Doctoral Thesis

Microscopic analysis of the spectra of aromatic guests in alkane hosts

Author(s):
Wallenborn, Ernst-Udo

Publication Date:
1999

Permanent Link:
https://doi.org/10.3929/ethz-a-003810429

Rights / License:
In Copyright - Non-Commercial Use Permitted
Microscopic Analysis of the Spectra of Aromatic Guests in Alkane Hosts

A dissertation submitted to the

SWISS FEDERAL INSTITUTE OF TECHNOLOGY ZÜRICH

For the degree of
Doctor of Natural Sciences

presented by
ERNST-UDO WALLENBORN
Dipl. Chem. ETH

born July 10, 1970
German citizen

Accepted on the recommendation of:
Prof. Dr. Urs P. Wild, Examiner
Prof. Dr. Ulrich W. Suter, Co-Examiner

1999
Contents

Acknowledgments ........................................ iii
Abstract ................................................. v
Zusammenfassung ........................................ vii
Symbols .................................................... ix

1 Introduction ........................................... 1
   1.1 Macroscopic Approaches ............................... 3
   1.2 Microscopic Approaches: Phenomenological Models .. 4
   1.3 Microscopic Approaches: Atomistic Methods ........... 5

2 Theory .................................................. 7
   2.1 Classical Mechanics .................................. 9
      2.1.1 The pcff forcefield ............................. 9
   2.2 Quantum Mechanics .................................. 13
      2.2.1 The SBEJ method ................................ 16
      2.2.2 Semiclassical Dynamical Theory of Lineshapes .... 27

3 Van-der-Waals Clusters ............................... 31
   3.1 Tetracene ............................................ 33
      3.1.1 Geometry and Vibrational Analysis ............... 34
      3.1.2 Choice of Parameters ........................... 36
      3.1.3 Molecular Dynamics .............................. 39
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1.4</td>
<td>Cluster Geometries</td>
<td>40</td>
</tr>
<tr>
<td>3.1.5</td>
<td>Spectra</td>
<td>45</td>
</tr>
<tr>
<td>4.1</td>
<td>Perylene as guest</td>
<td>70</td>
</tr>
<tr>
<td>4.1.1</td>
<td>Crystal structure</td>
<td>70</td>
</tr>
<tr>
<td>4.1.2</td>
<td>Vibrational analysis</td>
<td>74</td>
</tr>
<tr>
<td>4.1.3</td>
<td>ppcff force field</td>
<td>81</td>
</tr>
<tr>
<td>4.1.4</td>
<td>Functional form of ppcff</td>
<td>86</td>
</tr>
<tr>
<td>4.1.5</td>
<td>Error function of ppcff</td>
<td>89</td>
</tr>
<tr>
<td>4.1.6</td>
<td>Parameter optimization of ppcff</td>
<td>93</td>
</tr>
<tr>
<td>4.1.7</td>
<td>Parameters of ppcff</td>
<td>96</td>
</tr>
<tr>
<td>4.1.8</td>
<td>Performance of ppcff</td>
<td>96</td>
</tr>
<tr>
<td>4.2</td>
<td>Hexane as host</td>
<td>107</td>
</tr>
<tr>
<td>4.2.1</td>
<td>Crystal structure</td>
<td>107</td>
</tr>
<tr>
<td>4.2.2</td>
<td>Perylene in n-hexane</td>
<td>110</td>
</tr>
<tr>
<td>4.3</td>
<td>Heptane as host</td>
<td>118</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Crystal structure</td>
<td>118</td>
</tr>
<tr>
<td>4.3.2</td>
<td>Perylene in n-heptane</td>
<td>120</td>
</tr>
</tbody>
</table>

**Bibliography** 127

**Curriculum Vitae** 137

**Index** 138
Acknowledgments

First of all, I would like to thank Prof. Urs Wild for allowing me to work on this project, although it was slightly outside of the normal scope of his group.

Prof. Michel Orrit and Dr. Ross Brown at the University of Bordeaux invited me to work at Bordeaux for a month, a time which I greatly enjoyed and from which major parts of this work resulted. Ross contributed the entire molecular dynamics part of the calculations on tetracene-heptane clusters. His simulated annealing method solved the problem of proper ensemble representation which usually plagues studies like this.

The work of Prof. Ulrich Suter and Prof. Epameinondas Leontidis on chromophores in polymer matrices sparked the whole project. Minos suggested the new force field for perylene, which allowed the work on Shpol’skii systems. This work marked the first time the method described in this work was used for the successful assignment of spectral sites in these systems, yielding detailed information about these systems which currently not available experimentally.

The force field optimization benefitted greatly from the work of Marcel Zehnder and Serge Santos, whose standalone implementation of pcff drastically reduced the computational workload. Without their code, the force field optimization would have been possible only in a strongly reduced form.

Finally, I would like to thank Andreas Osterwalder, who helped by providing computational resources when it really counted.
Abstract

The calculation of electronic spectra of guest-host systems is a complex task. Several layers of problems have to be solved or circumvented to provide relevant information about the investigated system. There is one layer of geometrical and structural problems, one layer in which the nature of the intermolecular interaction has to be dealt with, and there is the layer in which a detailed and quantum-mechanically correct description of the guest molecule is necessary.

Atomistic models counter this layered problem structure with equally layered algorithms. The modularity introduced by this approach allows a flexible use of tools for the calculation of the spectra of guest-host systems. On each layer, the most accurate, most practical, or simply the computationally cheapest method may be chosen and combined with the methods from the other layers to form an algorithm highly adapted to the particular nature of the specific problem.

Of course, the fundamental assumption in all atomistic theories is that the layers are indeed independent of each other. Deviations from this assumption, for example a distortion of the guest molecule upon insertion, may demand a much more detailed treatment than desired, and, at worst, may render the atomistic approach computationally unfeasible.

Here, a general atomistic methodology is presented, which allows the calculation of spectral shifts and inhomogeneous lineshapes of general aromatic guest-alkane host systems. The algorithm contains an ab initio representation of the guest and a classical description of the host. It does not depend on a particular implementation of the encountered four-center integrals, although one such implementation (the Shalev-Ben-Horin-Even-Jortner method) is used exclusively here.
Using this methodology, two particular guest-host systems were investigated: the mass-selective Resonant Two-Photon Two-Color Ionization spectra of tetracene in n-heptane clusters, and Total Luminescence Spectra of the Shpol'skii systems perylene in n-hexane and n-heptane.

Slightly different approaches were chosen for the two studies. The clusters were simulated using standard force fields and simulation techniques. However, the clusters were generated using simulated annealing to obtain proper ensemble representations. Comparison of calculated and observed spectra allowed an unambiguous assignment of the isomers of small to medium-sized clusters. In particular, it could be confirmed that the bimodal distribution of the tetracene-heptane\textsubscript{2} cluster was caused by the presence of different isomers with significantly different bathochromic shifts and not due to the vibronic structure of just one isomer.

For the work on the two Shpol'skii systems, a different approach was necessary: Here the crystal structures and hence the possible inclusion sites were known beforehand and no special algorithms for their generation was necessary. However, the force field employed here was an augmented version of the standard pcff force field. The augmentation was necessary because of the default pcff's inability to model the guest molecule perylene correctly. In this study, it could be shown that the empirical parameters inherent to the SBEJ method are transferable between the two host systems. Furthermore, assignment of the most prominent lines was possible in both cases.

For perylene in n-hexane, the assignment differs from a previously published one. The assignment of the sites of perylene in n-heptane, however, agrees with measurements previously published by another group.
Zusammenfassung


Die Grundannahme aller atomistischen Theorien ist, daß diese Ebenen tatsächlich unabhängig voneinander behandelt werden können. Wenn diese Annahme nicht erfüllt ist, beispielsweise durch eine einschlußbedingte Verzerrung der Gastmolekülgewicht, wird eine deutlich detailliertere Behandlung erforderlich, die im schlimmsten Fall den atomistischen Ansatz unpraktikabel machen können.

Hier wird ein generelles atomistisches Modell zur Berechnung spektraler Verschiebungen und inhomogener Linienformen vorgestellt. Das Modell benutzt eine ab initio Darstellung des Gastmoleküls und eine klassische Beschreibung seiner Umgebung. Es ist unabhängig von der spezifischen Implementation der benötigten Vier-
zentrenintegrale, obwohl eine solche Implementation (die Shalev-Ben-Horin-Even-Jortner-Methode) in der vorliegenden Arbeit ausschließlich verwendet wird.


Symbols

Abbreviations and Acronyms

AO Atomic Orbital
CSF Configuration State Function
DFT Density Functional Theory
ESR Electron Spin Resonance
HOMO Highest occupied molecular orbital
IR Infra-Red spectroscopy
LUMO Lowest unoccupied molecular orbital
MD Molecular Dynamics
MM Molecular Modelling
MO Molecular Orbital
NpT Ensemble at constant number of atoms N, constant pressure p, and constant temperature T
NVT Ensemble at constant number of atoms N, constant volume V, and constant temperature T
occ Occupied orbital
PAH Polycyclic Aromatic Hydrocarbons
ppcff Perylene Polymer Consistent Force Field
QM Quantum Mechanics
R2P2CI Resonant Two-Photon Two-Colour Ionization spectroscopy
RSPT Rayleigh-Schrödinger Perturbation Theory
SAPT Symmetry-Adapted Perturbation Theory
SBEJ  Shalev-Ben-Horin-Even-Jortner-method

TLS  Total Luminescence Spectroscopy

virt  Virtual, i.e. empty orbital

**Mathematical and Physical Symbols**

\[\alpha_A\] Polarizability of \(A\)

\(A\)  Host group

\(\mathcal{A}\)  Antisymmetrization operator, \(\mathcal{A} = \sum_P (-1)^\chi(P) P\)

\(\chi(P)\)  Parity of permutation \(P\)

\(\chi\)  Atomic orbital of the guest

\(C_{\mu\nu}\)  MO coefficient of the \(\mu\)th AO in the \(\nu\)th MO of the guest

\((ar|ct)\)  Coulomb integral, \((ar|ct) = \int dr_1 \int dr_2 \psi_a(r_1)\phi_c(r_2) 1/r_12 \psi_r(r_1)\phi_t(r_2)\)

\(d(\vec{x}_1, \vec{x}_2)\)  Distance between atoms 1 and 2, in Ångström

\(d\)  Refractive index

\(\varepsilon\)  Dielectric constant of the host

\(\varepsilon_{ij}\)  Lennard-Jones parameter, \(\varepsilon_{ij} = 2\sqrt{\varepsilon_i \varepsilon_j} (r_i^6 + r_j^6) / (r_i^6 + r_j^6)\)

\(\eta\)  Adjustable parameter in SBEJ

\(\varepsilon^r_a\)  Energy difference between canonical orbitals \(r\) and \(a\)

\(E_{corr}\)  Long-range correction for the Lennard-Jones potential

\(\Delta E_i\)  Relative energy of the \(i\)th configuration, in kcal/mol

\(E_{RS}\)  Rayleigh-Schrödinger interaction energy

\(\Delta E\)  Solvation shift

\(\delta E^{(A)}\)  Contribution of group \(A\) to the total solvation shift

\(\Delta F_{ij}\)  \(j\)th Cartesian component of the force in the \(i\)th configuration, in kcal/molÅ

\(G^A_{mn}\)  Geometric factor in SBEJ, \(G^A_{mn} = (\vec{R}_{mA} \cdot \vec{R}_{nA}) / (R^3_{mA} R^3_{nA})\)

\(G\)  Green function

\(G_0(t)\)  Semiclassical Green function

\(\hbar\)  Planck constant divided by \(2\pi\), \(\hbar = 1.054 \cdot 10^{-34}\) Js
Symbols

$\mathcal{H}$ Hamiltonian

$I_A$ Ionization potential of $A$

$I(\omega)$ Spectral lineshape

$\kappa$ Adjustable parameter in SBEJ

$K_{nn}^A$ Electronic factor in SBEJ

$\mu_0$ Scalar transition dipole moment

$\langle \mu(0)\mu(t) \rangle$ Transition dipole moment autocorrelation function

$\mu$ Dipole moment

$\mu_c$ Transition dipole moment of the $c$th and $t$th MO of the host, $\mu_c = \langle \phi_c|\mu|\phi_t \rangle$

$M$ Guest molecule

m.p. Melting point

$\nu_i$ Frequency of the $i$th vibration, in cm$^{-1}$

$N$ Number of particles in a simulation

$\Phi$ Wavefunction of the host

$\phi$ Molecular orbital of the host

$\Psi$ Wavefunction of the guest

$\Psi_0$ Ground state wavefunction

$\Psi_R$ General excited state wavefunction

$\Psi_a$ Wavefunction with one electron excited from orbital $a$ to orbital $r$

$\psi$ Molecular orbital of the guest

$p$ Pressure

$\mathcal{P}$ Permutation operator

$Q_{ar}^m$ Transition monopole on the $m$th atom of the guest, $Q_{ar}^m = \sum_{\mu,\nu} C_{\mu a} C_{\nu r} S_{\mu\nu}$

$S_{\mu\nu}$ Overlap of the $\mu$th and $\nu$th atomic orbital, $S_{\mu\nu} = \int dr \chi_\mu(r)\chi_\nu(r)$

$r_{ij}$ Lennard-Jones parameter, $r_{ij} = (0.5 \cdot (r_i^6 + r_j^6))^{1/6}$

$\mathbf{R}_{mA}$ Vector from the $m$th atom of the guest to the $A$th group of the host

$R_{mA}$ Distance of the $m$th atom of the guest from the $A$th group of the host, $R_{mA} = |\mathbf{R}_{mA}|$
<table>
<thead>
<tr>
<th>Symbols</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Theta(t)$</td>
<td>Heaviside function, $\Theta(t) = 0$ if $t &lt; 0$, $1$ if $t \geq 0$</td>
</tr>
<tr>
<td>$T$</td>
<td>Temperature</td>
</tr>
<tr>
<td>$U$</td>
<td>Potential energy</td>
</tr>
<tr>
<td>$U(\tau)$</td>
<td>Shift function</td>
</tr>
<tr>
<td>$V$</td>
<td>Volume of the simulation box</td>
</tr>
<tr>
<td>$\vec{v}_i$</td>
<td>Normalized eigenvector of the $i$th vibration</td>
</tr>
<tr>
<td>$V$</td>
<td>Intermolecular interaction operator, $V = \sum_{ij} 1/r_{ij}$</td>
</tr>
<tr>
<td>$\vec{x}_i$</td>
<td>Cartesian coordinates of the $i$th atom</td>
</tr>
</tbody>
</table>
1. Introduction
Physics is the science of microscopic causes and their macroscopic effects. As a program, it aims to find the constituents of reality as well as the laws governing them. Ultimately, the properties of matter are to be described in terms of fundamental particles and fundamental laws. Of course, physics must not stop on a merely descriptive level: to be of more than merely explanatory value, these building blocks must be put together again to form numerical predictions and simulations, from the application of the simple law of gravity to the prediction of planetary motion [NEWTON 1686], to the complex simulations of biological systems that today routinely use more computational resources than were necessary to put a man on the moon thirty years ago.

The prediction of electronic structures and electronic spectra in solution has long resisted the trend towards microscopic theoretical analysis. The apparent successes of macroscopic reaction field or microscopic but phenomenological theories and the high complexity of atomistic descriptions for electronic spectra of solvated systems have led to the curious situation that quantum mechanics is virtually absent in some of the most successful theoretical accounts of spectroscopical observations that are intrinsically of quantum mechanical nature [REILLY AND SKINNER 1995].

### 1.1 Macroscopic Approaches

The simplest way to include the presence of a solvent in the quantum-mechanical description of a molecule is the so-called reaction field method [KIRKWOOD 1934, ONSAGER 1936]. The molecule is considered to be in the center of a cavity (radius $R$) in the solvent (dielectric constant $\varepsilon$), and the Hamiltonian of the system is written as

$$H = H_0 + H_1$$

where $H_0$ is the unperturbed molecular Hamiltonian, and $H_1 = -\mu \cdot \vec{R}$ is the interaction due to the dipole moment of the molecule, $\mu$, and $\vec{R}$, the reaction field created by the molecule in the solvent outside the cavity. The calculations then proceed in a straightforward manner: since the reaction field depends on the solution
of the Schrödinger equation for the molecule, first an eigenstate of $\mathcal{H}_0$ is found, $\tilde{R}$ is then calculated and inserted in the Hamiltonian, and the calculation is repeated iteratively until convergence. Today, this particular algorithm (cf. [WONG ET AL. 1991]) is a standard option in quantum-chemical packages, like Gaussian or Zindo.

1.2 Microscopic Approaches: Phenomenological Models

To overcome the weaknesses of purely macroscopic models, such as the lack of a proper representation of inhomogeneities or the neglect of dynamic aspects in the system, several phenomenological models have been suggested. Most successful in the context of low-temperature spectroscopy is the two-level-system model, which assumes that the guest molecule is surrounded by two-level systems, i.e. local quantum systems for which, at low temperatures, only the two lowest states are accessible. The physical nature of these states is irrelevant to the theory. The model is general enough that any possible kind of two-level-system can be assumed: flipping middle rings in $p$-terphenyl, local minima on the potential energy surface in reptationally moving polymer chains, and even methyl group rotational tunneling have been suggested [CROCI 1997]. In fact, over the years it has acquired mantra-like status in the community that the physical nature of two-level systems is unknown.

The lack of insight on the atomic level contrasts strongly with the success of two-level-system theories in describing dynamic properties, in particular the phenomenon of spectral diffusion (e.g. of single molecule fluorescence lines of pentacene in $p$-terphenyl, cf. [REILLY AND SKINNER 1995]). However, the evidence resulting from phenomenological analyses has often been circumstantial. Given the overall complexity of these models and the number of (generally unchecked) assumptions hidden in them, it is not surprising that progress in this area has been slow recently.
1.3 Microscopic Approaches: Atomistic Methods

Atomistic models offer an alternative which is not plagued by these problems. There have been several such studies. [Köhler AND WOEHL 1995] for example have calculated Stark splittings of \((1, 3E, 5E, 7)\)-octatetraene in solid \(n\)-hexane. They obtained the optimized geometry by a standard molecular mechanics simulation. It is worth to note that [Köhler AND WOEHL 1995] consider it important that the employed force field reproduces not only the crystal structure of the host but also that of the guest\(^1\). The electronic structure of the octatetraene guest is described by a simple Hückel-type model developed earlier by [Köhler 1990], and the Stark splittings are calculated using a simplified three-state Hamiltonian. Neither the classical molecular mechanics part, nor the Hückel-type electronic structure description, nor the empirical three-state Hamiltonian depend on other parts of the calculation; each could be replaced easily (the Hückel model for octatetraene for example by a thorough and quantum-chemically sound configuration interaction wave function). This is a general feature of atomistic models.

In 1991 [Shalev et al. 1991] presented a method to calculate the solvation shifts of aromatic/rare-gas clusters perturbatively. Their method relied on a Hückel-type description of the isolated systems, and was able to describe only rare-gas hosts. It was successfully used to simulate spectra of several aromatic molecules in jet-expansion clusters with rare-gas hosts.

This method was later modified to treat more general excitations as well as other hosts [Leontidis ET al. 1995]. In above publication the spectral shift of \(s\)-tetrazine immersed in completely amorphous atactic polypropylene was investigated, and the inhomogeneous broadening and its dependency on factors such as host density and packing of the immediate molecular environment was analyzed. It was found that the method was indeed able to model the solvation shift microscopically. However, the necessarily short simulation times of molecular dynamics runs prohibited an investigation of phenomena on a longer time scale, such as, for

---

\(^1\)Some care must be exercised in choosing a molecular mechanics package since some of them ... are not capable of preserving bond alternation in the conjugated octatetraene and predict an octatetraene crystal structure that has a qualitatively different packing than the real crystals" [Köhler AND WOEHL 1995].
example, spectral diffusion [CROCI 1997]. Thus, the study was restricted to the detailed microscopic investigation of nine different \( s \)-tetrazine/polypropylene simulation runs, which was barely sufficient to obtain statistically sound data about the inhomogeneous structure of the absorption line. This, of course, is a problem of polymeric systems in general.

To circumvent this problem, [GEISSINGER ET AL. 1998] investigated "the inhomogeneous broadening of the electronic transitions of the linear polyene octatetraene doped into mixed crystal matrices of various \( n \)-alkanes with added halogenated alkanes" [GEISSINGER ET AL. 1998]. The authors assumed a stochastic distribution of the halogenated alkanes in the regular alkane matrix, evaluated the spectral shifts using the SBEJ method described below, and calculated the inhomogeneous linewidths as a function of the concentration of halogenated alkanes. Again, the guest molecule was described using a simplified Hückel-type wave-function. The shifts were calculated with the original SBEJ algorithm.

Hence, modern studies focused on systems where the "ensemble limitation" intrinsic to the treatment of polymeric systems is not present. [WALLENBORN ET AL. 1997] for example analyzed the lineshapes of tetracene/\( n \)-heptane clusters as measured using resonant two-photon two-colour ionization spectroscopy (R2P2CI) in supersonic jet-expansion experiments by [BEN-HORIN ET AL. 1991]. These experiments allowed resolution of the cluster spectra according to cluster size. By applying the combined molecular mechanics/quantum mechanics method described here, [WALLENBORN ET AL. 1997] were able to analyze the tetracene/heptane spectra and assign the observed lines to different cluster isomers.

Another way to avoid the "ensemble problem" is to focus on systems characterized by a small number of isomers, such as systems where aromatic guests are inserted in alkane hosts in a fast-cooling procedure. These systems, generally known as Shpol'skii systems are characterized by the presence of sharp lines indicating a small number of different inclusion sites. Investigation of such a small number of isomers is feasible with today's computational resources and provides valuable insights into the principles which guide the structural and electronic properties of Shpol'skii systems [WALLENBORN ET AL. 1999].
2. Theory
2.1 Classical Mechanics

For brevity, no attempt is made here to give an introduction into or describe the principles and algorithms of classical mechanics. Such detailed descriptions can be found in various excellent publications, e.g. [Allen and Tildesley 1990], [Chandler 1987], or [Frenkel and Smit 1996]. Suffice to say that the fundamental idea of molecular dynamics is to solve Newton’s equations of motion

\[ F_i = m_i a_i = -\frac{\partial U(x_1, \ldots, x_n)}{\partial x_i} \]

(2.1)

under certain boundary conditions, like constant temperature, constant volume of the simulation box, or constant pressure, creating ensembles which are usually named after these constants (e.g. \( NpT \) for a simulation with constant particle number \( N \), pressure \( p \), and temperature \( T \)). Usually, periodic boundary conditions are assumed and for particles near the boundary of the simulation box interactions with the nearest images of the other particles in the simulation box are also taken into account.

It is important to note that the potential energy function \( U \), usually called the force field, is crucial for an accurate description of the system. Thus, the particular force field used in later chapters needs to be described in detail here.

2.1.1 The pcff forcefield

The pcff force field [BIOSYM/MSI 1997], as described in [Sun 1995], is a member of the family of Class II force fields\(^1\), like for example CFF91, CFF93 and QMFF [Maple et al. 1988, Maple et al. 1994, Hwang et al. 1994]. All these force fields are so-called \textit{ab initio} force fields, which means that the error function used in their parametrization contained \textit{ab initio} data on relative energies, forces and

\(^1\)The term “Class II force fields” denotes a collection of properties of the force field, most notably that it is a general-purpose atomic force field as opposed to united-atom force fields or force fields specially optimized for certain classes of compounds, as for example GROMOS [van Gunsteren and Berendsen 1987]. Furthermore it means that the force field is markedly an-harmonic and was optimized using a balanced error criterion, as opposed to, for example, just a set of vibrational frequencies [Hwang et al. 1994].
Hessian matrix elements of either minimized molecular geometries, or better a set of randomly distorted geometries, a procedure developed by [MAPLE ET AL. 1988].

![Figure 2.1: Definitions of the variables used in the description of the force field.](image)

In pcff, the potential energy consists of the following terms (see Fig. (2.1)):

$$E_{pot} = \sum_b \left[ K_2(b - b_0)^2 + K_3(b - b_0)^3 + K_4(b - b_0)^4 \right]$$  \hspace{1cm} (2.2a)

$$+ \sum_\theta \left[ H_2(\theta - \theta_0)^2 + H_3(\theta - \theta_0)^3 + H_4(\theta - \theta_0)^4 \right]$$  \hspace{1cm} (2.2b)

$$+ \sum_\phi \sum_{n=1}^3 V_n \left\{ 1 + \cos(n\phi - \phi_n^0) \right\}$$  \hspace{1cm} (2.2c)

$$+ \sum_x K_x \chi^2$$  \hspace{1cm} (2.2d)

$$+ \sum_{b,b'} F_{bb'}(b - b_0)(b' - b_0')$$  \hspace{1cm} (2.2e)

$$+ \sum_{b,b'} F'_{bb'}(b - b_0)(b' - b_0')$$  \hspace{1cm} (2.2f)

$$+ \sum_{b,\theta} F_{bb}(b - b_0)(\theta - \theta_0)$$  \hspace{1cm} (2.2g)
The terms in Eq. (2.2) are described below.

**quartic-bond, Eq. (2.2a):** The sum is over all bonds between atoms $i$ and $j$, $b$ being the distance between $i$ and $j$, $b_0$ is the equilibrium bond length. The energy is given by a power series expansion up to fourth order with the force constants as parameters.

**quartic-angle, Eq. (2.2b):** Analogously, the sum is over all angles $\theta$ (equilibrium value $\theta_0$) between adjacent atoms $i$, $j$ and $k$. The energy is again given by a power series expansion up to fourth order.

**torsion-3, Eq. (2.2c):** The torsional potential for four adjacent atoms $i$, $j$, $k$ and $l$ is given by a (cosine) Fourier expansion up to third order. The equilibrium torsional angles are usually (but not necessarily) set to $\pi$, hence the $\phi^0$ are usually zero.

**wilson-oop, Eq. (2.2d):** Given four atoms $i$, $j$, $k$ and $l$ such that $i$, $k$ and $l$ are all bonded to $j$, the $\chi$ in the definition of the Wilson-out-of-plane term is the angle between the $ij$ bond and the plane of $j$, $k$ and $l.$
**bond-bond**, Eq. (2.2e): This term describes the coupling between adjacent bonds $b$ and $b'$.

**bond-bond-1-3**, Eq. (2.2f): This term describes the coupling between 1-3 bonds $b$ and $b'$, such that $b$ and $b'$ are linked by a third (middle) bond.

**bond-angle**, Eq. (2.2f): This represents the coupling between a bond $b$ and an adjacent angle $\theta$.

**angle-angle**, Eq. (2.2h): This term describes the coupling between two angles $\theta$ and $\theta'$ linked by a bond.

**end-bond-torsion-3**, Eq. (2.2h): This is the coupling between a torsional angle $\phi$ and each of the (end) bonds $b$ and $b'$.

**middle-bond-torsion-3**, Eq. (2.2j): This is the coupling between a torsional angle $\phi$ and the (middle) bond $b$.

**angle-torsion-3**, Eq. (2.2k): The coupling between a torsional angle $\phi$ and each of the adjacent angles $\theta$ and $\theta'$.

**angle-angle-torsion-1**, Eq. (2.2l): The coupling between a torsional angle $\phi$ and both of the adjacent angles $\theta$ and $\theta'$.

**nonbond-9-6**, Eq. (2.2m): The Lennard-Jones 6-9 potential between two atoms $i$ and $j$, where

$$
\varepsilon_{ij} = 2 \sqrt{\varepsilon_i \cdot \varepsilon_j} \frac{r_i^3 r_j^3}{r_i^6 + r_j^6}
$$

and

$$
r_{ij} = \left( \frac{r_i^6 + r_j^6}{2} \right)^{\frac{1}{6}}
$$

and $\varepsilon$ and $r$ are the Lennard-Jones-parameters of the two atoms.

**nonbond-coulomb**, Eq. (2.2n): This is the familiar electrostatic interaction between two atoms $i$ and $j$. $r_{ij}$ is the distance between the two atoms and $\varepsilon$ is the dielectric constant of the solvent.
2.2 Quantum Mechanics

The perturbational formulation of intermolecular interactions is easily obtained and very old [London 1937, Longuet-Higgins 1956]. Although empirical expressions like the Lennard-Jones-Potential [Lennard-Jones 1924] actually predate the first quantum-mechanically sound work of [EisenSCHITZ AND London 1930], the perturbational theory of intermolecular interactions is almost as old as quantum mechanics itself.

In the simplest formulation, the Hamiltonian of the combined system consisting of two molecules 1 and 2 is written in terms of the Hamiltonians of the two constituents, \( \mathcal{H}^{(1)} \) and \( \mathcal{H}^{(2)} \), as

\[
\mathcal{H} = \mathcal{H}^{(1)} + \mathcal{H}^{(2)} + \mathcal{V}
\]

where \( \mathcal{V} \) is the intermolecular interaction operator\(^2\). If \( \Psi \) and \( \Phi \) are the solutions for separate molecules, \textit{i.e.}

\[
\mathcal{H}^{(1)} |\Psi\rangle = E^{(1)} |\Psi\rangle \quad \text{and} \quad \mathcal{H}^{(2)} |\Phi\rangle = E^{(2)} |\Phi\rangle
\]

the wavefunction of the combined system is written as a simple Hartree product of the constituents' wavefunctions. The Schrödinger equation then becomes:

\[
\mathcal{H} |\Psi\Phi\rangle = E |\Psi\Phi\rangle
\]

\(^2\mathcal{V} \) is usually taken to be \( \mathcal{V} = \sum_{ij} 1/r_{ij} \), where the sum is over all electrons \( i \) on molecule 1 and \( j \) on molecule 2. Of course, this is by no means the only possible choice: ideally all interactions between nuclei and electrons of 1 and 2 should be included. However, in practical applications of the theory most of these cancel for neutral molecules.
and Rayleigh-Schrödinger perturbation theory (up to second order) yields the familiar expression [Claverie 1978]:

\[
E_{RS} = \langle \Psi_0 \Phi_0 | \mathcal{V} | \Psi_0 \Phi_0 \rangle \\
- \sum_{I \neq 0} \frac{|\langle \Psi_0 \Phi_0 | \mathcal{V} | \Psi_I \Phi_0 \rangle|^2}{E_I - E_0} \\
- \sum_{J \neq 0} \frac{|\langle \Psi_0 \Phi_0 | \mathcal{V} | \Psi_0 \Phi_J \rangle|^2}{F_J - F_0} \\
- \sum_{I \neq 0} \sum_{J \neq 0} \frac{|\langle \Psi_0 \Phi_0 | \mathcal{V} | \Psi_I \Phi_J \rangle|^2}{(E_I - E_0) + (F_J - F_0)}
\]

+ higher order terms

(2.7)

where \(E_{RS}\) is the perturbative interaction energy, \(\Psi_0\) and \(\Phi_0\) are the ground state wavefunctions of the separated molecules with ground state energies \(E_0\) and \(F_0\). \(E_I\) and \(F_J\) are the energies of the excited states \(I\) and \(J\) of molecule 1 and 2, respectively, and all sums are over all excitations.

The first term in Eq. (2.7) is the classical electrostatic interaction between two charge clouds, the next two terms are the expressions for the induction or polarization energy (response of one molecule to a given charge distribution on the other molecule) and the fourth term represents the dispersion energy, which is usually identified with the classical van-der-Waals interaction.

As has been noted by [Claverie 1978], this formulation neglects a few aspects: most importantly, the Hartree product in Eq. (2.6) is not antisymmetric under permutation of electrons from different molecules. As a result, the subspace of the Hilbert space sampled by solutions of the form in Eq. (2.6) is different from the subspace of Pauli-principle-allowed state vectors and consequently Eq. (2.7) yields incorrect results. The usual way to correct this is the introduction of an anti-
symmetrized wavefunction $\mathcal{A} |\Psi\Phi\rangle$ with

$$\mathcal{A} = \sum_{\mathcal{P}} (-1)^{\chi(\mathcal{P})} \mathcal{P}$$

where the sum is over all permutations $\mathcal{P}$ which exchange an arbitrary number of electrons on molecule 1 with the same number of electrons on molecule 2, and the function $\chi(\mathcal{P})$ measures the number of elementary permutations occurring in $\mathcal{P}$. However, anti-symmetrization in this manner does not lead to a closed formula for the interaction. Evaluation of Eq. (2.7) is possible only approximatively, most frequently achieved through a series expansion for $\mathcal{A}$ [Claverie 1978] (cf. [Stone 1996] for a discussion of the various anti-symmetrization recipes). The resulting procedure, called symmetry-adapted perturbation theory (SAPT) has been shown to yield excellent results for the interaction in very small systems, such as the water dimer [Jeziorski and van Hemert 1976], or the He–HF adduct [Moszynski et al. 1994]. Simplified versions of SAPT have been successful in predicting structures of small clusters containing aromatic molecules, e.g., [Brenner et al. 1992] (but comparison of the suggested benzene dimer minimum geometry with the best available ab initio supermolecule calculation in [Karlström et al. 1983] shows a significant discrepancy in the orientation of the molecules).

Despite these successes, this long-known theoretical framework has not led to a corresponding unified numerical methodology. To this day, standard implementations of SAPT have remained patchwork solutions of selected problems and are far from being a usable approach to general nontrivial problems. In particular, attempts to use the SAPT formalism for the prediction of solvation shifts in guest-host-systems have not been successful at all. The set of approximations and algorithms which SAPT introduces, effectively constitute a semi-empirical set of parameters which are optimized for the prediction of interaction energies between small neutral molecules with simple charge distributions. As a general method for the theoretical investigation of nontrivial systems, SAPT has, so far, not succeeded.

$^3(-1)^{\chi(\mathcal{P})}$ is usually called the parity of $\mathcal{P}$. 
With the ever-increasing computational power available in recent years, and the nature of the relevant integrals and advanced methods of their evaluation now well understood [Gill 1994], methods based on brute force calculation of the relevant integrals and/or intelligent selection of the most relevant terms, combined with approximative methods for their evaluation can safely be expected to surpass state-of-the-art SAPT calculations in range of applicability and usefulness in the near future. The necessary steps will involve a reevaluation of the anti-symmetrization procedures described above combined with ab initio evaluation of the resulting terms, thus avoiding the SAPT patchwork.

While a general properly anti-symmetrized ab initio intermolecular/molecular crystal interaction package is not likely to be a readily available option soon, a slightly different approach focusing exclusively on solvatochromistic effects has already been developed by [Shalev et al. 1991] and [Leontidis et al. 1995] into a usable framework, which shall be presented below.

### 2.2.1 The SBEJ method

The RSPT expression for the interaction between guest and host molecules, Eq. (2.7), can of course be written for both the ground state of the guest molecule

\[
E_{RS}^{(g)} = \langle \Psi_0 \Phi_0 \vert \mathcal{V} \vert \Psi_0 \Phi_0 \rangle \\
- \sum_{I \neq 0} \frac{\vert \langle \Psi_0 \Phi_0 \vert \mathcal{V} \vert \Psi_I \Phi_0 \rangle \vert^2}{E_I - E_0} \\
- \sum_{J \neq 0} \frac{\vert \langle \Psi_0 \Phi_0 \vert \mathcal{V} \vert \Psi_0 \Phi_J \rangle \vert^2}{F_J - F_0} \\
- \sum_{I \neq 0} \sum_{J \neq 0} \frac{\vert \langle \Psi_0 \Phi_0 \vert \mathcal{V} \vert \Psi_I \Phi_J \rangle \vert^2}{(E_I - E_0) + (F_J - F_0)} \\
+ \text{higher order terms}
\]  

(2.8)
as well as for any excited state $R$ of the guest (assuming the guest’s wave function is $\Psi$):

$$E_{RS}^{(cz)} = \langle \Psi_R \Phi_0 | V | \Psi_R \Phi_0 \rangle$$

$$- \sum_{\substack{I \neq R \neq J \neq 0}} \frac{|\langle \Psi_R \Phi_0 | V | \Psi_I \Phi_J \rangle|^2}{E_I - E_R}$$

$$- \sum_{J \neq 0} \frac{|\langle \Psi_R \Phi_0 | V | \Psi_R \Phi_J \rangle|^2}{F_J - F_0}$$

$$- \sum_{I \neq R} \sum_{J \neq 0} \frac{|\langle \Psi_R \Phi_0 | V | \Psi_I \Phi_J \rangle|^2}{(E_I - E_R) + (F_J - F_0)}$$

$$+ \text{higher order terms} \quad (2.9)$$

Of course, the difference between Eq. (2.9) and Eq. (2.8) is the excitation energy of the combined system. This difference contains zeroth-order terms, which consist of the excitation energy of the unperturbed guest, and terms describing the host molecules and their interactions. The latter obviously cancel each other. If the former is ignored, too, the remaining terms $E_{RS}^{(cz)} - E_{RS}^{(g)}$ represent the solvation shift.

[Leontidis et al. 1995] have shown that for the system s-tetrazine in atactic polypropylene the contributions of the electrostatic and induction energies may be neglected in the calculation of the solvation shift. In less polar systems these terms can safely be expected to be of even lesser importance, and may consequently be neglected. Hence the essential part of the SBEJ concentrated solely on the evaluation of the dispersive contribution to the solvation shift.

First of all, the aliphatic host (an alkane cluster or crystal or a polymer) is considered to be a collection of polarizable methyl and methylene groups. These groups are treated as if they were polarizable rare-gas atoms. The dispersive interaction is

---

4"[They amount on average to] less than three percent of the experimentally observed shift" [Leontidis et al. 1995].
additive [LONDON 1937, MARGENAU 1939], and so the total shift is just the sum of the shifts caused by each group:

$$\Delta E = \sum_A \delta E^{(A)}$$  \hspace{1cm} (2.10)

where $\delta E^{(A)}$ is the shift due to the polarizable group $A$, and the sum is over all groups of the host. Let now the wavefunction of the guest be denoted by $\Psi$ and assume the host group $A$ could be assigned a wavefunction $\Phi^5$. Let likewise the energies in the denominators be denoted by $E$ for the guest and $F$ for the host groups. Then for each host group $A$, the contribution $\delta E^{(A)}$ is given by:

$$\delta E^{(A)} = \sum_{I \neq R} \sum_{J \neq 0} \frac{|\langle \Psi_R \Phi_0 | V | \Psi_I \Phi_J \rangle|^2}{(E_I - E_R) + (F_J - F_0)}$$

$$- \sum_{I \neq 0} \sum_{J \neq 0} \frac{|\langle \Psi_0 \Phi_I | V | \Psi_I \Phi_J \rangle|^2}{(E_I - E_0) + (F_J - F_0)}$$  \hspace{1cm} (2.11)

The only term in $V$ relevant to the spectral shift is the Coulomb repulsion between electrons on guest molecule $M$ and host group $A$:

$$V = \sum_{i \text{ on } M} \sum_{j \text{ on } A} \frac{1}{r_{ij}}$$  \hspace{1cm} (2.12)

Let the molecules now be described by single determinant wavefunctions and the excitations be written in the form $\Psi_R = \Phi^a_r$ for an excitation of an electron from orbital $a$ to orbital $r$ on the guest and analogously $\Phi_J = \Phi^c_t$ for an excitation from orbital $c$ to orbital $t$ on the host group. Let furthermore the excited state of the guest molecule be described simply by $\Phi^a_t$, which is correct for simple HOMO-LUMO transitions$^6$. Then the standard Slater rules for the evaluation of matrix elements

$^5$Strictly spoken, $A$ is a methyl or methylene group on a host molecule and has no separate wavefunction. Since the method will finally represent $A$ only by its position in space, its polarizability and its mean excitation energy, this problem is not a problem of assigning a wavefunction to a part of the host molecule but assigning these physical quantities to a part of the molecule. [LEONTIDIS ET AL. 1995] discuss this question in detail.

$^6$This is not always the case. In linear polyenes, for example, usually two low-lying singlet states are found, one of which corresponds to a two-electron excitation. In this case the wave function
(cf. [SZABO AND OSTLUND 1996]) state that in the dispersive terms of Eq. (2.8) and Eq. (2.9) only those integrals remain, in which $V$ "connects" exactly one single excitation on $M$ with exactly one single excitation on $A$. Thus

\[ E_{\text{disp}}^{(ex)} = - \sum_c \sum_t \left[ \frac{|\langle \Psi_0^e \Phi_0^t | V | \Psi_0^e \Phi_0^t \rangle|^2}{e_a^e + e_c^e} + \sum_s \frac{|\langle \Psi_0^e \Phi_0^t | V | \Psi_0^e \Phi_0^t \rangle|^2}{e_s^e + e_t^e} \right] + \sum_s \frac{|\langle \Psi_0^e \Phi_0^t | V | \Psi_0^e \Phi_0^t \rangle|^2}{e_s^e + e_t^e} \]

\[ + \sum_{b \neq a} \frac{|\langle \Psi_0^e \Phi_0^t | V | A \Psi_{ab}^s \Phi_c^t \rangle|^2}{e_b^e + e_c^e} \]

\[ + \sum_{b \neq a} \frac{|\langle \Psi_0^e \Phi_0^t | V | B \Psi_{ab}^s \Phi_c^t \rangle|^2}{e_b^e + e_c^e} \]

(2.13)

and

\[ E_{\text{disp}}^{(g)} = - \sum_c \sum_t \left[ \frac{|\langle \Psi_0^e \Phi_0^t | V | \Psi_0^e \Phi_0^t \rangle|^2}{e_a^e + e_c^e} + \sum_s \frac{|\langle \Psi_0^e \Phi_0^t | V | \Psi_0^e \Phi_0^t \rangle|^2}{e_s^e + e_t^e} \right] + \sum_s \frac{|\langle \Psi_0^e \Phi_0^t | V | \Psi_0^e \Phi_0^t \rangle|^2}{e_s^e + e_t^e} \]

\[ + \sum_{b \neq a} \frac{|\langle \Psi_0^e \Phi_0^t | V | B \Psi_{ab}^s \Phi_c^t \rangle|^2}{e_b^e + e_c^e} \]

\[ + \sum_{b \neq a} \frac{|\langle \Psi_0^e \Phi_0^t | V | B \Psi_{ab}^s \Phi_c^t \rangle|^2}{e_b^e + e_c^e} \]

(2.14)

where $e_a^e = e_r - e_a = -e_r^a$ is the energy difference between orbitals $r$ and $a$. The fact that $\Psi_{ab}^{rs}$ appears twice in Eq. (2.13), namely as $A \Psi_{ab}^{rs}$ and $B \Psi_{ab}^{rs}$ accounts for the degeneracy of the corresponding irreducible representation of the permutation group $S_4$ [SZABO AND OSTLUND 1996].

The next step consists in the evaluation of the configuration state functions in terms of slater determinants. Algorithms to construct these can be found in [PAUNCZ 1979], and a summary of the results is given in Table (2.7) of [SZABO AND OSTLUND 1996], which contains Slater-determinant representations of doubly-excited singlet

for the excited state and consequently the expression for $E_{\text{disp}}^{(ex)}$ must be modified, cf. [GEISSINGER ET AL. 1998]
CSFs:

\[
|^{1}\Psi_{rr}^{\alpha}\rangle = |\Psi_{a\alpha}^{\alpha}\rangle \quad \text{(2.15a)}
\]

\[
|^{1}\Psi_{rr}^{\alpha\beta}\rangle = \frac{1}{\sqrt{2}}(|\Psi_{a\alpha}^{\alpha\beta}\rangle + |\Psi_{a\alpha}^{\alpha\beta}\rangle) \quad \text{(2.15b)}
\]

\[
|^{1}\Psi_{ab}^{\alpha\beta}\rangle = \frac{1}{\sqrt{2}}(|\Psi_{a\alpha}^{\alpha\beta}\rangle + |\Psi_{b\beta}^{\alpha\beta}\rangle) \quad \text{(2.15c)}
\]

\[
|^{A}\Psi_{ab}^{\alpha\beta}\rangle = \frac{1}{2\sqrt{3}}(2|\Psi_{ab}^{\alpha\beta}\rangle + 2|\Psi_{ab}^{\alpha\beta}\rangle + |\Psi_{a\alpha}^{\alpha\beta}\rangle - |\Psi_{a\alpha}^{\alpha\beta}\rangle - |\Psi_{b\beta}^{\alpha\beta}\rangle - |\Psi_{b\beta}^{\alpha\beta}\rangle) \quad \text{(2.15d)}
\]

\[
|^{B}\Psi_{ab}^{\alpha\beta}\rangle = \frac{1}{2}(|\Psi_{ab}^{\alpha\beta}\rangle + |\Psi_{a\alpha}^{\alpha\beta}\rangle + |\Psi_{b\beta}^{\alpha\beta}\rangle + |\Psi_{b\beta}^{\alpha\beta}\rangle) \quad \text{(2.15e)}
\]

where \( r \) denotes a spin-up electron in orbital \( r \) and \( \bar{r} \) a spin-down electron in orbital \( r \). Hence for example the first integral in Eq. (2.13) becomes:

\[
\langle \Psi_{a}^{\alpha} \Phi_{0} | V | \Psi_{0} \Phi_{c}^{\prime} \rangle = \frac{1}{2} \langle (\Psi_{a}^{\alpha} + \Psi_{a}^{\alpha}) \Phi_{0} | V | \Psi_{0} (\Phi_{c}^{\prime} + \Phi_{c}^{\prime}) \rangle \quad \text{(2.16)}
\]

But

\[
\langle \Psi_{a}^{\alpha} \Phi_{0} | V | \Psi_{0} \Phi_{c}^{\prime} \rangle = \langle \Psi_{a}^{\alpha} \Phi_{0} | V | \Psi_{0} \Phi_{c}^{\prime} \rangle = \langle \Psi_{a}^{\alpha} \Phi_{0} | V | \Psi_{0} \Phi_{c}^{\prime} \rangle = \langle \Psi_{a}^{\alpha} \Phi_{0} | V | \Psi_{0} \Phi_{c}^{\prime} \rangle
\]

\[
= \int dr_{1} \int dr_{2} \psi_{a}(r_{1}) \phi_{c}(r_{2}) \frac{1}{r_{12}} \psi_{a}(r_{1}) \phi_{c}(r_{2})
\]

\[
= \int dr_{1} \int dr_{2} \psi_{a}(r_{1}) \psi_{a}(r_{1}) \frac{1}{r_{12}} \phi_{c}(r_{2}) \phi_{c}(r_{2})
\]

\[
= (ar|ct) \quad \text{(2.17)}
\]

where \( \psi \) and \( \phi \) are just the molecular orbitals (MOs) on the guest and the host and
The SBEJ method

\( (ar|ct) \) is the familiar Coulomb integral [SZABO AND OSTLUND 1996]. Hence:

\[
\langle \Psi^r_a \Phi_0 | \mathcal{V} | \Psi^s_a \Phi_c \rangle = 2 (ar|ct)
\]  \hspace{1cm} (2.18)

Correspondingly the second integral in Eq. (2.13) evaluates to:

\[
\langle \Psi^r_a \Phi_0 | \mathcal{V} | \Psi^s_a \Phi_c \rangle = \frac{1}{2\sqrt{2}} \langle (\Psi^r_a + \Psi^r_a) \Phi_0 | \mathcal{V} | (\Psi^s_a + \Psi^s_a) (\Phi_c + \Phi_c) \rangle
\]  \hspace{1cm} (2.19)

Here, however

\[
\langle \Psi^r_a \Phi_0 | \mathcal{V} | \Psi^s_a \Phi_c \rangle = 0
\]  \hspace{1cm} (2.20)

because this would correspond to a double excitation \( r \rightarrow \bar{a} \) and \( a \rightarrow s \), and so this and all similar matrix elements must vanish. The remaining terms yield:

\[
\langle \Psi^r_a \Phi_0 | \mathcal{V} | \Psi^s_a \Phi_c \rangle = \sqrt{2} (rs|ct)
\]  \hspace{1cm} (2.21)

Summarizing the results, one obtains [LEONTIDIS ET AL. 1995]:

\[
\langle \Psi^r_a \Phi_0 | \mathcal{V} | \Psi^s_a \Phi_c \rangle = 2 (ar|ct)
\]  \hspace{1cm} (2.22a)

\[
\langle \Psi^r_a \Phi_0 | \mathcal{V} | \Psi^s_a \Phi_c \rangle = 2 (ar|ct)
\]  \hspace{1cm} (2.22b)

\[
\langle \Psi^r_a \Phi_0 | \mathcal{V} | \Psi^s_a \Phi_c \rangle = 2 (ar|ct)
\]  \hspace{1cm} (2.22c)

\[
\langle \Psi^r_a \Phi_0 | \mathcal{V} | \Psi^s_a \Phi_c \rangle = \sqrt{2} (ar|ct)
\]  \hspace{1cm} (2.22d)

\[
\langle \Psi^r_a \Phi_0 | \mathcal{V} | \Psi^s_a \Phi_c \rangle = \sqrt{2} (ar|ct)
\]  \hspace{1cm} (2.22e)

\[
\langle \Psi^r_a \Phi_0 | \mathcal{V} | \Psi^s_a \Phi_c \rangle = \sqrt{2} (ab|ct)
\]  \hspace{1cm} (2.22f)

\[
\langle \Psi^r_a \Phi_0 | \mathcal{V} | \Psi^s_a \Phi_c \rangle = \sqrt{2} (rs|ct)
\]  \hspace{1cm} (2.22g)

\[
\langle \Psi^r_a \Phi_0 | \mathcal{V} | \Psi^s_a \Phi_c \rangle = \sqrt{3} (bs|ct)
\]  \hspace{1cm} (2.22h)

\[
\langle \Psi^r_a \Phi_0 | \mathcal{V} | \Psi^s_a \Phi_c \rangle = (bs|ct)
\]  \hspace{1cm} (2.22i)

\[
\langle \Psi^r_a \Phi_0 | \mathcal{V} | \Psi^s_a \Phi_c \rangle = (bs|ct)
\]  \hspace{1cm} (2.22j)
Insertion of these relations into Eq. (2.13) and Eq. (2.14) yields:

\[
E^{(ex)}_{\text{disp}} = - \sum_{c} \sum_{t} \left[ \frac{4 (ar|ct)^2}{e^c_r + e^c_t} + 2 \sum_{s \neq r} \frac{(rs|ct)^2}{e^s_r + e^s_t} \right]
+ 2 \sum_{b \neq a} \frac{(ab|ct)^2}{e^b_r + e^b_t} + 2 \sum_{b \neq a} \frac{(br|ct)^2}{e^b_r + e^b_t}
+ 3 \sum_{b \neq a} \sum_{s \neq r} \frac{(bs|ct)^2}{e^b_r + e^b_t}
+ \sum_{b \neq a} \sum_{s \neq r} \frac{(bs|ct)^2}{e^b_r + e^b_t}
\] (2.23)

and

\[
E^{(g)}_{\text{disp}} = - \sum_{c} \sum_{t} \left[ \frac{4 (ar|ct)^2}{e^c_r + e^c_t} + 4 \sum_{s \neq r} \frac{(as|ct)^2}{e^s_r + e^s_t} \right]
+ 4 \sum_{b \neq a} \frac{(br|ct)^2}{e^b_r + e^b_t} + 4 \sum_{b \neq a} \frac{(bs|ct)^2}{e^b_r + e^b_t}
\] (2.24)

Thus, the contribution of host group \( A \) to the dispersive shift is finally given by:

\[
\delta E(A) = - 2 \sum_{c} \sum_{t} \left[ 2 \frac{(ar|ct)^2}{e^c_r + e^c_t} + 4 \sum_{s \neq r} \frac{(as|ct)^2}{e^s_r + e^s_t} \right]
- \sum_{s \neq r} \frac{(as|ct)^2}{e^s_r + e^s_t}
+ \sum_{b \neq a} \frac{(ab|ct)^2}{e^b_r + e^b_t} - \sum_{b \neq a} \frac{(br|ct)^2}{e^b_r + e^b_t}
\] (2.25)

The evaluation of the Coulomb integrals in Eq. (2.25) then proceeds using a distributed multipole expansion for the interaction operator \( 1/r_{12} \), where the transition dipole of the guest is represented by a collection of transition monopoles located at the atoms of \( M \). This expansion is conveniently truncated after the monopole-dipole term.
First, the Coulomb integrals are expressed in terms of the atomic orbitals. One molecular orbital is of course a linear combination of several atomic orbitals, so that

$$\psi_a(r_1) = \sum_{\alpha \text{ on } M} C_{\alpha \alpha} \chi_a(r_1)$$
or, regrouped,

$$\psi_a(r_1) = \sum_{m \text{ on } M} \sum_{\mu \text{ on } m} C_{\mu \alpha} \chi_{\alpha}(r_1).$$

Consequently

$$(ar|ct) = \sum_{m \text{ on } M} \sum_{\mu \text{ on } m} \sum_{\nu \text{ on } n} \sum_{\rho \text{ on } n} C_{\mu \nu} C_{\rho \sigma}(\mu \nu | ct)$$

$$\approx \sum_{m \text{ on } M} \sum_{\mu \text{ on } m} C_{\mu \nu} C_{\rho \sigma}(\mu \nu | ct)$$

(2.26)
since the overlap between atomic orbitals on different atoms $m, n$ on $M$ is small, and may thus be neglected. This approximation allows the expansion of the interaction in terms of distributed monopoles located on all atoms on $M$.

Figure 2.2: Multipole Expansion. Definition of the vectors used in the text. $\vec{R}_m$ is the position of atom $m$ on the guest, $\vec{R}_A$ the position of the host group $A$. $\xi_1$ and $\xi_2$ are the expansion variables, hence $1/|\vec{r}_1 - \vec{r}_2| \approx 1/|\vec{R}_m| + (\xi_1 - \xi_2) \cdot \vec{R}_m A/|\vec{R}_m A|^3$. 
In the next step, \(1/r_{12}\) is expanded in a Taylor series according to Fig. (2.2), with the origin set to \(\vec{R}_m\), hence \(\vec{r}_1 = \vec{\xi}_1\), \(\vec{r}_2 = \vec{R}_mA + \vec{\xi}_2\), and:

\[
V = \frac{1}{r_{12}} = \frac{1}{|\vec{r}_2 - \vec{r}_1|} = \frac{1}{\sqrt{(\vec{r}_2 - \vec{r}_1)^2}}
\]
\[
= \frac{1}{\sqrt{(R_{mA} + \xi_2 - \xi_1)^2}}
\]
\[
= \frac{1}{\sqrt{(R_{mA x} + \xi_2 x - \xi_1 x)^2 + (R_{mA y} + \xi_2 y - \xi_1 y)^2 + (R_{mA z} + \xi_2 z - \xi_1 z)^2}}
\]

(2.27)

The Taylor expansion is truncated after the first-order terms:

\[
V(\vec{\xi}_1, \vec{\xi}_2) = V(0) + \xi_{1x} \frac{\partial V}{\partial \xi_{1x}} \bigg|_0 + \xi_{1y} \frac{\partial V}{\partial \xi_{1y}} \bigg|_0 + \xi_{1z} \frac{\partial V}{\partial \xi_{1z}} \bigg|_0 + \xi_{2x} \frac{\partial V}{\partial \xi_{2x}} \bigg|_0 + \xi_{2y} \frac{\partial V}{\partial \xi_{2y}} \bigg|_0 + \xi_{2z} \frac{\partial V}{\partial \xi_{2z}} \bigg|_0 + \ldots
\]

(2.28)

from which

\[
\frac{1}{r_{12}} = \frac{1}{|\vec{R}_mA|} + \frac{1}{|\vec{R}_mA|^3} (\vec{\xi}_1 - \vec{\xi}_2) \cdot \vec{R}_mA
\]

(2.29)

immediately follows. This is inserted in Eq. (2.26):

\[
(\mu\nu|\chi) = \left( \chi_{\mu}(r_1) \phi_c(r_2) \right) \left( \frac{1}{R_{mA}} + \frac{1}{R_{mA}^3} (\vec{\xi}_1 - \vec{\xi}_2) \cdot \vec{R}_mA \right) \left( \chi_{\nu}(r_1) \phi_t(r_2) \right)
\]
\[
+ \left( \chi_{\mu}(r_1) \right) \left( \frac{\vec{\xi}_1 \cdot \vec{R}_mA}{R_{mA}^3} \right) \left( \chi_{\nu}(r_1) \right) \left( \phi_c(r_2) \phi_t(r_2) \right)
\]
\[
- \left( \phi_c(r_2) \right) \left( \frac{\vec{\xi}_2 \cdot \vec{R}_mA}{R_{mA}^3} \right) \left( \chi_{\nu}(r_1) \right) \left( \chi_{\nu}(r_1) \right)
\]

(2.30)

\(\phi_c\) and \(\phi_t\) are two different molecular orbitals, hence \(\langle \phi_c(r_2) | \phi_t(r_2) \rangle = 0\). The atomic orbitals \(\chi_{\mu}\) and \(\chi_{\nu}\) however are, in general, not orthogonal; their overlap integral
The SBEJ method

\( S_{\mu \nu} = \langle \chi_{\mu}(r_1) | \chi_{\nu}(r_1) \rangle \) is, in general, nonzero\(^7\). Hence, with \( \langle \phi_c(r_2) | \tilde{\xi}_2 | \phi_t(r_2) \rangle = \tilde{\mu}_c^t \)

\[
(\mu \nu | ct) = -\frac{\tilde{\mu}_c^t \cdot \tilde{R}_{mA}^3}{R_{mA}^3} S_{\mu \nu}
\]  

(2.31)

and accordingly

\[
(\alpha r | ct) \approx \sum_{m \text{ on } M} \sum_{\mu, \nu \text{ on } m} \frac{\tilde{\mu}_c^t \cdot \tilde{R}_{mA}^3}{R_{mA}^3} C_{\alpha \mu} C_{\nu \nu} S_{\mu \nu}
\]

(2.32)

which for convenience may be rewritten in terms of the transition monopoles

\[
Q_{\alpha r}^m = \sum_{\mu, \nu \text{ on } m} C_{\alpha \mu} C_{\nu \nu} S_{\mu \nu}
\]

(2.33)

as

\[
(\alpha r | ct) \approx \sum_{m \text{ on } M} \frac{\tilde{\mu}_c^t \cdot \tilde{R}_{mA}^3}{R_{mA}^3} Q_{\alpha r}^m
\]

(2.34)

Hence the terms in Eq. (2.23) and Eq. (2.24) can be evaluated according to

\[
-4 \sum_{\alpha, \nu} \sum_{\nu} \frac{(\alpha r | ct)^2}{\epsilon_{\alpha} + \epsilon_{\nu}} \approx -4 \sum_{\alpha, \nu} \sum_{\nu} \frac{\tilde{\mu}_c^t \cdot \tilde{\mu}_c^t}{\epsilon_{\alpha} + \epsilon_{\nu}} \sum_{m \text{ on } M} \sum_{n \text{ on } M} \frac{\tilde{R}_{mA} \cdot \tilde{R}_{nA}}{R_{mA}^3 R_{nA}^3} Q_{\alpha r}^m Q_{\alpha r}^n
\]

(2.35)

It is customary to introduce at this point the mean excitation energy approximation\(^8\), which assumes that in the standard expression for the polarizability [ATKINS AND FRIEDMAN 1997]

\[
\alpha = \frac{2}{3} \sum_{n \neq 0} \frac{|\tilde{\mu}_{n0}|^2}{\Delta E_{n0}}
\]

(2.36)

\(^7\)The \( S_{\mu \nu} \) can easily be calculated with any standard quantum chemistry software package.

\(^8\)Sometimes, the mean excitation energy approximation is called the closure approximation [ATKINS AND FRIEDMAN 1997].
2. Theory

the denominators in the summation may be replaced by a suitable average value⁹, yielding

\[ \alpha = \frac{2}{3\Delta E_{n0}} \sum_{n\neq 0} |\bar{\mu}_{n0}|^2 \]  (2.37)

Usually this average is taken to be the ionization potential, but this is not necessarily the best choice. The same approach is to treat the average excitation energy as an adjustable parameter and write it as \( \eta I_A \). Hence¹⁰

\[ \sum_{c}^{\text{occ}} \sum_{t}^{\text{virt}} \bar{\mu}_c^t \cdot \bar{\mu}_c^t = \sum_{n\neq 0} |\bar{\mu}_{n0}|^2 = \eta I_A \alpha_A \]  (2.38)

Then Eq. (2.35) becomes (with all prefactors included in \( \eta \)):

\[-4 \sum_{c}^{\text{occ}} \sum_{t}^{\text{virt}} \frac{(ar|ct)^2}{e_a^t + e_c^t} \approx -\eta I_A \alpha_A \sum_{c}^{\text{occ}} \sum_{t}^{\text{virt}} \sum_{m,n \text{ on } M} G_{mn}^{A} Q_{ar}^m Q_{ar}^n \]  (2.39)

where the geometric factor

\[ G_{mn}^{A} = \frac{\vec{R}_{mA} \cdot \vec{R}_{nA}}{R_{mA}^3 R_{nA}^3} \]  (2.40)

is introduced as a shortcut. Now the mean excitation energy approximation is applied again, this time to the denominator in the last term on the right hand side. To be consistent with the notation in the reference literature [LEONTIDIS ET AL. 1995, WALLENBORN ET AL. 1997, GEISSINGER ET AL. 1998], in the summation \( e_c^t \) is replaced by \( \kappa I_A \) and the second adjustable parameter \( \kappa \) is also included in the first occurrence of \( I_A \). Thus

\[-4 \sum_{c}^{\text{occ}} \sum_{t}^{\text{virt}} \frac{(ar|ct)^2}{e_a^t + e_c^t} \approx -\eta \kappa I_A \alpha_A \sum_{m,n \text{ on } M} G_{mn}^{A} Q_{ar}^m Q_{ar}^n \]  (2.41)

⁹It is a trivial result of integral calculus that such a value must exist. Of course it does not have to be linked in any way to the ionization potential, it does not have to be the same value if the approximation is made several times, and it does not have to be constant under variation of the guest-host geometries.

¹⁰The ionization potential \( I_A \) appears here simply to keep continuity with the literature. \( I_A \) could easily be included in \( \eta \), without loss of generality or physical relevance.
After all terms in Eq. (2.25) are evaluated in this manner, the contribution of host group A to the dispersive shift is found to be

$$\delta E^{(A)} = -\eta \kappa I_A \alpha_A \sum_{m,n} K_{mn}^A G_{mn}$$

(2.42)

where the electronic factor $K_{mn}^A$ is

$$K_{mn}^A = d_{ar}^m + \frac{1}{2} \left[ \sum_{b \neq a} \left( d_{ab}^m - d_{rb}^n \right) + \sum_{s \neq r} \left( d_{sr}^m - d_{sa}^n \right) \right]$$

(2.43)

and

$$d_{ar}^m = \frac{Q_{ar}^m Q_{ar}^n}{e_a^r + \kappa I_A}$$

(2.44)

and so the final expression for the solvation shift is

$$\Delta E = -\sum_A \sum_{m,n \neq a} \eta \kappa I_A \alpha_A K_{mn}^A G_{mn}$$

(2.45)

This expression contains two adjustable parameters, $\eta$ and $\kappa$, due to the double application of the mean excitation energy approximation. It also contains the polarizabilities $\alpha_A$ and ionization potentials $I_A$ of the host groups\textsuperscript{11}. For the evaluation of the shift it is now only necessary to establish a proper geometric relation between guest and host (preferably using classical molecular dynamics) and calculate the molecular orbital coefficients $C_{am}$ and overlap integrals $S_{\mu\nu}$ (with any standard quantum chemistry software).

### 2.2.2 Semiclassical Dynamical Theory of Lineshapes

For cluster spectra it is possible to calculate the dynamical lineshape [Fried and Mukamel 1992, Mukamel 1982, Saven and Skinner 1993]. The basic idea is to

\textsuperscript{11}Strictly spoken, one could view these as additional parameters that allow the host groups to differ. From this point of view, methyl and methylene groups would share the same $\eta$ and $\kappa$, but differ in $\alpha_A$ and $I_A$. Technically these numbers could be included in the $\eta$ and $\kappa$, creating different parameters for different hosts and removing the slight awkwardness of defining an ionization potential for a part of a molecule.
write the absorption energy of the solvated guest as

$$U(\tau) = \Delta E(\tau) + \Delta E_{\text{vac}} - \hbar \omega_0$$

(2.46)

there $\Delta E_{\text{vac}}$ is the excitation energy of the isolated chromophore and $\omega_0$ is an arbitrary reference frequency, which is usually chosen such that the mean of $U(\tau)$ vanishes, $\langle U(\tau) \rangle = 0$ [Mukamel 1995].

In the next step the adiabatic Hamiltonian of the system in absence of any interaction with a radiative field is separated into a (two-level) electronic part and a (multi-state) nuclear part:

$$\mathcal{H}_0 = |g\rangle \mathcal{H}_g(\vec{q}) \langle g| + |e\rangle \mathcal{H}_e(\vec{q}) \langle e|$$

(2.47)

$\mathcal{H}_g(\vec{q})$ and $\mathcal{H}_e(\vec{q})$ are matrix elements in the electronic subspace but still operators in the subspace of the nuclear coordinates $\vec{q}$. The matrix elements of the Hamiltonian in Eq. (2.47) are evaluated by taking the trace over the electronic states analytically and replacing the trace over the states in the nuclear phase space by a classical ensemble average.

The absorption lineshape is proportional to the real part of the Fourier transform of the transition dipole moment autocorrelation function [Kubo and Tomita 1954]:

$$I(\omega) = \frac{1}{2\pi} \text{Re} \int_{-\infty}^{\infty} dt \ e^{-i\omega t} \langle \langle \mu(0)\mu(t) \rangle \rangle$$

(2.48)

The expectation value of the transition dipole moment autocorrelation function is in the Heisenberg picture given by [Mukamel 1995]

$$\langle \langle \mu(0)\mu(t) \rangle \rangle = \langle \langle \mu(0)\mathcal{G}(t)\mu(0)\mathcal{G}(t) \rangle \rangle$$

(2.49)

where the double brackets serve as a reminder that the expectation value has to be calculated as a trace over all nuclear and electronic degrees of freedom. The semiclassical approximation consists of the assumptions that
(1) the time-ordering in the usual definition of the Green function $G$ may be dropped, i.e. that the Born-Oppenheimer surfaces of ground and excited state are parallel and hence $G$ may be replaced by its semiclassical analog $G_0(t)$,

(2) the total Hamiltonian $\mathcal{H}$ in presence of the radiation field can be approximated by the unperturbed Hamiltonian defined in Eq. (2.47), and

(3) the transition dipole moment operator may be approximated by $\mu \approx \mu_0(\langle g | e \rangle + | e \rangle \langle g |)$, where $\mu_0$ is a scalar.

Then the semiclassical form of Eq. (2.49) is

$$\langle \langle \mu(0) | \mu(t) \rangle \rangle \approx \mu_0^2 \left\langle \text{Tr} \left[(\langle g | e \rangle + | e \rangle \langle g |)G_0^t(t)(\langle g | e \rangle + | e \rangle \langle g |)G_0(t)\right] \right\rangle_c$$

(2.50)

with the semiclassical Green function $G_0(t)$ and its adjoint $G_0^t(t)$ defined by:

$$G_0(t) = \Theta(t) \exp \left\{-\frac{i}{\hbar} \int_0^t d\tau \mathcal{H}_0(\tau) \right\}$$

(2.51)

and the initial density operator has been factored into an electronic part $| g \rangle \langle g |$ and a nuclear part, whose trace is replaced by the canonical ensemble average $\langle \ldots \rangle_c$.

The Heaviside function $\Theta(t)$ ensures that $G_0(t)$ vanishes for $t \leq 0$. Hence

$$\langle \langle \mu(0) | \mu(t) \rangle \rangle \approx \mu_0^2 \left\langle \langle g | \left[(\langle g | e \rangle + | e \rangle \langle g |)G_0^t(t)(\langle g | e \rangle + | e \rangle \langle g |)G_0(t)\right] | g \rangle \right\rangle_c$$

$$+ \left\langle \langle e | \left[(\langle g | e \rangle + | e \rangle \langle g |)G_0^t(t)(\langle g | e \rangle + | e \rangle \langle g |)G_0(t)\right] | e \rangle \right\rangle_c$$

$$\approx \mu_0^2 \left\langle \langle e | | G_0^t(t) e \rangle \langle G_0(t) | g \rangle \right\rangle_c$$

$$\approx \mu_0^2 \Theta(t) \left\langle \exp \left\{ \frac{i}{\hbar} \int_0^t d\tau \mathcal{H}_e(\tau) \right\} \times \exp \left\{ -\frac{i}{\hbar} \int_0^t d\tau \mathcal{H}_g(\tau) \right\} \right\rangle_c$$

$$\approx \mu_0^2 \Theta(t) \exp \left\{ i \omega_0 t \right\} \left\langle \frac{i}{\hbar} \int_0^t d\tau U(\tau) \right\rangle_c$$

(2.52)

where $U(\tau)$ and $\omega_0$ are the shift function and the reference frequency defined above.
3. Van-der-Waals Clusters
Of all gas-phase systems, molecular clusters as obtained in supersonic jet-expansion spectroscopy provide the most detailed insight into the origins of spectral shifts. Resonant two-photon two-color ionization (R2P2CI) spectroscopy provides mass-selected absorption spectra, which can be analyzed by theoretical means. Several such studies have been published to date by others [Fried and Mukamel 1992, Shalev et al. 1991, Heidenreich et al. 1994, Heidenreich and Jortner 1996], but all investigate rare-gas clusters, focusing on spectral line shapes in general or trivial isomers.

To apply the method developed above to a nontrivial aromatic guest/alkane host system, the R2P2CI spectra of tetracene in \textit{n}-heptane clusters of different masses as measured by [Ben-Horin et al. 1991] were investigated. The results of this investigation are presented in this chapter.

### 3.1 Tetracene

Van-der-Waals-complexes of tetracene (naphtacene), Fig. (3.1), are a popular toy system for experimental and theoretical spectroscopic investigation. Firstly, clusters are an intermediate between isolated molecules and droplets (and the liquid state in general), thus allowing information about liquid systems to be gained on simplified systems (for example, [Hartmann et al. 1998] recently studied tetracene-argon complexes in helium droplets of approximately 6000 atoms at very low (0.4 K) temperatures). In combination with resonant two-color two-photon ionization (R2C2PI) spectroscopy, mass-resolved spectra of small clusters can be obtained, which increases the amount of information available to develop and test microscopical the-
ories. Specifically, [BEN-HORIN ET AL. 1991] measured the R2P2CI spectra of tetracene in n-heptane clusters, which are shown in Fig. (3.2), cf. [WALLENBORN ET AL. 1997].

![Figure 3.2: Tetracene-heptane cluster spectra. From [BEN-HORIN ET AL. 1991].](image)

3.1.1 Geometry and Vibrational Analysis

Experimental and theoretical geometries of tetracene are compared in Tab. (3.1). The agreement of the geometries is excellent and the \textit{ab initio} geometries could be used to calculate the MO coefficients and overlap integrals of Eq. (2.33) These calculations were performed with the GAMESS(US) package [SCHMIDT ET AL. 1993], and the necessary MO coefficients and overlap integrals were extracted from the dictionary file of GAMESS. The frequencies of the first molecular vibrational modes of tetracene, as well as the irreducible representations they carry, are given
Table 3.1: Tetracene. Comparison of experimental and theoretical geometries.

<table>
<thead>
<tr>
<th>Bond lengths</th>
<th>STO-3G</th>
<th>6-31G*</th>
<th>6-31G**a</th>
<th>Exp. b</th>
</tr>
</thead>
<tbody>
<tr>
<td>C(5a)-C(11a)</td>
<td>1.431</td>
<td>1.428</td>
<td>1.428</td>
<td>1.460</td>
</tr>
<tr>
<td>C(5)-C(5a)</td>
<td>1.408</td>
<td>1.408</td>
<td>1.408</td>
<td>1.404</td>
</tr>
<tr>
<td>C(4a)-C(5)</td>
<td>1.377</td>
<td>1.374</td>
<td>1.373</td>
<td>1.390</td>
</tr>
<tr>
<td>C(4)-C(4a)</td>
<td>1.442</td>
<td>1.444</td>
<td>1.444</td>
<td>1.420</td>
</tr>
<tr>
<td>C(3)-C(4)</td>
<td>1.348</td>
<td>1.342</td>
<td>1.342</td>
<td>1.381</td>
</tr>
<tr>
<td>C(5)-H(5)</td>
<td>1.075</td>
<td>1.077</td>
<td>1.077</td>
<td></td>
</tr>
<tr>
<td>C(4)-H(4)</td>
<td>1.074</td>
<td>1.076</td>
<td>1.076</td>
<td></td>
</tr>
<tr>
<td>C(3)-H(3)</td>
<td>1.073</td>
<td>1.075</td>
<td>1.076</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Bond angles</th>
</tr>
</thead>
<tbody>
<tr>
<td>C(5)-C(5a)-C(11a)</td>
</tr>
<tr>
<td>C(4a)-C(5)-C(5a)</td>
</tr>
<tr>
<td>C(4)-C(4a)-C(5)</td>
</tr>
<tr>
<td>C(3)-C(4)-C(4a)</td>
</tr>
<tr>
<td>C(5a)-C(5)-H(5)</td>
</tr>
<tr>
<td>C(4a)-C(4)-H(4)</td>
</tr>
<tr>
<td>C(4)-C(3)-H(3)</td>
</tr>
</tbody>
</table>

*a* in [Cioslowski et al. 1991]

*b* in [Campbell et al. 1962, Kao and Allinger 1977]
in Tab. (3.4). Fig. (3.3) through Fig. (3.6) visualize the lowest four eigenmodes. The lowest mode is (as might be expected) the "butterfly" ($B_{3u}$) vibration depicted in Fig. (3.3). The second lowest mode is the torsional ($A_u$) mode in Fig. (3.4), followed by the chair-like ($B_{1g}$) vibration in Fig. (3.5). Tetracene also has a low-lying in-plane ($B_{1u}$) mode, which is the fourth lowest overall, shown in Fig. (3.6).

In an environment which does not break the molecular symmetry, all these modes constitute spectroscopically "forbidden" lines. However, in the lower symmetry of a cluster, each may become active. It could be expected that, since the $B_{3u}$ and $B_{1u}$ modes carry of the irreducible representation $A$ of the lowered point group $C_s$ of the cluster, their vibronic lines should be stronger than the other two. This will be discussed below.

![Figure 3.3: Tetracene. Mode 1 (B3LYP/6-31G*). $B_{3u}$ symmetry. 56 cm$^{-1}$.](image)

### 3.1.2 Choice of Parameters

As mentioned above, the choice of the group parameters for methyl and methylene is not based on clean physical grounds. Strictly spoken, neither ionization nor polarization are local phenomena. For a series of alkane homologues, however, the polarizability can be considered group-wise additive and [LEONTIDIS ET AL. 1995] have calculated group increments for the estimation of alkane polarizabilities. This of course does not necessarily mean that the alkane host can be described as a collect-
Figure 3.4: Tetracene. Mode 2 (B3LYP/6-31G*). $A_u$ symmetry. 56 cm$^{-1}$.

Figure 3.5: Tetracene. Mode 3 (B3LYP/6-31G*). $B_{1g}$ symmetry. 148 cm$^{-1}$. 
Figure 3.6: Tetracene. Mode 4 (B3LYP/6-31G*). $B_{1u}$ symmetry. 159 cm$^{-1}$.

tion of polarizable groups where the groups' polarizabilities are calculated according to the additivity property. On the other hand, [Amos and Crispin 1975] have shown that for alkanes above assumption is justified\textsuperscript{1}. [Leontidis et al. 1995] have suggested the values in Tab. (3.2) and these polarizabilities were used here. The concept of localized ionization potentials is even more elusive. But a look at Tab. (3.2) reveals that the numbers suggested by [Leontidis et al. 1995] do not differ by much, and (unsystematic) variation of the ionization potentials shows that the uncertainties in the other approximations (especially the mean excitation energy theorem) are larger than those induced by a variation of the ionization potentials\textsuperscript{2}. The particular choice used here was simply taken from [Leontidis et al. 1995].

Another issue is the choice of $\eta$ and $\kappa$ in Eq. (2.45). As adjustable parameters they may be fitted to actual spectral data. However, this method has drawbacks. The predictive value of the entire calculation is obviously reduced. Worse, since effectively one parameter, $\eta \cdot \kappa$, scales the entire shift, while the other one, $\kappa$, adjusts for the mean excitation approximation in the denominator of Eq. (2.44), there is a very real possibility of serious artefacts in the optimization of $\kappa$ later simply

\textsuperscript{1}For a host with nonlocal polarizabilities, such as naphthalene or p-terphenyl, this is different. Here, a description of the host as collection of polarizable groups located at the atomic positions of the carbon atoms will not work, and either a more elaborate selection of group positions or a brute force approach with direct evaluation of all integrals is necessary.

\textsuperscript{2}If, for example, both numbers are set to 10.0 eV, nearly identical results are obtained.
Table 3.2: Ionization potentials and polarizabilities of methyl and methylene groups.

<table>
<thead>
<tr>
<th>Group</th>
<th>Polarizability/Å</th>
<th>Ionization Potential/eV</th>
</tr>
</thead>
<tbody>
<tr>
<td>CH₂</td>
<td>2.23</td>
<td>11.5</td>
</tr>
<tr>
<td>CH₃</td>
<td>1.85</td>
<td>9.3</td>
</tr>
</tbody>
</table>

compensated for "on average" by $\eta$. Applied to a real system, these parameters would result in large errors. [Geissinger et al. 1998] have argued that $\kappa$ is best left untouched and only $\eta$ adjusted, such that a known spectral position is reproduced exactly.

To investigate this, two different approaches were chosen. The tetracene-heptane calculations were performed with *ab initio* data from RHF/STO-3G and RHF/6-31G* calculations. The parameters used in the STO-3G calculations were $\eta=0.29$ and $\kappa=3.5$ as recommended by [Leontidis et al. 1995]. For the 6-31G* calculations $\eta$ and $\kappa$ were optimized to minimize the total deviation, with respect to the experimental spectra, of the average shifts calculated for the first 100 frames of typical trajectories of tetracene-heptane₁ and tetracene-heptane₅ clusters (obtained as described below). These optimized parameters were $\eta = 0.17199$ and $\kappa = 1.8075$.

3.1.3 Molecular Dynamics

In the case of molecular clusters a sufficient sampling of the phase space is possible. Hence the MD part³ of the tetracene-heptane investigation consisted of two stages⁴:

---

³The MD calculations for the tetracene-heptane system were performed by R. Brown at the University of Bordeaux, using the DL POLY package, version 1.1 [Forester and Smith 1994].

⁴All simulations used the Verlet algorithm with a time step of 3 fs, and a cutoff of 25 Å for the long-range forces. The intermolecular potential was the all-atom Buckingham potential number IV optimized by Williams [Williams 1963] for aromatics and alkanes. All bond lengths were constrained with the Shake algorithm [Allen and Tildesley 1990]. Intramolecular bond bending and torsion potentials for tetracene were taken from the GROMOS87 forcefield [Van Gunsteren and Berendsen 1987]. The geometry and torsional potential of n-heptane were taken from [Ryckaert et al. 1989], in which the original cosine power series for torsions was replaced.
(1) Clusters of different masses were "grown" by simulated annealing from smaller clusters. Starting from tetracene, an additional heptane molecule and the existing cluster were randomly oriented and then launched at each other along the line connecting the centers of mass with an initial relative velocity of 100 m/s, corresponding to a "collision temperature" of about 100 K. The resulting bigger cluster was quenched by coupling to a Berendsen thermostat [BERENDSEN ET AL. 1984] with a coupling time constant of 0.1 ps, successively at 10 K and 1 K, for 9 ps each. The cluster was then minimized by setting the bath temperature to 0 K for a further 18 ps. Finally, convergence of the energy to a minimum, which is sluggish in the later stages of minimization, was accelerated by successive 10- and 100-fold increases of the potential energy function for 18 ps each. In this way, the total potential energy could be made to converge (usually) monotonically to a local minimum, to within 5 cm\(^{-1}\), in 24000 steps.

(2) The classical dynamics of the so obtained clusters was simulated in the canonical ensemble at 30 K. Of each cluster, trajectories of a length of 100 ps were sampled with a sampling interval of 16 fs. The atomic coordinates of these sampled points were then combined with the \textit{ab initio} information to compute \(U(\tau)\) and the autocorrelation function of the transition dipole moment as described in Sec. 2.2. The autocorrelation function was suitably apodized and finally Fourier transformed the obtain the spectrum of the trajectory. The shift and lineshape calculations were done using programs written in C and MATLAB on a Sparc 20 workstation at the ETH Zürich.

### 3.1.4 Cluster Geometries

The number of possible isomers grows rapidly with increasing number of solvating heptane molecules. Even for tetracene-heptane\(_n\) clusters with \(n = 1\) there are at least three local minima. The most stable one is shown in Fig. (3.7) and Fig. (3.8).

The molecular planes are parallel (in the nomenclature of [WALLENBORN ET AL. 1997] this is a "side-on" geometry). Three carbon atoms of the heptane molecule are

by an equivalent Fourier series suitable for DL_POLY. In addition the H–H distance in CH\(_2\) groups was constrained to its value in [RYYKAERT ET AL. 1989].
Figure 3.7: Tetracene-Heptane$_1$. Global minimum of the Buckingham IV potential. Top view.

Figure 3.8: Tetracene-Heptane$_1$. Global minimum of the Buckingham IV potential. Side view.
aligned with three carbon atoms of tetracene. The binding energy for this cluster as calculated with the Buckingham IV force field is $-10.00 \text{kcal/mol.}$

Another two isomers, both less stable by 0.93 kcal/mol and 1.59 kcal/mol, respectively, result when the molecular planes of tetracene and heptane are not parallel but orthogonal. One of them ("feet-on" in [WALLENBORN ET AL. 1997]) is displayed in Fig. (3.9) and Fig. (3.10). The other isomer, in which the heptane is rotated 180° around its long axis with respect to Fig. (3.10), is called "head-on" in [WALLENBORN ET AL. 1997] and Fig. (3.13) shows its LSD/3-21G* minimum.

![Figure 3.9: Tetracene-Heptane\textsubscript{1}. Local minimum of the Buckingham IV potential. Top view.](image)

![Figure 3.10: Tetracene-Heptane\textsubscript{1}. Local minimum of the Buckingham IV potential. Side view.](image)
For systems of this size, quantum-chemical calculations are still feasible. Fig. (3.11) through Fig. (3.13) show results of geometry optimizations on the LSD/3-21G* level. The obtained geometries are very similar to the force field minima. In the global minimum in Fig. (3.11), the same alignment is found as in Fig. (3.7), although here the heptane molecule is moved with respect to Fig. (3.7), such that bonds rather than Carbon atoms are aligned.

Figure 3.11: Tetracene-Heptane. LSD/3-21G* minimum at -958.806199768 au. Binding energy -17.71 kcal/mol, after counterpoise correction: -6.07 kcal/mol. Top view.

The binding energies reported in Fig. (3.11) through Fig. (3.13) are counterpoise corrected [Boys and Bernardi 1970], but may still misestimate the true binding energies by as much as 5 kcal/mol due to the small basis set and the local spin density approximation. In particular, the counterpoise correction produces a positive value for the binding energy of the isomer in Fig. (3.13), suggesting that this particular cluster is in the LSD approximation stable only due to basis set superposition. However, the energetic order of the DFT calculations and the force field calculations agree, and the relative binding energies of the two most stable isomers agree, too. Thus, the Buckingham IV binding energies are most likely correct up to 2-5 kcal/mol.

The second heptane molecule may be added either to the same side or to the opposite side of the tetracene molecule. In the former case a steric collision arises, and it is impossible for the heptane molecules to simultaneously occupy the ideal positions of Fig. (3.11) and Fig. (3.7). As shown in Fig. (3.14), the two heptane
Figure 3.12: Tetracene-Heptane. LSD/3-21G* minimum at -958.800035265 au. Binding energy -13.85 kcal/mol, after counterpoise correction: -4.66 kcal/mol. Side view.

Figure 3.13: Tetracene-Heptane. LSD/3-21G* minimum at -958.790415890 au. Binding energy -7.81 kcal/mol, after counterpoise correction: +2.38 kcal/mol. Side view.
molecules may avoid the steric problem by a rotation around the normal of the molecular plane, such that the interaction between each heptane and the tetracene is reduced. Here, the isomer which has heptane molecules both above and below the tetracene has a binding energy of $-20.29\text{kcal/mol}$, whereas the other isomer lies $-0.49\text{kcal/mol}$ higher. Since there is a positive correlation between binding interaction and shift [WALLENBORN ET AL. 1997], a smaller solvatochromic shift is expected for this isomer (this will be discussed in detail below).

![Tetracene-Heptane2. Local minimum of the Buckingham IV potential.](image)

Fig. (3.15) shows a typical geometry of tetracene-heptane6, after the completion of the first solvation shell. Six heptane molecules are just enough to fill the available space around one tetracene molecule. As can be seen in the figure, there appears to be considerable rotational freedom for the heptane molecules, hence many isomers (or even a broad continuum) are expected. Also, the increase of the spectral shift on addition of another molecule is expected to drop sharply after the completion of the first solvation shell. This is in agreement with Fig. (3.2).

### 3.1.5 Spectra

The solvation shifts of the clusters up to tetracene-heptane$_n$ with $1 \leq n \leq 12$, calculated as described above, are shown in Tab. (3.3). The shift of the most stable structure is given up to $n = 9$. For $n \geq 10$ the annealed clusters, quenched as
Figure 3.15: Tetracene-Heptane$_6$. Typical configuration of the Buckingham IV potential.
Table 3.3: Solvation shifts of tetracene-heptane\textsubscript{n} clusters. All numbers in cm\textsuperscript{-1}.

<table>
<thead>
<tr>
<th>(n)</th>
<th>Exp.</th>
<th>STO-3G</th>
<th>6-31G*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>251</td>
<td>238</td>
<td>224</td>
</tr>
<tr>
<td>2</td>
<td>421</td>
<td>391</td>
<td>381</td>
</tr>
<tr>
<td>2</td>
<td>498</td>
<td>479</td>
<td>451</td>
</tr>
<tr>
<td>3</td>
<td>654</td>
<td>633</td>
<td>611</td>
</tr>
<tr>
<td>4</td>
<td>812</td>
<td>798</td>
<td>774</td>
</tr>
<tr>
<td>5</td>
<td>814</td>
<td>807</td>
<td>785</td>
</tr>
<tr>
<td>5</td>
<td>829</td>
<td>848</td>
<td>828</td>
</tr>
<tr>
<td>6</td>
<td>850</td>
<td>834</td>
<td>816</td>
</tr>
<tr>
<td>6</td>
<td>1054</td>
<td>982</td>
<td>973</td>
</tr>
<tr>
<td>7</td>
<td>1059</td>
<td>1013</td>
<td>1011</td>
</tr>
<tr>
<td>8</td>
<td>1063</td>
<td>969</td>
<td>954</td>
</tr>
<tr>
<td>9</td>
<td>1105</td>
<td>992</td>
<td>977</td>
</tr>
<tr>
<td>10</td>
<td>1087</td>
<td>1043</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>1065</td>
<td>1040</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>1081</td>
<td>1089</td>
<td></td>
</tr>
</tbody>
</table>

\(a\) \(n \leq 9\): Shift of the most stable structure. \(n > 9\): Most probable shift.
described above were heated up to 80 K, equilibrated and then cooled back to 0 K. This was repeated several times and the final three rows of Tab. (3.3) contain the mean of these equilibrated shifts.

As can be seen from Tab. (3.3), the numbers from the STO-3G calculations with $\eta$ and $\kappa$ taken from [LEONTIDIS ET AL. 1995] in general fit the measured data better than those from calculations with the much larger basis set 6-31G* and optimized parameters. As discussed in Sec. 3.1.2, this is a strong indication that the algorithm used in the 6-31G* calculations to determine the adjustable parameters is suboptimal.

![Figure 3.16: Tetracene-Heptane$_n$. Real part of the averaged transition dipole autocorrelation function. The inset shows the short time behavior. Ab initio data: RHF/6-31G*](image)

The calculated transition dipole moment autocorrelation function decays faster for larger clusters (Fig. (3.16)), indicating an increase in the corresponding homo-
geneous linewidths. This is to be expected, as the number of independent contributions to the shift increases. The calculated decay times are, however, not short enough to explain the total linewidth of the heavier clusters: inhomogeneous broadening must be present. The data in Fig. (3.16) are averaged over all calculated spectra, decay times of the correlation functions of individual trajectories of isomers were much longer, typically tens of picoseconds, and also much longer than those of perylene/Ar clusters calculated by [HEIDENREICH ET AL. 1994]. This difference may be understood considering the that although the number of perturbing atoms is comparable, here their contributions to the shift are correlated by the rigidity of the heptane molecule. Additionally, for \( n = 1-3 \) the function is modulated with a period of approximately 0.4 ps (tetracene-heptane, which would correspond to a small sideband about 83 cm\(^{-1}\) from the center of the absorption line. This is most likely the effect of one of the intermolecular modes at 74 cm\(^{-1}\) and 88 cm\(^{-1}\) (Tab. (3.4)).

On the following pages, all calculated subspectra, their sums and comparisons with the measured spectra are presented. Since the lineshapes of the calculated 6-31G* and STO-3G results are almost identical and the only difference is the overall magnitude of the shift, only the former are shown. The tetracene-heptane\(_5\) cluster was a data point used in the determination of the parameters \( \eta \) and \( \kappa \), so it is clear that for this cluster the calculated shift reproduces the measured shift more accurately than for others. It should be noted that this is not the case for the other data point, \( n = 1 \), indicating again that the method used in the optimization of the parameters for the RHF/6-31G* wave function is not optimal.

Overall, the algorithm is surely a gross simplification of the real super-sonic jet, but overall the shifts and lineshapes are reproduced well, with the exception of \( t \cdot \text{h}_4 \). The discrepancy for \( n = 6 \) is likely to be caused by neglecting the clusters with one heptane in the second shell, which would produce shifts similar to those of the lowest \( t \cdot \text{h}_5 \) clusters.

\(^5\)A decay time of 10 ps (\( n = 1 \)) corresponds to a homogeneous linewidth of approximately 3 cm\(^{-1}\), whereas a decay time of 1 ps (\( n = 6 \)) translates into a linewidth of 33 cm\(^{-1}\).
Figure 3.17: Tetracene-Heptane$_1$. All calculated subspectra. The bottom curve is the measured spectrum [BEN-HORIN ET AL. 1991], the curve right above the measured spectrum is the sum of all subspectra. \textit{Ab initio} data: RHF/6-31G*. 

Excitation Energy in cm$^{-1}$
Figure 3.18: Tetracene-Heptane$_2$. All calculated subspectra. The bottom curve is the measured spectrum [BEN-HORIN ET AL. 1991], the curve right above the measured spectrum is the sum of all subspectra. \textit{Ab initio} data: RHF/6-31G*.
Figure 3.19: Tetracene-Heptane$_3$. All calculated subspectra. The bottom curve is the measured spectrum [Ben-Horin et al. 1991], the curve right above the measured spectrum is the sum of all subspectra. *Ab initio* data: RHF/6-31G*.
Figure 3.20: Tetracene-Heptane$_4$. All calculated subspectra. The bottom curve is the measured spectrum [BEN-HORIN ET AL. 1991], the curve right above the measured spectrum is the sum of all subspectra. *Ab initio* data: RHF/6-31G*.
Figure 3.21: Tetracene-Heptane$_5$. All calculated subspectra. The bottom curve is the measured spectrum [BEN-HORIN ET AL. 1991], the curve right above the measured spectrum is the sum of all subspectra. *Ab initio* data: RHF/6-31G*.
Figure 3.22: Tetracene-Heptane$_6$. All calculated subspectra. The bottom curve is the measured spectrum [Ben-Horin et al. 1991], the curve right above the measured spectrum is the sum of all subspectra. *Ab initio* data: RHF/6-31G*.
Tetracene-heptane$_1$

Fig. (3.23) displays the measured spectrum of tetracene-heptane$_1$ in greater detail. Several lines are well resolved and may be assigned to intermolecular vibronic transitions. The numbers representing tetracene-heptane$_1$ in Tab. (3.3) are for the side-on isomer. The other two isomers have shifts that are 32 cm$^{-1}$ (feet-on) and 44 cm$^{-1}$ (head-on) smaller (6-31G*). These two isomers may correspond to features 2 and 3 of the spectrum in Fig. (3.23), but most sidebands in Fig. (3.23) must be vibronic.

Tab. (3.4) is an attempt to assign the observed sidebands 1–10 (11 in Fig. (3.4) is the vacuum line of bare tetracene). As discussed on page 36, the $B_{3u}$ and the $B_{1u}$ modes are expected to be stronger than the other two of the four low-frequency
Table 3.4: Tetracene-heptane. Measured and calculated frequencies of the first vibrational modes.

<table>
<thead>
<tr>
<th>Tetracene-heptane</th>
<th>Tetracene</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exp.</td>
<td>Force Field</td>
</tr>
<tr>
<td>9</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>16</td>
</tr>
<tr>
<td>32</td>
<td></td>
</tr>
<tr>
<td>47</td>
<td>44</td>
</tr>
<tr>
<td>62</td>
<td></td>
</tr>
<tr>
<td>74</td>
<td>75</td>
</tr>
<tr>
<td>88</td>
<td>80</td>
</tr>
<tr>
<td>93</td>
<td>96</td>
</tr>
<tr>
<td>128</td>
<td>123</td>
</tr>
<tr>
<td>166</td>
<td></td>
</tr>
<tr>
<td>159</td>
<td>190 (B\textsubscript{2g})</td>
</tr>
<tr>
<td>257</td>
<td>267 (B\textsubscript{3u})</td>
</tr>
<tr>
<td>286</td>
<td>296 (B\textsubscript{3g})</td>
</tr>
</tbody>
</table>

\textsuperscript{a} From Fig. (3.2), [BEN-HORIN ET AL. 1991]

\textsuperscript{b} Buckingham IV force field
eigenmodes of tetracene. While the $B_{1u}$ mode at 159 cm$^{-1}$ may be identified with signal 8, the $B_{1g}$ mode at 148 cm$^{-1}$ is clearly absent from the spectrum. The $B_{3u}$ mode might be peak 3 (then the two isomers would not be 2 and 3, but 1 and 2. If that assignment was correct, lines at $21 + 56 = 77$ cm$^{-1}$ and $32 + 56 = 88$ cm$^{-1}$ should be present in the spectrum. Indeed, 4 and 5 could be just these lines, but then the measured peak at 93 cm$^{-1}$ would be the $A_u$ mode (predicted to be at 89 cm$^{-1}$). This would not be very convincing, because the DFT calculations on bare tetracene are expected to overpredict the frequencies. Hence the calculations on bare tetracene give some hints but not a definitive assignment of the measured signal.

The vibrational analysis of the entire cluster in the classical force field is somewhat more successful in assigning the peaks 4–7. However, the low-frequency vibrations predicted by the Buckingham IV frequency analysis of the cluster are clearly absent. Obviously the cluster is a region of the Born-Oppenheimer surface, where the harmonic approximation breaks down at least for these modes.

**Tetracene-heptane$_2$**

The spectrum of tetracene-heptane$_2$, Fig. (3.24), shows that this cluster consists of two isomers with peaks at approximately 21900 cm$^{-1}$ and 21980 cm$^{-1}$. As discussed on page 43, the signal with the larger redshift (at 21900 cm$^{-1}$) is presumably produced by an isomer where the two heptane molecules are on opposite sides of the tetracene molecule, whereas the other peak is likely to result from an isomer with both heptane molecules on the same side of the tetracene.

Furthermore, each of the two heptane molecules may be in any of the positions discussed in Sec. 3.1.4 relative to tetracene, creating a multitude of possible isomers. Those with low energies, however, will be most prominently visible in the spectrum$^6$. In Fig. (3.25), which is an analysis of the inhomogeneous structure of

$^6$This argument has to be applied with care. Clusters in supersonic jet expansions are by no means equilibrated ensembles. It was shown in Tab. (3.3) above that for large clusters the spectral shift of the energetic minimum deviates significantly from the position of the measured signal. On the other hand, for tetracene-heptane$_1$, the signals in the observed spectrum seem to agree with the energetic order of the three isomers. For the intermediate tetracene-heptane$_2$ the validity of the argument may tentatively be assumed.
Figure 3.24: Tetracene-Heptane$_2$. Dotted: Measured Spectrum [BEN-HORIN ET AL. 1991]. Solid: Calculated spectrum, shifted by 58 cm$^{-1}$ to facilitate the comparison. Ab initio data: RHF/6-31G*.
Figure 3.25: Tetracene-Heptane$_2$. (1|1) and (2|0) contribution to the total spectrum. The inset shows the contribution of different isomers to the (1|1) and (2|0) subspectra. *Ab initio* data: RHF/6-31G*.
the tetracene-heptane$_2$ spectrum, isomers with both heptane molecules on different sides of tetracene are denoted (1|1), those with both heptanes on the same side (2|0). The letters "s", "f", and "h" describe "side-on", "feet-on", and "head-on" heptane orientations relative to tetracene. As Fig. (3.25) demonstrates, the initial assignment of the peak at 21900 cm$^{-1}$ as originating from a multiplet of (1|1) isomers, seems correct. This is in agreement with [Heidenreich et al. 1994, Ben-Horin et al. 1991]. Both regions consist of different isomers with side, head and feet-on isomer spectra shown in the inset. These produce the structure in the center of the absorption line. Vibronic lines are present in the observed spectrum, too, but are, except for the wings, not pronounced and not reproduced by the calculation$^7$.

The transition dipole moment autocorrelation functions of the isomers of tetracene-heptane$_2$ are shown in Figure 3.26. In accordance with the preceding sections the decay time is shorter than in the $n = 1$ case, leading to a broader

$^7$In the tetracene-heptane$_n$ analysis, information for the assignment of the vibronic lines was not obtained from the shift calculation itself, either, but rather from conventional vibrational analysis.
band. Furthermore, the initial decay rate is larger for the (1|1) isomers, reflecting the broader bands on the blue side of the spectrum. The broad, structureless wings to the red of the 0-0 line in most spectra may be the contribution of absorption by "hot" clusters.

**Tetracene-heptane$_3$—6**

![Figure 3.27: Tetracene-Heptane$_3$. Dotted: Measured Spectrum [Ben-Horin et al. 1991]. Solid: Calculated spectrum, shifted by 55 cm$^{-1}$ to facilitate the comparison. Ab initio data: RHF/6-31G*.](image)

The simulated annealing approach to generate the clusters samples only physically relevant parts of the coordinate space. Thus, not all possible isomers of e.g. tetracene-heptane$_2$ have to be computed in the simulation displayed in Fig. (3.24)
and Fig. (3.25), and the relative intensities of the inhomogeneous contributions to
the overall absorption spectrum are at least qualitatively correct.

![Excitation Energy in cm⁻¹](image)

Figure 3.28: Tetracene-Heptane₄. Dotted: Measured Spectrum [BEN-HORIN ET AL. 1991]. Solid: Calculated spectrum, shifted by 62 cm⁻¹ to facilitate the comparison. \textit{Ab initio} data: RHF/6-31G*.

With growing cluster size the number of possible isomers grows as well. At the same time, the free energy differences of these isomers becomes smaller. Thus, the performance of the simulated annealing is expected to deteriorate for heavier clusters. Fig. (3.27) through Fig. (3.29) demonstrate this. The spectrum of tetracene-heptane₃ (Fig. (3.27)) is well reproduced, although the ratio of peak intensities is only qualitatively correct. For tetracene-heptane₅, the heaviest cluster whose spectrum could be calculated correctly, the calculation gives a good estimate for the inhomogeneous linewidth, but does not reproduce the inhomogeneous features.
Figure 3.29: Tetracene-Heptanes. Dotted: Measured Spectrum [Ben-Horin et al. 1991]. Solid: Calculated spectrum. Ab initio data: RHF/6-31G*.
Tetracene-heptane$_6$ (Fig. (3.30)) seems to be a special case. The observed spectrum is completely different from all others. Broad and diffusely structured, it seems to be a link between the small ($n \leq 5$) clusters, which exist in finite numbers of isomers and well defined geometries, and the large ($n \geq 7$) clusters, which may be considered droplets rather than clusters.


For tetracene-heptane$_6$, a defined cluster geometry still exists in the simulation (Fig. (3.15)), but it seems that the heptane molecules cannot move individually, but rather the solvation shell has to rearrange as a whole, creating entirely different local geometries and affecting the spectrum drastically. This is a possible explanation for
the strange shape of the observed spectrum, as well as the failure of the theory to reproduce even the most fundamental features.
4. Shpol’skii Systems
Shpol'skii systems present ideal systems for the investigation of low-temperature spectra in solid-state guest-host systems. First observed independently by [Bowen and Brocklehurst 1955] and [Shpol'skii and Klimova 1959], they are fast-cooled (usually, but not exclusively) n-alkane systems, doped with small amounts of fluorescing impurities. These impurities, usually polycondensed aromatic hydrocarbons, like pyrene, perylene, terrylene, coronene, and their derivatives, exhibit sharply defined fluorescence spectra with easily distinguishable and clearly separated features [Nakhimovsky et al. 1989]. Usually several spectral sites can be clearly identified by their vibronic progressions. The narrow fluorescence lines and the small number of spectral sites are usually attributed to the special way the guest molecules are inserted in Shpol'skii systems. Removal of a small number of host molecules from the crystal lattice creates a pocket, in which the guest may be included. The number of possible pockets as well as the number of favorable insertion modes is obviously small, and due to the fact that the guest molecules are surrounded by almost identical environments inhomogeneous broadening is virtually absent.

Investigations of Shpol'skii systems have concentrated on many of their special properties: The near-absence of inhomogeneous broadening besides that induced by crystal defects and similar perturbations, as well as the presence of zero-phonon lines at low temperatures have made Shpol'skii systems the most popular object of hole-burning spectroscopy [Moerner 1988].

Single-molecule spectroscopy [Basc hé et al. 1997] at first also focused on Shpol'skii systems. Shpol'skii systems were used to investigate molecular switching behavior [Moerner et al. 1994], and for experiments using polarized light to determine the orientation of the guest molecules within the host crystal [Göttler et al. 1993].

The spectral sites in Shpol'skii systems themselves are most easily identified and distinguished by their typical patterns in Total Luminescence Spectra (TLS) [Palewska et al. 1991], where the fluorescence emission wavelength of a Shpol'skii system is plotted against the excitation wavelength, creating a two-dimensional spectrum in which the sites are easily identified by their characteristic
wedge-like patterns. Information about the vibronic structure of the absorption and fluorescence lines are also easily extracted from the TLS.

4.1 Perylene as guest

Perylene (dibenz[de,kl]anthracene, Fig. (4.1)) is the smallest non-trivial compound in a series of pericondensed polynuclear aromatic hydrocarbons: naphtalene, perylene, terrylene, quaterrylene . . . , whose syntheses (e.g. [BUCHTA ET AL. 1957]) and crystal structures (e.g. [SHRIVASTAVA AND SPEAKMAN 1960]) have been known for about forty years. The fact that these compounds are by no means aromatic throughout the entire annealed system, but rather consist of naphthalenic subunits connected by what are essentially single bonds, was also known from the first successful crystal structure analysis [DONALDSON ET AL. 1953], and was soon afterwards studied theoretically using Hückel [GOODWIN 1960] and related Molecular Orbital [COULSON AND SKANCKE 1962] methods.

4.1.1 Crystal structure

The crystal structure of perylene was resolved first by [DONALDSON ET AL. 1953] and later refined by [CAMERMAN AND TROTTER 1964]. The crystal cell parameters
Table 4.1: Perylene. Measured crystal cell data. Space group: \( P_{21}/a \left( C_{2h}^5 \right) \), monoclinic, 4 molecules per unit cell.

<table>
<thead>
<tr>
<th>Perylene, ( \text{C}<em>{20}\text{H}</em>{12} )</th>
<th>Donaldson(^a)</th>
<th>Camerman(^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( m.p. /°C ) ( d ) ( a/\text{Å} ) ( b/\text{Å} ) ( c/\text{Å} ) ( \alpha/° ) ( \beta/° ) ( \gamma/° ) ( V/\text{Å}^3 )</td>
<td>266–268 1.322 11.35(2) 10.87(3) 10.31(3) 90.0 100.8(2) 90.0 1249</td>
<td>268 1.360 11.27(1) 10.82(1) 10.26(1) 90.0 100.55(2) 90.0 1231.78</td>
</tr>
</tbody>
</table>

\(^a\) in [Donaldson et al. 1953]
\(^b\) in [Camerman and Trotter 1964]

as determined by these groups are summarized in Tab. (4.1). As can be seen from the table, the crystal structure data vary slightly, but both groups agree on the essential issues. Perylene crystallizes in the space group\(^1\) \( P_{21}/a \). Four perylene molecules are found in one orthorhombic unit cell, forming two stacked pairs which crystallize in a herringbone pattern. This unusual crystal structure is depicted in Fig. (4.2), which shows a block of eight unit cells viewed in the \( c \)-direction. One unit cell is also sketched in the Figure.

[Camerman and Trotter 1964] discuss the geometry of perylene in detail with special emphasis on the question of the planarity of the molecule in the crystal. They find a slight but significant distortion of the molecule, best described as a slight bending around the long axis of perylene (irreducible representation \( B_{3u} \)), similar to the lowest \( B_{3u} \) mode depicted in Fig. (4.6). For the carbon atoms however, the largest deviation from the molecular plane is 0.0394 Å, so that the bending is almost negligible. The authors discuss possible reasons for the observed distortion

\(^1\)In earlier references this group is sometimes denoted by its Schönflies symbol \( C_{2h}^5 \). Sometimes even the outdated notation \( 2/m \) is encountered.
Figure 4.2: Perylene. Crystal structure from [CAMERMAN AND TROTTER 1964], view on the ab plane.
and conclude that it is entirely an effect of the crystallization rather than an intrinsic property of perylene. Hence it can safely be concluded that the geometry of perylene in vacuo is approximated very well by the crystal structure geometry if the distortion is neglected, and that this is also the geometry of perylene guest molecules in Shpol’skii systems.

Table 4.2: Perylene. Comparison of observed and calculated geometries. Bond lengths in Å.

<table>
<thead>
<tr>
<th>Bond</th>
<th>C&amp;T (^a)</th>
<th>DRW (^b)</th>
<th>B3LYP (^c)</th>
<th>LSD (^d)</th>
<th>pcf (^e)</th>
<th>ppcf (^f)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C(1)–C(1a)</td>
<td>1.396</td>
<td>1.38</td>
<td>1.391</td>
<td>1.388</td>
<td>1.420</td>
<td>1.390</td>
</tr>
<tr>
<td>C(1)–C(2)</td>
<td>1.417</td>
<td>1.45</td>
<td>1.405</td>
<td>1.399</td>
<td>1.383</td>
<td>1.398</td>
</tr>
<tr>
<td>C(2)–C(3)</td>
<td>1.368</td>
<td>1.38</td>
<td>1.375</td>
<td>1.374</td>
<td>1.378</td>
<td>1.373</td>
</tr>
<tr>
<td>C(3)–C(3a)</td>
<td>1.398</td>
<td>1.37</td>
<td>1.417</td>
<td>1.411</td>
<td>1.422</td>
<td>1.409</td>
</tr>
<tr>
<td>C(3a)–C(1b)</td>
<td>1.423</td>
<td>1.44</td>
<td>1.435</td>
<td>1.431</td>
<td>1.421</td>
<td>1.431</td>
</tr>
<tr>
<td>C(1a)–C(1b)</td>
<td>1.424</td>
<td>1.45</td>
<td>1.433</td>
<td>1.425</td>
<td>1.407</td>
<td>1.428</td>
</tr>
<tr>
<td>C(1a)–C(12a)</td>
<td>1.470</td>
<td>1.50</td>
<td>1.475</td>
<td>1.462</td>
<td>1.425</td>
<td>1.458</td>
</tr>
<tr>
<td>C(1)–H(1)</td>
<td>1.107</td>
<td></td>
<td>1.081</td>
<td>1.094</td>
<td>1.097</td>
<td>1.093</td>
</tr>
<tr>
<td>C(1)–H(2)</td>
<td>1.134</td>
<td></td>
<td>1.084</td>
<td>1.094</td>
<td>1.101</td>
<td>1.096</td>
</tr>
<tr>
<td>C(1)–H(3)</td>
<td>1.101</td>
<td></td>
<td>1.084</td>
<td>1.095</td>
<td>1.099</td>
<td>1.097</td>
</tr>
</tbody>
</table>

\(^a\) in [Camerman and Trotter 1964]  
\(^b\) in [Donaldson et al. 1953], bond lengths averaged over all symmetry equivalent bonds  
\(^c\) B3LYP/6-31G*  
\(^d\) LSD/3-21G*  
\(^e\) minimum not planar  
\(^f\) as described below

Tab. (4.2) displays the results of geometry optimizations using various theoretical methods and compares them to the two available experimental sources. As can be seen, both DFT methods yield bond lengths which are on average in excellent agreement with experiment.

\(^2\) The slight bending of the molecules allows the pair of perylene molecules to pack together more compactly, ... and there seems little doubt that the observed molecular distortion is a result of intermolecular forces.” [Camerman and Trotter 1964]
agreement with the measured structure\textsuperscript{3}. This is not the case for many forcefields. As shown in Tab. (4.2), the pcff force field strongly underpredicts the C(1a)–C(12a) bond length\textsuperscript{4} (1.425 \text{
A} instead of 1.470 \text{
A}, as found in the crystal structure [CAMERMAN AND TROTTER 1964]), but also finds an artificial nonplanar minimum. Both artifacts influence the shape and size of the perylene molecule and render the standard pcff force field inadequate for studies of guest-host systems as considered in this work. This issue will be treated in more detail in Section 4.1.3.

4.1.2 Vibrational analysis

The vibrational spectra of polycyclic aromatic hydrocarbon compounds have recently attracted considerable interest in astronomy, as it seems that previously unidentified infrared (IR) bands observed in signals from interstellar space may be identified with PAH infrared emission bands [LANGHOFF 1996], a suggestion which is usually called the \textit{interstellar polycyclic aromatic hydrocarbon hypothesis}, cf. [BRENNER AND BARKER 1992]. Hence, several recent publications analyze the vibrational modes of perylene both experimentally [HUDGINS AND SANDFORD 1998] and theoretically [LANGHOFF 1996, RUMI ET AL. 1998]. Although some of the recent theoretical analyses may be of questionable quality\textsuperscript{5}, they generally complement the authoritative works on the subject, [WHITMER ET AL. 1978] and [CYVIN ET AL. 1983] well. These older publications probably present the most accurate analysis of the vibrational spectra of perylene, as well as the most advanced attempt so far to create a customized force field to reproduce the vibrational modes of perylene as accurately as possible. Consequently they form the reference for calculations presented here.

\textsuperscript{3}The agreement is even better than for small basis set \textit{ab initio} calculations, as for example recently published in [RUMI ET AL. 1998].

\textsuperscript{4}This is a direct consequence of the fact that pcff considers bonds and bond orders a direct function of the atoms forming the bond. Hence a bond between two aromatic carbon atoms is an aromatic bond with a minimum of 1.4170 \text{
A} in the quartic-bond term of the force field.

\textsuperscript{5} [LANGHOFF 1996] for example identify their $b_{1u}$ vibration at 1091.4 cm\textsuperscript{-1} with an observed vibration at 1088 cm\textsuperscript{-1} or 1089 cm\textsuperscript{-1}, while [WHITMER ET AL. 1978] observe a $b_{2u}$ vibration at 1085 cm\textsuperscript{-1}, and their experimental data simply leave no room for an additional $b_{1u}$ mode in this region. A thorough separate assignment for each of the irreducible representations of $D_{2h}$ would have prevented this error.
Listed below are the positions of the $3 \cdot 32 - 6 = 90$ modes of perylene. The following tables compare the known observed positions (where information was available) with the theoretical positions from the Cyvin force field of [Whitmer et al. 1978], as well as the theoretical number calculated ab initio on the RHF/6-31G* level and using density functional methods on the B3LYP/6-31G* level of theory\textsuperscript{6}.

<table>
<thead>
<tr>
<th>WCC calc.\textsuperscript{a}</th>
<th>WCC obs.</th>
<th>RHF/6-31G*</th>
<th>B3LYP/6-31G*</th>
<th>ppcff\textsuperscript{b}</th>
</tr>
</thead>
<tbody>
<tr>
<td>3041</td>
<td>—</td>
<td>3426</td>
<td>3231</td>
<td>3124</td>
</tr>
<tr>
<td>3034</td>
<td>—</td>
<td>3382</td>
<td>3202</td>
<td>3033</td>
</tr>
<tr>
<td>3029 3031</td>
<td></td>
<td>3360</td>
<td>3184</td>
<td>3017</td>
</tr>
<tr>
<td>1752</td>
<td>1581</td>
<td>1801</td>
<td>1650</td>
<td>1597</td>
</tr>
<tr>
<td>1501 1574</td>
<td></td>
<td>1785</td>
<td>1620</td>
<td>1516</td>
</tr>
<tr>
<td>1440</td>
<td>—</td>
<td>1597</td>
<td>1487</td>
<td>1457</td>
</tr>
<tr>
<td>1406 1380</td>
<td></td>
<td>1507</td>
<td>1410</td>
<td>1378</td>
</tr>
<tr>
<td>1275 1369</td>
<td></td>
<td>1476</td>
<td>1400</td>
<td>1342</td>
</tr>
<tr>
<td>1198 1298</td>
<td></td>
<td>1439</td>
<td>1339</td>
<td>1281</td>
</tr>
<tr>
<td>1039 1104</td>
<td></td>
<td>1314</td>
<td>1225</td>
<td>1184</td>
</tr>
<tr>
<td>924</td>
<td>—</td>
<td>1207</td>
<td>1138</td>
<td>1053</td>
</tr>
<tr>
<td>884 973</td>
<td></td>
<td>1063</td>
<td>998</td>
<td>988</td>
</tr>
<tr>
<td>801 793</td>
<td></td>
<td>863</td>
<td>810</td>
<td>677</td>
</tr>
<tr>
<td>461</td>
<td>550</td>
<td>587</td>
<td>556</td>
<td>518</td>
</tr>
<tr>
<td>414</td>
<td>358</td>
<td>464</td>
<td>436</td>
<td>434</td>
</tr>
<tr>
<td>259</td>
<td>298</td>
<td>379</td>
<td>358</td>
<td>339</td>
</tr>
</tbody>
</table>

\textsuperscript{a} in [Whitmer et al. 1978]
\textsuperscript{b} as described below

The fully symmetric in-plane vibrations of $A_g$ symmetry, listed in Tab. (4.3), are inactive in both IR and Raman spectroscopy, but due to the quantum-mechanical selection rules the only observable vibronic lines in optical, especially total lumines-

\textsuperscript{6}The calculations were performed using the Gaussian 94 package [Frisch et al. 1995] on the workstations of the C4 cluster of the ETH Zürich.
## Table 4.4: Perylene. $B_{1g}$ modes.

<table>
<thead>
<tr>
<th>WCC calc.</th>
<th>WCC obs.</th>
<th>RHF/6-31G*</th>
<th>B3LYP/6-31G*</th>
<th>ppcf*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1115</td>
<td>960</td>
<td>1104</td>
<td>968</td>
<td>1003</td>
</tr>
<tr>
<td>961</td>
<td>900</td>
<td>1016</td>
<td>898</td>
<td>887</td>
</tr>
<tr>
<td>719</td>
<td>752</td>
<td>853</td>
<td>769</td>
<td>722</td>
</tr>
<tr>
<td>457</td>
<td>548</td>
<td>690</td>
<td>638</td>
<td>628</td>
</tr>
<tr>
<td>337</td>
<td>430</td>
<td>468</td>
<td>430</td>
<td>438</td>
</tr>
<tr>
<td>198</td>
<td>—</td>
<td>229</td>
<td>210</td>
<td>219</td>
</tr>
</tbody>
</table>

*a in [Whitmer et al. 1978]*

*b as described below*

## Table 4.5: Perylene. $B_{2g}$ modes.

<table>
<thead>
<tr>
<th>WCC calc.</th>
<th>WCC obs.</th>
<th>RHF/6-31G*</th>
<th>B3LYP/6-31G*</th>
<th>ppcf*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1341</td>
<td>—</td>
<td>1111</td>
<td>973</td>
<td>1020</td>
</tr>
<tr>
<td>1042</td>
<td>910</td>
<td>1024</td>
<td>902</td>
<td>907</td>
</tr>
<tr>
<td>927</td>
<td>889</td>
<td>938</td>
<td>860</td>
<td>823</td>
</tr>
<tr>
<td>719</td>
<td>791</td>
<td>845</td>
<td>767</td>
<td>733</td>
</tr>
<tr>
<td>596</td>
<td>625</td>
<td>721</td>
<td>664</td>
<td>671</td>
</tr>
<tr>
<td>389</td>
<td>—</td>
<td>497</td>
<td>462</td>
<td>411</td>
</tr>
<tr>
<td>215</td>
<td>249</td>
<td>327</td>
<td>303</td>
<td>270</td>
</tr>
<tr>
<td>136</td>
<td>—</td>
<td>132</td>
<td>129</td>
<td>92</td>
</tr>
</tbody>
</table>

*a in [Whitmer et al. 1978]*

*b as described below*
Table 4.6: Perylene. $B_{3g}$ modes.

<table>
<thead>
<tr>
<th>WCC calc.\textsuperscript{a}</th>
<th>WCC obs.\textsuperscript{b}</th>
<th>RHF/6-31G*</th>
<th>B3LYP/6-31G*</th>
<th>ppcff\textsuperscript{c}</th>
</tr>
</thead>
<tbody>
<tr>
<td>3041</td>
<td>—</td>
<td>3398</td>
<td>3215</td>
<td>3063</td>
</tr>
<tr>
<td>3033</td>
<td>—</td>
<td>3376</td>
<td>3198</td>
<td>3027</td>
</tr>
<tr>
<td>3029</td>
<td>—</td>
<td>3356</td>
<td>3181</td>
<td>3006</td>
</tr>
<tr>
<td>1742</td>
<td>—</td>
<td>1839</td>
<td>1676</td>
<td>1625</td>
</tr>
<tr>
<td>1578</td>
<td>—</td>
<td>1688</td>
<td>1574</td>
<td>1470</td>
</tr>
<tr>
<td>1473</td>
<td>—</td>
<td>1622</td>
<td>1502</td>
<td>1422</td>
</tr>
<tr>
<td>1357</td>
<td>—</td>
<td>1446</td>
<td>1391</td>
<td>1257</td>
</tr>
<tr>
<td>1212</td>
<td>—</td>
<td>1327</td>
<td>1257</td>
<td>1208</td>
</tr>
<tr>
<td>1059 (1222)</td>
<td>—</td>
<td>1314</td>
<td>1226</td>
<td>1155</td>
</tr>
<tr>
<td>1046</td>
<td>—</td>
<td>1231</td>
<td>1180</td>
<td>1083</td>
</tr>
<tr>
<td>978</td>
<td>—</td>
<td>1086</td>
<td>1086</td>
<td>997</td>
</tr>
<tr>
<td>863</td>
<td>—</td>
<td>1038</td>
<td>961</td>
<td>847</td>
</tr>
<tr>
<td>728</td>
<td>—</td>
<td>676</td>
<td>635</td>
<td>538</td>
</tr>
<tr>
<td>579</td>
<td>—</td>
<td>581</td>
<td>546</td>
<td>516</td>
</tr>
<tr>
<td>304</td>
<td>—</td>
<td>387</td>
<td>367</td>
<td>325</td>
</tr>
</tbody>
</table>

\textsuperscript{a} in [Whitmer et al. 1978]

\textsuperscript{b} numbers in parentheses from [Cyvin et al. 1983]

\textsuperscript{c} as described below
Table 4.7: Perylene. $A_u$ modes.

<table>
<thead>
<tr>
<th>WCC calc.</th>
<th>WCC obs.</th>
<th>RHF/6-31G*</th>
<th>B3LYP/6-31G*</th>
<th>pccff</th>
</tr>
</thead>
<tbody>
<tr>
<td>1120</td>
<td>960</td>
<td>1101</td>
<td>963</td>
<td>994</td>
</tr>
<tr>
<td>876</td>
<td>900</td>
<td>1006</td>
<td>881</td>
<td>874</td>
</tr>
<tr>
<td>724</td>
<td>752</td>
<td>868</td>
<td>788</td>
<td>745</td>
</tr>
<tr>
<td>578</td>
<td>658</td>
<td>709</td>
<td>660</td>
<td>612</td>
</tr>
<tr>
<td>388</td>
<td>(522)</td>
<td>584</td>
<td>537</td>
<td>491</td>
</tr>
<tr>
<td>244</td>
<td>(281)</td>
<td>264</td>
<td>239</td>
<td>243</td>
</tr>
<tr>
<td>77</td>
<td>(74)</td>
<td>10</td>
<td>28</td>
<td>51</td>
</tr>
</tbody>
</table>

- **a** in [Whitmer et al. 1978]
- **b** numbers in parentheses from [Cyvin ET AL. 1983]
- **c** as described below

Of the vibrations in the other irreducible representations, those with the lowest frequencies are the most important ones, because they are expected to show the strongest coupling to the hosts phonon bath. The lowest-frequency mode is shown in Fig. (4.3). It occurs at 74 cm$^{-1}$, is totally antisymmetric and belongs to the irreducible representation $A_u$. This is the mode which has an imaginary frequency in pcff, and the global pcff minimum for perylene is distorted in the way the arrows in Fig. (4.3) indicate. Indeed, this is one of the modes not well reproduced by the ab initio and DFT calculations, both of which underestimate the frequency of this mode considerably.

The next lowest mode, shown in Fig. (4.4) is the lowest $B_{3u}$ mode. Here the naphthalene subunits move out of the plane of the molecule as a whole and so this mode is generally known as the “butterfly mode” [Filliaux 1985, Wittmeyer and Topp 1991]. Fig. (4.5) and Fig. (4.6) show the next two modes of $B_{2g}$ and $B_{3u}$ symmetry, respectively. As noted above, the $B_{3u}$ distortion of mode four seems to be the optimal geometry of perylene in the crystal [Camerman and Trotter 1964]. These two modes are very similar and it is interesting to note that the “less
Table 4.8: Perylene. \( B_{1u} \) modes.

<table>
<thead>
<tr>
<th>WCC calc.(^a)</th>
<th>WCC obs.(^b)</th>
<th>RHF/6-31G*</th>
<th>B3LYP/6-31G*</th>
<th>ppcf(^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3041</td>
<td>3085</td>
<td>3398</td>
<td>3216</td>
<td>3063</td>
</tr>
<tr>
<td>3034</td>
<td>3054</td>
<td>3379</td>
<td>3199</td>
<td>3028</td>
</tr>
<tr>
<td>3029</td>
<td>3052</td>
<td>3359</td>
<td>3183</td>
<td>3016</td>
</tr>
<tr>
<td>1668</td>
<td>1592</td>
<td>1804</td>
<td>1650</td>
<td>1518</td>
</tr>
<tr>
<td>1499</td>
<td>1513</td>
<td>1790</td>
<td>1640</td>
<td>1470</td>
</tr>
<tr>
<td>1448</td>
<td>1382</td>
<td>1607</td>
<td>1490</td>
<td>1376</td>
</tr>
<tr>
<td>1394</td>
<td>1367</td>
<td>1531</td>
<td>1423</td>
<td>1343</td>
</tr>
<tr>
<td>1290</td>
<td>1288</td>
<td>1468</td>
<td>1401</td>
<td>1211</td>
</tr>
<tr>
<td>1215</td>
<td>1217</td>
<td>1341</td>
<td>1244</td>
<td>1153</td>
</tr>
<tr>
<td>1040</td>
<td>1151</td>
<td>1267</td>
<td>1178</td>
<td>1103</td>
</tr>
<tr>
<td>909</td>
<td>980</td>
<td>1191</td>
<td>1123</td>
<td>985</td>
</tr>
<tr>
<td>817</td>
<td>963</td>
<td>888</td>
<td>832</td>
<td>898</td>
</tr>
<tr>
<td>677</td>
<td>850</td>
<td>854</td>
<td>804</td>
<td>688</td>
</tr>
<tr>
<td>556</td>
<td>585</td>
<td>635</td>
<td>590</td>
<td>498</td>
</tr>
<tr>
<td>368</td>
<td>(417)</td>
<td>500</td>
<td>469</td>
<td>381</td>
</tr>
</tbody>
</table>

\(^a\) in [Whitmer et al. 1978]

\(^b\) numbers in parentheses from [Cyvin et al. 1983]

\(^c\) as described below
Table 4.9: Perylene. $B_{2u}$ modes.

<table>
<thead>
<tr>
<th>WCC calc.</th>
<th>WCC obs.</th>
<th>RHF/6-31G*</th>
<th>B3LYP/6-31G*</th>
<th>pccf*</th>
</tr>
</thead>
<tbody>
<tr>
<td>3041</td>
<td>3051</td>
<td>3426</td>
<td>3231</td>
<td>3125</td>
</tr>
<tr>
<td>3033</td>
<td>3038</td>
<td>3378</td>
<td>3200</td>
<td>3032</td>
</tr>
<tr>
<td>3029</td>
<td>3016</td>
<td>3357</td>
<td>3181</td>
<td>3006</td>
</tr>
<tr>
<td>1697</td>
<td>1605</td>
<td>1819</td>
<td>1663</td>
<td>1634</td>
</tr>
<tr>
<td>1594</td>
<td>(1580)</td>
<td>1674</td>
<td>1552</td>
<td>1593</td>
</tr>
<tr>
<td>1445</td>
<td>1495</td>
<td>1649</td>
<td>1522</td>
<td>1495</td>
</tr>
<tr>
<td>1403</td>
<td>1333</td>
<td>1457</td>
<td>1371</td>
<td>1413</td>
</tr>
<tr>
<td>1306</td>
<td>1281</td>
<td>1400</td>
<td>1319</td>
<td>1312</td>
</tr>
<tr>
<td>1157</td>
<td>1211</td>
<td>1313</td>
<td>1248</td>
<td>1233</td>
</tr>
<tr>
<td>1060</td>
<td>1127</td>
<td>1247</td>
<td>1225</td>
<td>1138</td>
</tr>
<tr>
<td>1025</td>
<td>1085</td>
<td>1216</td>
<td>1162</td>
<td>1069</td>
</tr>
<tr>
<td>862</td>
<td>1042</td>
<td>1121</td>
<td>1070</td>
<td>988</td>
</tr>
<tr>
<td>807</td>
<td>535</td>
<td>840</td>
<td>786</td>
<td>587</td>
</tr>
<tr>
<td>555</td>
<td>464</td>
<td>580</td>
<td>543</td>
<td>484</td>
</tr>
<tr>
<td>171</td>
<td>(178)</td>
<td>272</td>
<td>258</td>
<td>284</td>
</tr>
</tbody>
</table>

* in [Whitmer et al. 1978]

b numbers in parentheses from [Cyvin et al. 1983]

c as described below
Table 4.10: Perylene. $B_{3u}$ modes.

<table>
<thead>
<tr>
<th>WCC calc.</th>
<th>WCC obs.</th>
<th>RHF/6-31G*</th>
<th>B3LYP/6-31G*</th>
<th>ppcf</th>
</tr>
</thead>
<tbody>
<tr>
<td>1282</td>
<td>—</td>
<td>1114</td>
<td>977</td>
<td>1025</td>
</tr>
<tr>
<td>961</td>
<td>967</td>
<td>1033</td>
<td>916</td>
<td>929</td>
</tr>
<tr>
<td>914</td>
<td>813</td>
<td>917</td>
<td>831</td>
<td>803</td>
</tr>
<tr>
<td>713</td>
<td>768</td>
<td>857</td>
<td>782</td>
<td>748</td>
</tr>
<tr>
<td>525</td>
<td>544</td>
<td>610</td>
<td>563</td>
<td>578</td>
</tr>
<tr>
<td>390</td>
<td>(452)</td>
<td>519</td>
<td>482</td>
<td>428</td>
</tr>
<tr>
<td>162</td>
<td>(185)</td>
<td>193</td>
<td>183</td>
<td>143</td>
</tr>
<tr>
<td>57</td>
<td>(111)</td>
<td>105</td>
<td>97</td>
<td>101</td>
</tr>
</tbody>
</table>

a in [Whitmer et al. 1978]

b numbers in parentheses from [Cyvin et al. 1983]

c as described below

symmetric" mode $^1B_{2g}$ actually occurs at a lower frequency than the $^2B_{3u}$ mode, which is ungerade only with respect to the molecular plane.

4.1.3 ppcf force field

As noted above, the need for a modified force field, customized for the needs of an accurate MD simulation, arises from the inability of standard force fields, in particular pcff, to describe perylene accurately. This new force field should use the same functional form as pcff and share most of its parameters, but be able to reproduce the properties of perylene. The construction and parametrization of this new force field, called ppcf (perylene + polymer consistent force field), is described in detail below. The general procedure for any force field parametrization is

1. A functional form for the force field is chosen. This choice may depend on several considerations. A particular functional form may be chosen for its simplicity and corresponding computational efficiency [Cyvin et al. 1983], or for its support of special functional groups (unified atom force fields in
Figure 4.3: Perylene. Mode 1 (B3LYP/6-31G*). $A_u$ symmetry. 74 cm$^{-1}$.

Figure 4.4: Perylene. Mode 2 (B3LYP/6-31G*). $B_{3u}$ symmetry. 111 cm$^{-1}$.
Figure 4.5: Perylene. Mode 3 (B3LYP/6-31G*). $B_{2g}$ symmetry. 136 cm$^{-1}$ (Cyvin force field) resp. 129 cm$^{-1}$ (B3LYP/6-31G*).

Figure 4.6: Perylene. Mode 4 (B3LYP/6-31G*). $B_{3u}$ symmetry. 185 cm$^{-1}$. 
protein modelling) or, if the new force field is to be a derivative of an existing force field, the original functional form may be retained [HALGREN 1996].

2. A set of properties is chosen, which the force field should reproduce. Experimentally determined or calculated properties in this set typically include molecular structures of isolated molecules [LIFSON AND WARSHEL 1968], vibrational frequencies [CYVIN ET AL. 1983] or modes, torsional potentials, relative energies of modified geometries, or the forces acting on the atoms in these [SUN 1995]. In general, these properties should be quickly evaluated by the force field engine (which may be an existing MM/MD code or a separate implementation), for otherwise the optimization procedure would quickly become computationally unfeasible. Thus, condensed phase properties are not regularly used, except at special stages in the optimization.\footnote{In the development of the COMPASS force field [SUN 1998], for example, crystal structures and other condensed phase properties like density and cohesive energies were used at one point in the procedure to refine the Lennard-Jones parameters.}

3. The force field parameters are adjusted. This is the "black art" stage of the procedure. In the general case, the number of parameters which are to be optimized, lies between a few dozen and several thousand. At the same time, the set of properties defined in the previous step may contain many thousand data points\footnote{In the MMFF94 force field optimization in [HALGREN 1996] bond lengths and angles of around 1500 compounds, calculated by \textit{ab initio} and DFT means were used.} from different origins. The force field evaluation of these data

Figure 4.7: Perylene. Mode 9 (B3LYP/6-31G*). $A_g$ symmetry. 298 cm$^{-1}$. 
points may differ vastly in its computational cost. Hence, the use of a general, "off-the-shelf", minimization is usually not an option.

To remove this bottleneck, several strategies have been used in the past. Early works considered the parameters separately and replaced the global minimization (for example MM2 and its siblings were constructed like that, [Allinger 1976, Allinger 1977, Allinger et al. 1989]) by a series of local optimizations, most of which were not even done numerically. The few numerical minimizers used in early publications were usually fairly standard, conjugate-gradient [Press et al. 1992] or related methods, which compensated for the lack of sophistication in the optimizer engine by choosing data points whose evaluation was very cheap [Lifson and Warshel 1968]. Other groups searched for "silver bullet" solutions for the computational problem. [Das-Gupta et al. 1996], for example, presented a method which treats the minimization problem as a multidimensional least-squares polynomial-fit problem and suggest a singular-value-decomposition algorithm to solve it. Although the authors do present an example application with impressive convergence characteristics, their approach has serious drawbacks. The "modern" approach, exemplified in [Sun 1998], is more elastic. Its goal is to resolve the parameter and data spaces into separable, weakly-coupled subspaces, and to use specific optimization algorithms for each of these subspaces. [Sun 1998] for example optimized the partial charges in COMPASS first, then kept them fixed in a second step, in which the intramolecular potential parameters were optimized in the same way as in the original CFF development [Maple et al. 1988]. The intramolecular parameters were then refined by repeated adjustment of the parameter responsible for the largest error. After final optimization of the Lennard-Jones parameters, the consistency of the previous optimizations was verified.

Based on this, it is obvious that the new ppcff force field needs to:

---

9Whereas a geometry minimization of a small molecule has an almost negligible computational cost and the complete optimization of a set of several hundred or more molecules is a matter of seconds on modern computers, the calculation of crystal structure parameters or condensed phase densities is considerably more costly, and the use of dynamical properties as viscosities or (auto-)diffusion coefficients is clearly prohibitive except for special cases.
1. be based on an existing force field, its functional form and its parametrization. Specifically, since the goal is to study perylene in aliphatic hosts, the pcff force field [SUN 1995] is best suited for this.¹⁰

2. employ a balanced error function in the optimization. This means that the optimization criteria should not only be based on the molecular geometry of perylene and its vibrational spectrum, but also take relative energies of an ensemble of distorted structures into account.

3. use a flexible optimization procedure. Since analytical gradients are in general not available in complex optimizations like this, a fast and robust minimizer which does not require the evaluation of gradients is needed. But because minimizers in this class (of which the Simplex algorithm is probably the most prominent example) are in general very weak in certain regions of the optimization space, this method must be complemented by a gradient-based (Newton-Raphson-like, if feasible, at least conjugate-gradient-like, if not) method. Ideally, both methods would run in parallel, compete with each other, and would in regular intervals exchange their information, thus preventing a slowdown of the procedure due to an eventual inadequacy of one of them in a particular region of the optimization space.

The procedure based on these conclusions and the ppcff force field constructed with it are described below.

### 4.1.4 Functional form of ppcff

The general form of ppcff is identical to pcff (cf. Sec. 2.1.1), in particular the aliphatic part was left unchanged. In addition to the atom types and parameters of pcff, ppcff contains a new atom type “cp2”, which is used for the C(1a), C(7a), C(8a) and C(12a) atoms. Parameters describing the interaction of atoms of this new type with “cp” and “h” atoms were added to the ppcff definition.¹¹ In the next step, all

---

¹⁰ The name ppcff is by no means accidental.

¹¹ This was done by simply duplicating all appropriate lines in the force field description file.
parameters involving atoms from any subset of \{"cp", "cp2", "h"\} were declared as adjustable parameters\(^{12}\).

![Figure 4.8: Perylene. ppcff atom types. Only symmetry unique atoms are labeled. All hydrogen atoms (not shown) have type "h".](image)

This constituted a set of 174 adjustable parameters, grouped into:

1. 16 parameters for the quartic-bond terms in Eq. (2.2a), namely \(b_0, K_2, K_3\) and \(K_4\) for each of the bonds "cp–cp", "cp–h", "cp–cp2" and "cp2–cp2" which are the only bond types in perylene.

2. 21 parameters for the quartic-angle terms in Eq. (2.2b). All \(K_4\) terms for these parameters are zero in pcff, remained at zero during an initial 281-parameter test optimization and thus were kept at zero here. This left \(\theta_0, K_2\) and \(K_3\) for each of "cp–cp–cp", "cp–cp–cp2", "cp–cp2–cp", "cp–cp2–cp2", "cp2–cp–cp2", "cp–cp–h" and "cp2–cp–h".

\(^{12}\)There are four exceptions. Firstly, the Wilson-out-of-plane terms were kept at their pcff values due to computational problems in the optimization procedure. Secondly, all parameters whose values wouldn't change in the optimization due to symmetry reasons were kept fixed (this includes the equilibrium torsional angles, whose values were set to zero, after a test run verified that no spontaneous symmetry breaking occurred). Thirdly, the Lennard-Jones parameters were kept fixed at their pcff values. This was done because the nonbonded interactions are long-range interactions and determine the results of condensed-phase calculations. An optimization routine which optimizes the force field to reproduce properties of the molecule in vacuo should not change parameters of long-range interactions at all. Finally, no angle-angle coupling parameters as in Eq. (2.2h) were introduced, because pcff has no such terms for "cp–cp–cp–cp" interactions either.
3. 19 parameters for the torsion terms in Eq. (2.2c). The complete set would have had six parameters for each of the twelve torsional angle types present in perylene. However, most of these 72 remained zero during the first optimization run and were thus kept fixed.

4. 0 parameters for the Wilson-out-of-plane term and the Lennard-Jones 6-9 interaction in Eq. (2.2d) and Eq. (2.2m). For the reasons given above these parameters were not included.

5. 7 parameters for the bond-bond coupling in Eq. (2.2e), namely the $F_{bb'}$ for the seven possible combinations of adjacent bonds in perylene, “cp—cp—cp”, “cp—cp—h”, “cp—cp—cp2”, “cp—cp2—h”, “cp—cp2—cp”, “cp2—cp—cp2” and “cp2—cp2—cp”.


7. 11 parameters for the bond-angle coupling in Eq. (2.2f). The same seven bond-bond combinations as in Eq. (2.2e), but four of them, “cp—cp—h”, “cp—cp—cp2”, “cp2—cp—h” and “cp2—cp2—cp” are asymmetric and hence carry different parameters for each of the two possible bonds.

8. 0 parameters for the angle-angle coupling in Eq. (2.2h). All angle-angle coupling parameters were found to remain zero during the initial optimization.

9. 33 parameters for the end-bond-torsion-3 in Eq. (2.2h). This is one of the largest groups. Many of the 72 parameters are zero by symmetry, and many others remained zero during the test run.

10. 19 parameters for the middle-bond-torsion-3 in Eq. (2.2j).

11. 33 parameters for the angle-torsion-3 in Eq. (2.2k). The same 72 possibilities as for Eq. (2.2h) are possible here, and the same pattern of zeros emerged.

12. 5 parameters for the angle-angle-torsion-1 in Eq. (2.2l). Here, too, twelve parameters are possible, one for each of “cp—cp—cp—cp”, “cp—cp—cp—cp2”, “cp—cp—cp2—cp2”, “cp—cp2—cp—cp2”, “cp—cp2—cp—cp2”, “cp2—cp—cp—cp2”, “cp2—cp—cp—cp2”, “cp—cp—cp2—

cp–h". Only those involving hydrogen, however, were found to be non-zero.

Table 4.11: Perylene. Partial Charges used in this work. Numbering as in Fig. (4.1).

<table>
<thead>
<tr>
<th>Partial charges</th>
<th>pcff¹</th>
<th>R&amp;G²</th>
</tr>
</thead>
<tbody>
<tr>
<td>C(1)</td>
<td>-0.1268</td>
<td>-0.1140</td>
</tr>
<tr>
<td>C(1a)</td>
<td>0.0000</td>
<td>0.0060</td>
</tr>
<tr>
<td>C(1b)</td>
<td>0.0000</td>
<td>0.0590</td>
</tr>
<tr>
<td>C(2)</td>
<td>-0.1268</td>
<td>-0.1320</td>
</tr>
<tr>
<td>C(3)</td>
<td>-0.1268</td>
<td>-0.1210</td>
</tr>
<tr>
<td>C(3a)</td>
<td>0.0000</td>
<td>0.0460</td>
</tr>
<tr>
<td>H(1)</td>
<td>0.1268</td>
<td>0.0970</td>
</tr>
<tr>
<td>H(2)</td>
<td>0.1268</td>
<td>0.1110</td>
</tr>
<tr>
<td>H(3)</td>
<td>0.1268</td>
<td>0.1010</td>
</tr>
</tbody>
</table>

¹ in [BIOSYM/MSI 1997]
² in [RAPPE AND GODDARD III 1991]

The partial charges of the carbon and hydrogen atoms on perylene were not determined by pcff's built-in algorithm but rather externally calculated according to [RAPPE AND GODDARD III 1991]¹³. These partial charges are displayed in Tab. (4.11).

4.1.5 Error function of ppcff

One of the characteristic features of Class II force fields is the use of a balanced optimization procedure. In the process, not only single-point data, such as minimum geometry and vibrational frequencies, but also data from different geometries,

¹³It was found (see below) that these partial charges combined with the Lennard-Jones parameters of pcff led to excellent reproduction of the crystal properties of perylene. However, there is a drawback: These charges result in non-neutral charge groups, which forbids the use of group cutoffs and forces the evaluation of electrostatic interactions via the Ewald sum algorithm, which is significantly more expensive.
different conformers, or more typically, from quantum chemical calculations. The advantage of such broad and balanced criteria is obvious: single-point information can at best yield a Taylor expansion up to second order in the vicinity of the minimum geometry. If a third or fourth order functional is fitted to such second order data, numerical instability is the usual result: force fields parametrized only with such single point data tend to reproduce the geometry and vibrational frequencies rather well, but in actual MD runs strong distortions and unphysical behavior result as soon as the geometry leaves the region where the original Taylor expansion is valid. Class II force fields avoid this problem, and the ppcff force field follows this approach.

Hence, the ppcff error function consists of five terms:

\[
f_{\text{tot}} = f_{\text{geometry}} + f_{\text{frequencies}} + f_{\text{modes}} + f_{\text{energies}} + f_{\text{forces}}
\] (4.1)

where

- \(f_{\text{geometry}} = 10000 \cdot \sum_{i=1}^{32} d(x_i^\text{ff}, x_i^\text{LSD})^2\) (4.2a)
- \(f_{\text{frequencies}} = \sum_{i=1}^{90} \frac{|v_i^\text{ff} - v_i^\text{B3LYP}|^3}{v_i^\text{B3LYP}^2}\) (4.2b)
- \(f_{\text{modes}} = 0.1 \cdot \sum_{i=1}^{90} (1 - |\hat{v}_i^\text{ff} \cdot \hat{v}_i^\text{B3LYP}|)\) (4.2c)
- \(f_{\text{energies}} = \sum_{i=1}^{100} (\Delta E_i^\text{ff} - \Delta E_i^\text{LSD})^2\) (4.2d)
- \(f_{\text{forces}} = 0.001 \cdot \sum_{i=1}^{100} \sum_{j=1}^{96} (F_{ij}^\text{ff} - F_{ij}^\text{LSD})^2\) (4.2e)

The physical meaning of these terms and the rationale for their choice is described below:

**Molecular Geometry, Eq. (4.2a):** The reason ppcff became necessary was the failure of other force fields, especially pcff, to reproduce the geometry of perylene correctly. Hence this term is essential. \(f_{\text{geometry}}\) was calculated by compar-
ison of a minimized pcff geometry\textsuperscript{14} with a suitable reference. As reference the DFT minimum on the LSD/3-21G* level of theory was chosen\textsuperscript{15}. In Eq. (4.2a), $d(\mathbf{x}_i^{\text{ppcf}}, \mathbf{x}_i^{\text{LSD}})^2$ is the Cartesian distance in Å between corresponding atoms in the ppcff minimum and in the reference geometry\textsuperscript{16}.

**Vibrational Frequencies, Eq. (4.2b):** It could be argued that, since the eigenvalues and eigenvectors of the molecular Hessian constitute a Taylor expansion of the Born-Oppenheimer surface up to second order at one point, whereas the ensemble data below are linear expansions at 100 points in coordinate space, the information from the vibrational analysis could be neglected. Indeed, as explained above, strict optimization of these terms may lead to a numerically unstable force field. However, the observed vibrational frequencies of perylene are in general reproduced well by DFT calculations even in the harmonic approximation, and it is found that the vibrational frequencies of perylene in Shpol'skii systems do not differ significantly from the vacuum values. Hence the inclusion of these terms in the error function is justified, although a low weight is advisable. The actual error term in Eq. (4.2b) is the relative quadratic deviation of the ppcff eigenfrequency (in cm\textsuperscript{-1}) from the corresponding value in a vibrational analysis on the B3LYP/6-31G* level, weighted by the absolute value of that deviation. This is unusual, but warranted: during the optimization, the ppcff frequencies showed a tendency to be very inaccurate in the region > 2000 cm\textsuperscript{-1}, when only the relative quadratic deviations were used.

**Vibrational Modes, Eq. (4.2c):** The desired force field should also be expected to reproduce the eigenmodes of the vibrations at least for the low-frequency out-of-plane modes\textsuperscript{17}. Hence this term was included in the error function. The

\textsuperscript{14}It is important to use a tighter convergence criterion in the geometry optimization than Discover suggests. With the default criterion, pcff finds a planar perylene minimum whose Hessian has a negative eigenvalue.

\textsuperscript{15}This was necessary to keep the optimization consistent. The final two terms in Eq. (4.2d) and Eq. (4.2e) compare the energies and forces of an ensemble of geometries relative to the minimum with DFT energies and forces relative to their minimum. These two reference geometries must of course be identical.

\textsuperscript{16}Ideally, this term should be calculated after a suitable mapping of the molecular geometries onto each other. However, the force field minimization was done with the DFT result as initial guess, no center-of-mass translation or rotation was observed, and the coordinate mapping could be omitted.

\textsuperscript{17}Higher vibrational eigenmodes tend to have a complicated mixed-mode character and their robustness under small perturbations of the environment, as present in guest-host systems, may be
relevant quantity here is the overlap between the calculated normalized eigenvectors on the ppcf and B3LYP/6-31G* levels. The Hessian was calculated with Discover and the diagonalization and symmetry assignment of the eigenmodes was performed with a separate program. If the geometry optimization in the first step converged to a nonplanar minimum, this symmetry assignment failed for obvious reasons. Whenever this happened, \( f_{\text{frequencies}} \) and \( f_{\text{modes}} \) were both set to 1000, larger than any value found for \( f_{\text{frequencies}} \) or \( f_{\text{modes}} \) with planar minima.

**Ensemble Energies, Eq. (4.2d):** For the final two terms of the error function, an ensemble of non-minimum geometries was created from a MD simulation of an isolated perylene molecule in a **in vacuo** at 100 K using an intermediate version of the force field based on pcff, which reproduced the minimum geometry well. This simulation was, after an equilibration period of 50 ps (time-step 1 fs), sampled every 500 fs for a duration of 50 ps, yielding an ensemble of 100 different geometries. The energies of these geometries were then calculated on the LSD/3-21G* level of theory. In Eq. (4.2d), \( \Delta E_i^{\text{LSD}} \) are these energies in kcal/mol relative to the LSD/3-21G* minimum. The \( \Delta E_i^{\text{ff}} \) are the energies of the 100 geometries in ppcf, also in kcal/mol, relative to the ppcf minimum.

**Ensemble Forces, Eq. (4.2e):** If, in addition to the energies, the forces acting on the atoms for each geometry in the ensemble are included, each point in coordinate space is represented not only by its relative energy but also by gradient information. Thus the reference data in the geometry section are a set of linear approximations to the Born-Oppenheimer surface rather than just a set of points. In Eq. (4.2e), \( F_{ij}^{\text{ff}} \) represents the \( j \)th component of the ppcf

---

18The symmetry assignment routine expected a planar \( D_{2h} \) geometry. The irreducible representations of the eigenmodes of nonplanar minima could not be determined by the assignment routine.

19This approach is superior to the original pcff procedure, which, starting from the minimum geometry, generated a large number of geometries by random displacements of the atoms. The subspace sampled by the procedure suggested here is physically more relevant and thus a smaller number of geometries is sufficient.

20The energies and forces were calculated using code written by S. Santos and M. Zehnder, whose contribution is gratefully acknowledged.
force acting in the $i$th geometry in the ensemble (in kcal/molÅ). The sum is taken over all Cartesian components of all 32 atoms, thus each geometry contributes 96 terms. $F_{ij}^{LS}$ is the corresponding term in the DFT calculation.

The weights of the terms in Eq. (4.2) were rather arbitrarily chosen and refined during the optimization. The guiding principles behind the values finally chosen for them were robustness of the force field$^{21}$ as well as an accurate description of the molecular geometry. If, during the optimization, the global error function decreased due to a decrease of one of the terms in Eq. (4.2) on the expense of the others, the weights were adjusted to ensure balanced convergence of all terms.

### 4.1.6 Parameter optimization of ppcff

For the actual parameter optimization two different algorithms were adopted:

1. the generic conjugate direction set method suggested by Powell [Press et al. 1992] in the version modified by [Acton 1990]. This algorithm minimizes the error function first in all of a set of externally supplied directions, then replaces the direction in which the error function decreased most by the average direction of the minimization and repeats the process with the endpoint of the first round as new starting point and the modified set of directions as new starting set. The method is not quadratically convergent, but very robust, does not require gradient information and can thus utilize a cheap error function evaluation routine. It has, however, a few drawbacks: (a) the first set of directions is rather arbitrary and was chosen to be the set of Cartesian unit vectors in parameter space, scaled by 0.02 times the actual magnitude of the parameter. After each round of line minimizations, one new direction was chosen and each additional "good" direction required one more round of optimizations. Hence, the efficiency of this algorithm is initially very low in high-dimensional spaces, and increases over the optimization. (b) A related problem is the tendency of optimizing the error function in the space of just a few parameters early in

---

$^{21}$This robustness was tested with MD runs to verify the force field did not strongly distorted local minima.
the optimization. If the first parameters in each round are (as was the case) for example the quartic-bond terms, the higher force constants of these terms change by large amounts to compensate for deviations in other parameters. This can lead to instabilities in intermediate versions of the force field.

(2) the original Lifson/Warshel CFF algorithm from [LIFSON AND WARSHEL 1968], which is a conjugate-gradient method. This means that it does require calculation of the Hessian, which in this case is available only numerically, and, due to the high dimensionality of the parameter space, very expensive. The advantage of this algorithm is the complete absence of the problems discussed above. The Lifson/Warshel-method treats all parameters equally and is characterized by high efficiency in regions of parameters space which are described well by a quadratic approximation and characterized by well-conditioned Hessians.

Algorithm (1) was implemented on a Silicon Graphics O2 workstation at the ETH Zurich. The "driver" of the optimization routine was a program written in the Perl programming language, which contained only the minimizer and code that glued several "engines" together. These engines performed the actual workload of the error function evaluation. The engine for the geometry optimization and calculation of the Hessian was the program Discover [BIOSYM/MSI 1997]. The diagonalization of the Hessian and the assignment of the eigenmodes to the irreducible representations of $D_{2h}$ was performed using a Fortran program written in cooperation with Prof. E. Leontidis of the University of Cyprus at Nicosia. The engine for the energy and force evaluation was an implementation of pcff written in Fortran by S. Santos and M. Zehnder in the group of Prof. U. Suter at the ETH Zurich. This flexible approach allowed the use of the fastest implementation at each step of the error function evaluation and limited the requirements of the error function evaluation to approximately 10 s per evaluation. However, the minimizer of the Powell algorithm needed about 5 evaluations per line search, and since 174 line searches constituted one round of force field minimizations it usually took several days to exploit the advantages of the algorithm. Therefore the actual minimization started not with the full set of parameters but with a limited set of bond lengths, angles and torsions, which was, during the optimization, gradually expanded.
The implementation of algorithm (2) followed a different path. It was implemented by E. Leontidis on a Silicon Graphics O2 workstation at the University of Cyprus solely in Fortran and had its own implementation of the functional form of pcff. It did not calculate and assign vibrational frequencies and did not contain a geometry minimizer, but concentrated on the forces in the reference minimum and the energies and forces of the ensemble geometries.

Because the two algorithms complemented each other, and a comparison of their performance could only be done on one of the two machines\textsuperscript{22} the following approach was chosen:

1. The parameters from pcff were chosen as starting values, with the equilibrium bond distances set to their LSD/3-21G* values, for both algorithms.

2. In regular intervals (daily in the final production stage of the optimization) the intermediate force fields were compared using a test program based on the optimizer used on the ETH machine and the complete error function. The comparison focused mainly on the magnitude of the error function, but also on additional criteria, like uniform convergence (or the absence of it), overall fit of the vibrational modes, or the detailed distribution of the error terms in the $f_{\text{energies}}$ and $f_{\text{forces}}$ terms\textsuperscript{23}.

3. Based on the comparison, the program which had provided the better intermediate force field continued. The other program was stopped and restarted with the intermediate force field of the currently algorithm as new starting point.

4. A full optimization of the force field in the complete parameter space would have been computationally prohibitive. Thus, the optimization was stopped when the cost/performance ratio of the optimization reached a point where further optimization did not seem justified.

\textsuperscript{22}The error functions used by the two were only identical for the final two terms in Eq. (4.2).

\textsuperscript{23}To assess possible systematic errors in these terms, the energies and forces as calculated by the force field were plotted vs. those from the DFT calculations. Regression parameters like correlation coefficients, and the variance of the errors were used for the comparison as well as systematic deviations form the expected cigar-like shape of the point cloud.
During the optimization, the quartic-bond terms tended to diverge, causing force field instabilities and large unphysical changes in other parameters. Hence the equilibrium bond lengths were kept at their LSD/3-21G* values, and the higher-order force constants of the quartic-bond terms were kept at their pcff values, effectively reducing the size of the parameter set to 158.

The production stage of the combined algorithm required about two CPU-months.

4.1.7 Parameters of ppcff

The optimized ppcff parameters are listed in Tab. (4.12) through Tab. (4.15). Only those parameters are listed for which at least one final value was nonzero. Hence, all parameters not listed in the tables vanish.

4.1.8 Performance of ppcff

The final ppcff forcefield reproduced the LSD/3-21G* bond lengths of perylene to within less than 0.005 Å (for detailed results see Tab. (4.2)). The contribution of $f_{\text{geometry}}$ to the total error function was 3.831 (cf. Tab. (4.16)). This corresponds to an average root mean square deviation of the atomic positions in the ppcff force field minimum from those in the reference of 0.020 Å.\(^{24}\)

The average deviation of the calculated vibrational eigenfrequencies was 35.0 cm\(^{-1}\), which is considerably smaller than the average deviation of either \textit{ab initio} (RHF/6-31G*: 142.9 cm\(^{-1}\)) or DFT calculations (B3LYP/6-31G*: 63.1 cm\(^{-1}\)), and also smaller than the average deviation of the force field in [Whitmer et al. 1978] (65.1 cm\(^{-1}\)), which was customized for perylene, too (for details, see Tab. (4.3) through Tab. (4.10)). However, the overall fit of the vibrational frequencies was not very satisfactory. That the the contribution of this term to the total error was one

\(^{24}\)The term $f_{\text{geometry}}$ in Tab. (4.16) is for pcff five orders of magnitude larger than for ppcff, because of a strong helical distortion artificially induced by pcff. In the pcff minimum (found with a very tight convergence criterion) the torsional angle C(1)–C(1a)–C(12)–C(12a) is as large as 14.3°.
Table 4.12: ppcff. Final parameters I.

<table>
<thead>
<tr>
<th>quartic-bond</th>
<th>$b_0$</th>
<th>$K_2$</th>
<th>$K_3$</th>
<th>$K_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>cp–cp</td>
<td>1.4037</td>
<td>470.8361</td>
<td>-627.6179</td>
<td>1327.6345</td>
</tr>
<tr>
<td>cp–h</td>
<td>1.0941</td>
<td>372.8251</td>
<td>-803.4526</td>
<td>894.3173</td>
</tr>
<tr>
<td>cp–cp2</td>
<td>1.4067</td>
<td>470.8361</td>
<td>-627.6179</td>
<td>1327.6345</td>
</tr>
<tr>
<td>cp2–cp2</td>
<td>1.4620</td>
<td>470.8361</td>
<td>-627.6179</td>
<td>1327.6345</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>quartic-angle</th>
<th>$\theta_0$</th>
<th>$H_2$</th>
<th>$H_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>cp–cp–cp</td>
<td>116.7586</td>
<td>37.4703</td>
<td>-46.6178</td>
</tr>
<tr>
<td>cp–cp–cp2</td>
<td>120.0846</td>
<td>73.5034</td>
<td>-24.3015</td>
</tr>
<tr>
<td>cp–cp2–cp</td>
<td>125.7863</td>
<td>61.4213</td>
<td>-29.1984</td>
</tr>
<tr>
<td>cp–cp2–cp2</td>
<td>118.8330</td>
<td>43.9259</td>
<td>-29.9672</td>
</tr>
<tr>
<td>cp2–cp–cp2</td>
<td>110.2017</td>
<td>58.3224</td>
<td>-35.0340</td>
</tr>
<tr>
<td>cp–cp–h</td>
<td>118.7215</td>
<td>27.3700</td>
<td>-16.4581</td>
</tr>
<tr>
<td>cp2–cp–h</td>
<td>112.4751</td>
<td>34.7686</td>
<td>-12.1700</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>torsion-3</th>
<th>$V_1$</th>
<th>$V_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>cp–cp–cp–cp</td>
<td>26.1318</td>
<td>1.4455</td>
</tr>
<tr>
<td>cp–cp–cp–cp2</td>
<td>-8.9205</td>
<td>4.7891</td>
</tr>
<tr>
<td>cp–cp–cp2–cp</td>
<td>7.3582</td>
<td>2.3207</td>
</tr>
<tr>
<td>cp–cp–cp2–cp2</td>
<td>8.7943</td>
<td>5.3871</td>
</tr>
<tr>
<td>cp–cp2–cp2–cp</td>
<td>-3.2078</td>
<td>6.2373</td>
</tr>
<tr>
<td>cp–cp2–cp2–cp2</td>
<td>1.2404</td>
<td>1.6977</td>
</tr>
<tr>
<td>cp2–cp2–cp2–cp2</td>
<td>9.1848</td>
<td>0.6738</td>
</tr>
<tr>
<td>cp–cp–cp–h</td>
<td>0.0</td>
<td>3.7387</td>
</tr>
<tr>
<td>cp–cp2–cp–h</td>
<td>0.0</td>
<td>5.0718</td>
</tr>
<tr>
<td>cp2–cp–cp–h</td>
<td>0.0</td>
<td>4.3302</td>
</tr>
<tr>
<td>cp2–cp2–cp–h</td>
<td>0.0</td>
<td>5.6723</td>
</tr>
<tr>
<td>h–cp–cp–h</td>
<td>0.0</td>
<td>3.0977</td>
</tr>
<tr>
<td>bond-bond</td>
<td>$F_{bb}$</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>-----------</td>
<td></td>
</tr>
<tr>
<td>cp–cp–cp</td>
<td>122.4417</td>
<td></td>
</tr>
<tr>
<td>cp–cp–h</td>
<td>62.8443</td>
<td></td>
</tr>
<tr>
<td>cp–cp–cp2</td>
<td>105.9243</td>
<td></td>
</tr>
<tr>
<td>cp2–cp–h</td>
<td>53.9030</td>
<td></td>
</tr>
<tr>
<td>cp–cp2–cp</td>
<td>95.6640</td>
<td></td>
</tr>
<tr>
<td>cp2–cp–cp2</td>
<td>103.6602</td>
<td></td>
</tr>
<tr>
<td>cp2–cp2–cp</td>
<td>56.1173</td>
<td></td>
</tr>
<tr>
<td>bond-bond-1-3</td>
<td>$F_{bb}$</td>
<td></td>
</tr>
<tr>
<td>cp–cp–cp–cp</td>
<td>128.7038</td>
<td></td>
</tr>
<tr>
<td>cp–cp–cp–cp2</td>
<td>-40.3363</td>
<td></td>
</tr>
<tr>
<td>cp–cp–cp2–cp2</td>
<td>-29.7662</td>
<td></td>
</tr>
<tr>
<td>cp–cp2–cp–cp2</td>
<td>-15.3945</td>
<td></td>
</tr>
<tr>
<td>cp2–cp–cp2–cp2</td>
<td>29.6419</td>
<td></td>
</tr>
<tr>
<td>cp–cp–cp–h</td>
<td>-20.4450</td>
<td></td>
</tr>
<tr>
<td>cp2–cp–cp–h</td>
<td>87.5730</td>
<td></td>
</tr>
<tr>
<td>cp2–cp–cp–h</td>
<td>68.9221</td>
<td></td>
</tr>
<tr>
<td>cp–cp2–cp–h</td>
<td>-28.0472</td>
<td></td>
</tr>
<tr>
<td>h–cp–cp–h</td>
<td>-8.4911</td>
<td></td>
</tr>
<tr>
<td>bond-angle</td>
<td>$K_{bb}$</td>
<td>$K_{bb}$</td>
</tr>
<tr>
<td>cp–cp–cp</td>
<td>4.6592</td>
<td></td>
</tr>
<tr>
<td>cp–cp–h</td>
<td>62.4310</td>
<td>66.1697</td>
</tr>
<tr>
<td>cp2–cp–h</td>
<td>42.6635</td>
<td>30.0284</td>
</tr>
<tr>
<td>cp–cp–cp2</td>
<td>65.9070</td>
<td>-1.8949</td>
</tr>
<tr>
<td>cp–cp2–cp2</td>
<td>86.8482</td>
<td>88.2444</td>
</tr>
<tr>
<td>cp2–cp–cp2</td>
<td>21.3256</td>
<td></td>
</tr>
<tr>
<td>cp–cp2–cp</td>
<td>35.7466</td>
<td></td>
</tr>
</tbody>
</table>
Table 4.14: ppcf. Final parameters III.

<table>
<thead>
<tr>
<th>end-bond-torsion-3</th>
<th>$V_1$</th>
<th>$V_2$</th>
<th>$V_1$</th>
<th>$V_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>cp–cp–cp–cp</td>
<td>-1.0748</td>
<td>5.4595</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cp–cp–cp2–cp2</td>
<td>3.6966</td>
<td>4.1523</td>
<td>-0.1919</td>
<td>5.8665</td>
</tr>
<tr>
<td>cp–cp2–cp–cp2</td>
<td>-0.4650</td>
<td>5.9002</td>
<td>0.1531</td>
<td>5.5150</td>
</tr>
<tr>
<td>cp–cp2–cp2–cp</td>
<td>-125.4679</td>
<td>5.0836</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cp2–cp2–cp–cp2</td>
<td>-0.2196</td>
<td>6.4976</td>
<td>-4.2420</td>
<td>2.5439</td>
</tr>
<tr>
<td>cp–cp–cp2–cp</td>
<td>-3.4671</td>
<td>4.2862</td>
<td>-0.6652</td>
<td>5.2467</td>
</tr>
<tr>
<td>cp–cp–cp–h</td>
<td>0.0</td>
<td>-7.5841</td>
<td>0.0</td>
<td>-0.4423</td>
</tr>
<tr>
<td>cp–cp2–cp–h</td>
<td>0.0</td>
<td>-8.8232</td>
<td>0.0</td>
<td>-2.9996</td>
</tr>
<tr>
<td>cp2–cp–cp–h</td>
<td>0.0</td>
<td>-3.0074</td>
<td>0.0</td>
<td>2.4332</td>
</tr>
<tr>
<td>cp2–cp2–cp–h</td>
<td>0.0</td>
<td>-6.4524</td>
<td>0.0</td>
<td>-0.7478</td>
</tr>
<tr>
<td>h–cp–cp–h</td>
<td>0.0</td>
<td>-0.3143</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>middle-bond-torsion-3</th>
<th>$V_1$</th>
<th>$V_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>cp–cp–cp–cp</td>
<td>23.3839</td>
<td>-2.7121</td>
</tr>
<tr>
<td>cp–cp–cp–cp2</td>
<td>23.4579</td>
<td>-2.2443</td>
</tr>
<tr>
<td>cp–cp–cp2–cp2</td>
<td>28.1681</td>
<td>-2.1793</td>
</tr>
<tr>
<td>cp–cp2–cp–cp2</td>
<td>31.8323</td>
<td>-5.9547</td>
</tr>
<tr>
<td>cp–cp2–cp2–cp</td>
<td>-1.9458</td>
<td>-2.2445</td>
</tr>
<tr>
<td>cp2–cp2–cp2–cp2</td>
<td>23.4769</td>
<td>-6.4143</td>
</tr>
<tr>
<td>cp–cp–cp2–cp</td>
<td>27.0895</td>
<td>-1.8601</td>
</tr>
<tr>
<td>cp–cp–cp–h</td>
<td>0.0</td>
<td>-1.2005</td>
</tr>
<tr>
<td>cp–cp2–cp–h</td>
<td>0.0</td>
<td>2.4014</td>
</tr>
<tr>
<td>cp2–cp–cp–h</td>
<td>0.0</td>
<td>-5.0977</td>
</tr>
<tr>
<td>cp2–cp2–cp–h</td>
<td>0.0</td>
<td>1.6376</td>
</tr>
<tr>
<td>h–cp–cp–h</td>
<td>0.0</td>
<td>4.5670</td>
</tr>
</tbody>
</table>
Table 4.15: ppcff. Final parameters IV.

<table>
<thead>
<tr>
<th>angle-torsion-3</th>
<th>$V_1$</th>
<th>$V_2$</th>
<th>$V_1$</th>
<th>$V_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>cp–cp–cp–cp</td>
<td>1.9708</td>
<td>1.2783</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cp–cp–cp–cp2</td>
<td>0.7174</td>
<td>0.7686</td>
<td>-0.5677</td>
<td>5.2054</td>
</tr>
<tr>
<td>cp–cp–cp2–cp2</td>
<td>0.8638</td>
<td>1.8515</td>
<td>0.1376</td>
<td>0.8294</td>
</tr>
<tr>
<td>cp–cp2–cp–cp2</td>
<td>-0.5257</td>
<td>1.9691</td>
<td>-0.0402</td>
<td>3.2645</td>
</tr>
<tr>
<td>cp–cp2–cp2–cp</td>
<td>50.4082</td>
<td>1.2720</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cp2–cp–cp2–cp2</td>
<td>2.0514</td>
<td>0.2524</td>
<td>1.1118</td>
<td>2.9435</td>
</tr>
<tr>
<td>cp–cp–cp2–cp</td>
<td>4.0070</td>
<td>1.3310</td>
<td>2.9895</td>
<td>1.5914</td>
</tr>
<tr>
<td>cp–cp–cp–h</td>
<td>0.0</td>
<td>3.1590</td>
<td>0.0</td>
<td>3.0628</td>
</tr>
<tr>
<td>cp–cp2–cp–h</td>
<td>0.0</td>
<td>3.1321</td>
<td>0.0</td>
<td>3.9213</td>
</tr>
<tr>
<td>cp2–cp–cp–h</td>
<td>0.0</td>
<td>2.0509</td>
<td>0.0</td>
<td>2.9919</td>
</tr>
<tr>
<td>cp2–cp2–cp–h</td>
<td>0.0</td>
<td>10.0469</td>
<td>0.0</td>
<td>3.4910</td>
</tr>
<tr>
<td>h–cp–cp–h</td>
<td>0.0</td>
<td>1.6926</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>angle-angle-torsion-1</th>
<th>$K_{\phi\phi\phi}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>cp–cp–cp–cp</td>
<td>0.0</td>
</tr>
<tr>
<td>cp–cp–cp–cp2</td>
<td>0.0</td>
</tr>
<tr>
<td>cp–cp–cp2–cp2</td>
<td>0.0</td>
</tr>
<tr>
<td>cp–cp2–cp–cp2</td>
<td>0.0</td>
</tr>
<tr>
<td>cp–cp2–cp2–cp</td>
<td>0.0</td>
</tr>
<tr>
<td>cp–cp2–cp2–cp2</td>
<td>0.0</td>
</tr>
<tr>
<td>cp–cp2–cp–h</td>
<td>2.7639</td>
</tr>
<tr>
<td>cp–cp2–cp–h</td>
<td>-13.3469</td>
</tr>
<tr>
<td>cp2–cp–cp–h</td>
<td>40.7370</td>
</tr>
<tr>
<td>cp2–cp2–cp–h</td>
<td>18.1714</td>
</tr>
<tr>
<td>h–cp–cp–h</td>
<td>-8.0121</td>
</tr>
</tbody>
</table>
Table 4.16: ppcff. Error contributions.

<table>
<thead>
<tr>
<th>Term as in Eq. (4.2)</th>
<th>ppcff</th>
<th>pcff</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_{\text{geometry}}$</td>
<td>3.831</td>
<td>33523.023</td>
</tr>
<tr>
<td>$f_{\text{frequencies}}$</td>
<td>68.535</td>
<td>1000.000$^a$</td>
</tr>
<tr>
<td>$f_{\text{modes}}$</td>
<td>9.709</td>
<td>303.822</td>
</tr>
<tr>
<td>$f_{\text{energies}}$</td>
<td>3.657</td>
<td>244.505</td>
</tr>
<tr>
<td>$f_{\text{forces}}$</td>
<td>89.615</td>
<td>36071.349</td>
</tr>
</tbody>
</table>

$^a$ Penalty for nonplanar minimum

order of magnitude larger than that of the other terms in the final forcefield, may be attributed to:

(1) The harmonic approximation introduces a systematic error. Hence, a lower bound for this error term exists.

(2) Improvement in the $f_{\text{frequencies}}$ term could be forced by setting the corresponding weight in the optimization to a very large value. This, however, was found to happen on the expense of the other terms. In the balanced-force-field-approach, this is undesirable. Consequently, attempts to improve the frequency fit were undertaken only to a very limited extent.

Additionally, the eigenvectors of the Hessian, as calculated with ppcff, coincide very well with the B3LYP/6-31G* eigenvectors for the out-of-plane modes in the irreducible representations $B_{1g}$, $B_{2g}$, $A_u$, and $B_{3u}$ (the average absolute value for the scalar product of the normalized eigenvectors of B3LYP/6-31G* and ppcff is 0.96). The in-plane eigenmodes, however, are not well reproduced, but it should be noted that an accurate modelling of the out-of-plane modes is likely to be more important for studies of guest/host systems.

The average error in the relative energy of the 100 geometries was 0.239 kcal/mol. Fig. (4.9) displays the relative ppcff energies plotted vs. the relative DFT reference energies. From Fig. (4.9) it can be seen that both the slope and
the intercept of the linear regression differ from the expected values (1 and 0) by roughly one standard deviation. The root mean square deviation of the ppcff energies from the regression line is 0.097 kcal/mol. The residuals of the linear regression are displayed in the bottom half of Fig. (4.9) and do not indicate any systematic error. The quantile plot of the residuals from Fig. (4.9) vs. the \( \frac{1}{101}, \frac{2}{101}, \ldots, \frac{100}{101} \) quantiles of a normal distribution, scaled by the standard deviation of the residuals, is shown in Fig. (4.10). As expected, the points (with the exception of a few outliers) lie on a straight line with slope one through the origin, indicating normal distribution of the residuals. If the calculated energies are grouped into bins of width 0.2 kcal/mol, further statistical analysis is possible [WEISBERG 1985, COOK
Figure 4.10: Perylene. ppcf energies vs. DFT energies. Quantile plot of the regression for the data in Fig. (4.9). Regression parameters: slope 1.017(20), intercept 0.000(6) kcal/mol, correlation coefficient 0.982.
AND WEISBERG 1994: Let the bins be denoted by \( x_1, \ldots, x_n \) and the corresponding calculated ppcef energies by \( y_{11}, \ldots, y_{m1}, \ldots, y_{1n}, \ldots, y_{mn} \) with means \( \bar{y}_1, \ldots, \bar{y}_n \), and the pure-error sum of squares and the lack-of-fit sum of squares be denoted by

\[
PESS = \sum_{i=1}^{n} \sum_{j=1}^{m_i} (y_{ji} - \bar{y}_i)^2 \tag{4.3a}
\]

and

\[
LOFSS = \sum_{i=1}^{n} (\bar{y}_i - y(x_i))^2 \tag{4.3b}
\]

where \( y(x_i) = a + bx_i \) is the linear predictor for \( y \), then it can be shown that both sums of squares are \( \chi^2 \)-distributed random variables with

\[
df_{PE} = \sum_{i=1}^{n} (m_i - 1) \tag{4.4a}
\]

and

\[
df_{LOF} = N - df_{PE} - 2 \tag{4.4b}
\]

degrees of freedom, where \( N \) is the total number of data points and 2 degrees of freedom are spent in the linear regression. Furthermore, the ratio

\[
F = \frac{LOFSS \cdot df_{PE}}{df_{LOF} \cdot PESS} \tag{4.5}
\]

must be a \( F(df_{LOF}, df_{PE}) \)-distributed random variable [STAHEL 1995]. In this case, with \( df_{LOF} = 17 \) and \( df_{PE} = 7926 \), and \( F = 0.197 \), well below the critical \( (p = 0.95) \) value of 1.754. Again, this indicates the absence of systematic errors in Fig. (4.9).

The average difference between the components of the forces acting on the atoms of these 100 geometries of the optimized forcefield and the DFT calculation,

\[\text{In this case after the binning the } x_i \text{ are themselves means of } \sum_{j=1}^{m_i} x_{ji}/m_i, \text{ hence a robust linear regression algorithm with provision for nonzero variance in } x \text{ must be used, as described in [MEYER 1975].}
\]

\[\text{The sum is smaller than 98, because several bins contained exactly 1 data point and were discarded due to insufficient statistics.} \]
respectively, was 2.330 kcal/mol. Fig. (4.11) displays the comparison of the ppcff forces with their LSD/3-21G* analogs. For better visibility only the binned data are shown (bin width: 1 kcal/mol Å), with the error bars denoting the standard deviation of the ppcff forces within each bin and the residuals inserted at the bottom of Fig. (4.11). For the forces the regression on the quantile plot (not shown) is less satisfactory (the forces are not very independent of each other): the slope of the linear regression is 0.980(2) and the intercept 0.000(7), with a correlation coefficient of 0.979. However, it can be seen from the plot that here the residuals do not

---

27 These residuals represent the residuals of the means, and should not be mistaken for a measure of the residual sum of squares of the forces. The average residual of the 9600 force components is 2.303 kcal/mol Å and the square root of the average residual sum of squares is 3.294 kcal/mol Å.
follow a normal distribution, which can be explained either by the presence of a systematic error or by the assumption of correlation of the forces. The $F$-test, which would reveal a systematic error, yields no indication thereof. Here, $LOFSS = 164$, $df_{LOF} = 83$, $PESS = 103564$, and $df_{PE} = 9505$, which results in $F = 0.181$, again sufficiently small to retain the null-hypothesis that systematic errors are absent.

ppcff reproduces the crystal structure of perylene (cf. Fig. (4.2)) well. Tab. (4.17) presents a comparison of measured crystal structure parameters with the results of a NpT minimization using Discover, where the long range forces were treated using the Ewald sum method [BIOSYM/MSI 1997] (cf. footnote 13 on page 89).

| Table 4.17: Perylene. Crystal structure. Comparison of measured and calculated data. |
|---------------------------------|----------------|----------------|----------------|----------------|----------------|
| $a$/$\text{\AA}$ | $b$/$\text{\AA}$ | $c$/$\text{\AA}$ | $\alpha$/° | $\beta$/° | $\gamma$/° | $\rho$/g cm$^{-3}$ |
| Exp.$^a$ | 11.27 | 10.82 | 10.26 | 90.0 | 100.55 | 90.0 | 1.360 |
| ppcff | 10.79 | 11.42 | 10.00 | 90.0 | 97.42 | 90.0 | 1.372 |

$^a$ in [CAMERMAN AND TROTTER 1964]

The lattice energy (as calculated with the EnergyAnalysis tool in Discover) of the minimized structure was $-29.408$ kcal/mol (the enthalpy of sublimation as reported in [NASS ET AL. 1995] is $29.446$ kcal/mol). Although the calculation neglected the contributions of lattice vibrations and differences in molecular vibrations [WARSHEL AND LIFSON 1970], and did not contain a correction for the $pdV$-term, the agreement is very good.
4.2 Hexane as host

As host in Shpol’skii systems, hexane has not been studied widely. The most recent studies concentrated on anomalous emission from cyclopenta[c,d]pyrene [GOOIJER ET AL. 1998] or on pressure shift and pressure broadenings in hole-burning experiments [FRIEBEL ET AL. 1998]. The most recent relevant work on perylene in hexane is [PFISTER 1973], who found two spectral sites, with a splitting of about 60 cm⁻¹.

4.2.1 Crystal structure

\[\text{Figure 4.12: } n\text{-Hexane. Crystal structure viewed in the } \vec{a}\text{-direction.}\]

\(n\text{-Hexane crystallizes in a triclinic structure (space group } P\bar{1} \text{), shown in Fig. (4.12), Fig. (4.13), and Fig. (4.14). The crystal structure was first determined by [NORMAN AND MATHISEN 1960], and the atomic positions were resolved by [NORMAN AND MATHISEN 1961]. The unit cell parameters are shown in Tab. (4.18). The table also contains the results of MD simulations of a } 5 \times 5 \times 3 \text{ unit cell box in the } NpT \text{ ensemble. Shown is the pcff minimum, as well as the average of a } 50 \text{ ps run (time step } 1 \text{ fs) at } 10^5 \text{ Pa, which, after an equilibration time of } 50 \text{ ps was sampled every } 100 \text{ fs. Long-range interactions were treated via the charge group method with a cutoff of } 9.5 \text{ Å.}\)
Figure 4.13: *n*-Hexane. Crystal structure viewed in the \( \vec{a} \)-direction.

Figure 4.14: *n*-Hexane. Crystal structure viewed in the \( \vec{c} \)-direction.
Table 4.18: Hexane. Crystal structure data.

<table>
<thead>
<tr>
<th></th>
<th>(a/\text{Å})</th>
<th>(b/\text{Å})</th>
<th>(c/\text{Å})</th>
<th>(\alpha/\text{°})</th>
<th>(\beta/\text{°})</th>
<th>(\gamma/\text{°})</th>
<th>(\rho/\text{g cm}^{-3})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exp.(^a)</td>
<td>4.17</td>
<td>4.70</td>
<td>8.57</td>
<td>96.6</td>
<td>87.2</td>
<td>105.0</td>
<td>0.888</td>
</tr>
<tr>
<td>pcff minimum</td>
<td>4.21</td>
<td>4.51</td>
<td>8.68</td>
<td>96.3</td>
<td>88.7</td>
<td>103.1</td>
<td>0.897</td>
</tr>
<tr>
<td>pcff 20 K (NpT)</td>
<td>4.25</td>
<td>4.54</td>
<td>8.69</td>
<td>96.4</td>
<td>88.8</td>
<td>103.5</td>
<td>0.884</td>
</tr>
</tbody>
</table>

\(^a\) in [Norman and Mathisen 1961]

The pcff lattice energy is \(-11.57\text{kcal/mol}\) at 20 K. This includes a long-range correction of \(-1.01\text{kcal/mol}\) calculated according to [Allen and Tildesley 1990], which for a Lennard-Jones 9-6 potential has to be modified and results in

\[
E_{\text{corr}} = \sum_{ij} 2\pi N_i \rho_j \int_{r_c}^{\infty} r^2 v_{ij}(r) \, dr
\]

\[
= \sum_{ij} 2\pi N_i \rho_j \int_{r_c}^{\infty} r^2 \varepsilon_{ij} \left[ 2 \left( \frac{r^*_{ij}}{r_{ij}} \right)^9 - 3 \left( \frac{r^*_{ij}}{r_{ij}} \right)^6 \right] \, dr
\]

\[
= \sum_{ij} 2\pi N_i \rho_j \sigma_{ij}^3 \varepsilon_{ij} \left[ \frac{1}{3} \left( \frac{r^*_{ij}}{r_c} \right)^6 - \left( \frac{r^*_{ij}}{r_c} \right)^3 \right]
\]

where the sum is over all atoms \(i\) on one molecule and \(j\) on all others, \(r_c\) is the cutoff radius, and all other terms in the definition of the Lennard-Jones 9-6 potential are given in Eq. (2.2m) and Eq. (2.3).

This lattice energy should be compared to the semi-experimental value of the enthalpy of sublimation of \(12.15\text{kcal/mol}\) at 178 K, as published by [Bondi 1963]\(^{28}\).

\(^{28}\)As has been pointed out by [Warshel and Lifson 1970], this comparison neglects the contributions of lattice vibrations and differences in molecular vibrations, which amount to about 1 kcal/mol for hexane [Warshel and Lifson 1970]. Additionally, the correction for \(p\Delta V\) amounts to 0.35 kcal/mol at 178 K, hence a reasonable estimate for the sublimation enthalpy based on the pcff calculations would be approximately \(-10.85\text{kcal/mol}\).
4.2.2 Perylene in n-hexane

From Fig. (4.12) it seems that the most natural way to insert a perylene molecule into a pocket in n-hexane is the removal of two heptane molecules adjacent in the $\vec{b}$ direction of the crystal, i.e. removal of a hexane molecule from an arbitrary crystal position $(a, b, c)$, and another one from $(a, b + 1, c)$. In fact, this is the only pocket considered by [PFISTER 1973]. A perylene molecule might be inserted into this pocket with its long axis either aligned with the $\vec{b}$-axis or the $\vec{c}$-axis, creating two possible crystal sites, depicted in Fig. (4.15) and Fig. (4.16).

This is clearly not the only possible pocket. The next obvious insertion site is obtained, if two hexane molecules are removed from the $(a, b, c)$ and $(a + 1, b, c)$

---

29 Strictly spoken, the perylene long axis is in general never perfectly aligned with either axis [LAMOTTE AND JOUSSOT-DUBIEN 1974]. But since the perylene molecule has a roughly rectangular shape and the pocket is an almost rectangular parallelogram, the two possible inclusion sites may for convenience be denoted by the unit cell vector whose angle with the long axis of perylene is the smallest.

30 Fig. (4.15) through Fig. (4.20) show results of NVT minimizations using the ppcff force field developed in Sec. 4.1.3. The size of the simulation box was $5 \times 5 \times 3$ unit cells, as visible in the figures, long range forces were treated by group based cutoffs (cutoff 9.5 Å). If no convergence occurred within 1000 steps, the minimization was terminated. A simulation box of $5 \times 5 \times 3$ unit cells resp. $20.85 \times 23.5 \times 25.71$ Å is very small; it barely fulfills the usual condition of “twice the cutoff distance”. Indeed the shift calculations below were found to require box dimensions at least twice as large (the “production runs” were performed using a $9 \times 9 \times 5$ unit cell simulation box).
crystal positions, creating a pocket in the ac-plane. Again, the perylene molecule might be inserted in two different ways, its long axis aligned with either the \( \vec{a} \) or \( \vec{c} \) direction. These two sites, referred to as Site 3 and Site 4, are shown in Fig. (4.17) and Fig. (4.18). A close look at Fig. (4.18) reveals a significant distortion of the hexane molecules adjacent to perylene, indicating that Site 4 is disruptive to the immediately surrounding crystal structure, and hence not stable enough to be populated significantly. Indeed, in the “production runs” described below, Site 4 proved unstable, and the minimization Fig. (4.18) refers to did succeed only because of the artificially small simulation box (cf. the discussion in Footnote 30 on page 110).

Additionally, one might ask whether it is possible to insert a perylene molecule in a plane not identical with one of the faces of the unit cell. In particular, an insertion diagonal to the unit cell faces seems possible from Fig. (4.14): here a pocket in the \( \vec{a} - \vec{b} \) direction can be created by removal of a hexane molecule at the crystal coordinates \((a, b, c)\), and another at \((a + 1, b - 1, c)\), creating a “diagonal” pocket. In the \( \vec{a} + \vec{b} \) direction, two molecules removed from the \((a, b, c)\) and \((a + 1, b + 1, c)\) positions, will not produce a pocket, as can be seen from Fig. (4.14). Again, perylene might be inserted into the \((a + 1, b - 1, c)\)-pocket. The two possible insertion sites, Site 5 and Site 6, are depicted in Fig. (4.19) and Fig. (4.20).
Figure 4.17: Perylene in n-hexane. Site 3.

Figure 4.18: Perylene in n-hexane. Site 4.
Figure 4.19: Perylene in n-hexane. Site 5.

Figure 4.20: Perylene in n-hexane. Site 6.
The spectral positions corresponding to these inclusion sites can conveniently be investigated using Total Luminescence Spectroscopy [PALEWSKA ET AL. 1991]. In TLS, emission spectra are measured as a function of the excitation energy, yielding two-dimensional plots in which different sites can be easily distinguished by their characteristic overlapping wedge-like patterns, generated by the vibronic sidebands of the 0–0 lines.

Figure 4.21 shows the TL spectrum of perylene in crystalline n-hexane, measured at 6 K. Three different sites can be easily distinguished and are marked by straight horizontal and vertical lines in the figure. The points of intersection of these lines fall on the $E_{ex} = E_{em}$ diagonal, which is also drawn as a dashed diagonal line in the upper left corner of the figure.

The small signals on the very left of Figure 4.21 can not be assigned to any of the major sites. If these signals constituted a fourth site its corresponding 0–n-progression would coincide with the n–1 series of Site 1 and thus be very difficult to observe. Hence these signals can not be assigned to a fourth site.
The positions of the intersection points were determined from the TL spectrum by fitting Gaussian functions to the signals visible in appropriate cuts through the TLS, and are found to be 22413 cm\(^{-1}\), 22474 cm\(^{-1}\), and 22538 cm\(^{-1}\), respectively. The position of the weak signal mentioned above is 22757 cm\(^{-1}\). The accuracy of these numbers is of course given by the resolution of the spectrometer and the achievable accuracy of the fit and is estimated to be ±5 cm\(^{-1}\). The 0–0 line of perylene in vacuo as determined from jet expansion spectroscopy [Sonnen schein et al. 1984] is 24070 cm\(^{-1}\), hence the bathochromic solvation shifts for the three main sites are −1657 cm\(^{-1}\), −1596 cm\(^{-1}\), and −1532 cm\(^{-1}\) (the fourth site would be at −1313 cm\(^{-1}\)).

Table 4.19: Perylene in Hexane. Energies and Spectral shifts. Standard deviations were calculated according to Ref. [Flyvbjerg and Petersen 1989].

<table>
<thead>
<tr>
<th></th>
<th>Site 1</th>
<th>Site 2</th>
<th>Site 3</th>
<th>Site 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exp. shift/cm(^{-1})</td>
<td>−1657</td>
<td>−1596</td>
<td>−1532</td>
<td></td>
</tr>
<tr>
<td>Calc. shift/cm(^{-1})</td>
<td>−1657.0(3)</td>
<td>−1592.6(3)</td>
<td>−1531.9(4)</td>
<td>−1609.8(3)</td>
</tr>
<tr>
<td>Calc. FWHM/cm(^{-1})</td>
<td>37</td>
<td>35</td>
<td>43</td>
<td>39</td>
</tr>
<tr>
<td>Energy/kcal mol(^{-1})</td>
<td>−19.686(4)</td>
<td>−18.944(4)</td>
<td>−19.005(12)</td>
<td>−17.997(4)</td>
</tr>
</tbody>
</table>

The shifts of Sites 1–6 were calculated the following way: in a hexane simulation box of 9 × 5 × 5 unit cells two hexane molecules were replaced by perylene. Then the system was equilibrated in a NVT-ensemble simulation at 20 K for 20000 steps (time step 1 fs, using the ppcff forcefield described in Sec. 4.1.3. Long-range forces were treated by using group-based cutoffs with a cutoff length of 9.5 Å\(^{31}\). Then during additional 20000 steps the lattice energy was calculated every 100 steps using the AnalyzeEnergy tool in Discover. Finally, during the actual production run of 50000 steps the coordinates were written out every 20 steps, creating 251 snapshots of the dynamics. The spectral shifts of these snapshots were then calculated as described in Sec. 2.2.1, with \(\kappa = 3.5; \eta\), following [Geissinger et al. 1998], was adjusted

\(^{31}\)Because the non-neutrality of the perylene charge groups introduces numerical problems when group-based cutoffs are applied, long-range forces should be treated by more sophisticated methods. These, however are very expensive, and would considerably shorten the maximal length of a simulation and finally invalidate the results. Hence group-based cutoffs were used here. No instabilities were observed in the calculation, and the results justify the tradeoff.
so that the largest calculated shift (Site 1) matched the largest observed shift. The results of the calculations are summarized in Tab. (4.19).

The table compares measured values with calculated average shifts and allows an assignment of the three strongest sites. The standard deviations on the means in Tab. (4.19) were calculated according to [Flyvbjerg and Petersen 1989]32. The full widths at half maximum were calculated from the definition of the normal distribution:

\[ p(x) = \frac{1}{\sqrt{2\pi}\sigma^2} \exp\left\{ -\frac{(x - \mu)^2}{2\sigma^2} \right\} \]  
\[ \frac{1}{2} = \frac{p(x)}{p(\mu)} = \exp\left\{ -\frac{(x - \mu)^2}{2\sigma^2} \right\} \]  
\[ (x - \mu)^2 = 2\sigma^2 \ln 2 \]  
\[ FWHM = 2|x - \mu| = 2\sqrt{2}\sigma \ln 2 \]

instead of fitting a normal distribution to a histogram of the shifts as in [Wallenborn et al. 1999].

Only results from Sites 1–3 and 6 are presented in Tab. (4.19). Insertion of a perylene molecule in Sites 4 (Fig. (4.18)) and 5 (Fig. (4.19)) yielded unstable structures which showed strong distortions of the hexane lattice and, in the case of Site 5, even aborted in the molecular dynamics stage.

Tab. (4.19) also contains the calculated intermolecular interactions of the inserted perylene with the hexane host molecules. The order of these energies coincides with the order of the signal intensities in Fig. (4.21), providing independent evidence that the assignment in Tab. (4.19) is correct33.

---

32 The algorithm in [Flyvbjerg and Petersen 1989] is easily seen to be mathematically equivalent to the more frequently used "statistical inefficiency" method of [Allen and Tildesley 1990].

33 Shpol'skii systems are non-equilibrium systems, so the lowest energy inclusion site does not necessarily have to correspond to the strongest signal. But if the orders of signal intensities and
The origin of the weak signal on the left of Fig. (4.21) remains unclear. As seen in Tab. (4.19), can be excluded as a candidate for the origin of this signal. Since the measured shift for this site (if it is a real site) is more than 200 cm\(^{-1}\) smaller than that of Site 3, the packing of the hexane molecules around the perylene guest is expected to be significantly less dense than in the other sites. A pocket created by removal of three hexane molecules is hence the most likely explanation for the weak signal. From the investigation of the tetracene-heptane cluster spectra above it may be speculated that the fourth site may be created from an existing site by removal of a hexane that is atop the perylene rather than at the side (a simple \((a, b, c), (a, b + 1, c), (a, b + 2, c)\)-pocket would not lead to a shift which differs from that of Site 1 by such a large amount). Since there are several such possibilities for each of the sites 1–6, the number of possible pockets is rather large\(^{34}\).

The assignment suggested here differs from the previously accepted one. [Pfister 1973] argue that shift and intermolecular interaction are expected to be positively correlated, and that Site 2 should be energetically favored and thus exhibit the larger bathochromic shift\(^{35}\). The results presented here seem to support the first part of the argument (cf. [Wallenborn et al. 1997]), but not the latter.

\(^{34}\)For example, one such site could be created from Site 5, Fig. (4.19) by removal of an additional hexane at the \((a + 1, b, c)\)-position, turning the “diagonal” pocket into a triangle-shaped pocket.

\(^{35}\)In the arrangement labelled [Site] 2, the interaction between the \(\pi\) electrons delocalised in the two ‘naphthalenic units’ is maximum with the alkanes since one actually finds here the favourable relative disposition of naphthalene in the liquid or solid alkane (for which the long axis of the naphthalene molecule is parallel to the alkane chain with an appreciable overlap)” [Pfister 1973].
4.3 Heptane as host

[Lamotte and Joussot-Dubien 1974] presented the most extensive study of the inclusion sites of several PAHs in heptane. Coronene, perylene, and pyrene were co-crystallized in a heptane single crystal, and the fluorescence polarization measurements allowed a clear assignment of the inclusion site (see below).

The multiplet nature of spectra of Shpol’skii systems has been thoroughly studied for solid solutions in \( n \)-heptane, [Merlé et al. 1977] for example analyze the case of coronene in \( n \)-heptane, employing a electron spin resonance (ESR) technique and conclude that the coronene is inserted approximately in the \( bc \) plane\(^{36}\). This publication also contains the most recent X-ray structure analysis for heptane.

4.3.1 Crystal structure

The crystal structure of \( n \)-heptane is triclinic, its space group \( P\bar{1} \), and was first measured by [Norman and Mathisen 1960]. [Merlé et al. 1977] later determined the atomic positions in the unit cell\(^{37}\).

\[...\]

Figure 4.22: \( n \)-Heptane. Crystal structure viewed in the \( \bar{a} \)-direction.

\(^{36}\)One can then conclude with good confidence that the main component of the ESR multiplet at 77 K is due to coronene molecules inserted in the substitutional plane of the lattice defined by the \( c \) axis and the heptane axis by replacing three alkane chains\(^{9}\) [Merlé et al. 1977].

\(^{37}\)In [Merlé et al. 1977] there are two typographic errors in Table 1. The \( Y \) coordinate of C(4) must be 0.2560 instead of 0.2360, and the \( Y \) coordinate of C(6) is 0.3776, not 0.5776.
Figure 4.23: n-Heptane. Crystal structure viewed in the $\bar{b}$-direction.

This crystal structure is shown in Fig. (4.22) through Fig. (4.24). The unit cell parameters as determined by [MERLE ET AL. 1977] are displayed in Tab. (4.20), along with the results of MD simulations of a $5 \times 5 \times 3$ unit cell box in the $NpT$ ensemble. Shown is the pcff minimum as well as the average of a 50 ps run (time step 1 fs) at $10^5$ Pa, which, after an equilibration time of 50 ps, was sampled every 100 fs. Long-range interactions were treated via the charge group method with a cutoff of 9.5 Å.

Figure 4.24: n-Heptane. Crystal structure viewed in the $\bar{c}$-direction.
The calculated lattice energy of \( n \)-heptane is \(-12.80 \text{kcal/mol} \) (including a long-range correction of \(-1.17 \text{kcal/mol} \). [Bondi 1963] report 13.8 kcal/mol as the enthalpy of sublimation of \( n \)-heptane.

Table 4.20: Heptane. Crystal structure data.

<table>
<thead>
<tr>
<th></th>
<th>( a/\text{Å} )</th>
<th>( b/\text{Å} )</th>
<th>( c/\text{Å} )</th>
<th>( \alpha/\text{°} )</th>
<th>( \beta/\text{°} )</th>
<th>( \gamma/\text{°} )</th>
<th>( \rho/\text{g cm}^{-3} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exp.(^a)</td>
<td>4.15</td>
<td>19.97</td>
<td>4.69</td>
<td>91.3</td>
<td>74.3</td>
<td>85.1</td>
<td>0.890</td>
</tr>
<tr>
<td>pcff minimum</td>
<td>4.21</td>
<td>20.18</td>
<td>4.48</td>
<td>91.0</td>
<td>76.4</td>
<td>87.1</td>
<td>0.902</td>
</tr>
<tr>
<td>pcff ( 20 \text{K NpT} )</td>
<td>4.24</td>
<td>20.22</td>
<td>4.52</td>
<td>91.1</td>
<td>75.8</td>
<td>87.3</td>
<td>0.887</td>
</tr>
</tbody>
</table>

\(^a\) in [Merle et al. 1977]

4.3.2 Perylene in \( n \)-heptane

As mentioned above, [Lamotte and Joussot-Dubien 1974] measured the orientation of perylene included in mixed crystals of perylene in single crystalline \( n \)-heptane. According to them, perylene is included in the \( bc \)-plane\(^{38}\), most likely replacing two heptane molecules in \((a, b, c)\) and \((a, b, c + 1)\) positions. The direction of the transition dipole moments of the two lowest observable transitions of perylene, which, as the authors assume, coincide with the long and short axes of the molecule, were determined to be approximately aligned with the \( b \)- and \( c \)-axes of the crystal. Tab. (4.21) contains data from a ZINDO CI calculation using the Rumer-CI method and the CNDO/1 parameter set. The relative energies of the excited states are as usual overpredicted by several thousand cm\(^{-1}\), but an unambiguous assignment of the states with the observed absorption lines at 445 nm and 265 nm [Lamotte and Joussot-Dubien 1974] is still possible and allows the statement that the transition observed at 445 nm is indeed a \( B_{1u} \)-transition with the transition dipole moment aligned with the long axis of the molecule. The second transition observed by [Lamotte and Joussot-Dubien 1974] is most likely the \( 2B_{2u} \) transition, calculated to

\(^{38}\)The commonly accepted assignment of the crystal axes changed once, cf. [Lamotte et al. 1977]. In earlier publications the \( b \) and \( c \) directions are interchanged. Here the newer convention of [Merle et al. 1977] is used.
be at 237 nm, since the transition dipole moments for the other candidates, especially the \(^1B_{2u}\) transition, are vanishingly small, cf. Tab. (4.21).

Table 4.21: Perylene. ZINDO Rumer-CI (CNDO/1) calculation of several excited singlet states. The molecule was oriented in the \(yz\) plane, the long molecular axis coinciding with the \(y\) axis.

<table>
<thead>
<tr>
<th>State</th>
<th>Irrep</th>
<th>Rel. Energy/cm(^{-1}) (Wavelength/nm)</th>
<th>Transition dipole moment/D</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>(x)</td>
</tr>
<tr>
<td>1</td>
<td>(B_{1u})</td>
<td>27100 (369)</td>
<td>0.00</td>
</tr>
<tr>
<td>3</td>
<td>(B_{2u})</td>
<td>31000 (322)</td>
<td>0.00</td>
</tr>
<tr>
<td>10</td>
<td>(B_{2u})</td>
<td>42100 (237)</td>
<td>0.00</td>
</tr>
</tbody>
</table>

[LAMOTTE AND JOUSSOT-DUBIEN 1974] determined the angles between the projections of the two transition dipole moment vectors and the two most easily cleaved planes, \(ab\), and \(ac\) (named “Face II” and “Face I”, respectively, cf. the discussion of the assignment in [LAMOTTE ET AL. 1977]). According to their measurements, in the slowly cooled mixed crystal they investigated, the perylene occupied only one site. The angle between the transition dipole moment vector of the transition observed at 445 nm and the normal of the \(ac\) plane was determined to be approximately 10\(^\circ\), and the corresponding angle between the transition dipole moment of the second transition at 265 nm and a vector orthogonal to the \(ab\)-plane was found to be approximately 20\(^\circ\).39

Fig. (4.25) through Fig. (4.28) display the four possible insertion sites in the \(ab\) and \(bc\) planes when two heptane molecules are removed from the crystal to form the pocket (cf. Footnote 30 on page 110). It is immediately clear that Site 4 is not stable under minimization of the energy40. Hence Site 4 is unlikely to be present in the Shpol’skii system. This is not very surprising. Comparison with Fig. (4.18) shows that the analogous site in \(n\)-hexane, which has an even smaller unit cell,

\(^{39}\) The long axis of molecules appears to point out of face I making an angle of about 10\(^\circ\) with the normal, whereas the short axis makes an angle of about 20\(^\circ\) with the normal face II. The angles calculated from dichroic ratio are known within ±10\(^\circ\) of error due to many sources of light depolarization and, therefore, a more precise determination of the molecule orientation is not possible." [LAMOTTE AND JOUSSOT-DUBIEN 1974].

\(^{40}\) In fact, the production MD runs for Site 4 never completed successfully.
is also unstable. The detailed description in [Lamotte and Joussot-Dubien 1974] determines the most stable perylene site to be in the bc-plane, with the long axis of perylene approximately orthogonal to the ac-plane (this would be Site 1). The angles of the two transition dipole moment vectors of perylene (and thus the long and short axes of the molecule) with the normals to the ac and ab planes as calculated from the minimized configuration in Fig. (4.25) are 22° and 14°.

Figure 4.25: Perylene in n-heptane. Site 1.

Figure 4.26: Perylene in n-heptane. Site 2.

An alternative assignment would consider Site 2 as the site observed by [Lamotte and Joussot-Dubien 1974]. This would be possible if the literature contained an error in the assignment of crystal axes. As mentioned in Footnote 38 above, the commonly accepted assignment changed once and since [Lamotte and Joussot-Dubien 1974] never mention the crystal structure but rather refer to the cleavage planes exclusively as "Face I" and "Face II", the possibility of such an error is real. However, the measurements of the principal dielectric axes in [Lamotte
Figure 4.27: Perylene in n-heptane. Site 3.

Figure 4.28: Perylene in n-heptane. Site 4.
AND JOUSSOT-DUBIEN 1974] and [LAMOTTE ET AL. 1977] seem to indicate that such an error did not occur.

If Site 2 was the one observed by [LAMOTTE AND JOUSSOT-DUBIEN 1974], then the relevant angles are: long molecular axis with the normal to the $ab$-plane, and short molecular axis with the $ac$-plane. These angles, calculated from the minimized geometry in Fig. (4.26) are $17^\circ$ and $21^\circ$. Hence, the calculated angles of the transition dipole moments with the crystal plane normals are within the experimental error of the measurements in [LAMOTTE AND JOUSSOT-DUBIEN 1974] and thus can not be used to decide between the two assignments.

The sums of the intermolecular interactions are (see below) $-20.5\text{kcal/mol}$ for Site 1 and $-18.8\text{kcal/mol}$ for Site 2. As discussed above, the site observed in a slowly cooled sample should be the most stable site: this would be Site 1.

![Figure 4.29: Perylene in n-heptane. Total Luminescence Spectrum.](image)

Fig. (4.29) shows the TL spectrum of perylene in $n$-heptane. Three sites can be distinguished and are marked with lines in Fig. (4.29). The positions of these sites relative to the 0–0 line of perylene are $-1600\text{cm}^{-1}$, $-1500\text{cm}^{-1}$, and $-1390\text{cm}^{-1}$,
respectively. An absorption measured at 445 nm would correspond to a site at 
\(-1598 \text{ cm}^{-1}\), and thus to the site at 22470 cm\(^{-1}\), rightmost in Fig. (4.29). Hence, if
the reasoning above is correct, this site should be expected to be identical to Site 1
in Fig. (4.25).

<table>
<thead>
<tr>
<th>Site 1</th>
<th>Site 2</th>
<th>Site 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exp. shift/cm(^{-1})</td>
<td>-1600</td>
<td>-1500</td>
</tr>
<tr>
<td>Calc. shift/cm(^{-1})</td>
<td>-1579.0(3)</td>
<td>-1641.9(4)</td>
</tr>
<tr>
<td>FWHM/cm(^{-1})</td>
<td>38</td>
<td>39</td>
</tr>
<tr>
<td>Energy/kcal mol(^{-1})</td>
<td>-20.545(3)</td>
<td>-18.761(2)</td>
</tr>
</tbody>
</table>

Tab. (4.22) summarizes the results of shift calculations on Sites 1–3. These
calculations were performed using the same values for the adjustable parameters
\(\eta = 0.30884, \kappa = 3.5\) as in the calculations for perylene in hexane above. Thus,
the parameters were not adjusted to reproduce a particular observed shift of perylene
in n-heptane.

The results allow an tentative identification of Site 1 with the observed site at 
\(-1600 \text{ cm}^{-1}\). The calculation reproduces the shift to within 20 cm\(^{-1}\). Additionally,
Site 1 is indeed the most stable site and expected to be the only site visible in slowly
cooled samples. Furthermore, the orientation of the included perylene agrees with
the orientation observed by [LAMOTTE AND JOUSSOT-DUBIEN 1974].

Site 3 may be identified with the weak signal in the middle of the three sites in
Fig. (4.29). The arguments for this assignment are not as strong as for Site 1, however.
The shift is reproduced to within 30 cm\(^{-1}\), and the calculated intermolecular
energy indicates that Site 3 is indeed visible in the TLS\(^{41}\).

\(^{41}\)Site 3 is 0.85 kcal/mol less stable than Site 1. In the case of perylene in n-hexane, an energy
difference of 0.74 kcal/mol between the sites assigned to the strongest and the weakest observed
signals resulted in a noticeably weak but still clearly visible signal. Apparently, about 1 kcal/mol
One calculated site, Site 2, was not observed in the TLS. This might be due to energetic reasons. On the other hand, the pocket of Sites 1 and 2 is rather large, and it might be possible that rotational motion of perylene in Site 2, changing it into Site 1, is allowed.

One prominent signal in the TL spectrum could not be assigned. Apparently, just as in the case of perylene in n-hexane, additional pockets are possible, and it may be assumed that the observed site at $-1390 \text{ cm}^{-1}$ is a sign of such a pocket (cf. the discussion on Page 117).
Bibliography


Curriculum Vitae

Name: Ernst-Udo Wallenborn
Date of birth: July 10, 1970
Place of birth: Trier (Germany)
Nationality: German
Parents: Christoph Wallenborn and Helga Wallenborn (b. Barzen)

Education

July 1976 – June 1979 attended elementary school at Cologne, Birresborn and Gerolstein (all Germany)

July 1979 – June 1988 attended the St. Matthias Gymansium Gerolstein (Germany).
Abitur in June 1988

Studies

graduated as Diplom-Chemiker (ETH).
diploma thesis in the group of Prof. W. van Gunsteren about grand-canonical Monte-Carlo-Molecular Dynamics-Simulations on the parallel computer MUSIC.


Feb. 1994 – Mar. 1999 postgraduate studies in the group of Prof. Urs P. Wild at the physical chemistry laboratory of the ETH.
# Index

antisymmetrization operator, 15

Class II force fields, 9
closure approximation, see mean excitation energy approximation

cluster growth
  simulated annealing, 40
configuration state function, 19
  Slater determinant expansion, 20
Coulomb integral, 20
counterpoise correction, 43
crystal structure
  of heptane, 118
  of hexane, 107
  of perylene, 70
CSF, see configuration state function

force field optimization, 81

Green function
  semiclassical form, 29

heptane, 118
  as host in Shpol'skii systems, 118
assignment of crystal axes, 120
crystal structure, 118
density, 120
lattice energy, 119
perylene in
  TLS, 124
  unit cell dimensions, 120
hexane, 107
  as host in Shpol'skii systems, 107
crystal structure, 107
density, 109

enthalpy of sublimation, 109
lattice energy, 107
octatetraene in, 5
perylene in
  TLS, 114
  unit cell dimensions, 109

intermolecular interaction
dispersive term
  additivity, 18
  definition, 14
electrostatic interaction, 12, 14
Lennard-Jones potential, 12
long-range correction, 109
multipole expansion, 23
operator, 13
perturbational theory, 13
van-der-Waals interaction, 14
interstellar polycyclic aromatic hydrocarbon hypothesis, 74
ionization potential
  calculated shift as a function of, 38
  of methyl and methylene groups, 38

mean excitation energy approximation, 25

pcff, 9
  definition of terms, 10
  description of terms, 11
perylene distortion in, 78
perylene, 70
  crystal structure, 70
density, 106
Index

enthalpy of sublimation, 106
in n-heptane
  spectral shifts, 125
  TLS, 124
in n-hexane
  spectral shifts, 115
  TLS, 114
molecular geometry, 73
non-aromaticity of the middle ring, 70
pcff distortion, 78
planarity of, 72
ppcff force field for, 81
transition dipole moment, 120
unit cell dimensions, 106
vibrational analysis, 74
  $A_g$ modes, 75
  $A_u$ modes, 78
  $B_{1g}$ modes, 76
  $B_{1u}$ modes, 79
  $B_{2g}$ modes, 76
  $B_{2u}$ modes, 80
  $B_{3g}$ modes, 77
  $B_{3u}$ modes, 81
butterfly mode, 81
  Mode 1, $A_u$, 82
  Mode 2, $B_{3u}$, 82
  Mode 3, $B_{2g}$, 83
  Mode 4, $B_{3u}$, 83
  Mode 9, $A_g$, 84
WCC force field for, 74
polarizability
  additivity in alkanes, 36
  of methyl and methylene groups, 38
  quantum mechanical expression, 25
ppcff, 81
  error function, 89
  functional form, 86
optimization of, 93
parameter definition, 87
parameters, 96
partial charges, 89
performance of, 96
crystal structure, 106
error function terms, 96
lattice energy, 106
Rayleigh-Schrödinger Perturbation Theory, 14
reaction field method, 3
RSPT, see Rayleigh-Schrödinger Perturbation Theory
SAPT, see symmetry-adapted perturbation theory
SBEJ
  influence of the ionization potential on the shift, 38
  parameter choice, 46
SBEJ method, 16
dispersive shift, 27
electronic factor, 27
excited state interaction, 19
geometric factor, 26
ground state interaction, 19
parameter choice, 39
shift by a single group, 18, 22, 27
Shpol’skii systems, 6, 69
heptane as host, 118
hexane as host, 107
simulated annealing
  cluster growth by, 40
Slater rules, 19
  integration over spin space, 21
symmetry-adapted perturbation theory, 15
tetracene, 33
  heptane clusters
calculated shifts, 46
DFT minima, 42
force field minima, 41
measured spectra, 33
molecular dynamics, 39
transition dipole moment autocorrelation function, 40, 46
molecular geometry, 35
vibrational analysis, 35
  Mode 1, $B_{3u}$, 36
  Mode 2, $A_u$, 37
  Mode 3, $B_{1g}$, 37
  Mode 4, $B_{1u}$, 38
tetracene-heptane$_1$, 49
density functional theory
  global minimum, 43
  local minimum, 44
force field
  global minimum, 41
  local minimum, 42
spectrum
  intensity of vibronic lines, 36, 49
  sideband structure, 49
tetracene-heptane$_2$, 58
force field
  local minimum, 45
spectrum
  inhomogeneous structure, 61
  transition dipole moment autocorrelation function, 61
tetracene-heptane$_{3-6}$, 62
TLS
  of perylene
    in $n$-heptane, 124
    in $n$-hexane, 114
Total Luminescence Spectroscopy, 69
transition dipole moment
  autocorrelation function, 28
  semiclassical form, 29
transition monopoles, 22, 25