Each of the point charges takes over the interaction of “its” water molecule with the rest of the Universe. Hence, I have permitted myself (with the necessary *licentia poetica*) to use the *yin* and *yang* symbols – the two basic elements of ancient Chinese philosophy.

Right between the molecules, e.g., in the middle of the OO separation, we place two point charges $q > 0$ and $-q$, i.e. we place nothing, since the charges cancel each other (please compare a similar trick on p. 492). We therefore have just two water molecules. Now we start our game. We say that the charges are *real*: one belongs to one of the molecules and the other to the second (Fig. 13.1.b). The charge q could be anything, but we want to use it for a very special goal: to construct the two subsystems in a *more natural* way than just two water molecules. It is interesting that after the choice is made, any of the subsystems has *lower* energy than that of isolated water, since the molecules are oriented in such a way as to attract each other. This means that the value of q can be chosen from an interval making the choice of subsystems more natural. For a certain $q = q_{\text{opt}}$ we would obtain as

the interaction energy of the new subsystems: $E'_{\text{int}} = 0$. This certainly would be the most natural choice,² with the “dressed” water molecules, not seeing each other.³

Below in this chapter we will not use any fictitious charges.

13.2 BINDING ENERGY

Interaction energy can be calculated at any configuration \mathbf{R} of the nuclei. We may ask whether any “privileged” configuration exists with respect to the interaction energy. This was our subject in Chapter 7, and it turned out that the electronic energy may have many minima (equilibria) as a function of \mathbf{R} . For each of such configurations we may define the *binding energy with respect to a particular dissociation channel* as the difference of the corresponding interaction energies (all subsystems at the optimal positions $\mathbf{R}_{\text{opt}(j)}$ of the nuclei with respect to the electronic energy E_j):

equilibrium
geometry

$$E_{\text{bind}} = E_{ABC\dots}(\mathbf{R}_{\text{opt}(\text{tot})}) - \sum_{j=A,B,C,\dots} E_j(\mathbf{R}_{\text{opt}(j)}). \quad (13.2)$$

At a given configuration $\mathbf{R}_{\text{opt}(\text{tot})}$ we usually have many dissociation channels.

13.3 DISSOCIATION ENERGY

The calculated interaction energy of eq. (13.1), as well as the binding energies are only theoretical quantities and cannot be measured. What is measurable is the closely connected *dissociation energy*

$$E_{\text{diss}} = E_{\text{bind}} - \left[\Delta E_{0,\text{tot}} - \sum_{j=A,B,C,\dots} \Delta E_{0,j} \right], \quad (13.3)$$

where $\Delta E_{0,\text{tot}}$ stands for what is known as the *zero vibration energy* of the total system (cf. p. 304) at the equilibrium geometry $\mathbf{R}_{\text{opt}(\text{tot})}$, and $\Delta E_{0,j}$ for $j = A, B, C, \dots$ representing the zero vibration energies for the subsystems. In the harmonic approximation $\Delta E_{0,\text{tot}} = \frac{1}{2} \sum_i h\nu_{i,\text{tot}}$, $\Delta E_{0,A} = \frac{1}{2} \sum_i h\nu_{i,A}$, $\Delta E_{0,B} = \frac{1}{2} \sum_i h\nu_{i,B}$, $\Delta E_{0,C} = \frac{1}{2} \sum_i h\nu_{i,C}$, \dots at their equilibrium geometries $\mathbf{R}_{\text{opt}(A)}$, $\mathbf{R}_{\text{opt}(B)}$, $\mathbf{R}_{\text{opt}(C)}$, \dots , respectively.

13.4 DISSOCIATION BARRIER

If a molecule receives dissociation energy it is most often a sufficient condition for its dissociation. Sometimes however the energy is too low, and the reason is that there is an energy barrier to be overcome. Sometimes the barrier is very high and

²Although not unique, since the charges could be chosen at different points in space, and we could also use point multipoles, etc.

³Allusion to the elementary particles “dressed” by interactions, see section “What is it all about?”, Chapter 8. It is worth noting that we have to superpose the subsystems *first* (then the fictitious charges disappear), and *then* calculate the interaction energy of the water molecules deformed by the charges.

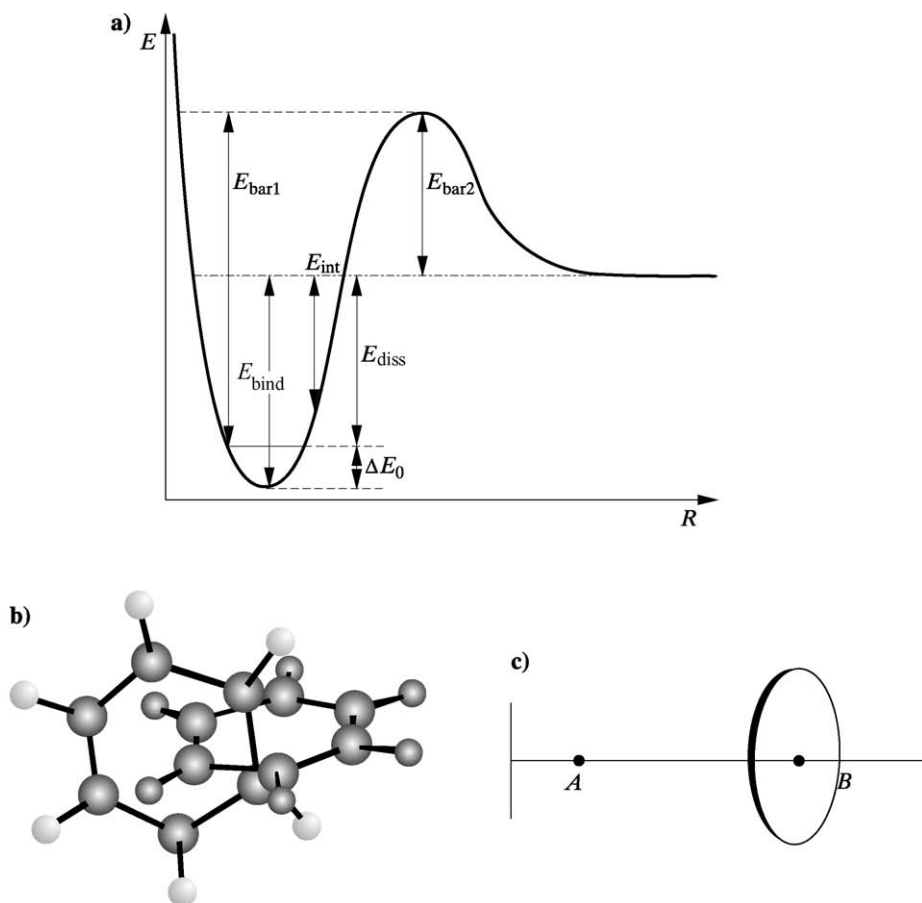


Fig. 13.2. (a) Interaction energy E_{int} , binding energy E_{bind} , dissociation energy E_{diss} and barrier energy E_{bar} . ΔE_0 is the zero vibration energy. Please note that the height of the barrier is different (E_{bar1} or E_{bar2}) depending on the starting point considered. The first is connected with the energy cost required to go from jail to freedom (large), while the second is the energetic price for going to jail (much easier). Figs. (b), (c), (d) represent rather exceptional cases of intermolecular interactions, when a part of the total system is somehow confined (Fig. (a)) by the rest of the system. The catenane shown in Fig. (b) consists of two intertwined rings, the rotaxan (scheme) is composed of a “molecular stick” with a molecular ring on it, Fig. (c), the latter having two stable positions (“stations A and B” to be used in future molecular computers). Fig. (d) shows what is called an endohedral complex. In this particular case a water molecule is confined in the fullerene cage. The systems have been trapped in metastable states, as they were formed. To free the subsystems, high energy has to be used to break some chemical bonds. One of the advantages of theory is that we can consider compounds which sometimes would be difficult to obtain in experiment.

the system is stable even if the dissociation products have (much) lower energy. Catenans, rotaxans and endohedral complexes shown in Fig. 13.2 may serve as examples.

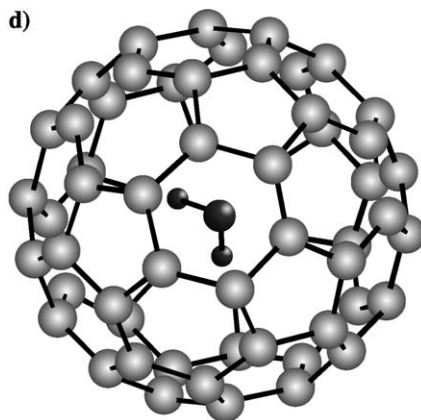


Fig. 13.2. *Continued.*

The energy necessary to overcome the barrier from the trap side is equal to

$$E_{\text{bar}} = E_{\#} - \left(E_{\text{min}} + \frac{1}{2} \sum_j h\nu_j \right),$$

where E_{min} is the energy of the bottom of the well, $E_{\#}$ represents the barrier top energy, and $\frac{1}{2} \sum_j h\nu_j$ is the zero vibration energy of the well.

13.5 SUPERMOLECULAR APPROACH

In the supermolecular method the interaction energy is calculated from its definition (13.1) using any reliable method of electronic energy calculations. For the sake of brevity we will consider the interaction of two subsystems: A and B .

13.5.1 ACCURACY SHOULD BE THE SAME

There is a problem though. The trouble is that we are unable to solve the Schrödinger equation exactly either for AB or for A or for B . We have to use approximations. If so, we have to worry about the same accuracy of calculation for AB as for A and B . From this we may expect that

in determining E_{AB} as well as E_A and E_B the same theoretical method should be preferred, because any method introduces its own systematic error and we may hope that these errors will cancel at least partially in the above formula.

This problem is already encountered at the stage of basis set choice. For example, suppose we have decided to carry out the calculations within the Hartree–Fock method in the LCAO MO approximation, p. 360. The same method has to be used for AB , A and B . However what does this really mean? Should we use the following protocol:

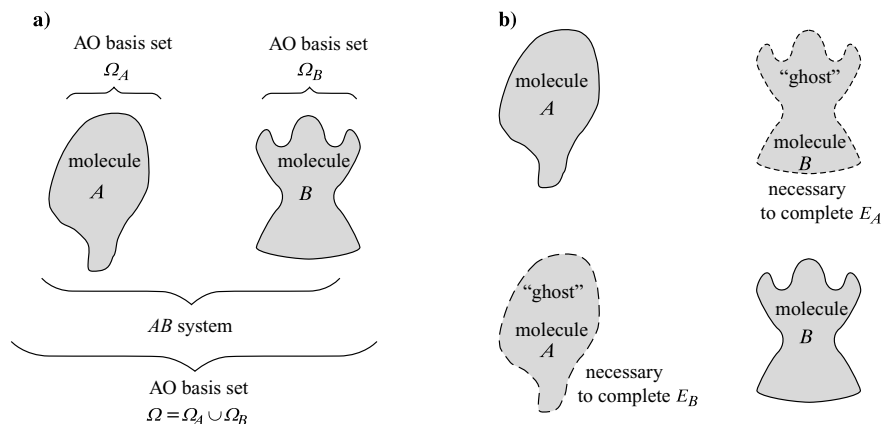


Fig. 13.3. (a) Basis set superposition problem (BSSE). Each of the molecules offers its own atomic orbitals to the total basis set $\Omega = \Omega_A \cup \Omega_B$. Fig. (b) illustrates the counter-poise method, in which the calculations for a single subsystem are performed within the full atomic basis set Ω : the atomic orbitals centred on it and what are called ghost orbitals centred on the partner.

1. Consider the atomic basis set Ω that consists of the atomic orbitals centred on the nuclei of A (set Ω_A) and on the nuclei of B (set Ω_B), i.e. $\Omega = \Omega_A \cup \Omega_B$.
2. Calculate E_{AB} using Ω , E_A using Ω_A and E_B using Ω_B (Fig. 13.3a). Apparently everything looks logical, but *we did not use the same method* when calculating the energies of AB , A and B . The basis set used has been different depending on what we wanted to calculate.

Thus, it seems more appropriate to calculate all three quantities using the same basis set Ω .

13.5.2 BASIS SET SUPERPOSITION ERROR (BSSE)

Such an approach is supported by the following reasoning. When the calculations are performed for E_{AB} within the basis set Ω we calculate implicitly not only the interaction energy, but also we allow the *individual subsystems* to lower their energy. Conclusion: by subtracting from E_{AB} the energies: E_A calculated with Ω_A and E_B with Ω_B , we are left not only with the interaction energy (as should be), but also with an unwanted and non-physical extra term (an error) connected with the artificial lowering of the subsystems' energies, when calculating E_{AB} . This error is called the BSSE (*Basis Set Superposition Error*).

... and the remedy

To remove the BSSE we may consider the use of the basis set Ω not only for E_{AB} but also for E_A and E_B . This procedure called the *counter-poise method*, was first introduced by Boys and Bernardi.⁴ Application of the full basis set Ω when calculating E_A results in the wave function of A containing not only its own atomic orbitals, but also the atomic orbitals of the ("absent") partner B, the "*ghost orbitals*" (Fig. 13.3b). As a by-product, the charge density of A exhibits broken symmetry

counter-poise
method

"ghost orbitals"

⁴S.F. Boys, F. Bernardi, *Mol. Phys.* 19 (1970) 553.

with respect to the symmetry of A itself (if any), e.g., the helium atom would have a small dipole moment, etc. Unfortunately, the counter-poise method depends on which channel of dissociation is considered. If several channels are considered at one time, not only are we confronted with an ambiguity, but also this inevitably leads to discontinuities in the calculated energies. This problem is not yet solved in the literature.

13.5.3 GOOD AND BAD NEWS ABOUT THE SUPERMOLECULAR METHOD

Two deficiencies

When performing the subtraction in formula (13.1), we obtain a number representing the interaction energy at a certain distance and orientation of the two sub-systems.

The resulting E_{int} has two disadvantages: it is less precise than E_{AB} , E_A and E_B , and it does not tell us anything about why the particular value is obtained.

The first disadvantage could be compared (following Coulson⁵) to weighing the captain's hat by first weighing the ship with the captain wearing his hat and the ship with the captain without his hat (Fig. 13.4). Formally everything is perfect, but there is a cancellation of significant digits in E_{AB} and $(E_A + E_B)$, that may lead to a very poor interaction energy.

The second deficiency deals with the fact that the interaction energy obtained is just a number and we will have no idea why the number is of such magnitude.⁶ Both deficiencies will be removed in the perturbational approach to intermolecular

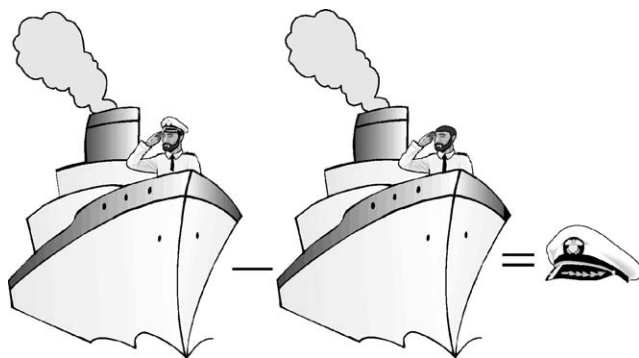


Fig. 13.4. In the supermolecular method we subtract two large numbers that differ only slightly and lose accuracy in this way. It resembles determining the weight of the captain's hat by weighing first the ship with the captain wearing his hat, then repeating the same with the captain without his hat and subtracting the two results. In order not to obtain a result like 240 kg or so, we have to have at our disposal a very accurate method of weighing things.

⁵C.A. Coulson, "Valence", Oxford University Press (1952).

⁶The severity of this can be diminished by analyzing the supermolecular interaction energy expression (using molecular orbitals of A and B) and identifying the physically distinguishable terms by the kind of

interaction. Then, the interaction energy will be calculated directly and we will be able to tell which physical contributions it consists of.

Important advantage

A big advantage of the supermolecular method is its applicability at any intermolecular distance, i.e. independently of how strong the interaction is.

13.6 PERTURBATIONAL APPROACH

13.6.1 INTERMOLECULAR DISTANCE – WHAT DOES IT MEAN?

What is the distance (in kilometers) between the Polish and German populations, or what does the distance between two buses mean? Because of the non-zero dimensions of both objects, it is difficult to tell what the distance could be and any measure of it will be arbitrary. It is the same story with molecules. Up till now we did not need a notion for the intermolecular distance, the positions of the nuclei were sufficient. At the beginning we need only an infinite distance and therefore any definition will be acceptable. Later, however, we will be forced to specify the intermolecular distance (cf. p. 698 and Appendix X on p. 1038). The final numerical values should not depend on this choice, but intermediate results could depend on it. It will turn out that despite the existing arbitrariness, we will prefer those definitions which are based upon the charge barycentre distance or similar.

13.6.2 POLARIZATION APPROXIMATION (TWO MOLECULES)

According to the Rayleigh–Schrödinger perturbation theory (Chapter 5) the unperturbed Hamiltonian $\hat{H}^{(0)}$ is a sum of the isolated molecules' Hamiltonians: $\hat{H}^{(0)} = \hat{H}_A + \hat{H}_B$. Following quantum theory tradition in the present chapter the symbol for the perturbation operator will be changed (when compared to Chapter 5): $\hat{H}^{(1)} \equiv V$.

Despite the fact that we may also formulate the perturbation theory for excited states, we will assume that we are dealing with the ground state (and denote it by subscript “0”). In what is called the *polarization approximation*, the zeroth-order wave function will be taken as a product

$$\psi_0^{(0)} = \psi_{A,0} \psi_{B,0}, \quad (13.4)$$

where $\psi_{A,0}$ and $\psi_{B,0}$ are the exact ground state wave functions for the isolated molecules A and B , respectively, i.e.

$$\hat{H}_A \psi_{A,0} = E_{A,0} \psi_{A,0},$$

$$\hat{H}_B \psi_{B,0} = E_{B,0} \psi_{B,0}.$$

molecular integrals of which they are composed (K. Kitaura, K. Morokuma, *Intern. J. Quantum Chem.* 10 (1976) 325).

We will assume that, because of the large separation of the two molecules, the electrons of molecule *A* are distinguishable from the electrons of molecule *B*. We have to stress the classical flavour of this approximation. Secondly, we assume that the exact wave functions of both isolated molecules:⁷ $\psi_{A,0}$ and $\psi_{B,0}$ are at our disposal.

Of course, function $\psi_0^{(0)}$ is only an approximation to the exact wave function of the total system. Intuition tells us that this approximation is probably very good, because we assume the perturbation is small and the product function $\psi_0^{(0)} = \psi_{A,0} \psi_{B,0}$ is an exact wave function for the *non-interacting* system.

The chosen $\psi_0^{(0)}$ has a wonderful feature, namely it represents an eigenfunction of the $\hat{H}^{(0)}$ operator, as is required by the Rayleigh–Schrödinger perturbation theory (Chapter 5).

The function has also an unpleasant feature: it differs from the exact wave function by symmetry. For example, it is easy to see that

the function $\psi_0^{(0)}$ is *not* antisymmetric with respect to the electron exchanges between molecules, while the exact function has to be antisymmetric with respect to any exchange of electron labels.

This deficiency exists for any intermolecular distance.⁸ We will soon pay a high price for this.

First order effect: electrostatic energy

The first order correction (see eq. (5.22), p. 207)

$$E_0^{(1)} \equiv E_{\text{elst}} \equiv E_{\text{pol}}^{(1)} = \langle \psi_0^{(0)} | V | \psi_0^{(0)} \rangle, \quad (13.5)$$

⁷We will eliminate an additional complication which sometimes may occur. The *n*-th state of the two non-interacting molecules comes, of course, from *some* states of the isolated molecules *A* and *B*. It may happen (most often when the two molecules are identical), that two different sets of the states give the same energy $E_n^{(0)}$, typically, this may happen upon the exchange of excitations of both molecules. Then, $\psi_n^{(0)}$ has to be taken as a linear combination of these two possibilities, which leads to profound changes of the formulae with respect to the usual cases. Such an effect is called the *resonance interaction* (R.S. Mulliken, *Phys. Rev.* 120 (1960) 1674). The resulting interaction decays with the distance as R^{-3} , i.e. quite slowly, making possible an excitation energy transfer through long distances between the interacting molecules. The resonance interaction turns out to be very important (e.g., in biology). An interested reader may find more in the review article J.O. Hirschfelder, W.J. Meath, *Advan. Chem. Phys.* 12 (1967) 3.

resonance
interaction

⁸We may say that the range of the Pauli principle is infinity. If somebody paints some electrons green and others red (distinguishable electrons, we do this in the perturbational method), they are in no man's land, between the classical and quantum worlds. Since the wave function $\psi_0^{(0)}$ does not have the proper symmetry, the corresponding operator $\hat{H}^{(0)} = \hat{H}_A + \hat{H}_B$ is just a mathematical object having little relation to the total system under study.

electrostatic
energy

represents what is called the electrostatic interaction energy (E_{elst}). To stress that E_{elst} is the first-order correction to the energy in the polarization approximation, the quantity will alternatively be denoted by $E_{\text{pol}}^{(1)}$. The electrostatic energy represents the Coulombic interaction of two “frozen” charge distributions corresponding to the isolated molecules A and B , because it is the mean value of the Coulombic interaction energy operator V calculated with the wave function $\psi_0^{(0)}$ being the product of the wave functions of the isolated molecules⁹ $\psi_0^{(0)} = \psi_{A,0}\psi_{B,0}$.

Second-order energy: induction and dispersion energies

The second-order energy (p. 208) in the polarization approximation approach can be expressed in a slightly different way. The n -th state of the total system at long intermolecular distances corresponds to some states n_A and n_B of the individual molecules, i.e.

$$\psi_n^{(0)} = \psi_{A,n_A}\psi_{B,n_B} \quad (13.6)$$

and¹⁰

$$E_n^{(0)} = E_{A,n_A} + E_{B,n_B}. \quad (13.7)$$

Using this assumption, the second-order correction to the ground-state energy (we assume $n = 0$ and $\psi_0^{(0)} = \psi_{A,0}\psi_{B,0}$) can be expressed as (see Chapter 5, p. 208)

$$E_0^{(2)} = \sum_{n_A} \sum'_{n_B} \frac{|\langle \psi_{A,n_A}\psi_{B,n_B} | V \psi_{A,0}\psi_{B,0} \rangle|^2}{(E_{A,0} - E_{A,n_A}) + (E_{B,0} - E_{B,n_B})}, \quad (13.8)$$

where “prime” in the summation means excluding $n = 0$, or $(n_A, n_B) = (0, 0)$. The quantity $E_0^{(2)}$ can be divided in the following way

$$E_0^{(2)} = \sum_{n_A} \sum'_{n_B} \dots = \sum_{(n_A=0, n_B \neq 0)} \dots + \sum_{(n_A \neq 0, n_B=0)} \dots + \sum_{(n_A \neq 0, n_B \neq 0)} \dots \quad (13.9)$$

Let us construct a matrix A (of infinite dimension) composed of the element $A_{00} = 0$ and the other elements calculated from the formula

$$A_{n_A, n_B} = \frac{|\langle \psi_{A,n_A}\psi_{B,n_B} | V \psi_{A,0}\psi_{B,0} \rangle|^2}{(E_{A,0} - E_{A,n_A}) + (E_{B,0} - E_{B,n_B})} \quad (13.10)$$

and divide it into the following parts (I, II, III on the scheme)

⁹We will see later that taking the zero-order wave function with the proper symmetry leads to the first order energy containing what is called the valence repulsion, besides the $E_{\text{pol}}^{(1)}$ term.

¹⁰Also in this case we exclude the resonance interaction.

		$n_A \rightarrow$						
		0	1	2	3	4	5	...
$n_B \downarrow$	0	0	II					
	1							
	2							
	3	I			III			
	4							
	5							
	\vdots							

The quantity $E_0^{(2)}$ is a sum of all the elements of A . This summation will be carried out in three steps. First, the sum of all the elements of column 0 (part I, $n_A = 0$) represents the *induction energy* associated with forcing a change in the charge distribution of the molecule B by the charge distribution of the isolated (“frozen”) molecule A . Second, the sum of all the elements of row 0 (part II, $n_B = 0$) has a similar meaning, but the roles of the molecules are interchanged. Finally, the sum of all the elements of the “interior” of the matrix (part III, n_A and n_B not equal to zero) represents the *dispersion energy*. Therefore,

$$E_0^{(2)} = \underbrace{E_{\text{ind}}(A \rightarrow B)}_{\text{I}} + \underbrace{E_{\text{ind}}(B \rightarrow A)}_{\text{II}} + \underbrace{E_{\text{disp}}}_{\text{III}}, \quad (13.11)$$

where

$$\begin{aligned} E_{\text{ind}}(A \rightarrow B) &= \sum'_{n_B} \frac{|\langle \psi_{A,0} \psi_{B,n_B} | V \psi_{A,0} \psi_{B,0} \rangle|^2}{(E_{B,0} - E_{B,n_B})}, \\ E_{\text{ind}}(B \rightarrow A) &= \sum'_{n_A} \frac{|\langle \psi_{A,n_A} \psi_{B,0} | V \psi_{A,0} \psi_{B,0} \rangle|^2}{(E_{A,0} - E_{A,n_A})}, \\ E_{\text{disp}} &= \sum'_{n_A} \sum'_{n_B} \frac{|\langle \psi_{A,n_A} \psi_{B,n_B} | V \psi_{A,0} \psi_{B,0} \rangle|^2}{(E_{A,0} - E_{A,n_A}) + (E_{B,0} - E_{B,n_B})}. \end{aligned} \quad (13.12)$$

What do these formulae tell us?

One thing has to be made clear. In formula (13.12) we sometimes see arguments for the interacting molecules undergoing excitations. We have to recall however that all the time we are interested in the ground state of the total system, and calculating its energy and wave function. The excited state wave functions appearing in the formulas are the consequence of the fact that the first-order correction to the wave function is expanded in a complete basis set chosen deliberately as $\{\psi_n^{(0)}\}$. If we took another basis set, e.g., the wave functions of another isoelectronic molecule, we would obtain the same numerical results (although formulae (13.12) will not hold), but the argument would be removed. From the mathematical point of view, the very essence of the perturbation theory means a small deformation of the

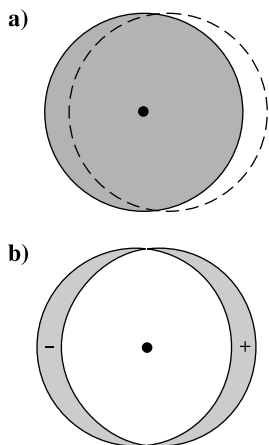


Fig. 13.5. A perturbation of the wave function is a *small* correction. Fig. (a) shows in a schematic way, how a wave function, spherically symmetric with respect to the nucleus, can be transformed into a function that is shifted off the nucleus. The function representing the correction is shown schematically in Fig. (b). Please note the function has symmetry of a *p* orbital.

starting $\psi_0^{(0)}$ function. This tiny deformation is the target of the expansion in the basis set $\{\psi_n^{(0)}\}$. In other words, the perturbation theory involves just a cosmetic adjustment of the $\psi_0^{(0)}$: add a small hump here (Fig. 13.5), subtract a small function there, etc. Therefore, the presence of the excited wave functions in the formulae is not an argument for observing some physical excitations. We may say that the system took what we have prepared for it, and we have prepared excited states.

This does not mean that the energy eigenvalues of the molecule have no influence on its induction or dispersion interactions with other molecules.¹¹ However this is a different story. It has to do with whether the small deformation we have been talking about does or does not depend on the energy eigenvalues spectrum of the individual molecules. The denominators in the expressions for the induction and dispersion energies suggest that the lower excitation energies of the molecules, the larger their deformation, induction and dispersion energies.

13.6.3 INTERMOLECULAR INTERACTIONS: PHYSICAL INTERPRETATION

Now the author would like to recommend the reader to study the multipole expansion concept (Appendix X on p. 1038, also cf. Chapter 12, p. 624).

“intermolecular distance”

The very essence of the multipole expansion is a replacement of the Coulombic interaction of two particles (one from molecule *A*, the other from the molecule *B*) by an infinite sum of interactions of what are called multipoles, where each interaction term has in the denominator an integer power of the distance (called the intermolecular distance *R*) between the origins of the two coordinate systems localized in the individual molecules.

¹¹The smaller the gap between the ground and excited states of the molecule, the larger the polarizability, see Chapter 12.

In other words, multipole expansion describes the intermolecular interaction of two non-spherically symmetric, distant objects by the “interaction” of deviations (multipoles) from spherical symmetry.

To prepare ourselves for the application of the multipole expansion, let us introduce two Cartesian coordinate systems with x and y axes in one system parallel to the corresponding axes in the other system, and with the z axes collinear (see Fig. X.1 on p. 1039). One of the systems is connected to molecule A , the other one to molecule B , and the distance between the origins is R (“intermolecular distance”).¹²

The operator V of the interaction energy of two molecules may be written as

$$V = - \sum_j \sum_a \frac{Z_a}{r_{aj}} - \sum_i \sum_b \frac{Z_b}{r_{bi}} + \sum_{ij} \frac{1}{r_{ij}} + \sum_a \sum_b \frac{Z_a Z_b}{R_{ab}}, \quad (13.13)$$

where we have used the convention that the summations over i and a correspond to all electrons and nuclei of molecule A , and over j and b of molecule B . Since the molecules are assumed to be distant, we have a *practical* guarantee that the interacting particles are distant too. In V many terms with inverse interparticle distance are present. For any such term we may write the corresponding multipole expansion (Appendix X, p. 1039, s is smaller of numbers k and l):

$$\begin{aligned} -\frac{Z_a}{r_{aj}} &= \sum_{k=0} \sum_{l=0} \sum_{m=-s}^{m=s} A_{kl|m|} R^{-(k+l+1)} \hat{M}_A^{(k,m)}(a) * \hat{M}_B^{(l,m)}(j), \\ -\frac{Z_b}{r_{bi}} &= \sum_{k=0} \sum_{l=0} \sum_{m=-s}^{m=s} A_{kl|m|} R^{-(k+l+1)} \hat{M}_A^{(k,m)}(i) * \hat{M}_B^{(l,m)}(b), \\ \frac{1}{r_{ij}} &= \sum_{k=0} \sum_{l=0} \sum_{m=-s}^{m=s} A_{kl|m|} R^{-(k+l+1)} \hat{M}_A^{(k,m)}(i) * \hat{M}_B^{(l,m)}(j), \\ \frac{Z_a Z_b}{R_{ab}} &= \sum_{k=0} \sum_{l=0} \sum_{m=-s}^{m=s} A_{kl|m|} R^{-(k+l+1)} \hat{M}_A^{(k,m)}(a) * \hat{M}_B^{(l,m)}(b), \end{aligned}$$

where

$$A_{kl|m|} = (-1)^{l+m} \frac{(k+l)!}{(k+|m|)!(l+|m|)!}, \quad (13.14)$$

¹²A sufficient condition for the multipole expansion convergence is such a separation of the charge distributions of both molecules, that they could be enclosed in two non-penetrating spheres located at the origins of the two coordinate systems. This condition cannot be fulfilled with molecules, because their electronic charge density distribution extends to infinity. The consequences of this are described in Appendix X. However, the better the sphere condition is fulfilled (by a proper choice of the origins) the more effective in describing the interaction energy are the first terms of the multipole expansion.

The very fact that we use closed sets (like the spheres) in the theory, indicates that in the polarization approximation we are in no man’s land between the quantum and classical worlds.

and the multipole moment $M_C^{(k,m)}(n)$ pertains to particle n and is calculated in "its" coordinate system $C = A, B$. For example,

$$\hat{M}_A^{(k,m)}(a) = Z_a R_a^k P_k^{|m|}(\cos \theta_a) \exp(im\phi_a), \quad (13.15)$$

where R_a, θ_a, ϕ_a are the polar coordinates of nucleus a (with charge Z_a) of molecule A taken in the coordinate system of molecule A . When all such expansions are inserted into the formula for V , we may perform the following chain of transformations

$$\begin{aligned} V &= - \sum_j \sum_a \frac{Z_a}{r_{aj}} - \sum_i \sum_b \frac{Z_b}{r_{bi}} + \sum_{ij} \frac{1}{r_{ij}} + \sum_a \sum_b \frac{Z_a Z_b}{R_{ab}} \\ &\cong \sum_j \sum_a \sum_{k=0} \sum_{l=0} \sum_{m=-s}^{m=s} A_{kl|m|} R^{-(k+l+1)} \hat{M}_A^{(k,m)}(a)^* \hat{M}_B^{(l,m)}(j) \\ &\quad + \sum_i \sum_b \sum_{k=0} \sum_{l=0} \sum_{m=-s}^{m=s} A_{kl|m|} R^{-(k+l+1)} \hat{M}_A^{(k,m)}(i)^* \hat{M}_B^{(l,m)}(b) \\ &\quad + \sum_{ij} \sum_{k=0} \sum_{l=0} \sum_{m=-s}^{m=s} A_{kl|m|} R^{-(k+l+1)} \hat{M}_A^{(k,m)}(i)^* \hat{M}_B^{(l,m)}(j) \\ &\quad + \sum_a \sum_b \sum_{k=0} \sum_{l=0} \sum_{m=-s}^{m=s} A_{kl|m|} R^{-(k+l+1)} \hat{M}_A^{(k,m)}(a)^* \hat{M}_B^{(l,m)}(b) \\ &= \sum_{k=0} \sum_{l=0} \sum_{m=-s}^{m=s} A_{kl|m|} R^{-(k+l+1)} \left\{ \left[\sum_a \hat{M}_A^{(k,m)}(a) \right]^* \left[\sum_j \hat{M}_B^{(l,m)}(j) \right] \right. \\ &\quad + \left[\sum_i \hat{M}_A^{(k,m)}(i) \right]^* \left[\sum_b \hat{M}_B^{(l,m)}(b) \right] + \left[\sum_i \hat{M}_A^{(k,m)}(i) \right]^* \left[\sum_j \hat{M}_B^{(l,m)}(j) \right] \\ &\quad \left. + \left[\sum_a \hat{M}_A^{(k,m)}(a) \right]^* \left[\sum_b \hat{M}_B^{(l,m)}(b) \right] \right\} \\ &= \sum_{k=0} \sum_{l=0} \sum_{m=-s}^{m=s} A_{kl|m|} R^{-(k+l+1)} \left[\sum_a \hat{M}_A^{(k,m)}(a) + \sum_i \hat{M}_A^{(k,m)}(i) \right]^* \\ &\quad \times \left[\sum_b \hat{M}_B^{(l,m)}(b) + \sum_j \hat{M}_B^{(l,m)}(j) \right] \\ &= \sum_{k=0} \sum_{l=0} \sum_{m=-s}^{m=s} A_{kl|m|} R^{-(k+l+1)} \hat{M}_A^{(k,m)*} \hat{M}_B^{(l,m)}. \end{aligned} \quad (13.16)$$

In the square brackets we can recognize the multipole moment operators for the total molecules calculated in “their” coordinate systems

$$\begin{aligned}\hat{M}_A^{(k,m)} &= \sum_a \hat{M}_A^{(k,m)}(a) + \sum_i \hat{M}_A^{(k,m)}(i), \\ \hat{M}_B^{(l,m)} &= \sum_b \hat{M}_B^{(l,m)}(b) + \sum_j \hat{M}_B^{(l,m)}(j).\end{aligned}$$

Eq. (13.16) has the form of a single multipole expansion, but this time the multipole moment operators correspond to *entire molecules*.

Using the table of multipoles (p. 1042), we may easily write down the multipole operators for the individual molecules. The lowest moment is the net charge (monopole) of the molecules

$$\begin{aligned}\hat{M}_A^{(0,0)} &= q_A = (Z_A - n_A), \\ \hat{M}_B^{(0,0)} &= q_B = (Z_B - n_B),\end{aligned}$$

where Z_A is the sum of all the nuclear charges of molecule A , and n_A is its number of electrons (similarly for B). The next moment is $\hat{M}_A^{(1,0)}$, which is a component of the dipole operator equal to

$$\hat{M}_A^{(1,0)} = - \sum_i z_i + \sum_a Z_a z_a, \quad (13.17)$$

where the small letters z denote the z coordinates of the corresponding particles measured in the coordinate system A (the capital Z denotes the nuclear charge). Similarly, we could very easily write other multipole moments and the operator V takes the form (see Appendix X)

$$\begin{aligned}V &= \frac{q_A q_B}{R} - R^{-2}(q_A \hat{\mu}_{Bz} - q_B \hat{\mu}_{Az}) + R^{-3}(\hat{\mu}_{Ax} \hat{\mu}_{Bx} + \hat{\mu}_{Ay} \hat{\mu}_{By} - 2\hat{\mu}_{Az} \hat{\mu}_{Bz}) \\ &\quad + R^{-3}(q_A \hat{Q}_{B,z^2} + q_B \hat{Q}_{A,z^2}) + \dots,\end{aligned}$$

where

$$\begin{aligned}\hat{\mu}_{Ax} &= - \sum_i x_i + \sum_a Z_a x_a, \\ \hat{Q}_{A,z^2} &= - \sum_i \frac{1}{2}(3z_i^2 - r_i^2) + \sum_a Z_a \frac{1}{2}(3z_a^2 - R_a^2)\end{aligned}$$

and symbol A means that all these moments are measured in coordinate system A . The other quantities have similar definitions, and are easy to derive. There is one thing that may bother us, namely that $\hat{\mu}_{Bz}$ and $\hat{\mu}_{Az}$ appear in the charge–dipole interaction terms with opposite signs, so are not on equal footing. The reason is that the two coordinate systems are also not on equal footing, because the z coordinate of the coordinate system A points to B , whereas the opposite is not true (see Appendix X).

13.6.4 ELECTROSTATIC ENERGY IN THE MULTIPOLE REPRESENTATION AND THE PENETRATION ENERGY

Electrostatic energy (p. 693) represents the first-order correction in polarization perturbational theory and is the mean value of V with the product wave function $\psi_0^{(0)} = \psi_{A,0}\psi_{B,0}$. Because we have the multipole representation of V , we may insert it into formula (13.5).

Let us stress, for the sake of clarity, that V is an *operator* that contains the *operators* of the molecular multipole moments, and that the integration is, as usual, carried out over the x, y, z, σ coordinates of all electrons (the nuclei have positions fixed in space according to the Born–Oppenheimer approximation), i.e. over the coordinates of electrons 1, 2, 3, etc. Since in the polarization approximation we know perfectly well which electrons belong to molecule A (“we have painted them green”), and which belong to B (“red”), therefore we perform the integration separately over the electrons of molecule A and those of molecule B . We have a comfortable situation, because every term in V represents a *product* of an operator depending on the coordinates of the electrons belonging to A and of an operator depending on the coordinates of the electrons of molecule B . This (together with the fact that in the integral we have a *product* of $|\psi_{A,0}|^2$ and $|\psi_{B,0}|^2$) results in a product of two integrals: one over the electronic coordinates of A and the other one over the electronic coordinates of B . This is the reason why we like multipoles so much.

Therefore,

the expression for $E_0^{(1)} = E_{\text{elst}}$ *formally* has to be of exactly the same form as the multipole representation of V , the only difference being that in V we have the molecular multipole *operators*, whereas in E_{elst} we have the molecular multipoles themselves as the *mean values* of the corresponding molecular multipole operators in the ground state (the index “0” has been omitted on the right-hand side).

However, the operator V from the formula (13.13) and the operator in the multipole form (13.16) are equivalent only when the multipole form converges. It does so when the interacting objects are non-overlapping, which is *not* the case here. The electronic charge distributions penetrate and this causes a small difference (*penetration energy* E_{penetr}) between the E_{elst} calculated with and without the multipole expansion. The penetration energy vanishes very fast with intermolecular distance R , cf. Appendix R, p. 1009.

$$E_{\text{elst}} = E_{\text{multipol}} + E_{\text{penetr}}, \quad (13.18)$$

where E_{multipol} contains all the terms of the multipole expansion

$$E_{\text{multipol}} = \frac{q_A q_B}{R} - R^{-2}(q_A \mu_{Bz} - q_B \mu_{Az}) \\ + R^{-3}(\mu_{Ax} \mu_{Bx} + \mu_{Ay} \mu_{By} - 2\mu_{Az} \mu_{Bz}) \\ + R^{-3}(q_A Q_{B,z^2} + q_B Q_{A,z^2}) + \dots$$

The molecular multipoles are

$$q_A = \langle \psi_{A,0} | - \sum_i 1 + \sum_a Z_a | \psi_{A,0} \rangle = \left(- \sum_i 1 + \sum_a Z_a \right) \langle \psi_{A,0} | \psi_{A,0} \rangle \\ = \sum_a Z_a - n_A = \text{the same as operator } q_A, \\ \mu_{Ax} = \langle \psi_{A,0} | \hat{\mu}_{Ax} | \psi_{A,0} \rangle = \langle \psi_{A,0} | - \sum_i x_i + \sum_a Z_a x_a | \psi_{A,0} \rangle \\ = \langle \psi_{A,0} | - \sum_i x_i | \psi_{A,0} \rangle + \sum_a Z_a x_a \quad (13.19)$$

and similarly the other multipoles.

Since the multipoles in the formula for E_{multipol} pertain to the isolated molecules, we may say that the electrostatic interaction represents the interaction of the permanent multipoles.

permanent
multipoles

The above multipole expansion also represents a useful source for the expressions for particular multipole–multipole interactions.

Dipole–dipole

Let us take as an example of the important case of the dipole–dipole interaction.

From the above formulae the dipole–dipole interaction reads as

$$E_{\text{dip-dip}} = \frac{1}{R^3}(\mu_{Ax} \mu_{Bx} + \mu_{Ay} \mu_{By} - 2\mu_{Az} \mu_{Bz}).$$

This is a short and easy to memorize formula, and we might be completely satisfied in using it *provided we always remember the particular coordinate system used for its derivation*. This may end up badly one day for those who have a short memory. Therefore, we will write down the same formula in a “waterproof” form.

Taking into account our coordinate system, the vector (pointing the coordinate system origin a from b) is $\mathbf{R} = (0, 0, R)$. Then we can express $E_{\text{dip-dip}}$ in a very useful form *independent of any choice of coordinate system* (cf., e.g., pp. 131, 655):

DIPOLE–DIPOLE INTERACTION:

$$E_{\text{dip-dip}} = \frac{\boldsymbol{\mu}_A \cdot \boldsymbol{\mu}_B}{R^3} - 3 \frac{(\boldsymbol{\mu}_A \cdot \mathbf{R})(\boldsymbol{\mu}_B \cdot \mathbf{R})}{R^5}. \quad (13.20)$$

This form of the dipole–dipole interaction has been used in Chapters 3 and 12.

Is the electrostatic interaction important?

Electrostatic interaction can be attractive or repulsive. For example, in the electrostatic interaction of Na^+ and Cl^- the main role will be played by the charge–

charge interaction, which is negative and therefore represents attraction, while for $\text{Na}^+ \dots \text{Na}^+$ the electrostatic energy will be positive (repulsion). For neutral molecules the electrostatic interaction may depend on their *orientation* to such an extent that the sign may change. This is an exceptional feature peculiar only to electrostatic interaction.

When the distance R is small when compared to size of the interacting subsystems, multipole expansion gives bad results. To overcome this the total charge distribution may be divided into *atomic* segments (Appendix S). Each atom would carry its charge and other multipoles, and the electrostatic energy would be the sum of the atom-atom contributions, any of which would represent a series similar¹³ to $E_0^{(1)}$.

Reality or fantasy?

In principle, this part (about electrostatic interactions) may be considered as completed. I am tempted, however, to enter some “obvious” subjects, which will turn out to lead us far away from the usual track of intermolecular interactions.

Let us consider the Coulomb interaction of two point charges q_1 on molecule A and q_2 on molecule B , both charges separated by distance r

$$E_{\text{elst}} = \frac{q_1 q_2}{r}. \quad (13.21)$$

This is an outstanding formula:

- first of all we have the amazing exponent of the *exact* value -1 ;
- second, change of the charge *sign* does not make any profound changes in the formula, except the change of *sign* of the interaction energy;
- third, the formula is bound to be false (it has to be only an approximation), since instantaneous interaction is assumed, whereas the interaction has to have time to travel between the interacting objects and during that time the objects change their distance (see Chapter 3, p. 131).

From these remarks follow some apparently obvious observations, that E_{elst} is invariant with respect to the following operations:

- II $q'_1 = -q_1, \quad q'_2 = -q_2$ (charge conjugation, Chapter 2, 2.1.8),
- III $q'_1 = q_2, \quad q'_2 = q_1$ (*exchange* of charge positions),
- IV $q'_1 = -q_2, \quad q'_2 = -q_1$ (charge conjugation *and* exchange of charge positions).

These invariance relations, when treated literally and rigorously, are not of particular usefulness in theoretical chemistry. They may, however, open new possibilities when considered as some limiting cases. Chemical reaction mechanisms very often involve the interaction of molecular ions. Suppose we have a particular reaction mechanism. Now, let us make the charge conjugation of all the objects involved in the reaction (this would require the change of matter to antimatter).

¹³A.J. Stone, *Chem. Phys. Lett.* 83 (1981) 233; A.J. Stone, M. Alderton, *Mol. Phys.* 56 (1985) 1047; W.A. Sokalski, R. Poirier, *Chem. Phys. Lett.* 98 (1983) 86; W.A. Sokalski, A. Sawaryn, *J. Chem. Phys.* 87 (1987) 526.

This will preserve the reaction mechanism. We cannot do such changes in chemistry. However, we may think of some *other molecular systems*, which have similar geometry but opposite overall charge pattern (“counter pattern”). The new reaction has a chance to run in a similar direction as before. This concept is parallel to the idea of *Umpolung* functioning in organic chemistry. It seems that nobody has looked, from that point of view, at all known reaction mechanisms.¹⁴

Umpolung

13.6.5 INDUCTION ENERGY IN THE MULTIPOLE REPRESENTATION

The induction energy contribution consists of two parts: $E_{\text{ind}}(A \rightarrow B)$ and $E_{\text{ind}}(B \rightarrow A)$ or, respectively, the polarization energy of molecule B in the electric field of the unperturbed molecule A and *vice versa*.

The goal of the present section is to take apart the induction mechanism by showing its multipole components. If we insert the multipole representation of V into the induction energy $E_{\text{ind}}(A \rightarrow B)$ then

$$\begin{aligned}
 E_{\text{ind}}(A \rightarrow B) &= \sum_{n_B}' \frac{|\langle \psi_{A,0} \psi_{B,n_B} | V \psi_{A,0} \psi_{B,0} \rangle|^2}{E_{B,0} - E_{B,n_B}} \\
 &= \sum_{n_B}' \frac{1}{E_{B,0} - E_{B,n_B}} \{ |R^{-1} q_A \cdot 0 - R^{-2} q_A \langle \psi_{B,n_B} | \hat{\mu}_{Bz} \psi_{B,0} \rangle + R^{-2} \cdot 0 \\
 &\quad + R^{-3} [\mu_{Ax} \langle \psi_{B,n_B} | \hat{\mu}_{Bx} \psi_{B,0} \rangle + \mu_{Ay} \langle \psi_{B,n_B} | \hat{\mu}_{By} \psi_{B,0} \rangle \\
 &\quad - 2\mu_{Az} \langle \psi_{B,n_B} | \hat{\mu}_{Bz} \psi_{B,0} \rangle] + \dots \|^2 \\
 &= \sum_{n_B}' \frac{1}{E_{B,0} - E_{B,n_B}} \{ |-R^{-2} q_A \langle \psi_{B,n_B} | \hat{\mu}_{Bz} \psi_{B,0} \rangle \\
 &\quad + R^{-3} [\mu_{Ax} \langle \psi_{B,n_B} | \hat{\mu}_{Bx} \psi_{B,0} \rangle + \mu_{Ay} \langle \psi_{B,n_B} | \hat{\mu}_{By} \psi_{B,0} \rangle \\
 &\quad - 2\mu_{Az} \langle \psi_{B,n_B} | \hat{\mu}_{Bz} \psi_{B,0} \rangle] + \dots \|^2 \\
 &= -\frac{1}{2} \frac{1}{R^4} q_A^2 \alpha_{B,zz} + \dots,
 \end{aligned}$$

where

- the zeros appearing in the first part of the derivation come from the orthogonality of the eigenstates of the isolated molecule B ,
- symbol “+...” stands for higher powers of R^{-1} ,
- $\alpha_{B,zz}$ represents the zz component of the dipole polarizability tensor of the molecule B , which absorbed the summation over the excited states of B according to definition (12.40).

¹⁴The author is aware of only a single example of such a pair of counter patterns: the Friedel–Crafts reaction and what is called the vicarious nucleophilic substitution discovered by Mieczysław Mąkosza (M. Mąkosza, A. Kwast, *J. Phys. Org. Chem.* 11 (1998) 341).

A molecule in the electric field of another molecule

Note that $\frac{1}{R^4}q_A^2$ represents the square of the electric field intensity $\mathcal{E}_z(A \rightarrow B) = \frac{q_A}{R^2}$ measured on molecule B and created by the net charge of molecule A . Therefore, we have

$$E_{\text{ind}}(A \rightarrow B) = -\frac{1}{2}\alpha_{B,zz}\mathcal{E}_z^2(A \rightarrow B) + \dots$$

according to formula (12.24) describing the molecule in an electric field. For molecule B its partner – molecule A (and *vice versa*...) represents an external world creating the electric field, and molecule B has to behave as described in Chapter 12. The net charge of A created the electric field $\mathcal{E}_z(A \rightarrow B)$ on molecule B , which as a consequence induced on B a dipole moment $\mu_{B,\text{ind}} = \alpha_{B,zz}\mathcal{E}_z(A \rightarrow B)$ according to formula (12.19). This is associated with the interaction energy term $-\frac{1}{2}\alpha_{B,zz}\mathcal{E}_z^2(A \rightarrow B)$, see eq. (12.24), p. 628.

There is however a small problem. Why is the induced moment proportional only to the net charge of molecule A ? This would be absurd. Molecule B does not know anything about multipoles of molecule A , it only knows about the *local* electric field that acts on it and has to react to that field by a suitable polarization. Everything is all right, though. The rest of the problem is in the formula for $E_{\text{ind}}(A \rightarrow B)$. So far we have analyzed the electric field on B coming from the net charge of A , but the other terms of the formula will give contributions to the electric field coming from *all other* multipole moments of A . Then, the response of B will pertain to the total electric field created by “frozen” A on B , as it should be. A similar story can be given for $E_{\text{ind}}(B \rightarrow A)$. This is all we have in the induction energy (second-order perturbation theory). Interaction of the induced multipoles of A and B is a subject of the third-order terms.

13.6.6 DISPERSION ENERGY IN THE MULTIPOLE REPRESENTATION

After inserting V in the multipole representation (p. 701) into the expression for the dispersion energy we obtain

$$\begin{aligned} E_{\text{disp}} &= \sum'_{n_A} \sum'_{n_B} \frac{1}{(E_{A,0} - E_{A,n_A}) + (E_{B,0} - E_{B,n_B})} \\ &\quad \times \left| R^{-1}q_A q_B \cdot 0 \cdot 0 - R^{-2}q_A \cdot 0 \cdot (\mu_{Bz})_{n_B,0} \right. \\ &\quad \left. - R^{-2}q_B \cdot 0 \cdot (\mu_{Az})_{n_A,0} + R^{-3}[(\mu_{Ax})_{n_A,0}(\mu_{Bx})_{n_B,0} + (\mu_{Ay})_{n_A,0}(\mu_{By})_{n_B,0} \right. \\ &\quad \left. - 2(\mu_{Az})_{n_A,0}(\mu_{Bz})_{n_B,0}] + \dots \right|^2 \\ &= \sum'_{n_A} \sum'_{n_B} \left| R^{-3}[(\mu_{Ax})_{n_A,0}(\mu_{Bx})_{n_B,0} + (\mu_{Ay})_{n_A,0}(\mu_{By})_{n_B,0} \right. \\ &\quad \left. - 2(\mu_{Az})_{n_A,0}(\mu_{Bz})_{n_B,0}] + \dots \right|^2 [(E_{A,0} - E_{A,n_A}) + (E_{B,0} - E_{B,n_B})]^{-1} \end{aligned}$$

where $(\mu_{Ax})_{n_A,0} = \langle \psi_{A,n_A} | \hat{\mu}_{Ax} | \psi_{A,0} \rangle$, $(\mu_{Bx})_{n_B,0} = \langle \psi_{B,n_B} | \hat{\mu}_{Bx} | \psi_{B,0} \rangle$ and similarly the other quantities. The zeros in the first part of the equality chain come from the orthogonality of the eigenstates of each of the molecules.

The square in the formula pertains to all terms. The other terms, not shown in the formula, have the powers of R^{-1} higher than R^{-3} .

Hence, if we squared the total expression, the most important term would be the dipole–dipole contribution with the asymptotic R^{-6} distance dependence.

As we can see from formula (13.12), its calculation requires *double* electronic excitations (one on the first, the other one on the second interacting molecules), and these already belong to the correlation effect (cf. Chapter 10, p. 558).

The dispersion interaction is a pure correlation effect and therefore the methods used in a supermolecular approach, that do not take into account the electronic correlation (as for example the Hartree–Fock method) are unable to produce any non-zero dispersion contribution.

Where does this physical effect come from?

Imagine we have two hydrogen atoms, each in its ground state, i.e. $1s$ state, and with a long internuclear distance R . Let us simplify things as much as possible and give only the possibility of two positions for each of the two electrons: one closer to the other proton and the opposite (crosses in Fig. 13.6), the electron–proton distance being $a \ll R$. Let us calculate the instantaneous dipole–dipole interactions for all four possible situations from formula (13.20) assuming the local coordinate systems on the protons (Table 13.1).

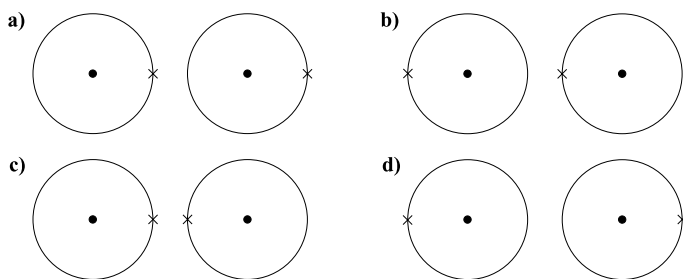


Fig. 13.6. Dispersion energy origin shown schematically for two hydrogen atoms. A popular explanation for the dispersion interaction is that, due to electron repulsion: the situations (a) and (b) occur more often than situation (c) and this is why the dispersion interaction represents a net attraction of dipoles. The positions of the electrons that correspond to (a) and (b) represent two favourable instantaneous dipole – instantaneous dipole interactions, while (c) corresponds to a non-favourable instantaneous dipole – instantaneous dipole interaction. The trouble with this explanation is that there is also the possibility of having electrons far apart as in (d). This most favourable situation (the longest distance between the electrons) means, however, repulsion of the resulting dipoles. It may be shown, though, that the net result (dispersion interaction) is still an attraction (see the text) as it should be.

Table 13.1.

Situation, i	Fig. 13.6	Interaction energy $E_{\text{int}}(i)$
1	a	$-2\frac{\mu^2}{R^3}$
2	b	$-2\frac{\mu^2}{R^3}$
3	c	$+2\frac{\mu^2}{R^3}$
4	d	$+2\frac{\mu^2}{R^3}$

Here $\mu = (0, 0, \pm a)$ for electrons $i = 1, 2$ according to definition (13.17), and $\mu \equiv a$ in a.u. Note that if we assume the same probability for each situation, the net energy would be zero, i.e. $\sum_i E_{\text{int}}(i) = 0$. These situations have, however, different probabilities (p_i), because the electrons repel each other, and the total potential energy depends on where they actually are. *Note, that the probabilities should be different only because of the electron correlation.* In this total energy, there is a common contribution, identical in all the four situations: the interaction within the individual atoms [the remainder is the interaction energy $E_{\text{int}}(i)$]. If we could somehow guess these probabilities p_i , $i = 1, 2, 3, 4$, then we could calculate the mean interaction energy of our model one-dimensional atoms as $\bar{E}_{\text{int}} = \sum_i p_i E_{\text{int}}(i)$. In this way we could see whether it corresponds to net attraction ($\bar{E}_{\text{int}} < 0$) or repulsion ($\bar{E}_{\text{int}} > 0$), which is most interesting for us. Well, but how to calculate them?¹⁵ We may suspect that for the ground state (we are interested in the ground state of our system) the lower the potential energy $V(x)$ the higher the probability density $p(x)$. This is what happens for the harmonic oscillator, for the Morse oscillator, for the hydrogen-like atom, etc. Is there any tip that could help us work out what such a dependence might be? If you do not know where to begin, then think of the harmonic oscillator model as a starting point! This is what people usually do as a first guess. As seen from eq. (4.16), the ground-state wave function for the harmonic oscillator may be written as $\psi_0 = A \exp[-BV(x)]$, where $B > 0$, and $V(x)$ stands for the potential energy for the harmonic oscillator. Therefore the probability density changes as $A^2 \exp[-2BV(x)]$. Interesting... Let us assume that a similar thing happens¹⁶ for the probabilities p_i of finding the electrons 1 and 2 in small cubes of volumes dV_1 and dV_2 , respectively, i.e. they may be reasonably estimated as

$$p_i = N A^2 \exp[-2BE_{\text{int}}(i)] dV_1 dV_2,$$

where E_{int} plays a role of potential energy, and

$$N = 1 / \left(\sum_i A^2 \exp[-2BE_{\text{int}}(i)] dV_1 dV_2 \right)$$

¹⁵In principle we could look at what people have calculated in the most sophisticated calculations for the hydrogen molecule at a large R , and assign the p_i 's as the squares of the wave function value for the corresponding four positions of both electrons. Since these wave functions are awfully complex, we leave this path without regret.

¹⁶This is like having the electron attached to the nucleus by a harmonic spring (instead of Coulombic attraction).

is the normalization constant assuring that in our model $\sum_i p_i = 1$. For long distances R [small $E_{\text{int}}(i)$] we may expand this expression in a Taylor series and obtain

$$\begin{aligned} p_i &= \frac{A^2[1 - 2BE_{\text{int}}(i)]dV_1 dV_2}{\sum_j A^2 \exp[-2BE_{\text{int}}(j)]dV_1 dV_2} \approx \frac{1 - 2BE_{\text{int}}(i) + \dots}{\sum_j (1 - 2BE_{\text{int}}(j) + \dots)} \\ &= \frac{1 - 2BE_{\text{int}}(i) + \dots}{4 - 2B \cdot 0 + \sum_j \frac{1}{2}[2BE_{\text{int}}(j)]^2 + \dots} \approx \frac{1}{4} - \frac{B}{2}E_{\text{int}}(i), \end{aligned}$$

where the Taylor series has been truncated to the accuracy of the linear terms in the interaction. Then, the mean interaction energy

$$\begin{aligned} \bar{E}_{\text{int}} &= \sum_i p_i E_{\text{int}}(i) \approx \sum_i \left[\frac{1}{4} - \frac{B}{2}E_{\text{int}}(i) \right] E_{\text{int}}(i) \\ &= \frac{1}{4} \sum_i E_{\text{int}}(i) - \frac{B}{2} \sum_i [E_{\text{int}}(i)]^2 = 0 - \frac{B}{2} \frac{16\mu^4}{R^6} = -8B \frac{\mu^4}{R^6} < 0. \end{aligned}$$

We may not expect our approximation to be extremely accurate, but it is worth noting that we have grasped two important features of the correct dispersion energy: that it corresponds to attractive interaction and that it vanishes with distance as R^{-6} .

Examples

The electrostatic interaction energy of two molecules can be calculated from formula (13.5). However, it is very important for a chemist to be able to predict the main features of the electrostatic interaction *without any calculation* at all, based on some general rules. This will create chemical intuition or chemical common sense so important in planning, performing and understanding experiments. The data of Table 13.2 were obtained assuming a long intermolecular distance and the molecular orientations as shown in the table.

In composing Table 13.2 some helpful rules have been used:

- *Induction and dispersion energies always represent attraction, except in some special cases when they are zero.* These special cases are obvious, e.g., it is impossible to induce some changes on molecule *B*, if molecule *A* does not have any non-zero permanent multipoles. Also, the dispersion energy is zero if an interacting subsystem has no electrons on it.
- *Electrostatic energy is non-zero, if both interacting molecules have some non-zero permanent multipoles.*
- *Electrostatic energy is negative (positive), if the lowest non-vanishing multipoles of the interacting partners attract (repel) themselves.*¹⁷ How to recognize that a particular multipole–multipole interaction represents attraction or repulsion? First we replace the molecules by their lowest non-zero multipoles represented by

¹⁷This statement is true for sufficiently long distances.

Table 13.2. The table pertains to two molecules in their electronic ground states. For each pair of molecules a short characteristic of their electrostatic, induction and dispersion interactions is given. It consists of the sign of the corresponding interaction type (the minus sign means attraction, the plus sign means repulsion and 0 corresponds to the absence of such an interaction, the penetration terms have been neglected)

System	Electrost.	Induc.	Disper.
He...He	0	0	–
He...H ⁺	0	–	0
He...HCl	0	–	–
H ⁺ ...HCl	+	–	0
HCl...ClH	+	–	–
HCl...HCl	–	–	–
H–H...He	0	–	–
H–H...H–H	+	–	–
$\begin{smallmatrix} \text{H} \\ \\ \text{H} \end{smallmatrix} \dots \text{H} - \text{H}$	–	–	–
$\begin{smallmatrix} \text{H} \\ \diagdown \\ \text{H} \end{smallmatrix} \text{O} \dots \text{H} - \text{O} \diagup \text{H}$	–	–	–
$\begin{smallmatrix} \text{H} \\ \diagdown \\ \text{H} \end{smallmatrix} \text{O} \dots \text{O} \diagup \begin{smallmatrix} \text{H} \\ \\ \text{H} \end{smallmatrix}$	+	–	–

Table 13.3. The exponent m in the asymptotic dependence R^{-m} of the electrostatic (column 2), induction (column 3) and dispersion (column 4) contributions for the systems given in column 1. Zero denotes that the corresponding contribution is equal to zero in the multipole approximation

System	Electrost.	Induc.	Disper.
He...He	0	0	6
He...H ⁺	0	4	0
He...HCl	0	6	6
H ⁺ ...HCl	2	4	0
HCl...ClH	3	6	6
HCl...HCl	3	6	6
H–H...He	0	8	6
H–H...H–H	5	8	6
$\begin{smallmatrix} \text{H} \\ \\ \text{H} \end{smallmatrix} \dots \text{H} - \text{H}$	5	8	6
$\begin{smallmatrix} \text{H} \\ \diagdown \\ \text{H} \end{smallmatrix} \text{O} \dots \text{H} - \text{O} \diagup \text{H}$	3	6	6
$\begin{smallmatrix} \text{H} \\ \diagdown \\ \text{H} \end{smallmatrix} \text{O} \dots \text{O} \diagup \begin{smallmatrix} \text{H} \\ \\ \text{H} \end{smallmatrix}$	3	6	6

point charges, e.g., ions by + or –, dipolar molecules by +–, quadrupoles by +–+, etc. In order to do this we have to know which atoms are electronegative and which electropositive.¹⁸ After doing this we replace the two molecules by the multipoles. If the nearest neighbour charges in the two multipoles are of opposite sign, the multipoles attract each other, otherwise they repel (Fig. 13.7).

¹⁸This is common knowledge in chemistry and is derived from experiments as well as from quantum mechanical calculations. The later provides the partial atomic charges from what is called population analysis (see Appendix S). Despite its non-uniqueness it would satisfy our needs. A unique and elegant method of calculation of atomic partial charges is related to the Bader analysis described on p. 573.

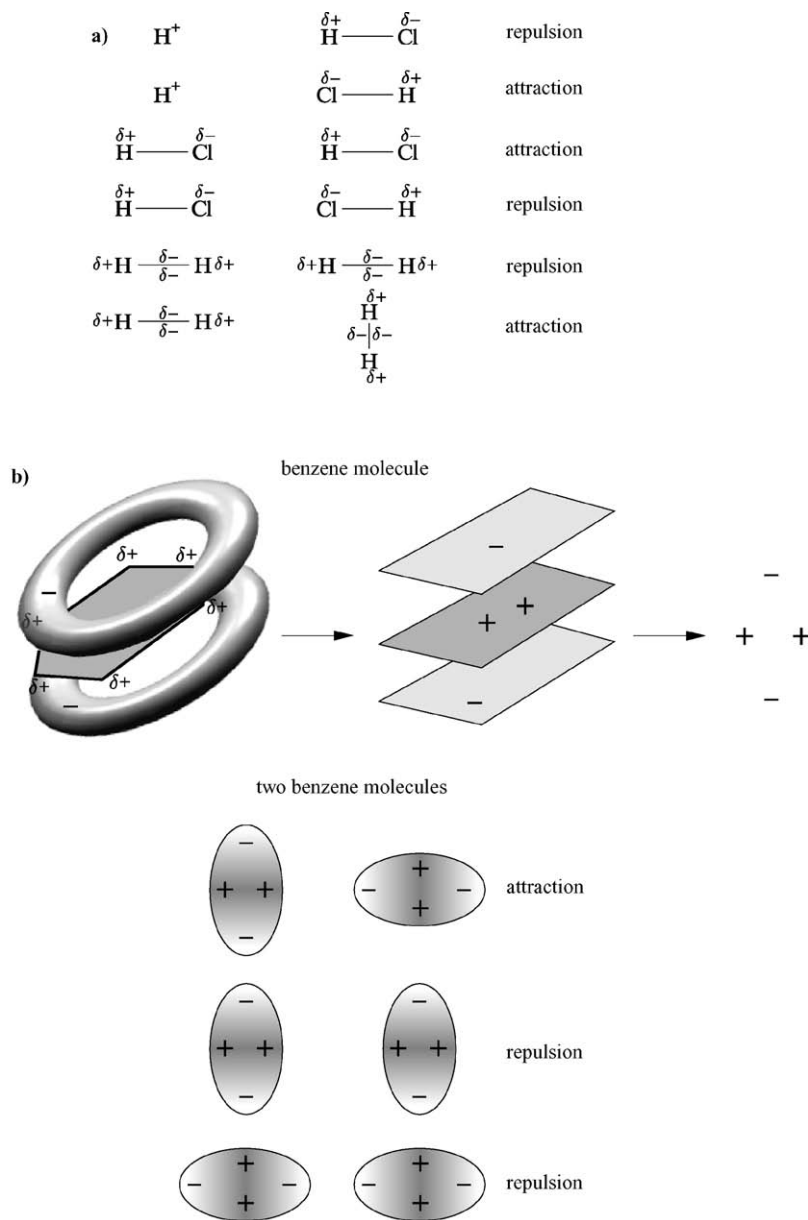


Fig. 13.7. For sufficiently large intermolecular separations the interaction of the lowest non-vanishing multipoles dominates. Whether this is an attraction or repulsion can be recognized by representing the molecular charge distributions by non-point-like multipoles (clusters of point charges). If such multipoles point to each other by point charges of the opposite (same) sign, then the electrostatic interaction of the molecules is attraction (repulsion). (a) A few examples of simple molecules and the atomic partial charges. (b) Even the interaction of the two benzene molecules obeys this rule: in the face-to-face configuration they repel, while they attract each other in the perpendicular configuration.

Since we can establish which effect dominates, its asymptotic dependence (Table 13.3), as the intermolecular distance R tends to ∞ , can be established.

Table 13.3 was composed using a few simple and useful rules:

1. The dispersion energy always decays as R^{-6} .
2. The electrostatic energy vanishes as $R^{-(k+l+1)}$, where the 2^k -pole and 2^l -pole represent the lowest non-vanishing multipoles of the interacting subsystems.
3. The induction energy vanishes as $R^{-2(k+2)}$, where the 2^k -pole is the lower of the two lowest non-zero permanent multipoles of the molecules A and B . The formula is easy to understand if we take into account that the lowest *induced* multipole is always a dipole ($l = 1$), and that the induction effect is of the second order (hence 2 in the exponent).

13.7 SYMMETRY ADAPTED PERTURBATION THEORIES (SAPT)

The SAPT approach is applicable for intermediate intermolecular separations, where the electron clouds of both molecules overlap to such an extent, that

- the polarization approximation, i.e. ignoring the Pauli principle (p. 692), becomes a very poor approximation,
- the multipole expansion becomes invalid.

13.7.1 POLARIZATION APPROXIMATION IS ILLEGAL

First, the polarization approximation zero-order wave function $\psi_{A,0}\psi_{B,0}$ will be deprived of the privilege of being the unperturbed function $\psi_0^{(0)}$ in a perturbation theory. Since it will still play an important role in the theory, let us denote it by $\varphi^{(0)} = \psi_{A,0}\psi_{B,0}$.

The polarization approximation seems to have (at first glimpse) a very strong foundation, because at long intermolecular distances R , the zero-order *energy* is close to the exact one. The trouble is, however, that a similar statement is not true for the zero-order *wave function* $\varphi^{(0)}$ and the exact wave-function at any intermolecular distance (even at infinity).

Let us take an example of two ground-state hydrogen atoms. The polarization approximation zero-order wave function

$$\varphi^{(0)}(1, 2) = 1s_a(1)\alpha(1)1s_b(2)\beta(2), \quad (13.22)$$

where the spin functions have been introduced (the Pauli principle is ignored¹⁹)

This function is neither symmetric (since $\varphi^{(0)}(1, 2) \neq \varphi^{(0)}(2, 1)$), nor anti-symmetric (since $\varphi^{(0)}(1, 2) \neq -\varphi^{(0)}(2, 1)$), and therefore is “illegal” and in principle not acceptable.

¹⁹This is the essence of the polarization approximation.

13.7.2 CONSTRUCTING A SYMMETRY ADAPTED FUNCTION

In the Born–Oppenheimer approximation the electronic ground-state wave function of H_2 has to be the eigenfunction of the nuclear inversion symmetry operator \hat{I} interchanging nuclei a and b (cf. Appendix C). Since $\hat{I}^2 = 1$, the eigenvalues can be either -1 (called u symmetry) or $+1$ (g symmetry).²⁰ The ground-state is of g symmetry, therefore the projection operator $\frac{1}{2}(1 + \hat{I})$ will take care of that (it says: make fifty-fifty combination of a function and its counterpart coming from the exchange of nuclei a and b).²¹ On top of this, the wave function has to fulfil the Pauli exclusion principle, which we will ensure with the antisymmetrizer \hat{A} (cf. p. 986). Altogether the proper symmetry will be assured by projecting $\varphi^{(0)}$ using the projection operator

zero-order wave function

$$\hat{A} = \frac{1}{2}(1 + \hat{I})\hat{A}. \quad (13.23)$$

We obtain as a zero-order approximation to the wave function (N ensures normalization)

$$\begin{aligned} \psi_0^{(0)} &= N \hat{A} \frac{1}{2}(1 + \hat{I})\varphi^{(0)} = \frac{1}{2!} N \frac{1}{2}(1 + \hat{I}) \sum_P (-1)^P \hat{P} [1s_a(1)\alpha(1)1s_b(2)\beta(2)] \\ &= \frac{1}{2} N \frac{1}{2}(1 + \hat{I}) [1s_a(1)\alpha(1)1s_b(2)\beta(2) - 1s_a(2)\alpha(2)1s_b(1)\beta(1)] \\ &= \frac{1}{2} N \frac{1}{2} [1s_a(1)\alpha(1)1s_b(2)\beta(2) - 1s_a(2)\alpha(2)1s_b(1)\beta(1) \\ &\quad + 1s_b(1)\alpha(1)1s_a(2)\beta(2) - 1s_b(2)\alpha(2)1s_a(1)\beta(1)] \\ &= N \frac{1}{2} [1s_a(1)1s_b(2) + 1s_a(2)1s_b(1)] \left\{ \frac{1}{2} [\alpha(1)\beta(2) - \alpha(2)\beta(1)] \right\}. \end{aligned}$$

This is precisely the Heitler–London wave function from p. 521, where its important role in chemistry has been highlighted:

Heitler–London wave function

$$\psi_{\text{HL}} \equiv \psi_0^{(0)} = N [1s_a(1)1s_b(2) + 1s_a(2)1s_b(1)] \left\{ \frac{1}{2} [\alpha(1)\beta(2) - \alpha(2)\beta(1)] \right\}. \quad (13.24)$$

The function is of the same symmetry as the exact solution to the Schrödinger equation (antisymmetric with respect to the exchange of electrons and symmetric with respect to the exchange of protons). It is easy to calculate,²² that normaliza-

²⁰The symbols come from German: g or *gerade* (even) and u or *ungerade* (odd).

²¹We ignore the proton spins.

²²

$$\begin{aligned} \int |\psi_0^{(0)}|^2 d\tau_1 d\tau_2 &= |N|^2 \frac{1}{4} \frac{1}{2} \sum_{\sigma_1} \sum_{\sigma_2} \frac{1}{2} [\alpha(1)\beta(2) - \alpha(2)\beta(1)]^2 [2 + 2S^2] \\ &= |N|^2 \frac{1}{4} (1 + S^2) = 1, \end{aligned} \quad (13.25)$$

tion of $\psi^{(0)}$ means $N = 2[(1 + S^2)]^{-1/2}$, where $S = (1s_a|1s_b)$ stands for the overlap integral of the atomic orbitals $1s_a$ and $1s_b$.

13.7.3 THE PERTURBATION IS ALWAYS LARGE IN POLARIZATION APPROXIMATION

Let us check (Appendix B) how distant are functions $\varphi^{(0)}$ and $\psi^{(0)}$ in the Hilbert space (they are both normalized, i.e. they are unit vectors in the Hilbert space). We will calculate the norm of difference $\varphi^{(0)} - \psi_0^{(0)}$. If the norm were small, then the two functions would be close in the Hilbert space. Let us see:

$$\begin{aligned}
 & \|\varphi^{(0)} - \psi_0^{(0)}\| \\
 &= \left[\int (\varphi^{(0)} - \psi_0^{(0)})^* (\varphi^{(0)} - \psi_0^{(0)}) d\tau \right]^{\frac{1}{2}} = \left[1 + 1 - 2 \int \psi_0^{(0)} \varphi^{(0)} d\tau \right]^{\frac{1}{2}} \\
 &= \left\{ 2 - 2 \int [1s_a(1)\alpha(1)1s_b(2)\beta(2)] N \frac{1}{2} [1s_a(1)1s_b(2) + 1s_a(2)1s_b(1)] \right. \\
 &\quad \times \left. \left\{ \frac{1}{2} [\alpha(1)\beta(2) - \alpha(2)\beta(1)] \right\} d\tau \right\}^{\frac{1}{2}} \\
 &= \left\{ 2 - N \frac{1}{2} \int [1s_a(1)1s_b(2)] [1s_a(1)1s_b(2) + 1s_a(2)1s_b(1)] dv \right\}^{\frac{1}{2}} \\
 &= \left\{ 2 - \frac{1}{\sqrt{1+S^2}} (1+S^2) \right\}^{\frac{1}{2}} = \{2 - \sqrt{1+S^2}\}^{1/2}
 \end{aligned}$$

where we have assumed that the functions are real. When $R \rightarrow \infty$, then $S \rightarrow 0$ and

$$\lim_{R \rightarrow \infty} \|\varphi^{(0)} - \psi_0^{(0)}\| = 1 \neq 0. \quad (13.27)$$

Thus, the Heitler–London wave function differs from $\varphi^{(0)}$, this difference is huge and *does not vanish*, when $R \rightarrow \infty$.

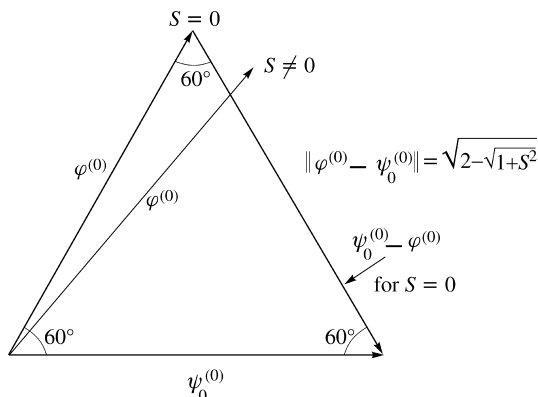
The two normalized functions $\varphi^{(0)}$ and $\psi_0^{(0)}$ represent two unit vectors in the Hilbert space. The scalar product of the two unit vectors $\langle \varphi^{(0)} | \psi_0^{(0)} \rangle$ is equal to $\cos \theta$. Let us calculate this angle θ_{lim} which corresponds to R tending to ∞ . The quantity

$$\begin{aligned}
 \lim_{R \rightarrow \infty} \|\varphi^{(0)} - \psi_0^{(0)}\|^2 &= \lim_{R \rightarrow \infty} \int (\varphi^{(0)} - \psi_0^{(0)})^* (\varphi^{(0)} - \psi_0^{(0)}) d\tau \\
 &= \lim_{R \rightarrow \infty} [2 - 2 \cos \theta] = 1.
 \end{aligned}$$

$$N = \frac{2}{\sqrt{1+S^2}}. \quad (13.26)$$

In a moment we will need function $\psi_0^{(0)}$ with the intermediate normalization with respect to $\varphi^{(0)}$, i.e. satisfying $\langle \psi_0^{(0)} | \varphi^{(0)} \rangle = 1$. Then N will be different and equal to $\langle \varphi^{(0)} | \hat{\mathcal{A}} \varphi^{(0)} \rangle^{-1}$.

Fig. 13.8. The normalized functions $\varphi^{(0)}$ and $\psi_0^{(0)}$ for the hydrogen molecule as unit vectors belonging to the Hilbert space. The functions differ widely at any intermolecular distance R . For $S = 0$, i.e. for long internuclear distances the difference $\psi_0^{(0)} - \varphi^{(0)}$ represents a vector of the Hilbert space having the length 1. Therefore, for $R = \infty$ the three vectors $\varphi^{(0)}$, $\psi_0^{(0)}$ and $\psi_0^{(0)} - \varphi^{(0)}$ form an equilateral angle. For shorter distances the angle between $\varphi^{(0)}$ and $\psi_0^{(0)}$ becomes smaller than 60° .



Hence, $\cos \theta_{\text{lim}} = \frac{1}{2}$, and therefore $\theta_{\text{lim}} = 60^\circ$, see Fig. 13.8. This means that the three unit vectors: $\varphi^{(0)}$, $\psi_0^{(0)}$ and $\varphi^{(0)} - \psi_0^{(0)}$ for $R \rightarrow \infty$ form an equilateral triangle, and therefore, $\varphi^{(0)}$ represents a highly “handicapped” function, which lacks about a half with respect to a function of the proper symmetry.²³ This is certainly bad news.

Therefore, the perturbation V has to be treated as *always large*, because it is responsible for a huge wave function change: from the unperturbed one of bad symmetry to the exact one of the correct symmetry.

In contrast to this, there would be no problem at all with the vanishing of the $\|\psi_0^{(0)} - \psi_0\|$ as $R \rightarrow \infty$, where ψ_0 represents the ground state solution of the Schrödinger equation. Indeed, $\psi_0^{(0)}$ correctly describes the dissociation of the molecule into two hydrogen atoms (both in the 1s state), as well as both functions having the same symmetry for all interatomic distances. Therefore,

the Heitler–London wave function represents a good approximation to the exact function for long (and we hope medium) intermolecular distances. Unfortunately, it is not the eigenfunction of the $\hat{H}^{(0)}$ and therefore we cannot construct the usual Rayleigh–Schrödinger perturbation theory.

And this is the second item of bad news today...

13.7.4 ITERATIVE SCHEME OF THE SYMMETRY ADAPTED PERTURBATION THEORY

We now have two issues: either to construct another zero-order Hamiltonian, for which the $\psi_0^{(0)}$ function would be an eigenfunction (then the perturbation would be small and the Rayleigh–Schrödinger perturbation theory might be applied), or to abandon any Rayleigh–Schrödinger perturbation scheme and replace it by

²³In Appendix Y, p. 1050, we show, how the charge distribution changes when the Pauli exclusion principle is forced by a proper projection of the $\varphi^{(0)}$ wave function.

something else. The first of these possibilities was developed intensively in many laboratories. The approach had the deficiency that the operators appearing in the theories depended explicitly on the basis set used, and therefore there was no guarantee that a basis independent theory exists.

The second possibility relies on an iterative solution of the Schrödinger equation, forcing the proper symmetry of the intermediate functions. The method was proposed mainly by Bogumił Jeziorski and Włodzimierz Kołos.

Claude Bloch was probably the first to write the Schrödinger equation in the form shown in formulae²⁴ (10.76) and (10.59). Let us recall them in a notation adapted to the present situation:

Bloch equations

$$\begin{aligned}\psi_0 &= \varphi^{(0)} + \hat{R}_0(E_0^{(0)} - E_0 + V)\psi_0, \\ E_0 &= E_0^{(0)} + \langle \varphi^{(0)} | V \psi_0 \rangle,\end{aligned}$$

where we assume that $\varphi^{(0)}$ satisfies

$$\hat{H}^{(0)} \varphi^{(0)} = E_0^{(0)} \varphi^{(0)}$$

with the eigenvalues of the unperturbed Hamiltonian $\hat{H}^{(0)} = \hat{H}_A + \hat{H}_B$ given as the sum of the energies of the isolated molecules A and B :

$$E_0^{(0)} = E_{A,0} + E_{B,0},$$

and ψ_0 is the exact ground-state solution to the Schrödinger equation with the total non-relativistic Hamiltonian \hat{H} of the system:

$$\hat{H} \psi_0 = E_0 \psi_0.$$

We focus our attention on the difference \mathcal{E}_0 between E_0 , which is our target and $E_0^{(0)}$, which is at our disposal as the unperturbed energy. We may write the Bloch equations in a form exposing the interaction energy $\mathcal{E}_0 = E_0 - E_0^{(0)}$

$$\begin{aligned}\psi_0 &= \varphi^{(0)} + \hat{R}_0(-\mathcal{E}_0 + V)\psi_0, \\ \mathcal{E}_0 &= \langle \varphi^{(0)} | V \psi_0 \rangle,\end{aligned}$$

the equations are valid for intermediate normalization $\langle \varphi^{(0)} | \psi_0 \rangle = 1$. This system of equations for \mathcal{E}_0 and ψ_0 might be solved by an iterative method:²⁵

ITERATIVE SCHEME:

$$\psi_0(n) = \varphi^{(0)} + \hat{R}_0[-\mathcal{E}(n) + V]\psi_0(n-1), \quad (13.28)$$

$$\mathcal{E}_0(n) = \langle \varphi^{(0)} | V \psi_0(n-1) \rangle, \quad (13.29)$$

where the *iteration* number n is in the parentheses.

²⁴C. Bloch, *Nucl. Phys.* 6 (1958) 329.

²⁵In such a method we have freedom in choosing the starting point – this is one of its most beautiful features.

Polarization scheme replaced

We start in the zeroth iteration with $\psi_0(0) = \varphi^{(0)}$.

When repeating the above iterative scheme and grouping the individual terms according to the powers of V , at each turn we obtain the exact expression appearing in the Rayleigh–Schrödinger polarization approximation (Chapter 5) plus some higher order terms.

It is worth noting that $\mathcal{E}_0(n)$ is the *sum* of corrections of the Rayleigh–Schrödinger *up to the n -th order* with respect to V (*not* the n -th perturbation correction). For large R , the quantity $\mathcal{E}_0(n)$ is an arbitrarily good approximation of the exact interaction energy.

Of course, the rate, at which the iterative procedure converges depends very much on the starting point chosen. From this point of view, the start from $\psi_0(0) = \varphi^{(0)}$ is particularly unfortunate, because the remaining (roughly) 50% of the *wave function* has to be restored by the hard work of the perturbational series (high-order corrections are needed). This will be especially pronounced for long intermolecular distances, where the exchange interaction energy will not be obtained in any finite order.

Murrell–Shaw and Musher–Amos (MS–MA) perturbation theory

A much more promising starting point in eq. (13.28) seems to be $\psi_0(0) = \psi_0^{(0)}$, because the symmetry of the wave function is already correct. For convenience the intermediate normalization is used (see p. 204) $\langle \varphi^{(0)} | \psi_0^{(0)} \rangle = 1$, i.e. $\psi_0^{(0)} = N \hat{\mathcal{A}} \varphi^{(0)}$ with $N = \langle \varphi^{(0)} | \hat{\mathcal{A}} \varphi^{(0)} \rangle^{-1}$. The first iteration of eqs. (13.28) and (13.29) gives the first-order correction to the energy

intermediate
normalization

$$\begin{aligned}\mathcal{E}_0(1) &= N \langle \varphi^{(0)} | V \hat{\mathcal{A}} \varphi^{(0)} \rangle = E_{\text{pol}}^{(1)} + E_{\text{exch}}^{(1)}, \\ E_{\text{pol}}^{(1)} &\equiv E_{\text{elst}} = \langle \varphi^{(0)} | V \varphi^{(0)} \rangle.\end{aligned}$$

We have obtained the electrostatic energy already known plus a correction $E_{\text{exch}}^{(1)}$ which we will discuss in a minute.

The first-iteration wave function will be obtained in the following way. First, we will use the commutation relation $\hat{\mathcal{A}}\hat{H} = \hat{H}\hat{\mathcal{A}}$ or

$$\hat{\mathcal{A}}(\hat{H}^{(0)} + V) = (\hat{H}_0 + V)\hat{\mathcal{A}}. \quad (13.30)$$

Of course

$$\hat{\mathcal{A}}(\hat{H}^{(0)} - E_0^{(0)} + V) = (\hat{H}^{(0)} - E_0^{(0)} + V)\hat{\mathcal{A}}, \quad (13.31)$$

which gives²⁶ $V\hat{\mathcal{A}} - \hat{\mathcal{A}}V = [\hat{\mathcal{A}}, \hat{H}^{(0)} - E_0^{(0)}]$, as well as $(V - \mathcal{E}_1)\hat{\mathcal{A}} = \hat{\mathcal{A}}(V - \mathcal{E}_1) +$

²⁶Let us stress *en passant* that the left-hand side is of the first order in V , while the right-hand side is of the zeroth order. Therefore, in symmetry adapted perturbation theory, the order is not a well defined quantity, its role is taken over by the iteration number.

$[\hat{A}, \hat{H}^0 - E_0^{(0)}]$. Now we are ready to use formula (13.28) with $n = 1$:

$$\begin{aligned}\psi_0(1) &= \varphi^{(0)} + \hat{R}_0(V - \mathcal{E}_0(1))\psi_0^{(0)} = \varphi^{(0)} + N\hat{R}_0(V - \mathcal{E}_0(1))\hat{A}\varphi^{(0)} \\ &= \varphi^{(0)} + N\hat{R}_0\{\hat{A}(V - \mathcal{E}_0(1)) + \hat{A}(\hat{H}^{(0)} - E_0^{(0)}) - (\hat{H}^{(0)} - E_0^{(0)})\hat{A}\}\varphi^{(0)} \\ &= \varphi^{(0)} + N\hat{R}_0\hat{A}(V - \mathcal{E}_0(1))\varphi^{(0)} + N\hat{R}_0\hat{A}(\hat{H}^{(0)} - E_0^{(0)})\varphi^{(0)} \\ &\quad - N\hat{R}_0(\hat{H}^{(0)} - E_0^{(0)})\hat{A}\varphi^{(0)}.\end{aligned}$$

The third term is equal to 0, because $\varphi^{(0)}$ is an eigenfunction of $\hat{H}^{(0)}$ with an eigenvalue $E_0^{(0)}$. The fourth term may be transformed by decomposing $\hat{A}\varphi^{(0)}$ into the vector (in the Hilbert space) parallel to $\varphi^{(0)}$ or $\langle\hat{A}\varphi^{(0)}|\varphi^{(0)}\rangle\varphi^{(0)}$ and the vector orthogonal to $\varphi^{(0)}$, or $(1 - |\varphi^{(0)}\rangle\langle\varphi^{(0)}|)\hat{A}\varphi^{(0)}$. The result of $\hat{R}_0(\hat{H}^{(0)} - E_0^{(0)})$ acting on the first vector is zero (p. 554), while the second vector gives $(1 - |\varphi^{(0)}\rangle\langle\varphi^{(0)}|)\hat{A}\varphi^{(0)}$. This gives as the first iteration ground-state wave function $\psi_0(1)$:

$$\begin{aligned}\psi_0(1) &= \varphi^{(0)} + N\hat{R}_0\hat{A}(V - \mathcal{E}_0(1))\varphi^{(0)} + N\hat{A}\varphi^{(0)} - N\langle\varphi^{(0)}|\hat{A}\varphi^{(0)}\rangle\varphi^{(0)} \\ &= \frac{\hat{A}\varphi^{(0)}}{\langle\varphi^{(0)}|\hat{A}\varphi^{(0)}\rangle} + N\hat{R}_0\hat{A}(V - \mathcal{E}_0(1))\varphi^{(0)} \\ &= \hat{B}\varphi^{(0)} - N\hat{R}_0\hat{A}(\mathcal{E}_0(1) - V)\varphi^{(0)},\end{aligned}$$

where

$$\hat{B}\varphi^{(0)} = \frac{\hat{A}\varphi^{(0)}}{\langle\varphi^{(0)}|\hat{A}\varphi^{(0)}\rangle}. \quad (13.32)$$

After inserting $\psi_0(1)$ into the iterative scheme (13.29) with $n = 2$ we obtain the second-iteration energy

$$\begin{aligned}\mathcal{E}_0(2) &= \langle\varphi^{(0)}|V\psi_0(1)\rangle \\ &= \frac{\langle\varphi^{(0)}|V\hat{A}\varphi^{(0)}\rangle}{\langle\varphi^{(0)}|\hat{A}\varphi^{(0)}\rangle} - N\langle\varphi^{(0)}|V\hat{R}_0\hat{A}[\mathcal{E}_0(1) - V]\varphi^{(0)}\rangle.\end{aligned} \quad (13.33)$$

These equations are identical to the corresponding corrections in perturbation theories derived by Murrell and Shaw²⁷ and by Musher and Amos²⁸ (MS-MA).

13.7.5 SYMMETRY FORCING

Finally, there is good news. It turns out that we may formulate a general iterative scheme which is able to produce various perturbation procedures, known and unknown in the literature. In addition the scheme has been designed by my nearest-neighbour colleagues (Jeziorski and Kołos). This scheme reads as:

²⁷J.N. Murrell, G. Shaw, *J. Chem. Phys.* 46 (1967) 1768.

²⁸J.I. Musher, A.T. Amos, *Phys. Rev.* 164 (1967) 31.

Table 13.4. Symmetry forcing in various perturbation schemes. The operator \hat{B} is defined by: $\hat{B}\chi = \hat{A}\chi / \langle \varphi^{(0)} | \hat{A}\chi \rangle$

Perturbation scheme	$\psi^{(0)}$	$\hat{\mathcal{F}}$	$\hat{\mathcal{G}}$
polarization	$\varphi^{(0)}$	1	1
symmetrized polarization ^a	$\varphi^{(0)}$	1	\hat{B}
MS–MA	$\hat{B}\varphi^{(0)}$	1	1
Jeziorski-Kołos scheme ^b	$\hat{B}\varphi^{(0)}$	\hat{A}	1
EL–HAV ^c	$\hat{B}\varphi^{(0)}$	\hat{A}	\hat{B}

^aB. Jeziorski, K. Szalewicz, G. Chałasiński, *Int. J. Quantum Chem.* 14 (1978) 271; in the expression for the energy in the polarization perturbation theory all corrections to the wave function are first subject to the operator \hat{B} .

^bB. Jeziorski, W. Kołos, *Int. J. Quantum Chem.* 12 (1977) 91.

^cEisenschitz–London and Hirschfelder–van der Avoird perturbation theory: R. Eisenschitz, F. London, *Zeit. Phys.* 60 (1930) 491; J.O. Hirschfelder, *Chem. Phys. Letters* 1 (1967) 363; A. van der Avoird, *J. Chem. Phys.* 47 (1967) 3649.

$$\begin{aligned}\psi_0(n) &= \varphi^{(0)} + \hat{R}_0[-\mathcal{E}_0(n) + V]\hat{\mathcal{F}}\psi_0(n-1), \\ \mathcal{E}_0(n) &= \langle \varphi^{(0)} | V \hat{\mathcal{G}} \psi_0(n-1) \rangle\end{aligned}$$

where in eqs. (13.28) and (13.29) we have inserted operators $\hat{\mathcal{F}}$ and $\hat{\mathcal{G}}$ which have to fulfil the obvious condition

$$\hat{\mathcal{F}}\psi_0 = \hat{\mathcal{G}}\psi_0 = \psi_0, \quad (13.34)$$

where ψ_0 is the solution to the Schrödinger equation.

WHY FORCE THE SYMMETRY?

At the end of the iterative scheme (convergence) the insertion of the operators $\hat{\mathcal{F}}$ and $\hat{\mathcal{G}}$ has no effect at all, but *before that* their presence may be crucial for the numerical convergence. This is the goal of symmetry forcing.

This method of generating perturbation theories has been called by the authors the *symmetry forcing* method in symmetry adapted perturbation theory (SAPT).

Polarization collapse removed

The corrections obtained in SAPT differ from those of the polarization perturbational method. The first-order energy correction is already different.

To show the relation between the results of the two approaches, let us first introduce some new quantities. The first is an idempotent antisymmetrizer

$$\hat{A} = C \hat{A}^A \hat{A}^B (1 + \hat{P}) \quad \text{with } C = \frac{N_A! N_B!}{(N_A + N_B)!},$$

where \hat{A}^A , \hat{A}^B are idempotent antisymmetrizers for molecules A and B , each molecule contributing N_A and N_B electrons. Permutation operator \hat{P} contains all the electron exchanges between molecules A and B :

$$\hat{P} = \hat{P}^{AB} + \hat{P}', \quad \hat{P}^{AB} = - \sum_{i \in A} \sum_{j \in B} \hat{P}_{ij},$$

with \hat{P}^{AB} denoting the single exchanges only, and \hat{P}' the rest of the permutations, i.e. the double, triple, etc. exchanges. Let us stress that $\varphi^{(0)} = \psi_{A,0} \psi_{B,0}$ represents a product of two antisymmetric functions²⁹ and therefore $\hat{A}\varphi^{(0)} = C(1 + \hat{P}^{AB} + \hat{P}')\psi_{A,0}\psi_{B,0}$. Taking into account the operator \hat{P} in $\langle \varphi^{(0)} | V \hat{A}\varphi^{(0)} \rangle$ and $\langle \varphi^{(0)} | \hat{A}\varphi^{(0)} \rangle$ produces (p. 715, $E^{(1)} \equiv \mathcal{E}_0(1)$):

$$E^{(1)} = \frac{\langle \psi_{A,0} \psi_{B,0} | V \psi_{A,0} \psi_{B,0} \rangle + \langle \psi_{A,0} \psi_{B,0} | V \hat{P}^{AB} \psi_{A,0} \psi_{B,0} \rangle + O(S^4)}{1 + \langle \psi_{A,0} \psi_{B,0} | \hat{P}^{AB} \psi_{A,0} \psi_{B,0} \rangle + O(S^4)}, \quad (13.35)$$

where the integrals with \hat{P}^{AB} are of the order³⁰ of S^2 .

exchange
interaction

In the polarization approximation

$$E_{\text{pol}}^{(1)} \equiv E_{\text{elst}} = \langle \varphi^{(0)} | V \varphi^{(0)} \rangle \quad (13.36)$$

while in the symmetry adapted perturbation theory

$$E^{(1)} = \frac{\langle \varphi^{(0)} | V \hat{A}\varphi^{(0)} \rangle}{\langle \varphi^{(0)} | \hat{A}\varphi^{(0)} \rangle}, \quad (13.37)$$

$$E^{(1)} = E_{\text{pol}}^{(1)} + E_{\text{exch}}^{(1)}, \quad (13.38)$$

where *the exchange interaction* in first-order perturbation theory

$$E_{\text{exch}}^{(1)} = \langle \psi_{A,0} \psi_{B,0} | V \hat{P}^{AB} \psi_{A,0} \psi_{B,0} \rangle - \langle \psi_{A,0} \psi_{B,0} | V \psi_{A,0} \psi_{B,0} \rangle \langle \psi_{A,0} \psi_{B,0} | \hat{P}^{AB} \psi_{A,0} \psi_{B,0} \rangle + O(S^4). \quad (13.39)$$

In the most commonly encountered interaction of closed shell molecules the $E_{\text{exch}}^{(1)}$ term represents the *valence repulsion*.

The symbol $O(S^4)$ stands for all the terms that vanish with the fourth power of the overlap integrals or faster. The valence repulsion already appears (besides the

valence
repulsion

²⁹The product itself does not have this symmetry.

³⁰This means that we also take into account such a decay in other than overlap integrals S , e.g., $(1s_a 1s_b | 1s_b 1s_a)$ is of the order S^2 , where $S = (1s_a | 1s_b)$. Thus the criterion is the differential overlap rather than the overlap integral.

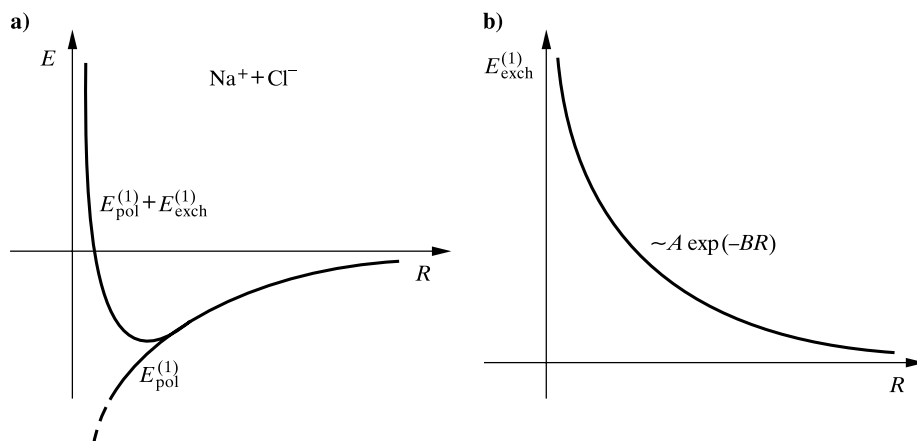


Fig. 13.9. Interaction energy of Na^+ and Cl^- . The polarization approximation gives an absurdity for small separations: the subsystems attract very strongly (mainly because of the electrostatic interaction), while they have had to repel very strongly. The absurdity is removed when the valence repulsion is taken into account (a). Fig. (b) shows the valence repulsion alone modelled by the term $A \exp(-BR)$, where A and B are positive constants.

electrostatic energy $E_{\text{pol}}^{(1)}$ in the first order of the perturbation theory as a result of the Pauli exclusion principle.³¹

We have gained a remarkable thing, which may be seen by taking the example of two interacting subsystems: Na^+ and Cl^- . In the polarization approximation the electrostatic, induction and dispersion contributions to the interaction energy are negative, the total energy will go down and we would soon have a catastrophe: both subsystems would occupy the same place in space and according to the energy calculated (Fig. 13.9) the system would be extremely happy (very low energy). This is absurd.

If this were true, we could not exist. Indeed, sitting safely on a chair we have an equilibrium of the gravitational force and... well, and what? First of all, the force coming from valence repulsion. It is claimed sometimes that quantum effects are peculiar to small objects (electrons, nuclei, atoms, molecules) and are visible only when dealing with such particles. We see, however, that we owe even sitting on a chair to the Pauli exclusion principle (a quantum effect).

The valence repulsion removes the absurdity of the polarization approximation, which made the collapse of the two subsystems possible.

³¹An intriguing idea: the polarization approximation should be an extremely good approximation for the interaction of a molecule with an antimolecule (built from antimatter). Indeed, in the molecule we have electrons, in the antimolecule positrons and no antisymmetrization (between the systems) is needed. Therefore a product wave function should be a very good starting point. No valence repulsion will appear, the two molecules will penetrate like ghosts. Soon after, the tremendous lightning will be seen and the terrible thunder of annihilation will be heard. The system will disappear.

13.7.6 A LINK TO THE VARIATIONAL METHOD – THE HEITLER–LONDON INTERACTION ENERGY

Since the $\hat{\mathcal{A}}\varphi^{(0)}$ wave function is a good approximation of the exact ground state wave function at high values of R , we may calculate what is called the Heitler–London interaction energy ($E_{\text{int}}^{\text{HL}}$) as the mean value of the total (electronic) Hamiltonian minus the energies of the isolated subsystems

$$E_{\text{int}}^{\text{HL}} = \frac{\langle \hat{\mathcal{A}}\varphi^{(0)} | \hat{H} \hat{\mathcal{A}}\varphi^{(0)} \rangle}{\langle \hat{\mathcal{A}}\varphi^{(0)} | \hat{\mathcal{A}}\varphi^{(0)} \rangle} - (E_{A,0} + E_{B,0}).$$

This expression may be transformed in the following way

$$\begin{aligned} E_{\text{int}}^{\text{HL}} &= \frac{\langle \varphi^{(0)} | \hat{H} \hat{\mathcal{A}}\varphi^{(0)} \rangle}{\langle \varphi^{(0)} | \hat{\mathcal{A}}\varphi^{(0)} \rangle} - (E_{A,0} + E_{B,0}) \\ &= \frac{\langle \varphi^{(0)} | \hat{H}^{(0)} \hat{\mathcal{A}}\varphi^{(0)} \rangle + \langle \varphi^{(0)} | V \hat{\mathcal{A}}\varphi^{(0)} \rangle}{\langle \varphi^{(0)} | \hat{\mathcal{A}}\varphi^{(0)} \rangle} - (E_{A,0} + E_{B,0}) \\ &= \frac{(E_{A,0} + E_{B,0}) \langle \varphi^{(0)} | \hat{\mathcal{A}}\varphi^{(0)} \rangle + \langle \varphi^{(0)} | V \hat{\mathcal{A}}\varphi^{(0)} \rangle}{\langle \varphi^{(0)} | \hat{\mathcal{A}}\varphi^{(0)} \rangle} - (E_{A,0} + E_{B,0}) \\ &= \frac{\langle \varphi^{(0)} | V \hat{\mathcal{A}}\varphi^{(0)} \rangle}{\langle \varphi^{(0)} | \hat{\mathcal{A}}\varphi^{(0)} \rangle}. \end{aligned}$$

Therefore, the Heitler–London interaction energy is equal to the first order SAPT energy

$$E_{\text{int}}^{\text{HL}} = E^{(1)}.$$

13.7.7 WHEN WE DO NOT HAVE AT OUR DISPOSAL THE IDEAL $\psi_{A,0}$ AND $\psi_{B,0}$

Up till now we have assumed that the ideal ground-state solutions of the Schrödinger equation for molecules A and B are at our disposal. In practice this will never happen. Instead of $\psi_{A,0}$ and $\psi_{B,0}$ we will have some approximate functions, $\tilde{\psi}_{A,0}$ and $\tilde{\psi}_{B,0}$, respectively. In such a case

$$E_{\text{int}}^{\text{HL}} \neq E^{(1)}.$$

Let us assume that $\tilde{\psi}_{A,0}$ and $\tilde{\psi}_{B,0}$, respectively, represent *Hartree–Fock solutions* for the subsystems A and B . Then the corresponding Heitler–London interaction energy equal to $\tilde{E}_{\text{int}}^{\text{HL}}$ may be written as

$$\tilde{E}_{\text{int}}^{\text{HL}} = \tilde{E}^{(1)} + \Delta_L + \Delta_M,$$

where $\tilde{E}^{(1)}$ is what the old formula gives in the new situation

$$\tilde{E}^{(1)} = \frac{\langle \tilde{\psi}_{A,0} \tilde{\psi}_{B,0} | V \hat{\mathcal{A}} \tilde{\psi}_{A,0} \tilde{\psi}_{B,0} \rangle}{\langle \tilde{\psi}_{A,0} \tilde{\psi}_{B,0} | \hat{\mathcal{A}} \tilde{\psi}_{A,0} \tilde{\psi}_{B,0} \rangle}$$

and Δ_L denotes a correction – called the *Landshoff delta*

Landshoff Δ

$$\Delta_L = \Delta_L^A + \Delta_L^B$$

with the *Landshoff's delta* for individual molecules³²

$$\Delta_L^A = \frac{\langle \tilde{\psi}_{A,0} \tilde{\psi}_{B,0} | \hat{A}(\hat{F}_A - \langle \hat{F}_A \rangle) \tilde{\psi}_{A,0} \tilde{\psi}_{B,0} \rangle}{\langle \tilde{\psi}_{A,0} \tilde{\psi}_{B,0} | \hat{A}(\tilde{\psi}_{A,0} \tilde{\psi}_{B,0}) \rangle}$$

and similar definition for Δ_L^B . The other correction – called the *Murrell delta*³³ – is defined as³⁴

Murrell Δ

$$\Delta_M = \Delta_M^A + \Delta_M^B$$

with

$$\Delta_M^A = \frac{\langle \tilde{\psi}_{A,0} \tilde{\psi}_{B,0} | \hat{A}(\hat{W}_A - \langle \hat{W}_A \rangle) \tilde{\psi}_{A,0} \tilde{\psi}_{B,0} \rangle}{\langle \tilde{\psi}_{A,0} \tilde{\psi}_{B,0} | \hat{A}(\tilde{\psi}_{A,0} \tilde{\psi}_{B,0}) \rangle}$$

where \hat{F}_A and \hat{F}_B are the sums of the Fock operators for molecules A and B , respectively, whereas $\hat{W}_A = \hat{H}_A - \hat{F}_A$ and $\hat{W}_B = \hat{H}_B - \hat{F}_B$ are the corresponding *fluctuation potentials* (see p. 558), i.e. $\hat{H}^{(0)} = \hat{F}_A + \hat{F}_B + \hat{W}_A + \hat{W}_B$. The symbols $\langle \hat{F}_A \rangle$ and $\langle \hat{W}_A \rangle$ denote the mean values of the corresponding operators calculated with the approximate wave functions: $\langle \hat{F}_A \rangle \equiv \langle \tilde{\psi}_{A,0} | \hat{F}_A \tilde{\psi}_{A,0} \rangle$ and $\langle \hat{W}_A \rangle \equiv \langle \tilde{\psi}_{A,0} | \hat{W}_A \tilde{\psi}_{A,0} \rangle$, and similarly for B .

fluctuation
potential

13.8 CONVERGENCE PROBLEMS

In perturbation theories all calculated corrections are simply *added* together. This may lead to partial sums that do not converge. This pertains also to the symmetry adapted perturbation theories. Why? Let us see Table 13.4. One of the perturbational schemes given there, namely that called the symmetrized polarization approximation, is based on the calculation of the wave function exactly as in the polarization approximation scheme, but just before the calculation of the corrections to the energy, the polarization wave function is projected on the antisymmetrized space. This procedure is bound to have trouble. The system changes its charge distribution without paying any attention to the Pauli exclusion principle (thus allow-

³²It has been shown that the Landshoff's deltas Δ_L^A and Δ_L^B vanish for the Hartree–Fock solutions for individual molecules A and B (R. Landshoff, *Zeit. Phys.* 102 (1936) 201). They vanish as well for the SCF solutions (i.e. for finite basis sets) for individual molecules calculated in the basis of *all atomic orbitals of the total system* (B. Jeziorski, M. Bulski, L. Piela, *Intern. J. Quantum Chem.* 10 (1976) 281; M. Gutowski, G. Chalaśiński, J. van Duijneveldt-van de Rijdt, *Intern. J. Quantum Chem.* 26 (1984) 971).

³³J.N. Murrell, A.J.C. Varandas, *Mol. Phys.* 30 (1975) 223.

³⁴It has been shown (B. Jeziorski, M. Bulski, L. Piela, *Intern. J. Quantum Chem.* 10 (1976) 281) that Δ_M^A and Δ_M^B are of the order of $O(S^4)$.

over-
polarization

Pauli blockade

ing it to polarize itself in a non-physical way), while it turns out that it has to fulfil a principle³⁵ (the Pauli principle). This may be described as “*overpolarization*”.

This became evident after a study called the *Pauli blockade*.³⁶ It was shown that, if the Pauli exclusion principle is not obeyed, the electrons of the subsystem *A* can flow, without any penalty and totally unphysically, to the low-energy orbitals of *B*. This may lead to occupation of that orbital by, e.g., four electrons, whereas the Pauli principle admits only a maximum of a double occupation.

Thus, any realistic deformation of the electron clouds has to take into account simultaneously the exchange interaction (valence repulsion), or the Pauli principle. Because of this, we have introduced what is called the *deformation–exchange interaction energy* as

deformation–
exchange
interaction
energy

$$E_{\text{def-exch}} = E^{(2)} - (E_{\text{elst}} + E_{\text{disp}}). \quad (13.40)$$

Padé approximants may improve convergence

Any perturbational correction carries information. Summing up (this is the way we calculate the total effect) these corrections means a certain processing of the information. We may ask an amazing question: *is there any possibility of taking the same corrections and squeezing out more information*³⁷ *than just making the sum*?

In 1892 Henri Padé³⁸ wrote his doctoral dissertation in mathematics and presented some fascinating results.

For a power series

$$f(x) = \sum_{j=0}^{\infty} a_j x^j \quad (13.41)$$

we may define a *Padé approximant* $[L/M]$ as the ratio of two polynomials:

$$[L/M] = \frac{P_L(x)}{Q_M(x)} \quad (13.42)$$

where $P_L(x)$ is a polynomial of at most L -th degree, while $Q_M(x)$ is a polynomial of M -th degree. The coefficients of the polynomials P_L and Q_M will be determined by the following condition

$$f(x) - [L/M] = \text{terms of higher degree than } x^{L+M}. \quad (13.43)$$

In this way it will be guaranteed that for $x = 0$ the Padé approximant $[L, M]$ will have the derivatives up to the $(L + M)$ -th degree identical with those of the original function $f(x)$. In other words,

³⁵This is similar to letting all plants grow as they want and just after harvesting everything selecting the wheat alone.

³⁶M. Gutowski, L. Piela, *Mol. Phys.* 64 (1988) 337.

³⁷That is, a more accurate result.

³⁸H. Padé, *Ann. Sci. Ecole Norm. Sup., Suppl.* [3] 9 (1892) 1.

the first $L + M$ terms of the Taylor expansion for a function $f(x)$ and for its Padé approximant are identical.

Since the nominator and denominator of the approximant can be harmlessly multiplied by any non-zero number, we may set, without losing anything, the following normalization condition

$$Q_M(0) = 1. \quad (13.44)$$

Let us assume also that $P_L(x)$ and $Q_M(x)$ do not have any common factor.

If we now write the polynomials as:

$$\begin{aligned} P_L(x) &= p_0 + p_1x + p_2x^2 + \cdots + p_Lx^L, \\ Q_M(x) &= 1 + q_1x + q_2x^2 + \cdots + q_Mx^M, \end{aligned}$$

then multiplying eq. (13.42) by Q_M and forcing the coefficients at the same powers of x being equal we obtain the following system of equations for the unknowns p_i and q_i (there are $L + M + 1$ of them, the number of equations is the same):

$$\begin{aligned} a_0 &= p_0, \\ a_1 + a_0q_1 &= p_1, \\ a_2 + a_1q_1 + a_0q_2 &= p_2, \\ a_L + a_{L-1}q_1 + \cdots + a_0q_L &= p_L, \\ a_{L+1} + a_Lq_1 + \cdots + a_{L-M+1}q_M &= 0, \\ \vdots & \\ a_{L+M} + a_{L+M-1}q_1 + \cdots + a_Lq_M &= 0. \end{aligned} \quad (13.45)$$

Note please, that the sum of the subscripts in either term is a constant {from the range $[0, L + M]$ }, which is connected to the above mentioned equal powers of x .

Example 1 (Mathematical). The method is best illustrated in action. Let us take a function

$$f(x) = \frac{1}{\sqrt{1-x}}. \quad (13.46)$$

Suppose we have an inspiration to calculate $f(\frac{1}{2})$. We get of course $\sqrt{2} = 1.414213562 \dots$ Let us expand f in a Taylor series:

$$f(x) = 1 + \frac{1}{2}x + \frac{3}{8}x^2 + \frac{5}{16}x^3 + \frac{35}{128}x^4 + \cdots. \quad (13.47)$$

Therefore, $a_0 = 1$; $a_1 = \frac{1}{2}$; $a_2 = \frac{3}{8}$; $a_3 = \frac{5}{16}$; $a_4 = \frac{35}{128}$. Now let us forget that these coefficients came from the Taylor expansion of $f(x)$. Many other functions may have the same *beginning* of the Taylor series. Let us calculate some partial sums of the right-hand side of eq. (13.47):

$f(\frac{1}{2})$	sum up to the n -th term
$n = 1$	1.00000
$n = 2$	1.25000
$n = 3$	1.34375
$n = 4$	1.38281
$n = 5$	1.39990

We see that the Taylor series “works very hard”, it succeeds but not without pain and effort.

Now let us check out how one of the simplest Padé approximants, namely, $[1/1]$, performs the same job. By definition

$$\frac{(p_0 + p_1 x)}{(1 + q_1 x)}. \quad (13.48)$$

Solving (13.45) gives as the approximant:³⁹

$$\frac{(1 - \frac{1}{4}x)}{(1 - \frac{3}{4}x)}. \quad (13.49)$$

Let us stress that information contained in the power series (13.41) has been limited to a_0, a_1, a_2 (all other coefficients have not been used). For $x = \frac{1}{2}$ the Padé approximant has the value

$$\frac{(1 - \frac{1}{4}\frac{1}{2})}{(1 - \frac{3}{4}\frac{1}{2})} = \frac{7}{5} = 1.4, \quad (13.50)$$

which is *more effective* than the painful efforts of the Taylor series that used a coefficients up to a_4 (this gave 1.39990). To be fair, we have to compare the Taylor series result that used only a_0, a_1, a_2 and this gives only 1.34375! Therefore, the approximant failed by 0.01, while the Taylor series failed by 0.07. The Padé approximant $[2/2]$ has the form:

$$[2, 2] = \frac{(1 - \frac{3}{4}x + \frac{1}{16}x^2)}{(1 - \frac{5}{4}x + \frac{5}{16}x^2)}. \quad (13.51)$$

For $x = \frac{1}{2}$ its value is equal to $\frac{41}{29} = 1.414$, which means accuracy of 10^{-4} , while *without Padé approximants, but using the same information contained in the coefficients, we get accuracy two orders of magnitude worse.*

Our procedure did not have the information that the function expanded is $(1 - x)^{-\frac{1}{2}}$, for we gave the first five terms of the Taylor expansion only. Despite this, the procedure determined, with high accuracy, what will give higher terms of the expansion.

³⁹Indeed, $L = M = 1$, and therefore the equations for the coefficients p and q are the following: $p_0 = 1, \frac{1}{2} + q_1 = p_1, \frac{3}{8} + \frac{1}{2}q_1 = 0$. This gives the solution: $p_0 = 1, q_1 = -\frac{3}{4}, p_1 = -\frac{1}{4}$.

Example 2 (Quantum mechanical). This is not the end of the story yet. The reader will see in a minute some things which will be even stranger. Perturbation theory also represents a power series (with respect to λ) with coefficients that are energy corrections. If perturbation is small, the corrections are small as well. In general the higher the perturbation order, the smaller the corrections. As a result, a partial sum of a few low-order corrections, usually gives sufficient accuracy. However, the higher the order the more difficult are the corrections to calculate. Therefore, we may ask if there is any possibility of obtaining good results and at a low price by using the Padé approximants. In Table 13.5 some results of a study by Jeziorski et al. are collected.⁴⁰

For $R = 12.5$ a.u., we see that the approximants had a very difficult task to do. First of all they “recognized” the series limit, only at about $2L + 1 = 17$. Before that, they have been less effective than the original series. It has to be stressed, however, that they “recognized” it *extremely* well (see $2L + 1 = 21$). In contrast to this, the (traditional) partial sums ceased to improve when L increased. This means that either the partial sum series converges to a false limit or it converges to the correct limit, but does it extremely slowly. We see from the variational result (the error is calculated with respect to this) that the convergence *is* false. If the variational result had not been known, we would say that the series has already converged. However, the Padé approximants said: “*no, this is a false convergence*” and they were right.

For $R = 3.0$ a.u. (see Table 13.5) the original series represents a real tragedy. For this distance, the perturbation is too large and the perturbational series just evidently *diverges*. The greater our effort, the greater the error of our result. The error is equal to 13% for $2L + 1 = 17$, then to 22% for $2L + 1 = 19$ and attains 36% for $2L + 1 = 21$. Despite of these hopeless results, it turns out that the problem

Table 13.5. Convergence of the MS–MA (p. 715) perturbational series for the hydrogen atom in the field of a proton (state $2p\sigma_u$) for internuclear distance R (a.u.). The error (in %) is given for the sum of the original perturbational series and for the Padé $[L + 1, L]$ approximant, and is calculated with respect to the variational method (i.e. the best for the basis set used)

$2L + 1$	$R = 12.5$		$R = 3.0$	
	pert. series	$[L + 1, L]$	pert. series	$[L + 1, L]$
3	0.287968	0.321460	0.265189	0.265736
5	0.080973	−0.303293	0.552202	−1.768582
7	0.012785	−0.003388	0.948070	0.184829
9	−0.000596	−0.004147	1.597343	0.003259
11	−0.003351	−0.004090	2.686945	0.002699
13	−0.003932	−0.004088	4.520280	0.000464
15	−0.004056	−0.004210	7.606607	0.000009
17	−0.004084	−0.001779	12.803908	0.000007
19	−0.004090	0.000337	21.558604	−0.000002
21	−0.004092	−0.000003	36.309897	0.000001

⁴⁰B. Jeziorski, K. Szalewicz, M. Jaszuński, *Chem. Phys. Letters* 61 (1979) 391.

represent “peanuts” for the Padé approximants.⁴¹ They were already much better for $L = 3$.

Why are the Padé approximants so effective?

The apparent garbage produced by the perturbational series represented for the Padé approximants precise information that the absurd perturbational corrections pertain the energy of the $2p\sigma_u$ state of the hydrogen atom in the electric field of the proton. How come? Low-order perturbational corrections, even if absolutely crazy, somehow carry information about the physics of the problem. The convergence properties of the Rayleigh–Schrödinger perturbation theory depend critically on the poles of the function approximated (see discussion on p. 210). A pole cannot be described by any power series (as happens in perturbation theories), whereas the Padé approximants have poles built in the very essence of their construction (the denominator as a polynomial). This is why they may fit so well with the nature of the problems under study.⁴²

13.9 NON-ADDITIVITY OF INTERMOLECULAR INTERACTIONS

Interaction energy represents the non-additivity of the total energy

The *total* energy of interacting molecules is not an additive quantity, i.e. does not represent the sum of the energies of the isolated molecules. The reason for this non-additivity is the interaction energy.

Let see, whether the interaction energy itself has some additive properties. First of all the interaction energy requires the declaration of which fragments of the total system we treat as (interacting) molecules (see beginning of this chapter). The only real system is the total system, not these fragments. The fragments or subsystems can be chosen in many ways (Fig. 13.10).

If the theory is exact, the total system can be described at any such choice (cf. p. 492).

A theory has to be invariant with respect to any choice of subsystems in the system under consideration. Such a choice (however in many cases apparently evident) represents an arbitrary operation, similar to the choice of coordinate system.

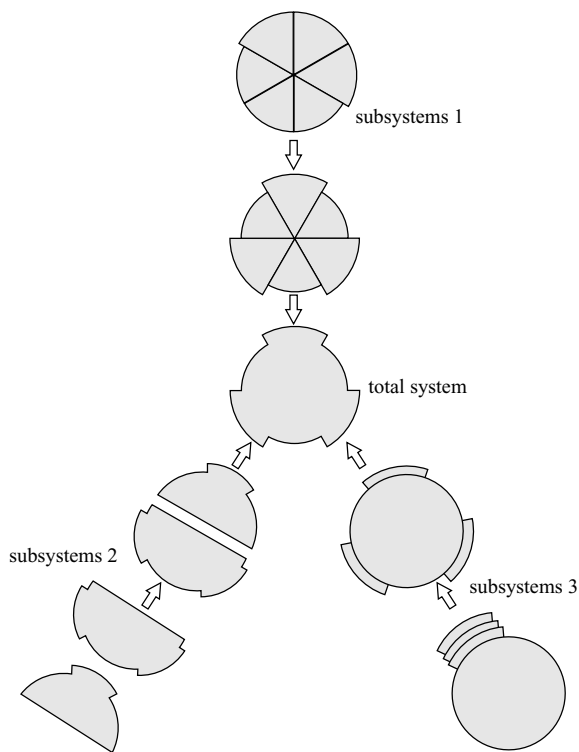
Only the supermolecular theory is invariant with respect to such choices.⁴³ The perturbation theory so far has no such powerful feature (this problem is not even raised in the literature), because it requires the intra and intermolecular interac-

⁴¹Similar findings are reported in T.M. Perrine, R.K. Chaudhuri, K.F. Freed, *Intern. J. Quantum Chem.* 105 (2005) 18.

⁴²There are cases, however, where Padé approximants may fail in a spectacular way.

⁴³However for rather trivial reasons, i.e. interaction energy represents a by-product of the method. The main goal is the total energy, which by definition is independent of the choice of subsystems.

Fig. 13.10. Schematic illustration of arbitrariness behind the selection of subsystems within the total system. The total system under study is in the centre of the figure and can be divided into subsystems in many different ways. The isolated subsystems may differ from those incorporated in the total system (e.g., by shape). Of course, the sum of the energies of the isolated molecules depends on the choice made. The rest of the energy represents the interaction energy and depends on choice too. A correct theory has to be invariant with respect to these choices, which is an extreme condition to fulfil. The problem is even more complex. Using isolated subsystems does not tell us anything about the kind of complex they are going to make. We may imagine several stable aggregates (our system in the centre of the figure is only one of them). In this way we encounter the fundamental and so far unsolved problem of the most stable structure (cf. Chapter 7).



tions to be treated on the same footing. However this is extremely difficult in such a theory, because the assumption that the perturbation is small is inherent to perturbational theories.⁴⁴ Of course, choice of subsystems as with choice of coordinate systems, influences very strongly the mathematical difficulties and therefore the economy of the solution to be reached. Before performing calculations, a scientist already has some intuitive knowledge as to which choice is the most realistic. The intuition is applied when considering different ways in which our system may disintegrate and concentrating on those that require the least energy. The smaller the changes in the subsystems when going from isolated to bound, the smaller the interaction energy and the easier the application of the perturbational theory (cf. p. 685). The smaller the intermolecular distance(s) the more difficult and ambiguous the problem of subsystem choice becomes. In Chapter 9 probably the only example of the invariance of a quantum mechanical method is described.

13.9.1 MANY-BODY EXPANSION OF INTERACTION ENERGY

A next question could be: *is the interaction energy pair-wise additive*, i.e.

is the interaction energy a sum of *pairwise* interactions?

⁴⁴It has to be an infinite order perturbation theory with a large radius of convergence.

If this were true, it would be sufficient to calculate all possible interactions of pairs of molecules in the configuration identical to that of the total system⁴⁵ and our problem would be solved.

For the time being let us take the example of a stone, a shoe and a foot. The owner of the foot will certainly remember the three-body interaction, while nothing special happens when you put a stone into the shoe, or your foot into the shoe, or a small stone on your foot (two-body interactions). The molecules behave like this –

their interactions are not pairwise additive.

In the case of three interacting molecules, there is an effect of a strictly three-body character, which cannot be reduced to any two-body interactions. Similarly for larger numbers of molecules, there is a non-zero four-body effect, because all cannot be calculated as two- and three-body interactions, etc.

In what is called the many-body expansion for N molecules A_1, A_2, \dots, A_N the interaction energy $E_{\text{int}}(A_1 A_2 \dots A_N)$, i.e. the difference between the total energy $E_{A_1 A_2 \dots A_N}$ and the sum of the energies of the isolated molecules $\sum_i E_{A_i}$ can be represented as a series of m -body terms $\Delta E(m, N)$, $m = 2, 3, \dots, N$:

$$E_{\text{int}} = E_{A_1 A_2 \dots A_N} - \sum_{i=1}^N E_{A_i} = \sum_{i>j}^N \Delta E_{A_i A_j}(2, N) + \sum_{i>j>k}^N \Delta E_{A_i A_j A_k}(3, N) + \dots + \Delta E_{A_1 A_2 \dots A_N}(N, N). \quad (13.52)$$

The $\Delta E(m, N)$ contribution to the interaction energy of N molecules ($m \leq N$) represents the sum of the interactions of m molecules (all possible combinations of m molecules among N molecules keeping their configurations fixed as in the total system) inexplicable by the interactions of $m' < m$ molecules.

One more question. Should we optimize the geometry, when calculating the individual many-body terms? In principle, we should not do this, because we are interested in the interaction energy at a given configuration of the nuclei. However, we may present the opposite point of view. For instance, we may be interested in how the geometry of the AB complex changes in the presence of molecule C. This is also a three-body interaction. These dilemmas have not yet been solved in the literature.

Example 3. Four molecules. The many-body expansion concept is easiest to understand by taking an example. Suppose we have four (point-like, for the sake of simplicity) molecules: A, B, C and D lying on a straight line. Their distances (in arbitrary units) are equal to the number of “stars”: $A^*B^{***}C^{**}D$. Let us assume

⁴⁵This would be much less expensive than the calculation for the total system.

Table 13.6.

Three molecules	Interaction energy	Pairwise interactions	Difference
A*B***C	−8	−10	+2
A*B*****D	−5	−6	+1
A****C**D	−7	−8	+1
B***C**D	−9	−10	+1

that the total energy calculated for this configuration equals to -3000 kcal/mol, whereas the sum of the energies of the isolated molecules is -2990 kcal/mol. Hence, the interaction energy of the four molecules is -10 kcal/mol. The negative sign means that the interaction corresponds to attraction, i.e. the system is stable (as far as the binding energy is concerned) with respect to dissociation on $A+B+C+D$. Now we want to analyze the many-body decomposition of this interaction energy. First, we calculate the two-body contribution, let us take all possible pairs of molecules and calculate the corresponding interaction energies (the results are in parentheses, kcal/mol): $A*B$ (-4), $A****C$ (-2), $A*****D$ (-1), $B***C$ (-4), $B*****D$ (-1), $C**D$ (-5). As we can see, the sum of all the pairwise interaction energies is $\Delta E(2, 4) = -17$ kcal/mol. We did not obtain -10 kcal/mol, because *the interactions are not pairwise additive*. Now let us turn to the three-body contribution $\Delta E(3, 4)$. To calculate this we consider all possible three-molecule systems in a configuration identical to that in the total system: $A*B***C$, $A*B*****D$, $A****C**D$, $B***C**D$, and calculate, in each case, the interaction energy of three molecules minus the interaction energies of all pairwise interactions involved. In Table 13.6 we list all the three-body systems possible and in each case give three numbers (in kcal/mol): the interaction energy of the three bodies (with respect to the isolated molecules), the sum of the pairwise interactions and the difference of these two numbers, i.e. the contribution of these three molecules to $\Delta E(3, 4)$.

Hence, the three-body contribution to the interaction energy $\Delta E(3, 4) = 2 + 1 + 1 + 1 = +5$ kcal/mol. The last step in the example is to calculate the four-body contribution. This can be done by subtracting from the interaction energy (-10) the two-body contribution (-17) and the three-body contribution ($+5$). We obtain $\Delta E(4, 4) = -10 + 17 - 5 = 2$ kcal/mol.

We may conclude that in our (fictitious) example, at the given configuration, the many-body expansion of the interaction energy $E_{\text{int}} = -10$ kcal/mol represents a series decaying rather quickly: $\Delta E(2, 4) = -17$ kcal/mol for the two-body, $\Delta E(3, 4) = +5$ for the three-body and $\Delta E(4, 4) = +2$ for the four-body interactions.

Are non-additivities large?

Already a vast experience has been accumulated and some generalizations are possible.⁴⁶ The many-body expansion usually converges faster than in our fictitious example.⁴⁷ For three argon atoms in an equilibrium configuration, the three-body

⁴⁶V. Lotrich, K. Szalewicz, *Phys. Rev. Letters* 79 (1997) 1301.

⁴⁷In quantum chemistry this almost always means a numerical convergence, i.e. a fast decay of individual contributions.

term is of the order of 1%. It should be noted, however, that in the argon *crystal* there is a lot of three-body interactions and the three-body effect increases to about 7%. On the other hand, for liquid water the three-body effect is of the order of 20%, and the higher contributions are about 5%. Three-body effects are sometimes able to determine the crystal structure and have significant influence on the physical properties of the system close to a phase transition (“critical region”).⁴⁸

In the case of the interaction of metal atoms, the non-additivity is much larger than that for the noble gases, and the three-body effects may attain a few tens of percent. This is important information since the force fields widely used in molecular mechanics (see p. 284) are based almost exclusively on effective pairwise interactions (neglecting the three- and more-body contributions).⁴⁹

Although the intermolecular interactions are non-additive, we may ask whether individual contributions to the interaction energy (electrostatic, induction, dispersion, valence repulsion) are additive?

Let us begin from the electrostatic interaction.

13.9.2 ADDITIVITY OF THE ELECTROSTATIC INTERACTION

Suppose we have three molecules A, B, C, intermolecular distances are long and therefore it is possible to use the polarization perturbation theory, in a very similar way to that presented in the case of two molecules (p. 692). In this approach, the unperturbed Hamiltonian $\hat{H}^{(0)}$ represents the sum of the Hamiltonians for the isolated molecules A, B, C. Let us change the abbreviations a little bit to be more concise for the case of three molecules. A *product* function $\psi_{A,n_A}\psi_{B,n_B}\psi_{C,n_C}$ will be denoted by $|n_A n_B n_C\rangle = |n_A\rangle|n_B\rangle|n_C\rangle$, where n_A, n_B, n_C ($= 0, 1, 2, \dots$) stand for the quantum numbers corresponding to the orthonormal wave functions for the molecules A, B, C, respectively. The functions $|n_A n_B n_C\rangle = |n_A\rangle|n_B\rangle|n_C\rangle$ are the eigenfunctions of $\hat{H}^{(0)}$:

$$\hat{H}^{(0)}|n_A n_B n_C\rangle = [E_A(n_A) + E_B(n_B) + E_C(n_C)]|n_A n_B n_C\rangle.$$

The perturbation is equal to $\hat{H} - \hat{H}^{(0)} = V = V_{AB} + V_{BC} + V_{AC}$, where the operators V_{XY} contain all the Coulomb interaction operators involving the nuclei and electrons of molecule X and those of molecule Y.

Let us recall that the electrostatic interaction energy $E_{\text{elst}}(ABC)$ of the ground-state ($n_A = 0, n_B = 0, n_C = 0$) molecules is defined as the first-order correction to the energy in the polarization approximation perturbation theory⁵⁰

⁴⁸R. Bukowski, K. Szalewicz, *J. Chem. Phys.* 114 (2001) 9518.

⁴⁹That is, the effectivity of a force field relies on such a choice of interaction parameters, that the experimental data are reproduced (in such a way the parameters implicitly contain part of the higher-order terms).

⁵⁰The $E_{\text{elst}}(ABC)$ term in symmetry-adapted perturbation theory represents only part of the first-order correction to the energy (the rest being the valence repulsion).

$$E_{\text{pol}}^{(1)} \equiv E_{\text{elst}}(ABC) = \langle 0_A 0_B 0_C | V | 0_A 0_B 0_C \rangle = \langle 0_A 0_B 0_C | V_{AB} + V_{BC} + V_{AC} | 0_A 0_B 0_C \rangle,$$

where the quantum numbers 000 have been supplied (maybe because of my excessive caution) by the redundant and self-explanatory indices ($0_A, 0_B, 0_C$).

The integration in the last formula goes over the coordinates of all electrons. In the polarization approximation, the electrons can be unambiguously divided into three groups: those belonging to A, B and C. Because the zero-order wave function $|0_A 0_B 0_C\rangle$ represents a product $|0_A\rangle|0_B\rangle|0_C\rangle$, the integration over the electron coordinates of one molecule can be easily performed and yields

$$E_{\text{elst}} = \langle 0_A 0_B | V_{AB} | 0_A 0_B \rangle + \langle 0_B 0_C | V_{BC} | 0_B 0_C \rangle + \langle 0_A 0_C | V_{AC} | 0_A 0_C \rangle,$$

where, in the first term, the integration was performed over the electrons of C, in the second over the electrons of B, and in the third over those of A.

Now, let us look at the last formula. We easily see that the individual terms simply represent the electrostatic interaction energies of pairs of molecules: AB, BC and AC, that we would obtain in the perturbational theory (within the polarization approximation) for the interaction of AB, BC and AC, respectively. Conclusion:

the electrostatic interaction is pairwise additive.

13.9.3 EXCHANGE NON-ADDITIVITY

What about the exchange contribution? This contribution *does not exist in the polarization approximation*. It appears only in symmetry-adapted perturbation theory, in pure form in the first-order energy correction and coupled to other effects in higher order energy corrections.⁵¹ The exchange interaction is difficult to interpret, because it appears as a result of the antisymmetry of the wave function (Pauli exclusion principle). The antisymmetry is forced by one of the postulates of quantum mechanics (see Chapter 1) and its immediate consequence is that the probability density of finding two electrons with the same spin and space coordinates is equal to zero.

A CONSEQUENCE OF THE PAULI EXCLUSION PRINCIPLE

In an atom or molecule, the Pauli exclusion principle results in a shell-like electronic structure (electrons with the same spin coordinates hate each other and try to occupy different regions in space). The valence repulsion may be seen as the same effect manifesting itself in the intermolecular interaction. Any attempt to make the molecular charge distributions overlap or occupy the same space (“pushing”) leads to a violent increase in the energy.

⁵¹Such terms are bound to appear. For example, the induction effect is connected to deformation of the electron density distribution. The interaction (electrostatic, exchange, dispersive, etc.) of such a deformed object will change with respect to that of the isolated object. The coupling terms take care of this change.

PAULI DEFORMATION

The Pauli exclusion principle leads to a *deformation* of the wave functions describing the two molecules (by projecting the product-like wave function by the antisymmetrizer \hat{A}) with respect to the product-like wave function. The Pauli deformation (cf. Appendix Y) appears in the zeroth order of perturbation theory, whereas in the polarization approximation, the deformation of the wave function appears in the first order and is not related to the Pauli exclusion principle.

The antisymmetrizer pertains to the permutation symmetry of the wave function with respect to the coordinates of all electrons and therefore is different for a pair of molecules and for a system of three molecules.

The expression for the three-body non-additivity of the valence repulsion⁵² [given by formula (13.39), based on definition (13.37) of the first-order correction in symmetry-adapted perturbation theory⁵³ and from definition (13.52) of the three-body contribution] is:

$$E_{\text{exch},ABC}^{(1)} = N_{ABC} \langle 0_A 0_B 0_C | V_{AB} + V_{BC} + V_{AC} | \hat{A}_{ABC} (0_A 0_B 0_C) \rangle - \sum_{(XY)=(AB),(AC),(BC)} N_{XY} \langle 0_X 0_Y | V_{XY} | \hat{A}_{XY} (0_X 0_Y) \rangle, \quad (13.53)$$

where $N_{ABC} \hat{A}_{ABC} | 0_A 0_B 0_C \rangle$ and $N_{AB} \hat{A}_{AB} | 0_A 0_B \rangle$, and so forth represent the normalized (N_{ABC} etc. are the normalization coefficients) antisymmetrized product-like wave function of the systems ABC, AB, etc. The antisymmetrizer \hat{A}_{ABC} pertains to subsystems A, B, C , similarly \hat{A}_{AB} pertains to A and B , etc., all antisymmetrizers containing only the intersystem electron exchanges and the summation goes over all pairs of molecules.

There is no chance of proving that the exchange interaction is additive, i.e. that eq. (13.53) gives 0. Let us consider the simplest possible example: each molecule has only a single electron: $| 0_A(1) 0_B(2) 0_C(3) \rangle$. The operator \hat{A}_{ABC} (see p. 986) makes (besides other terms) the following permutation:

$$\hat{A}_{ABC} | 0_A(1) 0_B(2) 0_C(3) \rangle = \dots - \frac{1}{(N_A + N_B + N_C)!} | 0_A(3) 0_B(2) 0_C(1) \rangle + \dots,$$

which according to eq. (13.53) leads to the integral

$$\begin{aligned} & - \frac{1}{(N_A + N_B + N_C)!} N_{ABC} \langle 0_A(1) 0_B(2) 0_C(3) | \frac{1}{r_{12}} | 0_A(3) 0_B(2) 0_C(1) \rangle \\ & = - \frac{1}{(N_A + N_B + N_C)!} N_{ABC} \langle 0_A(1) 0_B(2) | \frac{1}{r_{12}} | 0_B(2) 0_C(1) \rangle \langle 0_C(3) | 0_A(3) \rangle \end{aligned}$$

⁵²B. Jeziorski, M. Bulski, L. Piela, *Intern. J. Quantum Chem.* 10 (1976) 281.

⁵³Because, as we have already proved, the rest, i.e. the electrostatic energy, is an additive quantity.

involving the wave functions centred on A , B and C . This means that the term belongs to the *three-body* effect.

The permutation operators of which the \hat{A}_{ABC} operator is composed, correspond to the identity permutation⁵⁴ as well as to the exchange of one or more electrons between the interacting subsystems: $\hat{A}_{ABC} = 1 + \text{single exchanges} + \text{double exchanges} + \dots$.

It is easy to demonstrate,⁵⁵ that

the larger the number of electrons exchanged, the less important such exchanges are, because the resulting contributions would be proportional to higher and higher powers of the (as a rule small) overlap integrals (S).

Single Exchange (SE) Mechanism

The smallest non-zero number of electron exchanges in \hat{A}_{ABC} is equal to 1 (two electrons involved). Such an exchange may only take place between two molecules, say, AB.⁵⁶ This results in terms of the order of S^2 in the three-body expression. The third molecule *does not participate in the electron exchanges, but is not just a spectator in the interaction* (Fig. 13.11.a,b,c). If it were, the interaction would not be three-body.

SE MECHANISM

Molecule C interacts electrostatically with the *Pauli deformation of molecules A and B* (i.e. with the *multipoles that represent the deformation*). Such a mixed interaction is called the SE mechanism.

It would be nice to have a simple formula which could replace the tedious calculations involving the above equations. The three-body energy may be approximated⁵⁷ by the product of the exponential term $b \exp(-aR_{AB})$ and the electric field produced by C, calculated, e.g., in the middle of the distance R_{AB} between molecules A and B. The goal of the exponential terms is to grasp the idea that the overlap integrals (and their squares) vanish exponentially with distance. The exponent a should depend on molecules A and B as well as on their mutual orientation and reflects the hardness of both molecules. These kind of model formulae have low scientific value but are of practical use.

⁵⁴The operator reproduces the polarization approximation expressions in SAPT.

⁵⁵ • First, we write down the exact expression for the first-order exchange non-additivity.

• Then, we expand the expression in the Taylor series with respect to those terms that arise from all electron exchanges except the identity permutation.

• Next, we see that the exchange non-additivity expression contains terms of the order of S^2 and higher, where S stands for the overlap integrals between the orbitals of the interacting molecules.

S decays very fast (exponentially), when intermolecular distance increases.

⁵⁶After that we have to consider AC and BC.

⁵⁷Three-body effects are difficult to calculate. Researchers would like to understand the main mechanism and then capture it by designing a simple approximate formula ready to use in complex situations.

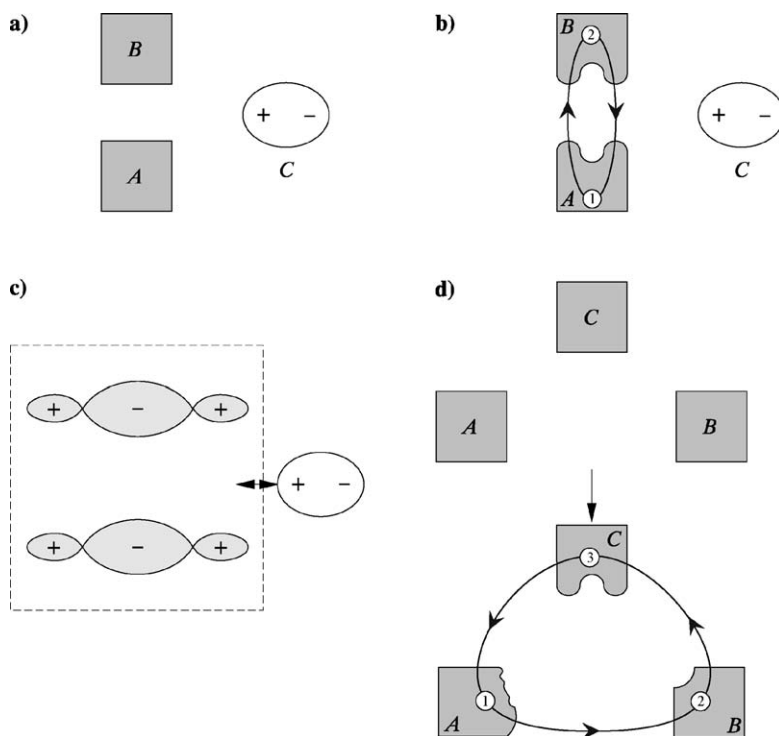


Fig. 13.11. A scheme of the SE and TE exchange non-additivities. Figs. (a), (b), (c) show the single exchange mechanism (SE). (a) Three non-interacting molecules (schematic representation of electron densities). (b) Pauli deformation of molecules A and B. (c) Electrostatic interaction of the Pauli deformation (resulting from exchange of electrons 1 and 2 between A and B) with the dipole moment of C. (d) The TE mechanism: molecules A and B exchange an electron with the mediation of molecule C.

When the double electron exchanges are switched on, we would obtain part of the three-body effect of the order of S^4 . Since S is usually of the order of 10^{-2} , this contribution is expected to be small, although caution is advised, because the *number* of such terms is much more important.

Triple Exchange (TE) Mechanism

Is there any contribution of the order of S^3 ? Yes. The antisymmetrizer \hat{A}_{ABC} is able to make the single electron exchange between, e.g., A and B, but by mediation of C. The situation is schematically depicted in Fig. 13.11.d.

TE MECHANISM

This effect is sometimes modelled as a product of three exponential functions: $\text{const} \exp(-a_{AB}R_{AB}) \exp(-a_{BC}R_{BC}) \exp(-a_{AC}R_{AC})$ and is misleadingly called a triple electron exchange. A molecule is involved in a single exchange with another molecule by mediation of a third.

Let us imagine that molecule B is very long and the configuration corresponds to: A B C. When C is far from A, the three-body effect is extremely small, because almost everything in the interaction is of the two-body character, Appendix Y. If molecule C approaches A and has some non-zero low-order multipoles, e.g., a charge, then it may interact by the SE mechanism even from a far. The TE mechanism operates only at short intermolecular distances.

The exchange interaction is non-additive, but the effects pertain to the contact region of both molecules.⁵⁸ The Pauli exclusion principle does not have any finite range in space, i.e. after being introduced it has serious implications for the wave function even at infinite intermolecular distance (cf. p. 712). Despite this, it always leads to the differential overlap of atomic orbitals (as in overlap or exchange integrals), which decays exponentially with increasing intermolecular distance (the SE mechanism has a partly long-range character).

13.9.4 INDUCTION ENERGY NON-ADDITIVITY

The non-additivity of the intermolecular interaction results mainly from the non-additivity of the induction contribution.

How do we convince ourselves about the non-additivity? This is very easy. It will be sufficient to write the expression for the induction energy for the case of three molecules and to see whether it gives the sum of the pairwise induction interactions. Before we do this, let us write the formula for the total second order energy correction (similar to the case of two molecules on p. 694):

$$E^{(2)}(ABC) = \sum'_{n_A, n_B, n_C} \frac{|\langle n_A n_B n_C | V | 0_A 0_B 0_C \rangle|^2}{[E_A(0_A) - E_A(n_A)] + [E_B(0_B) - E_B(n_B)] + [E_C(0_C) - E_C(n_C)]}. \quad (13.54)$$

According to perturbation theory, the term with all the indices equal to zero has to be omitted in the above expression. It is much better like this, because otherwise the denominator would “explode”. The terms with all non-zero indices are equal to zero. Indeed, let us recall that V is the sum of the Coulomb potentials corresponding to all three *pairs* of the three molecules. This is the reason why it is easy to perform the integration over the electron coordinates of the third molecule (not involved in the pair). A similar operation was performed for the electrostatic interaction. This time, however, the integration makes the term equal to zero, because of the orthogonality of the ground and excited states of the third molecule. All this leads to the conclusion that to have a non-zero term in the summation, among the three indices, *one or two of them have to be of zero value*. Let us perform the summation in two stages: all the terms with only-two-zeros (or a single

⁵⁸See Appendix Y.

non-zero) indices will make a contribution to $E_{\text{ind}}(ABC)$, while all the terms with only-one-zero (or, two non-zero) indices will sum to $E_{\text{disp}}(ABC)$:

$$E^{(2)}(ABC) = E_{\text{ind}}(ABC) + E_{\text{disp}}(ABC), \quad (13.55)$$

where the first term represents the *induction energy*:

$$E_{\text{ind}}(ABC) = E_{\text{ind}}(AB \rightarrow C) + E_{\text{ind}}(AC \rightarrow B) + E_{\text{ind}}(BC \rightarrow A),$$

where

$$E_{\text{ind}}(BC \rightarrow A) \equiv \sum_{n_A \neq 0} \frac{|\langle n_A 0_B 0_C | V | 0_A 0_B 0_C \rangle|^2}{[E_A(0_A) - E_A(n_A)]}$$

means that the “frozen” molecules B and C acting together polarize molecule A , etc. The second term in (13.55) represents the *dispersion energy* (this will be considered later on, see p. 740).

For the time being let us consider the induction energy $E_{\text{ind}}(ABC)$. Writing V as the sum of the Coulomb interactions of the pairs of molecules we have

$$\begin{aligned} E_{\text{ind}}(BC \rightarrow A) &= \sum_{n_A \neq 0} \langle n_A 0_B 0_C | V_{AB} + V_{BC} + V_{AC} | 0_A 0_B 0_C \rangle \langle 0_A 0_B 0_C | V_{AB} + V_{BC} + V_{AC} | n_A 0_B 0_C \rangle \\ &\quad \times [E_A(0_A) - E_A(n_A)]^{-1} \\ &= \sum_{n_A \neq 0} [\langle n_A 0_B | V_{AB} | 0_A 0_B \rangle + \langle n_A 0_C | V_{AC} | 0_A 0_C \rangle] \\ &\quad \times [\langle 0_A 0_B | V_{AB} | n_A 0_B \rangle + \langle 0_A 0_C | V_{AC} | n_A 0_C \rangle] \\ &\quad \times [E_A(0_A) - E_A(n_A)]^{-1}. \end{aligned}$$

Look at the product in the nominator. The induction non-additivity arises just because of this product. If the product (being the square of the absolute value of $\langle n_A 0_B | V_{AB} | 0_A 0_B \rangle + \langle n_A 0_C | V_{AC} | 0_A 0_C \rangle$) were equal to the square of the absolute values of the first and second component, the total expression shown explicitly would be equal to the induction energy corresponding to the polarization of A by the frozen charge distribution of B plus a similar term corresponding to the polarization of A by C , i.e. the polarization occurring *separately*. Together with the other terms in $E_{\text{ind}}(AB \rightarrow C) + E_{\text{ind}}(AC \rightarrow B)$ we would obtain the *additivity* of the induction energy $E_{\text{ind}}(ABC)$. However, besides the sum of squares we also have the mixed terms. They will produce the non-additivity of the induction energy:

$$E_{\text{ind}}(ABC) = E_{\text{ind}}(AB) + E_{\text{ind}}(BC) + E_{\text{ind}}(AC) + \Delta_{\text{ind}}(ABC). \quad (13.56)$$

Thus, we obtain the following expression for the *induction non-additivity* $\Delta_{\text{ind}}(ABC)$:

$$\Delta_{\text{ind}}(ABC) = 2 \operatorname{Re} \sum_{n_A \neq 0} \frac{\langle n_A 0_B | V_{AB} | 0_A 0_B \rangle \langle n_A 0_C | V_{AC} | 0_A 0_C \rangle}{[E_A(0_A) - E_A(n_A)]} + \dots, \quad (13.57)$$

where “ $+\dots$ ” stands for the non-additivities of $E_{\text{ind}}(AB \rightarrow C) + E_{\text{ind}}(AC \rightarrow B)$.

Example 4. Induction non-additivity. To show that the induction interaction of two molecules depends on the presence of the third molecule let us consider the system shown in Fig. 13.12.

Let molecule B be placed half-way between A^+ and C^+ , thus the configuration of the system is: $A^+ \dots\dots B \dots\dots C^+$ with long distances between the subsystems. In such a situation, the total interaction energy is practically represented by the induction contribution plus the constant electrostatic repulsion of A^+ and C^+ . Is the three-body term (induction non-additivity) large? We will show in a minute that this term is large and positive (destabilizing). Since the electric field intensities nearly cancel within molecule B, then despite the high polarizability of the latter, the induction energy will be small. On the contrary, the opposite is true when considering two-body interaction energies. Indeed, A^+ polarizes B very strongly, C^+ does the same, resulting in high stabilization due to high two-body induction energy. Since the total effect is nearly zero, the induction non-additivity is bound to be a large positive number.⁵⁹

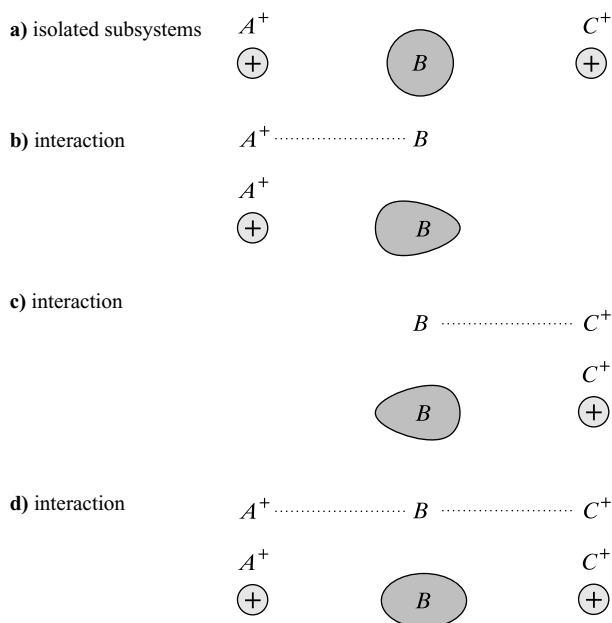


Fig. 13.12. The induction interaction may produce a large non-additivity. (a) Two distant non-polarizable cations: A^+ , C^+ and a small, polarizable neutral molecule B placed exactly in the middle between AC. (b) The two-body induction interaction A^+B , a strong polarization. (c) The two-body induction interaction BC^+ , a strong polarization. (d) The two cations polarize molecule B. Their electric field vectors cancel each other in the middle of B and give a small electric field intensity within B (a weak polarization).

⁵⁹If the intermolecular distances were small, B were not in the middle of AC or molecule B were of large spatial dimension, the strength of our conclusion would diminish.

Self-consistency and polarization catastrophe

The second-order induction effects pertain to polarization by the charge distributions corresponding to the isolated molecules. However, the induced multipoles introduce a change in the electric field and in this way contribute to further changes in charge distribution. These effects already belong to the third⁶⁰ and higher orders of perturbation theory.

It is therefore evident that a two-body interaction model cannot manage the induction interaction energy. This is because we have to ensure that any subsystem, e.g., A, should experience polarization in an electric field, which is the *vector sum* of the electric fields from all its partner subsystems (B, C, ...) calculated at the position of A. The calculated induced dipole moment of A (we focus on the lowest multipole) creates the electric field that produces some changes in the dipole moments of B, C, ..., which in turn change the electric field acting on all the molecules, including A. The circle closes and the polarization procedure has to be performed till self-consistency is reached. This can often be done, although such a simplified interaction model does not allow for geometry optimization, which may lead to a *polarization catastrophe* ending up with induction energy equal to $-\infty$ (due to excessive approach and lack of the Pauli blockade described on p. 722).

polarization
catastrophe

Three-body polarization amplifier

After recognizing that self-consistency may be achieved safely within a variational method (with the Pauli exclusion principle satisfied), this may end the story. However, it would be instructive to get a feeling for the polarization machinery. Let us pose a simple question: is it possible that the polarization of molecule B by molecule A, both separated by distance R , is amplified by the presence of molecule C which is unable by itself to polarize B? An interesting problem. If it appears that the mediator C might increase the polarization of A or B, C would play the role of an amplifier based on three-body induction non-additivity.

Suppose A is represented by a non-polarizable cation A^+ , molecule B separated from A by R , is medium-polarizable, and a strongly polarizable molecule C (to-be-amplifier) enters between A^+ and B: $A^+ \dots C \dots B$, Fig. 13.13.a. How do we measure the polarization of B due to the presence of C? We might propose the electric dipole moment induced on B (of course, there will be nothing induced on A^+). Let R_{BC} be the BC separation. A simple calculation of the electric field intensities gives the following result⁶¹ for the ratio of the induced dipole moments on B with and without the presence of C:

⁶⁰Each of the induced multipoles is proportional to V , their interaction introduces another V ; altogether this gives a term proportional to VVV , i.e. indeed of the third-order.

⁶¹Let us denote the distance AB by R . The axis x is directed from A to B, C is between A and B. If C possessed dipole moment μ_C , then the unit positive charge $+1$ on B would feel the potential (the charge-dipole term from the multipole expansion): $\frac{\mu_C}{R_{BC}^2}$. The corresponding electric field

$$\mathcal{E}_{C \rightarrow B} = -\frac{\partial}{\partial R_{BC}} \left(\frac{\mu_C}{R_{BC}^2} \right) = 2 \frac{\mu_C}{R_{BC}^3}.$$

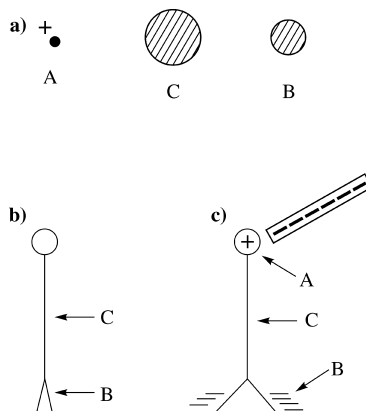


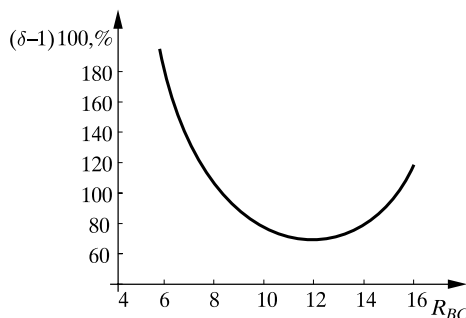
Fig. 13.13. A three-body polarization amplifier. (a) A represents a non-polarizable cation, B and C are polarizable molecules (C is the amplifier). (b) An electrostatic model. C is a metal wire, B represents two strips of paper. (c) The molecular amplifier resembles an electroscope with A being a positively charged metal ball, C a metal wire (i.e. a body with high polarizability) and B undergoes a huge charge redistribution due to the mediation of wire C.

$$\delta = 1 + 2\alpha_C \frac{R^2}{(R - R_{BC})^2 R_{BC}^3}. \quad (13.58)$$

For the polarizabilities chosen, polarization amplification ($\delta - 1$) takes place and at any R_{BC} exceeds 60%, see Fig. 13.14. When the amplifier is about in the middle of the AB distance the amplification is about 60%. When the amplifier approaches the electric field source (i.e. A^+), the amplification increases to about 100%. When the amplifier is close to B, it increases to about 200%.

This seems to be an interesting three-body effect we could investigate both theoretically and experimentally. Let us go a little crazy and assume C is made of a metal plate perpendicular to the AB line. Why a metal plate? Because the polarizability of a piece of metal is huge.⁶² There is trouble though. The dipoles induced

Fig. 13.14. Polarization amplification ($\delta - 1$ in %) on molecule B due to the mediation (three-body effect) of a polarizable molecule C ($\alpha_C = 100$ a.u.). The distance $R = 20$ a.u. The cation A^+ strongly polarizes molecule C. The dipole moment induced in this way on C, creates an additional electric field on B. This leads to polarization amplification on B.



Molecule C has the dipole moment (induced by the electric field from A^+). Let us calculate it as follows. The electric field created on mediator C by A^+ equals $\mathcal{E}_{A \rightarrow C} = \frac{1}{(R - R_{BC})^2}$, and therefore the corresponding induced dipole moment (component x) on C is $\mu_C = \alpha_C \frac{1}{(R - R_{BC})^2}$.

In the absence of the mediator C, the electric field on B would equal $\mathcal{E}_{A \rightarrow B} = \frac{1}{R^2}$, while with it (neglecting the self-consistency of the dipole moments on B and C) $\mathcal{E}_{A \rightarrow B} + \mathcal{E}_{C \rightarrow B}$. The ratio of the second and the first is given by eq. (13.58).

⁶²Let us recall the description of the metallic state of Chapter 9, p. 454 (HOMO-LUMO degeneracy) and then the definition of polarizability in Chapter 12, p. 635. Since the HOMO-LUMO separation is 0, the polarizability of a metal gives ∞ .

in the metal plate perpendicular to AB will be parallel to each other (side by side), which is energetically unfavourable. However, if the metal-plate is replaced by a metal wire oriented along line AB, everything would be amplified: the elementary dipoles would form a chain thus giving a big dipole within the wire. This means that the cation A^+ would attract a lot of the electrons within wire, so that on the opposite side of the wire we would have a sort of copy of it. Since the copy of A^+ would be very close to B, the polarization of B would increase very much.⁶³

13.9.5 ADDITIVITY OF THE SECOND-ORDER DISPERSION ENERGY

The dispersion energy is a second-order correction, eq. (13.12) on p. 695 gives the formula for the interaction of two molecules. For *three* molecules we obtain the following formula for the dispersion part of the second-order effect (cf. the discussion on the induction energy on p. 736)

$$E_{\text{disp}}(ABC) = \sum_{n_A, n_B \neq (0_A, 0_B)} \frac{|\langle n_A n_B 0_C | V_{AB} + V_{BC} + V_{AC} | 0_A 0_B 0_C \rangle|^2}{[E_A(0_A) - E_A(n_A)] + [E_B(0_B) - E_B(n_B)]} + \dots,$$

where $+\dots$ denotes analogous terms with summations over n_A, n_C as well as n_B, n_C . Among three integrals in the nominator only the first one will survive, since the other vanish due to the integration over the coordinates of the electrons of molecule Z not involved in the interaction V_{XY} :

$$\begin{aligned} E_{\text{disp}}(ABC) &= \sum_{n_A, n_B \neq (0_A, 0_B)} \frac{|\langle n_A n_B 0_C | V_{AB} | 0_A 0_B 0_C \rangle + 0 + 0|^2}{[E_A(0_A) - E_A(n_A)] + [E_B(0_B) - E_B(n_B)]} + \sum_{n_A, n_C \neq (0_A, 0_C)} \dots \\ &+ \sum_{n_B, n_C \neq (0_B, 0_C)} \dots \end{aligned}$$

In the first term we can integrate over the coordinates of C. Then the first term displayed in the above formula turns out to be the dispersion interaction of A and B,

$$\begin{aligned} E_{\text{disp}}(ABC) &= \sum_{n_A, n_B \neq (0, 0)} \frac{|\langle n_A n_B | V_{AB} | 0_A 0_B \rangle|^2}{[E_A(0_A) - E_A(n_A)] + [E_B(0_B) - E_B(n_B)]} \\ &+ \sum_{n_A, n_C \neq (0_A, 0_C)} \dots + \sum_{n_B, n_C \neq (0_B, 0_C)} \dots \\ &= E_{\text{disp}}(AB) + E_{\text{disp}}(AC) + E_{\text{disp}}(BC). \end{aligned}$$

⁶³Is it something (Fig. 13.13.b) you may recall from a lesson in physics with an electroscope in your school? A glass rod (Fig. 13.13.c) rubbed by fur acquires a charge (an analogue of A^+), then it approaches a metal (analogue of C) protruding from a glass vessel it causes repulsion of two pieces of paper attached to the metal in the vessel. The induction has reached distant regions of space. If, instead of the pieces of paper we have molecule B, it would exhibit a large induced dipole moment.

Thus, we have proved that

the dispersion interaction (second-order of the perturbation theory) is *additive*:

$$E_{\text{disp}}(ABC) = E_{\text{disp}}(AB) + E_{\text{disp}}(AC) + E_{\text{disp}}(BC).$$

13.9.6 NON-ADDITIVITY OF THE THIRD-ORDER DISPERSION INTERACTION

One of the third-order energy terms represents a correction to the dispersion energy. The correction as shown by Axilrod and Teller⁶⁴ has a three-body character. The part connected to the interaction of three distant instantaneous dipoles on A, B and C reads as

$$E_{\text{disp}}^{(3)} = 3C_{\text{ddd}}^{(3)} \frac{1 + 3 \cos \theta_A \cos \theta_B \cos \theta_C}{R_{AB}^3 R_{AC}^3 R_{BC}^3}, \quad (13.59)$$

where R_{XY} and θ_X denote the sides and the angles of the ABC triangle, and $C_{\text{ddd}}^{(3)} > 0$ represents a constant. The formula shows that

when the ABC system is in a linear configuration, the dispersion contribution is negative, i.e. *stabilizing*, while the equilateral triangle configuration corresponds to a *destabilization*.

ENGINEERING OF INTERMOLECULAR INTERACTIONS

13.10 NOBLE GAS INTERACTION

Theoretical description of the noble gas interaction requires quite advanced computational techniques, because here the binding effect comes from the dispersion interaction, which represents an electronic correlation effect. Such an effect is inaccessible in Hartree-Fock calculations. Some very expensive post-Hartree-Fock methods have to be used. The larger the number of electrons (N), the more expensive the calculations quickly become as N increases (as we have seen in Chapter 10): proportionally to N^5 for the MP2 method, and even as N^7 for the CCSD(T) method. Therefore, whereas He_2 CCSD(T) calculations would take a minute, similar Xe_2 calculations would take about $(\frac{108}{4})^7 = 26^7$ minutes, i.e. about 3000 years. No wonder, the xenon atom has 54 electrons, and in a system of 108 electrons there are plenty of events to correlate, but because of the 3000 years this

⁶⁴B.M. Axilrod, E. Teller, *J. Chem. Phys.* 11 (1943) 299.

is scary. To complete the horror, the calculations would have to be performed for many interatomic distances.

We may, however, make use of the following. First the calculations may be performed for He_2 , Ne_2 , Ar_2 , Kr_2 , Xe_2 using some reasonably poor basis sets. For each of the systems we obtain the equilibrium distance R_0 and the corresponding binding energy ε . Then, every curve $E(R)$ will be transformed (energy in ε units, distance in R_0 units) to $\frac{E(\frac{R}{R_0})}{\varepsilon}$. Every curve (independently of the system considered) has therefore depth 1 and minimum at $\frac{R}{R_0} = 1$.

It turns out that all the curves coincide to good accuracy.⁶⁵

Thus, all these objects are made out of the same matrix, despite the fact that this is so difficult to reveal using our computers. If we assume that this property were preserved for larger basis sets, we would be able to foresee the curve $E(R)$ for Xe_2 from good quality calculations for smaller noble gas dimers calculating $E(R_{\min})$.

13.11 VAN DER WAALS SURFACE AND RADII

It would be of practical importance to know how close two molecules can approach each other. We will not enter this question too seriously, because this problem cannot have an elegant solution: it depends on the direction of approach, and the atoms involved, as well as how strongly the two molecules collide. Searching for the effective radii of atoms would be nonsense, if the valence repulsion were not a sort of “soft wall” or if the atom sizes were very sensitive to molecular details. Fortunately, it turns out that an atom, despite different roles played in molecules, can be characterized by its approximate radius, called the *van der Waals radius*. The radius may be determined in a naive, but quite effective, way. For example, we may approach two HF molecules axially with the fluorine atoms heading on, then find the distance⁶⁶ R_{FF} at which the interaction energy is equal to, say, 5 kcal/mol (repulsion). The proposed fluorine atom radius would be $r_{\text{F}} = \frac{R_{\text{FF}}}{2}$. A similar procedure may be repeated with two HCl molecules with the resulting r_{Cl} . Now, let us consider an axial complex $\text{H-F}\cdots\text{Cl-H}$ with the intermolecular distance corresponding to 5 kcal/mol. What F \cdots Cl distance are we expecting? Of course, something close to $r_{\text{F}} + r_{\text{Cl}}$. It turns out that we are about right. This is why the atomic van der Waals radius concept is so attractive from the practical point of view.

We may define a superposition of atomic van der Waals spheres. This defines what is called the *van der Waals surface* of the molecule,⁶⁷ or a *molecular shape* – a concept of great importance and of the same arbitrariness as the radii themselves.

⁶⁵Similar results have been obtained for the noble gas atom and sulphur atom interactions [J. Klos, G. Chałasiński, R.V. Krems, A.A. Buchachenko, V. Aquilanti, F. Pirani, D. Cappelletti, *J. Chem. Phys.* 116 (2002) 9269].

⁶⁶Using a reliable quantum mechanical method.

⁶⁷The van der Waals surface of a molecule may sometimes be very complex, e.g., a molecule may have two or more surfaces (like fullerenes).

van der Waals
radius

molecular shape

In a similar way we may define *ionic radii*,⁶⁸ to reproduce the ion packing in ionic crystals, as well as *covalent radii* to foresee chemical bond lengths.

13.11.1 PAULI HARDNESS OF THE VAN DER WAALS SURFACE

How would an atom penetrate the van der Waals surface? It depends on the particular molecule, surface point and atom. The helium atom seems to be a good probe, because of its simplicity and small size. The question may be more specific: what is the value of the valence repulsion gradient or, alternatively, the interaction energy gradient, when the atomic probe penetrates perpendicularly at a given point of the van der Waals isosurface? Such hardness depends on the particular spot on the isosurface and exhibits the symmetry of the molecule.⁶⁹

The van der Waals surface might be modelled as one of the isosurfaces of the function $D(\mathbf{r}) = \sum_i A_i \exp(-B_i |\mathbf{r} - \mathbf{R}_i|)$, where the summation goes over the atoms of the molecule and the coefficients A_i and B_i depend not only on their kind (element), but also on their neighbourhood in the molecule. Therefore, we may propose

$$T(\mathbf{r}_0) = |(\nabla D)_{\mathbf{r}=\mathbf{r}_0}|$$

as the Pauli hardness at point \mathbf{r}_0 of the isosurface. Any point of the isosurface defined this way corresponds to a Pauli deformation of the wave function (Appendix Y) of the system: molecule and probe. *This represents another kind of deformation than that corresponding to the polarization of the molecule in an external electric field.* In one case the perturbation corresponds to a mechanical pushing, while in the other it pertains to the external electric field. The Pauli deformation will have complex anisotropic characteristic, when the probe penetrates the molecule. It is intriguing that, while the deformation due to the electric field results in an energy contribution of the second and higher orders, the Pauli deformation already appears in the first order energy correction.

13.11.2 QUANTUM CHEMISTRY OF CONFINED SPACE – THE NANOVESSELS

Molecules at long distances interact through the mediation of the electric fields created by them. The valence repulsion is of a different character, since it results from the Pauli exclusion principle, and may be interpreted as an energy penalty for an attempt by electrons of the same spin coordinate to occupy the same space (cf. Chapter 1 and p. 516).

Luty and Eckhardt⁷⁰ have highlighted the role of pushing one molecule by another. Let us imagine an atomic probe, e.g., a helium atom. The pushing by the probe deforms the molecular electronic wave function (Pauli deformation), but

⁶⁸This concept was introduced by Pauling, based on crystallographic data (L. Pauling, *J. Amer. Chem. Soc.* 49 (1927) 765).

⁶⁹Interestingly, water molecule is the hardest when approached in its plane about 44° off the OH direction, and the softest normal to the plane right above (and below) the oxygen atom. Data from E. Małolepsza, L. Piela, *J. Phys. Chem.* 107 (2003) 5356.

⁷⁰T. Luty, C.J. Eckhardt, in “*Reactivity of Molecular Solids*”, eds. E. Boldyreva, V. Boldyrev, Wiley, 1999, p. 51.

motion of the electrons is accompanied by the motion of the nuclei. Both motions may lead to dramatic events. For example, we may wonder how an explosive reaction takes place. Nothing happens during tens of years, and suddenly: boom! The spike hitting the material in its metastable chemical state is similar to the helium atom probe pushing a molecule. Due to the pushing, the molecule distorts to such an extent that the HOMO-LUMO separation vanishes and the system rolls down (see Chapter 14) to a deep potential energy minimum on the corresponding potential energy hypersurface. The HOMO-LUMO gap closing takes place within the reaction barrier. Since the total energy is conserved, the large reaction net energy gain goes to highly excited vibrational states (in the classical approximation corresponding to large amplitude vibrations). The amplitude may be sufficiently large to assure the pushing of the next molecules in the neighbourhood and a chain reaction starts with exponential growth.

Now imagine a lot of atomic probes confining the space (like a cage or template) available to a molecule under study. In such a case the molecule will behave differently from a free one. For example,

- a protein molecule, when confined, *will fold to another conformation*;⁷¹
- some photochemical reactions that require a space for the rearrangement of molecular fragments *will not occur*, if the space is not accessible;
- in a restricted space some other chemical reactions will take place (new chemistry – *chemistry in “nanovessels”*);
- some unstable molecules may *become stable* when enclosed in a nanovessel.

These are fascinating and little explored topics.

13.12 SYNTHONS AND SUPRAMOLECULAR CHEMISTRY

Alexandr Butlerov (1828–1886) Russian chemist, professor at the University of Kazan and Saint Petersburg. In 1861 Butlerov presented a concept of molecular *spatial structure*, where the atoms are bound by atom-to-atom chemical bonds, with properties characteristic for the atoms involved, an atom being able to bind only a few nearest neighbour atoms. Kazan University may be proud of several excellent scholars. Besides Butlerov, among others, there



are one of the founders of the non-Euclidean geometry Nicolai Lobachevsky as well as the inventor of electronic paramagnetic resonance Evgeniy Zavoiski.

To make complex chemical structures, synthetic chemists take advantage of the large scale of the atom–atom binding energies: from strong chemical bonds (of the order of 100 kcal/mol) to weak intermolecular interactions (of the order of a fraction of kcal/mol). For over one hundred and fifty years (since the time of Butlerov and Kekulé) chemists have used theory (of various levels) to plan and then build chemical structures with some chemical bonds to be broken and others to be created. Often the substances do not resemble the reagents, and the structure is held together by

⁷¹For example, in E. Małolepsza, M. Boniecki, A. Koliński, L. Piela, *Proc. Nat. Acad. Sciences* 102 (2005) 7835 a theoretical model of the conformational autocatalysis is investigated. The native conformation of a protein becomes unstable in presence of a misfolded conformation of another molecule of the protein. The native conformation unfolds and refolds to the metastable conformation.

strong chemical bonds, and therefore may be called “hard architecture”. The use of intermolecular interactions in synthesis (“soft architecture”, supramolecular chemistry) has arisen only during the last few decades (since Cram, Pedersen and Lehn⁷²). The supramolecular structures contain (as bricks) some loosely bound molecules, which therefore do not lose their individual properties.⁷³

It would seem that these “soft” structures are not interesting as they are unstable (it is sufficient to increase the temperature to make the structure disappear). The opposite is true, because such structures, after performing their function, may be destroyed without any significant energy expense.

13.12.1 BOUND OR NOT BOUND

Do the confined complexes such as catenans, rotaxans⁷⁴ and endohedral complexes (see Fig. 13.2, p. 688) represent intermolecular or intramolecular complexes? Certainly, when the distance between the subsystems, still within the structure of the complex, is large enough (this might be achieved by synthesis) the interaction is weak, as in any typical intermolecular interaction.

And what about the interaction of fragments of the same macromolecule that are close in space and at the same time distant, when walking through the framework of the chemical bonds? In this case we will also have some constraint for approaching two fragments, but chemists treat the interaction of two fragments of the DNA as if they were separate molecules. In such a way we have a coupling of the present section of the book with Chapter 7, where the force field contained the electrostatic interaction energy (of the net atomic charges, thus also taking into account higher-order molecular multipoles), valence repulsion and dispersion interaction (e.g., *via* terms r^{-12} and r^{-6} in the Lennard-Jones, p. 287). Among important contributions, only the induction energy is neglected in typical force fields.⁷⁵

Friedrich August Kekulé von Stradonitz (1829–1896), German organic and theoretical chemist, professor at the universities in Gent and Bonn. In 1858 Kekulé proved, that carbon has valency four and in 1865 proposed the correct ring-like formula for benzene after a peculiar dream about a serpent eating its own tail.



⁷²Three scholars shared the 1987 Nobel Prize in chemistry for creating supramolecular chemistry, in particular “for their development and use of molecules with structure-specific interactions of high selectivity”. Donald James Cram (b. 1919), American chemist, professor at the University of California–Berkeley; Charles John Pedersen (1904–1989), American chemist, employee of Dupont; Jean-Marie Lehn (b. 1939), French chemist, professor at the Université de Strasbourg and Collège de France in Paris.

⁷³Although small modifications still take place.

⁷⁴See a review article A.B. Braunschweig, B.H. Northrop, J.F. Stoddart, *J. Materials Chem.* 16 (2006) 32.

⁷⁵Although the new generation of force fields take it into account, see W.D. Cornell, P. Cieplak, C.I. Bayly, I.R. Gould, K.M. Merz, Jr., D.M. Ferguson, D.C. Spellmeyer, T. Fox, J.W. Caldwell, P.A. Kollman, *J. Am. Chem. Soc.* 117 (1995) 5179.

13.12.2 DISTINGUISHED ROLE OF THE ELECTROSTATIC INTERACTION AND THE VALENCE REPULSION

The electrostatic contribution plays a prominent role in intermolecular interaction. The electrostatic forces already operate effectively at long intermolecular distances (their range may, however, be reduced in polar solvents).

The induction and dispersion contributions, even if sometimes larger than the electrostatic interaction, usually play a less important role. This is because only the electrostatics may change the sign of the energy contribution when the molecules reorient, thus playing the pivotal role in the interaction energy.

The induction and dispersion contributions are negative (at any orientation of the molecules), and we may say, as a rule of thumb, that their role is to make the configurations (already being stabilized by the electrostatics) more stable.

The valence repulsion plays the role of a hard wall (covered by a “soft blanket”) that forbids the closed-shell molecules to approach too closely. This represents a very important factor, since those molecules that do not fit together receive an energy penalty.

13.12.3 HYDROGEN BOND

Among the electrostatic interactions, the most important are those having a strong dependence on orientation, the most representative being the *hydrogen bonds* $X-H \dots Y$, where an electronegative atom X plays the role of a proton donor, while an electronegative atom Y – plays the role of a proton acceptor. Most often the hydrogen bond $X-H \dots Y$ deviates only a little from linearity. Additionally, the XY separation usually falls into a narrow range: 2.5–3.2 Å, at least for the most important $X, Y \in \{O, N\}$. The hydrogen bond features are unique, because of the extraordinary properties of the hydrogen atom itself. This is the only atom which occasionally may attain the partial charge equal to +0.45 e, which means it represents a nucleus devoid to a large extent of an electron density. This is one of the reasons why the hydrogen bond is so strong when compared with other types of intermolecular interactions.

Example 5. Water–water dimer. Let us take the example of two water molecules to show the dominant role of electrostatics in the hydrogen bond.

As it is seen, while at the equilibrium distance $R_{OO} = 3.00$ Å all the contributions are of equal importance (although the electrostatics dominates), all the contributions except electrostatics, diminish considerably after increasing separation by only about 0.70 Å. For the largest separation ($R_{OO} = 4.76$), the electrostatics

Table 13.7. Energy contributions to the interaction energy E_{int} in the system $\text{HO}-\text{H}\dots\text{OH}_2$ (hydrogen bond) calculated^a within the SAPT method: electrostatic energy E_{elst} , valence repulsion energy $E_{\text{exch}}^{(1)}$, induction energy E_{ind} and dispersion energy E_{disp} for three O...O distances: equilibrium distance $R_{\text{OO}} = 3.00$ and two distances a little larger: medium 3.70 \AA and large 4.76 \AA

Contributions to E_{int} (in kcal/mol)				
$R_{\text{OO}} (\text{\AA})$	E_{elst}	$E_{\text{exch}}^{(1)}$	E_{ind}	E_{disp}
3.00	-7.12	4.90	-1.63	-1.54
3.70	-2.79	0.30	-0.18	-0.31
4.76	-1.12	0.00	-0.02	-0.05

^aB. Jeziorski, M. van Hemert, *Mol. Phys.* 31 (1976) 713.

dominates by far. This is why the hydrogen bond is said to have a mainly electrostatic character.⁷⁶

13.12.4 COORDINATION INTERACTION

Coordination interaction appears if an electronic pair of one subsystem (electron donor) lowers its energy by interacting⁷⁷ with an electron acceptor offering an empty orbital, e.g., a cation (acceptor) interacts with an atom or atoms (donors) offering lone electronic pairs. This may be also seen as a special kind of electrostatic interaction.⁷⁸ Fig. 13.15.a shows a derivative of porphyrin as well as a cryptand (the name comes from the ritual of burying the dead in crypts), Fig. 13.15.b, the compounds offering lone pairs for the interaction with a cation.

cryptands

When concentrating on the ligands we can see that in principle they represent a negatively charged cavity (lone pairs) waiting for a monoatomic cation with *dimensions of a certain range only*. The interaction of such a cation with the ligand would be exceptionally large and therefore “specific” for such a pair of interacting moieties, which is related to the selectivity of the interaction.

Let us consider a water solution containing ions: Li^+ , Na^+ , K^+ , Rb^+ , Cs^+ . After adding the above mentioned cryptand and after the equilibrium state is attained (ions/cryptand, ions/water and cryptand/water solvation), only for K^+ will the equilibrium be shifted towards the K^+ /cryptand complex. For the other ions the equilibrium will be shifted towards their association with water molecules, not the cryptand.⁷⁹ This is remarkable information.

⁷⁶It has been proved that covalent structures (cf. p. 520) also contribute to the properties of the hydrogen bond, but their role decreases dramatically when the molecules move apart.

⁷⁷Forming a molecular orbital.

⁷⁸A lone pair has a large dipole moment (see Appendix T), which interacts with the positive charge of the acceptor.

⁷⁹J.-M. Lehn, “*Supramolecular Chemistry*”, Institute of Physical Chemistry Publications, 1993, p. 88: the equilibrium constants of the ion/cryptand association reactions are: for Li^+ , Na^+ , K^+ , Rb^+ , Cs^+

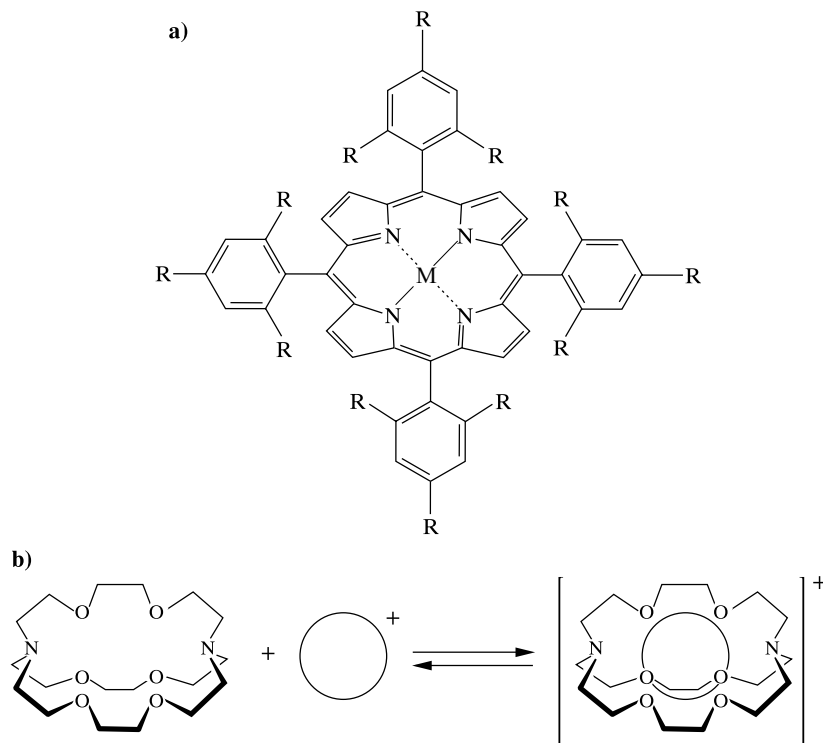


Fig. 13.15. A cation fits (a) the porphyrin ring or (b) the cryptand.

We are able to selectively extract objects of some particular shape and dimensions (recognition).

13.12.5 HYDROPHOBIC EFFECT

This is quite a peculiar type of interaction, which appears mainly (not only) in water solutions.⁸⁰ The hydrophobic interaction does not represent any particular new interaction (beyond those we have already considered), because at least potentially they could be explained by the electrostatic, induction, dispersion, valence repulsion and other interactions already discussed, cf. pp. 718 and 695.

The problem may be seen from a different point of view. The basic interactions have been derived as if operating in vacuum. However, in a medium the molecules interact with one another through the mediation of other molecules, including those of the solvent. In particular, a water medium creates a strong network of

(only the order of magnitude is given): 10^2 , 10^7 , 10^{10} , 10^8 , 10^4 , respectively. As seen the cryptand's cavity only fits well to the potassium cation.

⁸⁰W. Kauzmann, *Advan. Protein Chem.* 14 (1959) 1. A contemporary theory is given in K. Lum, D. Chandler, J.D. Weeks, *J. Phys. Chem.* 103 (1999) 4570.

the hydrogen bonds that surround the hydrophobic moieties expelling them from the solvent⁸¹ and pushing together which imitates their mutual attraction, resulting in the formation of a sort of “oil drop”.

We may say in a rather simplistic way that hydrophobic molecules aggregate not because they attract particularly strongly, but because water strongly prefers them to be out of its hydrogen bond net structure.

Hydrophobic interactions have a complex character and are not yet fully understood. The interaction depends strongly on the size of the hydrophobic synthons. For small sizes, e.g., such as two methane molecules in water, the hydrophobic interaction is small, increasing considerably for larger synthons. The hydrophobic effects become especially important for what is called the *amphiphilic* macromolecules with their van der Waals surfaces differing in hydrophobic character (hydrophobic/hydrophilic). The amphiphilic molecules are able to self-organize, forming structures up to the nanometer scale (“nanostructures”).

amphiphilic
molecules

self-
organization
nano-structures

Fig. 13.16 shows an example of the hierarchic (“multi-level”) character of a molecular architecture:

- The chemical binding of the amino acids into the oligopeptides is the first level (“hard architecture”).
- The second level (“soft architecture”) corresponds to a beautiful network of hydrogen bonds responsible for forming the α -helical conformation of each of the two oligopeptides.
- The third level corresponds to an extremely effective hydrophobic interaction, the *leucine-valine zipper*. Two α -helices form a very stable structure⁸² winding up around each other and thus forming a kind of a superhelix, known as *coiled-coil*, due to the hydrophobic leucine-valine zipper.⁸³

leucine-valine
zipper

coiled-coil

The molecular architecture described above was first planned by a chemist. The system fulfilled all the points of the plan and self-organized in a spontaneous process.⁸⁴

⁸¹Hydrophobic interactions involve not only the molecules on which we focus our attention, but also, to an important extent, the water molecules of the solvent. The hydrogen bond network keeps the hydrophobic objects together, as a shopping bag keeps lard slabs together.

The idea of solvent-dependent interactions represents a general and fascinating topic of research. Imagine the interaction of solutes in mercury, in liquid gallium, liquid sodium, in a highly polarizable organic solvent, etc. Due to the peculiarities of these solvents, we will have different chemistry going on in them.

⁸²B. Tripet, L. Yu, D.L. Bautista, W.Y. Wong, T.R. Irvin, R.S. Hodges, *Prot. Engin.* 9 (1996) 1029.

⁸³Leucine may be called the “flag ship” of the hydrophobic amino acids, although this is not the most polite compliment for a hydrophobe.

⁸⁴One day I said to my friend Leszek Stolarczyk: “If those organic chemists wanted to, they could synthesize anything you might dream of. They are even able to cook up in their flasks a molecule composed of the carbon atoms that would form the shape of a cavalry man on his horse”. Leszek answered: “Of course! And the cavalry man would have a little sabre, made of iron atoms.”

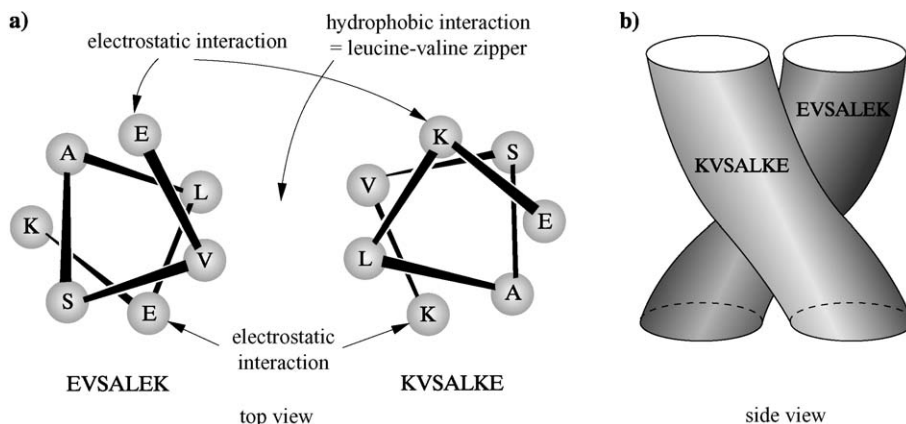


Fig. 13.16. An example of formation of the coiled-coil in the case of two oligopeptide chains (a): (EVSALEK)_n with (KVSALKE)_n, with E standing for the glutamic acid, V for valine, S for serine, A for alanine, L for leucine, K for lysine. This is an example of a multi-level molecular architecture. First, each of the two oligopeptide chains form α -helices, which afterwards form a strong hydrophobic complex due to a perfect matching (leucine and valine of one of the α -helices with valine and leucine of the second one, known as the leucine-valine zipper (b)). The complex is made stronger additionally by two salt bridges (COO^- and NH_3^+ electrostatic interaction) involving pairs of glutamic acid (E) and lysine (K). The resulting complex (b) is so strong that it serves in analytical chemistry for the separation of some proteins.

13.12.6 MOLECULAR RECOGNITION – SYNTHONS

Organic molecules often have quite a few donor and acceptor substituents. These names may pertain to donating/accepting electrons or protons (cf. the charge conjugation described on p. 702). Sometimes a particular side of a molecule displays a system of donors and acceptors. Such a system “awaiting” interaction with a complementary object is called a synthon,⁸⁵ and their matching represents the *molecular recognition*. The cryptand in Fig. 13.15.b therefore contains a synthon able to recognize a narrow class of cations (with sizes within a certain range).

In Fig. 13.17 we show another example of synthons based on hydrogen bonds. Due to the particular geometry of the molecules as well as to the above mentioned weak dependence of the XY distance on X and Y, both synthons are complementary. The example is of immense importance, because it pertains to guanine (G), cytosine (C), adenine (A) and thymine (T). Thanks to these two pairs of synthons (GC and AT) we exist, because the G, C together with the A and T represent the four letters which are sufficient to write the Book of Life word by word in a single molecule of DNA. The words, the sentences and the chapters of this Book decide the majority of the very essence of your (and my) personality. The whole DNA strand may be considered as a large single synthon. The synthon has its important counterpart which fits the DNA perfectly because of the complementarity. The molecular machine which synthesizes this counterpart molecule (a “negative”) is

⁸⁵G.R. Desiraju, “*Crystal Engineering, The Design of Organic Solids*”, Elsevier, Amsterdam, 1989.

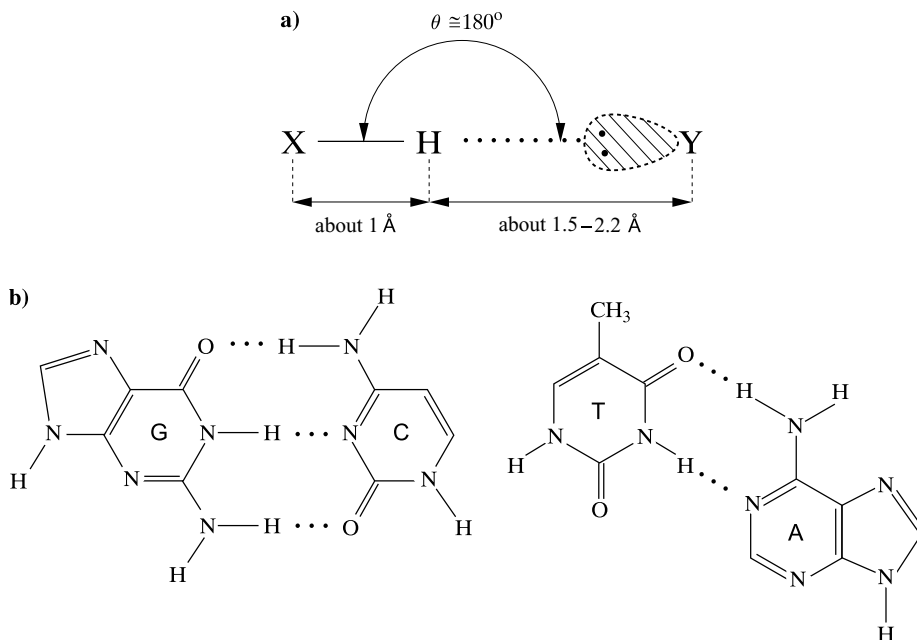


Fig. 13.17. Synthons are often based on a hydrogen bond pattern (a). The synthon of guanine (G) fits the synthon of cytosine (C), while the synthon of adenine (A) fits that of the thymine (T) (b).

the polymerase, a wonderful molecule (you will read about in Chapter 15). Any error in this complementarity results in a mutation.⁸⁶

13.12.7 “KEY-LOCK”, TEMPLATE AND “HAND-GLOVE” SYNTHON INTERACTIONS

The energy spectrum of a molecule represents something like its finger print. The particular energy levels correspond to various electronic, vibrational and rotational states (Chapter 6). Different electronic states⁸⁷ may be viewed as representing different chemical bond patterns. Different vibrational states⁸⁸ form series, each series for an energy well on the PES. The energy level pattern is completed by the rotational states of the molecule as a whole. Since the electronic excitations are of the highest energy, the PES of the ground electronic state is most important. For flexible molecules such a PES is characterized by a lot of potential energy wells corresponding to the conformational states. If the bottoms of the excited conformational wells are of high energy (with respect to the lowest-energy well, Fig. 13.18.a), then the molecule in its ground state may be called “rigid”, because high energy is needed to change the molecular conformation.

⁸⁶Representing a potential or real danger, as well as a chance for evolution.

⁸⁷In the Born–Oppenheimer approximation, each corresponding to a potential energy hypersurface, PES.

⁸⁸Including internal rotations, such as those of the methyl group.

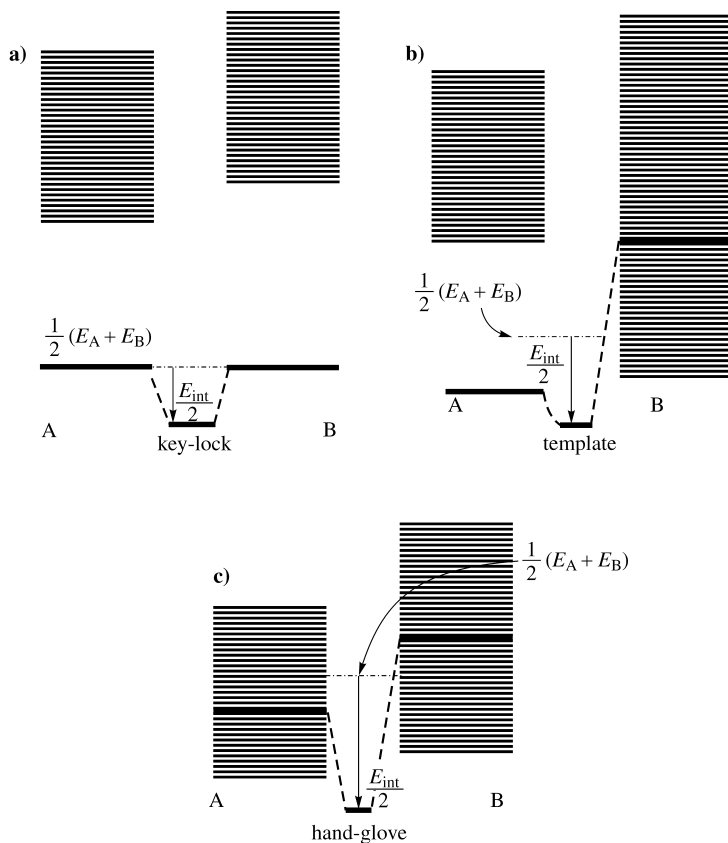


Fig. 13.18. The key-lock, template and hand-glove molecular recognition. Any molecule may be characterized by a spectrum of its energy states. (a) In the key-lock type interaction of two rigid molecules A and B their low-energy conformational states are separated from the quasi-continuum high-energy conformational states (including possibly those of some excited electronic states) by an energy gap, in general different for A and B. Due to the effective synthon interactions the energy per molecule lowers substantially with respect to that of the isolated molecules leading to the molecular recognition without significant changes of molecular shape. (b) In the template-like interaction one of the molecules is rigid (large energy gap), while the other one offers a quasi-continuum of conformational states. Among the later, there is one that (despite of being a conformational excited state), due to the perfect matching of synthons results in considerable energy lowering, much below the energy of isolated molecules. Thus, one of the molecules has to distort in order to get perfect matching. (c) In the hand-glove type of interaction the two interacting molecules offer quasi-continua of their conformational states. Two of the excited conformational states correspond to such molecular shapes as match each other perfectly and lower the total energy considerably. This lowering is so large that it is able to overcome the conformational excitation energy (an energy cost of molecular recognition).

If such rigid molecules A and B match each other, this corresponds to the *key-lock* type of molecular recognition. To match, the interacting molecules sometimes have only to orient properly in space when approaching one another and then dock (the AT or GC pairs may serve as an example). The key-lock concept of Fischer from 100 years ago (concerning enzyme–substrate interaction) is considered as

the foundation of supramolecular chemistry – the chemistry that deals with the complementarity and matching of molecules.

One of the molecules, if rigid enough, may serve as a *template* for another molecule, which is flexible, Fig. 13.18.b, and together they form a strong complex. Finally two flexible molecules (Fig. 13.18.c) may pay an energy penalty for acquiring higher-energy conformations, but such ones which lead to a very strong interaction of the molecules in the *hand-glove* type of molecular recognition.⁸⁹

Still another variation of this interaction comes into play, when during the approach, a new type of synthon appears, and the synthons match afterwards. For example in the Hodges superhelical structure (Fig. 13.16), only after formation of the α -helices does it turn out that the leucine and valine side chains of one helix match perfectly similar synthons of the second helix (“leucine-valine zipper”).

Nature has done it routinely for millions of years. Endonuclease (EcoRV) represents an enzyme whose function is selective chemical bond breaking between nucleotides (linking the adenine and thymine) in a single DNA strand. Fig. 13.19 shows a model of the complex of EcoRV with a fragment of DNA,⁹⁰ altogether about 62000 atoms. Fig. 13.19 highlights some aspects of the interaction.

Note the hierarchic structure of the *host-guest* complex (DNA-EcoRV). DNA is a double-helix (Fig. 13.19.a) and this shape results mainly from the intermolecular A...T and G...C interactions through mediation of the hydrogen bonds. The enzyme EcoRV (Fig. 13.19.b) also has a highly organized structure, in particular six α -helices and a few β strands exhibit their characteristic hydrogen bond patterns (not displayed in the figure), these secondary structure elements fit together through the mediation of hydrophobic interactions. As we can see, the cavity of the EcoRV is too small, but becomes larger when the guest molecule is accommodated (“hand-glove” effect), thus enabling an effective host-guest interaction. This example shows how important valence repulsion is. *If the EcoRV cavity differed much from that suitable to accommodate the guest molecule, the host when deforming would pay a too high an energetic price and the energetic gain connected to docking would become too small to compensate for this energy expense.*

Hermann Emil Fischer (1852–1919), German chemist, professor at the universities in Strasbourg, Munich, Erlangen, Würzburg, Berlin. Known mainly for his excellent works on the structure of sugar compounds. His (recognized decades later) correct determination of the absolute conformation of sugars was based solely on the analysis of their chemical properties. Even today this would require advanced physicochemical in-



vestigations. In 1902 Fischer received the Nobel Prize “for his work on sugar and purine syntheses”.

“host-guest”
complex

⁸⁹Our immunological system represents an excellent example. When a foreign agent enters the blood system, it is bound by an antibody that is *able to adapt its shape to practically any agent*. Moreover, a complex mechanism transmits the information about the agent’s size and shape, and all this results in mass production of antibodies with the particular shape needed to bind the invader.

⁹⁰L. Wróblewska, Master thesis, University of Warsaw, 2000.



Fig. 13.19. The DNA fragment (“guest”) fits the cavity in the enzyme EcoRV (“host”) structure very well. (a) A fragment of the double-strand DNA helix (side view). (b) EcoRV. (c) Host–guest complex (the DNA molecule shown in the top view). Besides the geometric fitting (i.e. a lack of considerable valence repulsion) there is also an electrostatic and amphiphilic fitting of both subsystems.

Another masterpiece of nature – self-organization of the tobacco virus is shown in Fig. 13.20. Such a complex system self-assembles, because its parts not only fit one another (synthons), but also found themselves in solution and made perfect matching accompanied by an energy gain. Even more spectacular is the structure and functioning of bacteriophage T (Fig. 13.21).

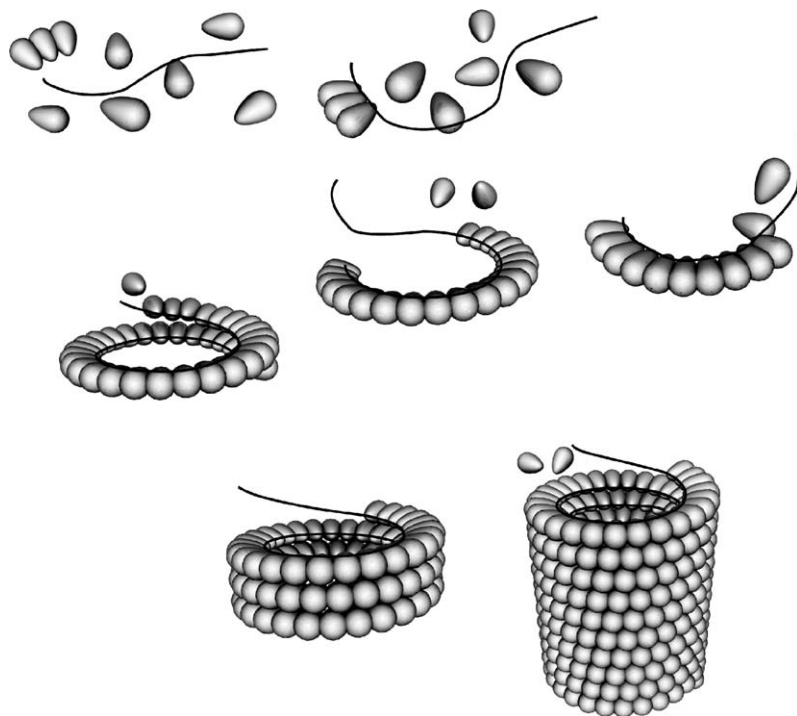


Fig. 13.20. Self-organization of the tobacco virus. The virus consists of an RNA helix (shown as a single strand) containing about 7000 nucleotides – sufficient genetic material to code the production of 5–10 proteins (first level of supramolecular self-organization). The RNA strand interacts very effectively with a certain protein (shown as a “drop”; the second level). The protein molecules associate with the RNA strand forming a kind of necklace, and then the system folds (third level) into a rod-like shape, typical for this virus. The rods are able to form a crystal (level four, not shown here), which melts after heating, but is restored when cooled down.

Summary

- Interaction energy of two molecules (at a given geometry) may be calculated within any reliable quantum mechanical method by subtracting from the total system energy the sum of the energies of the subsystems. This is called a *supermolecular method*.
- The supermolecular method has at least one important advantage: *it works independently of the interaction strength and of the intermolecular distance*. The method has the disadvantage that due to the subtraction, a loss of accuracy occurs and no information is obtained about the structure of the interaction energy.
- In the supermolecular method there is a need to compensate for what is called the basis set superposition error (BSSE). The error appears because due to the incompleteness of the atomic basis set (Ω_A, Ω_B), the individual subsystem A with the interaction switched off profits from the Ω_A basis set only, while when interacting lowers its energy due to the total $\Omega_A \cup \Omega_B$ basis set (the same pertains to any of the subsystems). As a result a part of the calculated interaction energy does not come from the interaction, but from the problem of the basis set used (BSSE) described above. A remedy is called the counterpoise method, in which all quantities (including the energies of the individual subsystems) are calculated within the $\Omega_A \cup \Omega_B$ basis set.

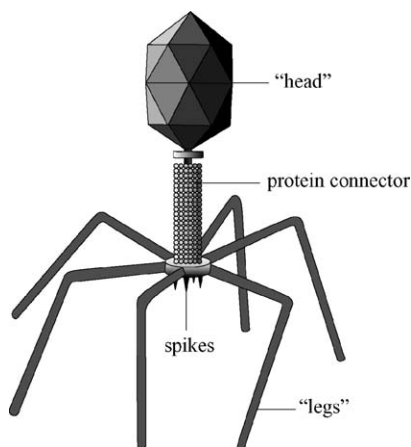


Fig. 13.21. Bacteriophage T represents a supramolecular construction that terrorizes bacteria. The hexagonal “head” contains a tightly packed double helix of DNA (the virus genetic material) wrapped in a coat build of protein subunits. The head is attached to a tube-like molecular connector built of 144 contractible protein molecules. On the other side of the connector there is a plate with six spikes protruding from it as well as six long, kinked “legs” made of several different protein molecules. The legs represent a “landing apparatus” which, using intermolecular interactions, attaches to a particular receptor on the bacterium cell wall. This reaction is reversible, but what happens next is highly irreversible. First, an enzyme belonging to the monster makes a hole in the cell wall of the bacterium. Then the 144 protein molecules contract probably at the expense of energy from hydrolysis of the ATP molecule (adenosine triphosphate – a universal energy source in biology), which the monster has at its disposal. This makes the head collapse and the whole monster serves as a syringe. The bacteriophage’s genetic material enters the bacterium body almost in no time. That is the end of the bacterium.

- Perturbational method has limited applicability:
 - at long intermolecular separations what is called the polarization approximation may be used,
 - at medium distances a more general formalism called the symmetry adapted perturbation theory may be applied,
 - at short distances (of the order of chemical bond lengths) perturbational approaches are inapplicable.
- One of the advantages of a low-order perturbational approach is the possibility of dividing the interaction energy into well defined physically distinct energy contributions.
- In a polarization approximation approach, the unperturbed wave function is assumed as a *product* of the exact wave functions of the individual subsystems: $\psi_0^{(0)} = \psi_{A,0}\psi_{B,0}$. The corresponding zero-order energy is $E_0^{(0)} = E_{A,0} + E_{B,0}$.
- Then, the first-order correction to the energy represents what is called the *electrostatic* interaction energy: $E_0^{(1)} = E_{\text{elst}} = \langle \psi_{A,0}\psi_{B,0} | V | \psi_{A,0}\psi_{B,0} \rangle$, which is the Coulombic interaction (at a given intermolecular distance) of the frozen charge density distributions of the individual, non-interacting molecules. After using the multipole expansion E_{elst} can be divided into the sum of the multipole–multipole interactions plus a remainder, called the penetration energy. A multipole–multipole interaction corresponds to the permanent multipoles of the isolated molecules. An individual multipole–multipole interaction term (2^k -pole with 2^l -pole) vanishes asymptotically as $R^{-(k+l+1)}$, e.g., the dipole–dipole term decreases as $R^{-(1+1+1)} = R^{-3}$.

- In the second order we obtain the sum of the induction and dispersion terms: $E^{(2)} = E_{\text{ind}} + E_{\text{disp}}$.
- The induction energy splits into

$$E_{\text{ind}}(A \rightarrow B) = \sum'_{n_B} \frac{|\langle \psi_{A,0} \psi_{B,n_B} | V | \psi_{A,0} \psi_{B,0} \rangle|^2}{(E_{B,0} - E_{B,n_B})},$$

which pertains to polarization of molecule B by the unperturbed molecule A , and

$$E_{\text{ind}}(B \rightarrow A) = \sum'_{n_A} \frac{|\langle \psi_{A,n_A} \psi_{B,0} | V | \psi_{A,0} \psi_{B,0} \rangle|^2}{(E_{A,0} - E_{A,n_A})},$$

with the roles of the molecules exchanged. The induction energy can be represented as the permanent multipole – induced multipole interaction, where the interaction of the 2^k -pole with the 2^l -pole vanishes as $R^{-2(k+l+1)}$.

- The dispersion energy is defined as

$$E_{\text{disp}} = \sum'_{n_A} \sum'_{n_B} \frac{|\langle \psi_{A,n_A} \psi_{B,n_B} | V | \psi_{A,0} \psi_{B,0} \rangle|^2}{(E_{A,0} - E_{A,n_A}) + (E_{B,0} - E_{B,n_B})}$$

and represents a result of the electronic correlation. After applying the multipole expansion, the effect can be described as a series of instantaneous multipole – instantaneous multipole interactions, with the individual terms decaying asymptotically as $R^{-2(k+l+1)}$. The most important contribution is the dipole–dipole ($k = l = 1$), which vanishes as R^{-6} .

- The polarization approximation fails for medium and short distances. For medium separations we may use *symmetry-adapted perturbation theory* (SAPT). The unperturbed wave function is symmetry-adapted, i.e. has the same symmetry as the exact function. This is not true for the polarization approximation, where the product-like $\varphi^{(0)}$ does not exhibit the proper symmetry with respect to electron exchanges between the interacting molecules. The symmetry-adaptation is achieved by a projection of $\varphi^{(0)}$.
- Symmetry-adapted perturbation theory *reproduces all the energy corrections that appear in the polarization approximation* ($E_{\text{elst}}, E_{\text{ind}}, E_{\text{disp}}, \dots$) *plus provides some exchange-type terms* (in any order of the perturbation except the zeroth).
- The most important exchange term is the valence repulsion appearing in the first-order correction to the energy:

$$E_{\text{exch}}^{(1)} = \langle \psi_{A,0} \psi_{B,0} | V \hat{P}^{AB} \psi_{A,0} \psi_{B,0} \rangle - \langle \psi_{A,0} \psi_{B,0} | V \psi_{A,0} \psi_{B,0} \rangle \langle \psi_{A,0} \psi_{B,0} | \hat{P}^{AB} \psi_{A,0} \psi_{B,0} \rangle + O(S^4),$$

where $O(S^4)$ represents all the terms decaying as the fourth power of the overlap integral(s) or faster, \hat{P}^{AB} stands for the single exchanges' permutation operator.

- The interaction energy of N molecules is *not pairwise additive*, i.e. is not the sum of the interactions of all possible pairs of molecules. Among the energy corrections up to the second order, the exchange and, first of all, the induction terms contribute to the non-additivity. The electrostatic and dispersion (in the second order) contributions are pairwise additive.
- The non-additivity is highlighted in what is called the *many-body expansion of the interaction energy*, where the interaction energy is expressed as the sum of two-body, three-body, etc. energy contributions. The m -body interaction is defined as that part of the interaction energy that is non-explicable by any interactions of $m' < m$ molecules, but explicable by the interactions among m molecules.

- The dispersion interaction in the third-order perturbation theory contributes to the three-body non-additivity and is called the Axilrod–Teller energy. The term represents a correlation effect. Note that the effect is negative for three bodies in a linear configuration.
- The most important contributions: electrostatic, valence repulsion, induction and dispersion lead to a richness of supramolecular structures.
- The electrostatic interaction plays a particularly important role, because it is of a long-range character as well as very sensible to relative orientation of the subsystems. The hydrogen bond $X-H \cdots Y$ represents an example of the domination of the electrostatic interaction. This results in its directionality, linearity and a small (as compared to typical chemical bonds) interaction energy of the order of 5 kcal/mol.
- Also the valence repulsion is one of the most important energy contributions, because it controls how the interacting molecules fit together in space.
- The induction and dispersion interactions for polar systems, although contributing significantly to the binding energy, in most cases do not have a decisive role and only slightly modify the geometry of the resulting structures.
- In aqueous solutions the solvent structure contributes very strongly to the intermolecular interaction, thus leading to what is called the hydrophobic effect. The effect expels the non-polar subsystems from the solvent, thus causing them to approach, which *looks* like an attraction.
- A molecule may have such a shape that it fits that of another molecule (synthons, small valence repulsion and a large number of attractive atom–atom interactions).
- In this way molecular recognition may be achieved by the key-lock-type fit (the molecules non-distorted), template fit (one molecule distorted) or by the hand-glove-type fit (both molecules distorted).
- Molecular recognition may be planned by chemists and used to build complex molecular architectures, in a way similar to that in which living matter operates.

Main concepts, new terms

interaction energy (p. 684)	polarization collapse (p. 717)
natural division (p. 684)	symmetrized polarization approximation (p. 717)
binding energy (p. 687)	MS–MA perturbation theory (p. 717)
dissociation energy (p. 687)	Jeziorski–Kołos perturbation theory (p. 717)
dissociation barrier (p. 687)	valence repulsion (p. 718)
catenans (p. 688)	Padé approximants (p. 722)
rotaxans (p. 688)	Pauli blockade (p. 722)
endohedral complexes (p. 688)	exchange–deformation interaction (p. 722)
supramolecular method (p. 689)	interaction non-additivity (p. 726)
basis set superposition error (BSSE) (p. 690)	many-body expansion (p. 727)
ghosts (p. 690)	SE mechanism (p. 733)
polarization perturbation theory (p. 692)	TE mechanism (p. 734)
electrostatic energy (p. 693)	polarization catastrophe (p. 738)
induction energy (p. 694)	three-body polarization amplifier (p. 738)
dispersion energy (p. 694)	Axilrod–Teller dispersion energy (p. 741)
multipole moments (p. 698)	van der Waals radius (p. 742)
permanent multipoles (p. 701)	van der Waals surface (p. 742)
SAPT (p. 710)	supramolecular chemistry (p. 744)
function with adapted symmetry (p. 711)	hydrogen bond (p. 746)
symmetry forcing (p. 716)	

hydrophobic effect (p. 748)
amphiphilicity (p. 749)
nanostructures (p. 749)
leucine-valine zipper (p. 749)

synthon (p. 750)
template (p. 751)
“key-lock” interaction (p. 751)
“hand-glove” interaction (p. 751)

From the research front

Intermolecular interactions influence any liquid and solid state measurements. Physicochemical measurement techniques give only some indications of the shape of a molecule, except NMR, X-ray and neutron analyses, which provide the atomic positions in space, but are very expensive. This is why there is a need for theoretical tools which may offer such information in a less-expensive way. For very large molecules, such an analysis uses the force fields described in Chapter 7. This is currently the most powerful theoretical tool for determining the approximate shape of molecules with numbers of atoms even of the order of hundreds of thousands. To obtain more reliable information about intermolecular interactions we may perform calculations within a supermolecular approach, necessarily of an *ab initio* type, because other methods give rather poor quality results. The DFT method popular nowadays fails at its present stage of development, because the intermolecular interactions area, especially the dispersion interaction, is a particularly weak point of the method. If the particular method chosen is the Hartree–Fock approach (currently limited to about 300 atoms), we have to remember that it cannot take into account any dispersion contribution to the interaction energy by definition.⁹¹ *Ab initio* calculations of the correlation energy still represent a challenge. Good quality calculations for a molecule with a dozen atoms may be carried out using the MP2 method. Still more time consuming are the CCSD(T) or SAPT calculations, which are feasible for only a few atom systems, but offer an accuracy of 1 kcal/mol required for chemical applications.

Ad futurum...

No doubt the computational techniques will continue to push the limits mentioned above. The more coarse the method used, the more spectacular this pushing will be. The most difficult to imagine would be a great progress in methods using explicitly correlated wave functions. It seems that pushing the experimental demands and calculation time required will cause experimentalists (they will perform the calculations⁹²) to prefer a rough estimation using primitive methods rather than wait too long for a precise result (still not very appropriate, because the results are obtained without taking the influence of solvent, etc. into account). It seems that in the near future we may expect theoretical methods exploiting the synthon concept. It is evident that a theoretician has to treat the synthons on an equal footing with other atoms, but a practice-oriented theoretician cannot do that, otherwise he would wait virtually forever for something to happen in the computer, while in reality the reaction takes only a picosecond or so. Still further into the future we will see the planning of hierarchic multi-level supramolecular systems, taking into account the kinetics and competitiveness among such structures. In the still more distant future, functions performed by such supramolecular structures, as well as their sensitivity to changing external conditions, will be investigated.

⁹¹The dispersion energy represents an electronic correlation effect, absent in the Hartree–Fock energy.

⁹²What will theoreticians do? My answer is given in Chapter 15.

Additional literature

J.O. Hirschfelder, C.F. Curtiss, R.B. Bird, "Molecular Theory of Gases and Liquids", Wiley, New York, 1964.

A thick "bible" (1249 pages) of intermolecular interactions. We find there everything before the advent of computers and of the symmetry-adapted perturbation theory.

H. Margenau, N.R. Kestner, "Theory of Intermolecular Forces", Pergamon, Oxford, 1969.

A lot of detailed derivations. There is also a chapter devoted to the non-additivity of the interaction energy – a *rara avis* in textbooks.

A.J. Stone, "The Theory of Intermolecular Forces", Oxford Univ. Press, Oxford, 1996.

The book contains the basic knowledge in the field of intermolecular interactions given in the language of perturbation theory as well as the multipole expansion (a lot of useful formulae for the electrostatic, induction and dispersion contributions). This very well written book presents many important problems in a clear and comprehensive way.

"Molecular Interactions", ed. S. Scheiner, Wiley, Chichester, 1997.

A selection of articles written by experts.

P. Hobza, R. Zahradnik, "Intermolecular Complexes", Elsevier, Amsterdam, 1988.

Many useful details.

I.G. Kaplan, "Intermolecular Interactions", Wiley, Chichester, 2006.

A very well written book, covering broad and fundamental aspects of the field.

Questions

1. In order to avoid the basis set superposition error using the counter-poise method:
 - a) all quantities have to be calculated within the total joint atomic basis set;
 - b) the energy of the total system has to be calculated within the total joint atomic basis set, while the subsystem energies have to be calculated within their individual atomic basis sets;
 - c) the total energy has to be calculated within a modest quality basis set, while for the subsystems we may allow ourselves a better quality basis set;
 - d) all the atomic orbitals have to be centred in a single point in space.
2. In the polarization approximation the zero-order wave function is:
 - a) the sum of the wave functions of the polarized subsystems;
 - b) the product of the wave functions of the polarized subsystems;
 - c) a linear combination of the wave functions for the isolated subsystems;
 - d) the product of the wave functions of the isolated subsystems.
3. Please find the false statement:
 - a) dissociation energy depends on the force constants of the normal modes of the molecule;
 - b) interaction energy calculated for the optimal geometry is called the binding energy;
 - c) the absolute value of the dissociation energy is larger than the absolute value of the binding energy;
 - d) some systems with energy higher than the dissociation limit may be experimentally observed as stable.

4. Please find the false statement. The induction energy (R stands for the intermolecular distance):
 - a) represents an electronic correlation effect;
 - b) for two water molecules decays asymptotically as R^{-6} ;
 - c) is always negative (attraction);
 - d) for the water molecule and the argon atom vanishes as R^{-6} .
5. Dispersion energy (R stands for the intermolecular distance):
 - a) is non-zero only for the noble gas atom interaction;
 - b) is equal to zero for the interaction of the two water molecules calculated with the minimal basis set;
 - c) represents an electronic correlation effect;
 - d) decays asymptotically as R^{-4} .
6. A particle has electric charge equal to q and Cartesian coordinates x, y, z . Please find the false statement:
 - a) the multipole moments of the particle depend in general on the choice of the coordinate system;
 - b) the operator of the z component of the dipole moment (μ_z) of the particle is equal to qz ;
 - c) a point-like particle does not have any dipole moment, therefore $\mu_z = 0$;
 - d) the octupole moment operator of the particle is a polynomial of the third degree.
7. In symmetry adapted perturbation theory for He_2 :
 - a) we obtain the dispersion energy in the first order of the perturbation theory;
 - b) the electronic energy exhibits a minimum as a function the interatomic distance;
 - c) the binding energy is non-zero and is caused by the induction interaction;
 - d) the unperturbed wave function is asymmetric (do not mix with antisymmetry) with respect to every electron permutation.
8. Please choose the false statement. In symmetry adapted perturbation theory:
 - a) the zeroth order wave function is a product of the wave functions of the individual molecules;
 - b) the electrostatic energy together with the valence repulsion appears in the first order;
 - c) the exchange corrections appear in all (non-zero) orders;
 - d) the Pauli exclusion principle is taken into account.
9. The non-additivity of the interaction energy has the following features:
 - a) the electrostatic energy is additive, while the induction energy is not;
 - b) the valence repulsion is additive for single exchanges;
 - c) the dispersion interaction being a correlation effect is non-additive;
 - d) the Axilrod–Teller interaction is additive.
10. Typical distances $\text{X}\dots\text{Y}$ and $\text{H}\dots\text{Y}$ for a hydrogen bond $\text{X}-\text{H}\dots\text{Y}$ are closest to the following values (\AA):
 - a) 1.5, 1; b) 2.8, 1.8; c) 3.5, 2.5; d) 2.8, 1.

Answers

1a, 2d, 3c, 4a, 5c, 6c, 7b, 8a, 9a, 10b