Calculation of molecular thermochemical data and their availability in databases

In:

Battin-Leclerc F;Simmie J E;Blurock E (Eds.)

Cleaner Combustion: Developing Detailed Chemical Kinetic Models.

London: Springer, 2013. pp. 515-547.

Chapter 20
Calculation of molecular thermochemical data and their availability in databases

Elke Goos[1] and György Lendvay[2]

Abstract Thermodynamic properties of molecules can be obtained by experiment, by statistical mechanics in conjunction with electronic structure theory and by empirical rules like group additivity. The latter two methods are briefly reviewed in this chapter. The overview of electronic structure methods is intended for readers less experienced in electronic structure theory and focuses on concepts without going into mathematical details. This is followed by a brief description of group additivity schemes; finally, an overview of databases listing reliable thermochemical data is given.

20.1 Introduction

Both our understanding and the reliability of predictions based on modeling complex reaction systems depend on the accuracy of the thermodynamic data of the molecules (stable molecules, radicals, individual atoms and possibly ions) present in the system. Measurements of the most important thermodynamic properties like enthalpy, entropy and specific heat are not easy. Enthalpies of formation are derived from calorimetric experiments (Marsh 2001), which determine heats of reaction yielding unique products whose enthalpies of formation are known. To determine heats of formation of a compound, one can use Haber-Born type thermodynamic cycles (Atkins, de Paula 2005) involving reactions of the selected compound in which heats of formation of all other

[1] E. Goos
Institute of Combustion Technology, Deutsches Zentrum für Luft- und Raumfahrt e.V. (German Aerospace Center (DLR)), Pfaffenwaldring 38-40, 70569 Stuttgart, Germany
e-mail: Elke.Goos@dlr.de

[2] G. Lendvay
Institute of Materials and Environmental Chemistry, Research Center for Natural Sciences, Hungarian Academy of Sciences, Pusztaszeri út 59-67, Budapest, H-1025, Hungary
e-mail: lendvay.gyorgy@ttk.mta.hu

reactants and products are known. By measuring the reaction enthalpies in the cycle, one can get the heat of formation of the compound of interest. Such "ideal" reactions are complete combustion of hydrocarbons producing the stoichiometric amount of H_2O and CO_2. These methods are, unfortunately, not universally applicable because, for molecules above a certain complexity (more than about 20 atoms and/or several functional groups), design of appropriate thermodynamic cycles with well defined stoichiometry is not simple. For example, there are species whose reactions (mostly those with multiple product channels) are very sensitive to the conditions so that their stoichiometry is hard to control. Under such conditions the measurable "reaction enthalpy" actually refers to a set of simultaneous reactions whose relative importance is not well defined.

Even more complicated is the case of reactive species like free radicals whose heats of formation cannot be determined directly in calorimetric measurements. Instead, a completely different kind of experiments, the measurement of reaction rate coefficients need to be performed, and heats of formation can be inferred indirectly from their data (Benson 1968).

A widely used method, the so-called Second Law analysis yields enthalpies and, in principle, entropies of reaction. The rate coefficients for the forward and reverse reaction are measured, from their ratio the equilibrium constant of the reaction and its temperature dependence is derived, and using the Van't Hoff equation enthalpies of reaction and of formation can be obtained.

The Third Law method requires the equilibrium constant at a single temperature but the reaction entropy needs to be known from independent sources. The temperature dependence of the thermodynamic properties can be obtained from heat capacities, the measurement of which is not easier than that of enthalpies, so that the most useful is to derive them from the temperature dependence of enthalpies of formation, obtained in direct experiments.

Certain thermodynamic quantities such as dissociation and ionization energies of individual molecules can be determined by spectroscopic methods. These methods provide probably the most accurate experimental information on thermodynamic quantities. Their applicability, however, is limited: one needs to know several other data to get the desired quantity. For example, to derive the enthalpy of formation of a fragmentation product, one needs to measure the dissociation (or fragmentation) energy of the parent compound, know the enthalpy of formation of the latter and of the other fragmentation product.. Spectroscopic measurements also provide molecular properties (vib-rotational energy levels etc.) from which entropies and heat capacities can be calculated.

The available thermodynamic data has been collected in databases. The earlier forms were simple printed tables (Cox and Pilcher 1970) while more recent versions are databases in electronic form with specified data formats (Gordon and McBride 1971, Kee et al. 1996) with the possibility of electronic data retrieval (Goos et al. 2013a, NIST-JANAF 2012). These tables are incomplete: the thermochemical properties are not available for all known chemical species. Even if there were experimental methods available for the determination of

thermodynamic properties of all kinds of compounds, the experiments would be hard to perform for all important species in a reactive system merely due to their enormous number: there are hundreds or thousands of species involved in a real chemical system like the combustion of fuels or even in that of a compound as simple as pure methane.

Theory can effectively help filling in the gaps in the thermochemical tables containing experimental data. Moreover, theoretical approaches can also allow correction of values that are not consistent with the rest of the data. There are two major approaches to determine thermodynamic properties of chemical species by applying computational methods. One direction is their calculation *ab initio* using up-to-date methods of electronic structure theory (quantum chemistry). The basis of these methods is that one considers the chemical system as an ensemble of atomic nuclei and electrons and solves the electronic Schrödinger equation numerically using sophisticated techniques and smart approximations. No chemical or other empirical information is used, and every molecule has to be calculated from scratch without any reference to the results obtained for related compounds. As a result, no generalization can be made except by handling the calculated data in the same way as experimental values and trying to observe tendencies and quantify them in "empirical" rules.

The other way of obtaining thermodynamic data without measurement is more empirical but very successful: it utilizes the fact that, especially in organic compounds, functional groups carry not only their chemical characteristics from one molecule to the other, but also numerical values of their physical properties. The values of the entire molecule's physical quantities can be quite closely approximated by summing the contributions of the groups constituting the molecule. The additivity is not exact, i.e. a group's contribution may depend on the environment, and a lot of empiricism is introduced for reliable predictions. As such, each additivity scheme is applicable only within the range of compounds it has been worked out for.

In the first part of this chapter the ab initio methods will be reviewed; the second will be devoted to an overview of group additivity techniques. In the third part the available databases will be surveyed.

20.2 Calculation of thermodynamic properties using electronic structure theory

Electronic structure theory currently is at the level that for simple cases it can produce results that match, or even exceed the accuracy of measurements. For example, the experimental spectroscopic parameters of the H_2 molecule and its deuterium-substituted isotopomers have been reproduced by theory with almost 10^{-4} cm^{-1} accuracy (Piszczatowski et al. 2009). A more complicated example is the vib-rotational spectrum of the water molecule whose spectral lines can be assigned

using the ab initio potential surface of the molecule calculated to better than a cm^{-1} accuracy (Polyansky et al. 2003). Another example, directly connected to combustion chemistry is the standard heat of formation of the OH radical whose generally accepted value was 2 kJ mol^{-1} too low before a joint experimental and theoretical effort yielded the accurate value of $\Delta_f H_0(OH) = 8.85 \pm 0.07$ kcal mol^{-1} (Ruscic et al. 2002). For molecules that are more complicated, ab initio calculation of thermodynamic properties is less straightforward: the accessible accuracy decreases with the size of the molecule, expressed not only by the number of atomic nuclei, rather by the number of electrons.

The basic idea of electronic structure calculations is that the molecule (using the term in a wider sense to include radicals and ions) is taken to be a physical system consisting of atomic nuclei and electrons that interact via electrostatic forces. The description of the motion of electrons requires the use of quantum mechanics; the motion of the nuclei is at the borderline of the ranges of validity of quantum and classical mechanics. Accordingly, the accurate treatment of a system consisting of electrons and nuclei requires the solution of the corresponding Schrödinger equation.

Setting up the Schrödinger equation is trivial: the Hamiltonian operator, whose eigenvalues and eigenfunctions are searched for, consists of the kinetic energy operators of the electrons and of the nuclei and the ensemble of attractive potential energy terms describing the nucleus-electron interaction, and repulsive term corresponding to the electron-electron as well as nucleus-nucleus repulsion (Figure 20.1).
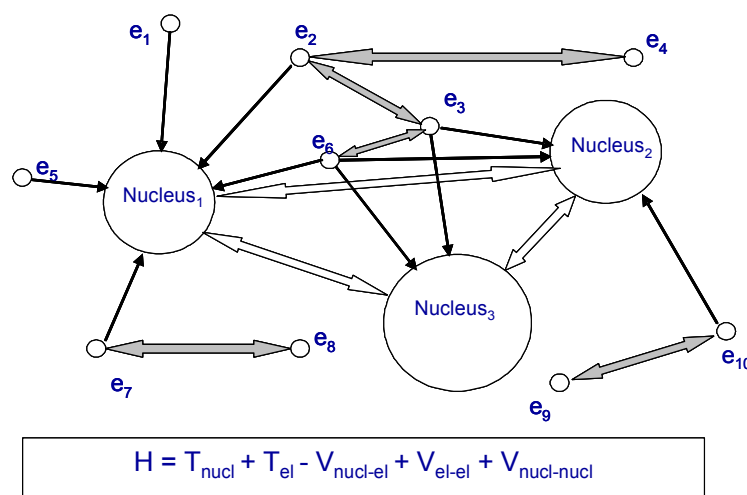

$$H = T_{nucl} + T_{el} - V_{nucl-el} + V_{el-el} + V_{nucl-nucl}$$

Figure 20.1 The energy terms in the Hamilton operator for a polyatomic molecule (water in this example: Nucleus$_1$=Nucleus$_2$=H; Nucleus$_3$=O). The black arrows denote attractive interactions, the gray and white arrows the electron-electron and

nucleus-nucleus repulsion. Only a few examples of each kind of interaction are shown.

The solution is much less trivial. No analytic solution can be written down even for the simplest systems, so that all methods of solving the molecular Schrödinger equation are based on approximate schemes. Some of the approximations are physical in nature, others are empirical in the sense that certain numerical contributions are neglected, or numerical tendencies are observed and based on them various approximate methods are constructed.

The most important approximation is the separation of nuclear and electronic motion (called *Born-Oppenheimer approximation*), based on the fact that the mass of electrons is much smaller than that of atomic nuclei (by a factor of about 2000 for the lightest atom, protium). The reasoning is that as the heavy particles, the nuclei move, the electrons, moving much faster, instantaneously adjust to the new Coulomb field of the nuclei. The detailed derivation shows that if certain conditions hold, the nuclear and electronic Schrödinger equations can be solved separately. Solving the electronic part at a fixed spatial arrangement (configuration) of nuclei, one obtains the energy eigenvalues corresponding to different electronic states. The eigenvalues, called electronic energies depend on the nuclear coordinates and, together with the nucleus-nucleus repulsion terms (that also depend on the nuclear arrangement) they constitute a potential energy function corresponding to each electronic state of the system. These scalar-vector functions are called potential energy surfaces (PES) and are the most important targets of electronic structure calculations (a schematic example is shown in Figure 20.2).

A PES shows the relative energy of a chemical system (a given set of atoms) in different arrangements that may correspond to different compounds, with the assumption that the nuclear motion is frozen. This is not what one can observe macroscopically: thermodynamic properties refer to ensembles of vibrating-rotating molecules. Accordingly, calculation of thermodynamic observables requires further work: the knowledge of energy levels corresponding to nuclear motion is also needed. The vibrational and rotational energy levels can be obtained by solving the nuclear Schrödinger equation in which the forces acting on the atoms, represented by the nuclei are derived from the PES. Generally, approximate methods are used to estimate the energy levels of molecular vibrations and rotations.

The main purpose of electronic structure theory, often called *quantum chemistry*[3], is then to provide a quantum mechanical description of electronic

---

[3] In fields where the needed molecular properties are calculated using electronic structure theory methods, one often finds the reference "quantum mechanical" to emphasize that the parameters are calculated using electronic structure theory instead of being estimated using empirical rules. This term is quite misleading as it makes the impression that the nuclear motion is also handled using the methods of quantum mechanics, which is generally done in a very approximate (rigid

motion at fixed nuclear arrangements (referred to as nuclear or molecular geometry) and obtain the electronic energy. There are two paradigms applied for this purpose: the so-called wave-function based methods and the density-functional based methods. The *wave-function based methods* (wave function theory, WFT (Truhlar 2008)) are devoted to solve the Schrödinger equation, get the wave function and the energy eigenvalue corresponding to the given nuclear arrangement. The proper methods for the solution of the Schrödinger equation work "ab initio" (i.e. no empirical information is used) and are numerical (in contrast to being analytical).
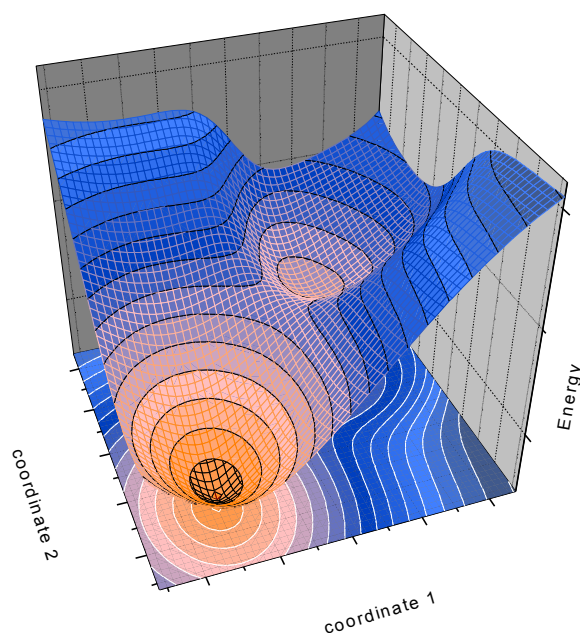

Figure 20.2 Schematic view of the section of a potential energy surface along two coordinates. Two minima corresponding to two isomers as well as a saddle point separating them is visible.

They apply the principle of successive approximation and are often variational. If a method is variational, the best solution is the one that provides the lowest energy, so that finding the right wave function is equivalent to solving a general optimization problem. Various approximations can be introduced to make the

---

rotor − harmonic oscillator) way. "Quantum chemistry" is a much better term to describe the methodology in such applications.

problem tractable, but these are "controlled" in the sense that we know what we omit or simplify. As a result, in principle, the solution can be made to approach the exact value by omitting step by step the simplifications and approximations.

Density functional theory (DFT) applies a conceptually different approach, based on the Hohenberg-Kohn theorem. This theorem claims that there is a one-to-one correspondence between the energy and the total electron density at every molecular geometry (in mathematical terms, the energy is a *functional* of the electron *density*). The purpose is then to find the correct electron density, from which one can get the electronic energy at the selected molecular geometry. The problem is that the general mathematical form of the energy functional is not known and seems not to be feasible to obtain except for its qualitative behavior in limiting cases.

Table 1. Some properties of ab initio and density functional theory-based quantum chemical methods

Property	Ab initio methods	Density functional theory methods
object to find	electron wave function	electron density function
way of improvement	systematic	empirical
computational expense	can be large	small
size of molecules routinely handled	very accurate: 6 nonhydrogen average: 20-30 nonhydrogen atoms	hundreds of atoms
expertise needed	very accurate: significant average: medium	little
reliability	very accurate: large average: medium	unpredictable
systems not possible to handle	large molecules, polyatomic transition metal complexes	molecules where dispersion interactions are important

There are a number of approximate functional forms that contain numerical parameters. The latter are generally optimized so that the energy calculated via the functional should match that obtained for selected systems from ab initio calculations considered to be exact. This is similar to introducing parameters that are taken from experiment, but here the external source of information is not experiment but accurate theory. Although DFT does not use empirical parameters, it can not be considered to be a truly ab initio method. This is why the term "first principles methods" (Truhlar 2008) for DFT has been introduced. A consequence of the way of selecting the numerical parameters is that the methods of DFT may work well for systems that are similar to those included in the parameter optimization set, but for those that differ significantly from the latter, the results provided by DFT may be far from good. There are numerous versions of DFT

applying different functionals. Oddly enough, although DFT does not refer to the wave function directly, when the density is needed, in the majority of codes the machinery of the wave-function based methods is applied, and the density is calculated from the wave function.

Finding the energy corresponding to the electronic motion is only one step on the way of getting thermodynamic properties of molecules. The absolute energy at a certain nuclear arrangement does not characterize a real molecule. Molecules are semi-rigid structures oscillating around some equilibrium geometry, rotating and flying in space. The amplitude of vibrations, the speed of rotations, i.e., the internal energy corresponding to the given quantum state, as well as the velocity of flight is generally different for each molecule in the ensemble. The macroscopically observable properties are averages over ensembles of molecules in thermal equilibrium, characterized by the temperature. The population of various quantum states then follows the Boltzmann distribution. Accordingly, when thermodynamic properties are calculated, the effects of intra- and intermolecular nuclear motion, the *thermal contributions* need also be taken into account.

The accurate description of molecular vibrations and rotations also requires the application of quantum mechanics: one has to solve the Schrödinger equation for nuclear motion. The nature of motion is determined by the potential surface, more precisely, by the shape of the PES near the equilibrium molecular structure. Any distortion of the molecule's geometry from this nuclear arrangement gives rise to forces preventing further distortion and directing the molecule back to the equilibrium. Taking into account that forces are the negative derivatives of the potential energy surface, this means that all derivatives are zero at the molecular geometry corresponding to the equilibrium and positive in its neighborhood, which means the energy is minimum at this structure.

The first step of electronic structure calculations is then finding the minimum on the potential energy surface, more precisely, the molecular geometry corresponding to it (the *equilibrium geometry*), which in numerical mathematical terms is a multivariate optimization procedure. Once the minimum is found, one has to characterize the shape of the potential surface whose parameters go into the nuclear Schrödinger equation. Commercially available quantum chemistry packages are all equipped with programs handling the minimum search, often referred to as "geometry optimization".

Intramolecular motion is bound (in contrast to being free, like the motion of the center of mass of the molecule, which is free flight), and quantum mechanics tells that such systems have quantized energy levels. Molecules have a number of vib-rotational states having distinct energy levels and the population of these states varies depending on the temperature. These states and energy levels have to be determined in the next step on the way to thermodynamic quantities. Once the energy levels are known, the formulas of statistical mechanics can be applied to determine the values of the thermodynamic properties characterizing the molecule, the central quantity being the *partition function*. In a general case there is no

analytical formula for the partition function, but various simplified models (for example, the normal mode approximation) allow one to derive such formulas. In critical cases, however, the magnitude of inaccuracy of these schemes is the same or larger than the actual effect to be calculated, so that it is necessary to apply corrections or obtain the energy levels from scratch and calculate the partition function numerically.

The molecular vib-rotational levels for relatively low-energy states are arranged systematically. The rules governing the spacing between them are generally interpreted using simplified models. These models are devoted to make possible the *assignment* of spectral transitions so that one can tell which vibrational and rotational degree of freedom changes its quantum number from which value to which other. The most generally applicable approximate or zeroth-order model is the normal mode picture. This is based on the approximation that, if the amplitude of molecular motion is small, only that part of the PES is visited that is very close to the equilibrium geometry. Then, in the nuclear Schrödinger equation this part of the PES will determine the energy of the lowest vibrational states. Being close to a minimum, the potential surface can then be approximated as a set of quadratic terms. The kinetic energy is also a sum of quadratic terms, and for such systems with an appropriate coordinate transformation the Hamiltonian can be decomposed into terms corresponding to separate harmonic oscillators that are called *vibrational modes*.

For an *N*-atomic molecule there are 3*N*-6 such modes (3*N*-5 if the molecule is linear). For the normal-mode harmonic oscillators the energy levels are equally spaced, the spacing being *h* (Planck's constant) times the frequency *v* that the oscillator had if it was classical, starting from a nonzero energy level 1/2 *hv*.

At absolute zero temperature all oscillators in a thermodynamic (canonical) ensemble occupy the lowest energy level (for this it is called the zero-point energy). For harmonic oscillators there is a simple analytical formula for the partition function and other thermodynamic properties. It should be mentioned that the normal-mode picture is an approximation and in reality the normal vibrational modes are "coupled" to each other because of the failure of the quadratic approximation of the PES. This means that within the amplitude of vibration the PES deviates from the quadratic shape, higher-order terms arise that first destroy the harmonicity of the vibration (levels are not equally spaced) then the separability of normal modes (meaning that the frequency of one vibration depends on the excitation of the other).

An additional complication is that molecules also rotate. At low excitation, rotation and vibrations can be considered separable (the rigid-rotor-harmonic oscillator, RRHO approximation). Then the partition function is also separable into terms corresponding to various degrees of freedom: translational, vibrational, rotational and electronic contributions that can be calculated from analytical formulas. It has to be kept in mind, however, that this picture is an approximation. The formulas can be found in many textbooks and reference works and are not presented here.

The most frequently occurring problem causing the failure of the RRHO approximation is that certain vibrational (most of the time torsional) modes have very low frequencies. It often happens that the torsional motion, a hindered internal rotation has a low barrier, so low that it is lower than the second (or sometimes even the first) harmonic vibrational energy level, causing the obvious failure of the approximation. There are several correction schemes for treating hindered rotors (Pitzer and Gwinn 1942, Ayala and Schlegel 1998, Ellingson et al. 2006, Vansteenkiste et al. 2006).

Most quantum chemistry codes are set up to perform the calculation of vibrational frequencies and the motion of atoms in the normal vibrational modes (the procedure is called *vibrational analysis*). From the equilibrium geometry of the molecule the principal moments of inertia are calculated, from which the (rigid-rotor) rotational contribution to the partition function is obtained. The thermodynamic parameters are generally calculated using the RRHO approximation. The corrections that go beyond the RRHO approximations are generally handled with the user's own codes, although some quantum chemistry codes do contain ways for the automatic calculation of hindered-rotor partition functions.

The block diagram of the overall procedure is shown in Figure 20.3. The need for the third block will become clear in the next section.
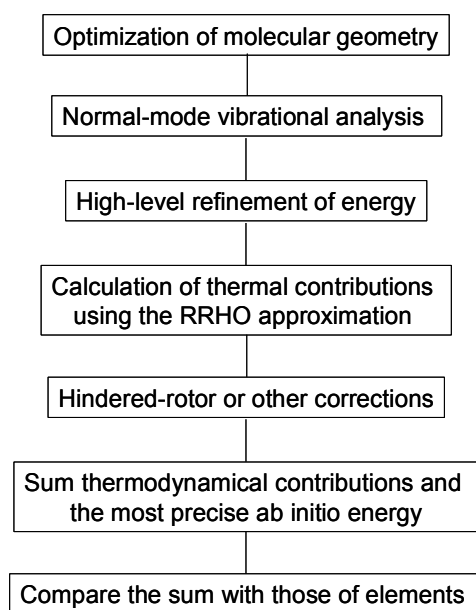

Figure 20.3 The flow chart of the calculation of molecular thermodynamic parameters ab initio

The theory described above is based on the Schrödinger equation written down for the electrons and nuclei of the molecule; no external interaction is taken into account. This is a good approximation for molecules in the gas phase. In condensed phase systems the interaction with the neighboring molecules is much stronger than in the gas phase, but it generally does not completely destroy the picture working for the internal motion, yet the thermodynamic parameters generally differ significantly from the gas phase case. If the molecule in question is embedded in a liquid and forms a dilute solution, the solvent's electrostatic interaction with the electrons of the molecule can be included into the electronic structure part of the calculation. The precise calculation of the thermodynamic properties and the handling of solutions that are not dilute goes beyond the possibilities of ab initio quantum chemistry. Similar is the case of solids, whose quantum mechanical treatment requires completely different approaches and technologies. We do not detail these methods in this chapter.

20.3 Electronic structure methods

The ab initio methods available for routine calculations are all approximate. The mathematical details are often very complicated. The present description is intended to introduce the reader not familiar with electronic structure theory to the principles only, without any mathematical formulas, and as such, it is not absolutely precise in certain aspects. The emphasis is placed more on the hierarchy of approximations and the basic concepts and principles of electronic structure calculations. For interested users, a list of textbooks is provided at the end of the section.

Table 20.1 The hierarchy of approximations of quantum chemistry

Approximation	Main feature not covered	Occurrence of failure
Born-Oppenheimer separation of electronic and nuclear motion	nonadiabatic effects	mostly away from equilibrium geometries
Nonrelativistic electronic Schrödinger equation	inner-shell relativistic effects; spin-orbit coupling	mostly in compounds of atoms with large atomic number
One-electron (mean field) approximation	dynamical and chemical electron correlation	everywhere
Expansion of molecular orbitals in terms of atomic orbital basis sets	accuracy because of incompleteness of basis set	everywhere

Beyond the Born-Oppenheimer separation described in the previous section, the second approximation in most of the electronic structure methods currently in routine use is that electronic motion is described by the nonrelativistic electronic Schrödinger equation. Application of relativistic quantum mechanics is desirable, especially for compounds containing heavy atoms because of the high speed of the innermost electrons in the latter. There are various approximate methods to remedy the missed effects.

The third approximation is that most ab initio methods are based on the one-electron approximation or Hartree-Fock (HF) method, and various corrections are applied *a posteriori*. The HF method is based on a formal decomposition of the polyelectronic Hamiltonian into terms each depending on the coordinates of only one electron. The eigenfunctions of such Hamiltonians are the products of one-electron wave functions called *molecular orbitals* (MO). The equations are made a bit complicated by the fact that electrons are fermions, and the Pauli principle requires that the wave function be antisymmetric with respect to exchange of any two electrons. In particular, the simple product of MOs does not fulfill this requirement; the antisymmetry is achieved by summing all permutations of electrons among the MOs each with the proper sign (which is equivalent to forming a determinant, called Slater determinant).

The one-electron Hamiltonian of electron i contains its kinetic energy operator, the attraction exerted on it by the nuclei of all atoms in the molecule, and the sum of the average repulsion due to each of the rest of the electrons (the *mean* electron repulsion *field*). The average interaction between the selected electron i and another electron j depends on the location of electron i but not on that of electron j because the potential energy is averaged over the spatial distribution of the latter. This is achieved by assuming that we know the wave function (the MO) of electron j; then its square will give the probability that electron j is located at any selected point in space, which, multiplied by the charge of the electron is a charge density.
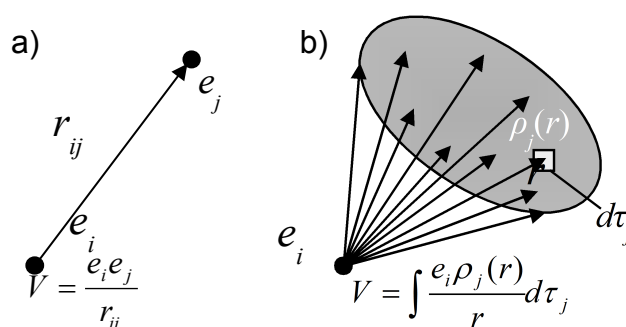

Figure 20.4 a) The real interaction of two electrons: repulsion of two point charges. b) The electron-electron interaction in the one-electron approximation:

electron e_i is repelled by a diffuse charge distribution $\rho_j(r)$ corresponding to electron e_j.

The overall interaction of the "cloud" of electron j with electron i (the latter being at a given point in space, Figure 20.4) is the average of the interaction of the "parts" of electron j at each space element with electron i (an integral over the coordinates of electron j). This is an approximation that makes possible the decomposition of the polyelectronic Hamiltonian into one-electron terms. What we miss this way from the exact Hamiltonian is that, as electron i "feels" only the average repulsion of electron j, it can penetrate the cloud of the latter, instead of continuously "watching" where the other electron actually is and adjusting its own motion to avoid getting close to the other. In other words, in this model the motion of electrons is not *dynamically correlated*.

The wave function of electron i is the eigenfunction of the Hartree-Fock Hamiltonian or Fock operator. The latter contains the wave functions of all other electrons in the average repulsion terms. Those wave functions, however, are eigenfunctions of the respective Fock operators that can be written down and solved if we know the wave functions of the rest of electrons, including electron i. This problem is solved iteratively, by successive approximation: a set of trial wave functions is generated (often called initial guess), then the electron-electron repulsion (and some other) integrals are calculated, and each one-electron Schrödinger equation is solved, yielding a set of wave functions each of which changes from the initial guess because they adjust themselves to match the mean field generated by the other electrons. This way a new repulsion field arises. These modified wave functions are fed again into the Fock operators, the eigenfunctions are found, and the procedure is repeated until the wave functions reproduce themselves, in other words, the Coulomb field is consistent with itself. This is called the *self consistent field* (SCF) method, often meant to be synonymous to the HF method.

The question is how one can generate MOs that are flexible so that

1. they should be able to describe electron density where it is needed (near atoms, with the possibility of distortion towards other atoms) and

2. be easy to modify during the iteration.

It is hard to find simple analytical functions that fulfill these requirements. What helps is a theorem of quantum mechanics, which states that all wave functions can be expanded in terms of any set of functions that constitute a so-called *complete basis*. This means that the desired wave function can be constructed as a linear combination of the basis functions each having the appropriate coefficient. (It should be noted that a complete basis is always infinite-dimensional). Selecting a basis set, the MOs can be modified by changing the weights of the basis functions in the linear combination. The basis sets we can actually use are never complete in the mathematical sense (we can never use an infinite number of functions). The larger the number of basis functions, the more computationally intense is the solution of the Schrödinger equation. The good basis functions are those that look

similar to what the desired orbitals look like, and only slight modifications need to be made. According to accurate numerical calculations, the one-electron wave functions of free atoms and, if they are in molecules, those of their inner shells resemble in shape those of the H atom (s, p, d etc. orbitals). This suggests that sets of wave functions, centered at nuclei and shaped like H-atom orbitals could be efficiently used as basis sets: from such atom-centered basis functions called *atomic orbitals* (AO) a relatively low number needs to be "mixed" to get a molecular wave function. Application of this principle constitutes the fourth approximation used in ab initio methods: the one-electron wave functions (the MOs) are constructed from AOs by linear combination (LCAO-MO method). The variation of the MOs during the SCF iteration is possible by adjusting the linear combination parameters, the AO coefficients. Thus, the solution of the polyelectronic integro-differential Schrödinger equation is converted into a linear algebraic, matrix eigenvalue problem for which well established and efficient numerical mathematical tools have been worked out.

The atomic orbital basis sets used in the expansion of MOs have to fulfill several criteria for making the solution of the Schrödinger equation efficient. There have been a number of basis sets worked out for LCAO-MO calculations and many of these are built into the commercial quantum chemistry codes or are available on the internet (EMSL Basis Set Library webpage 2013, Schuchardt et al. 2007). The AO basis sets consist of hydrogen-type functions with varying parameters. This guarantees that the shape of the AOs is "almost good" at the atomic cores. A good set of atomic core wave functions is transferable from molecule to molecule. What is critical for the correct description of chemical bonds in molecules is the selection of valence orbitals. The application of several functions similar in shape, differing only in how far they extend from the nucleus enables one to describe the electron density farther from atomic nuclei, in regions between atoms (*split-valence basis sets*).

For technical reasons, the AOs generally are built from Gaussian functions of the distance from the atomic nucleus (Gauss-type orbitals, GTOs). This has a historical origin: in the 1950s, analytical formulas (series expansions) for integrals expressing the interaction of electrons were easier to derive than those for the (more appropriately shaped) exponential functions (Slater-type orbitals, STOs). Each hydrogen-type AO consists of a linear combination of GTOs with different width parameter and weight in the sum. Often a group of GTOs are *contracted* so that their relative weights are kept constant and the group constitutes one AO.

Generally, the ensembles of sets of AOs for various atoms of the periodic table are referred to as basis set families. The GTO parameters for a family of AOs in a basis set are optimized to reproduce some target. In the early days of ab initio calculations every researcher developed his own basis set for the molecule to be studied. Later "universal" basis set families were developed.

In the case of the so-called Pople-type basis sets (Hehre et al. 1971, the target is a set of experimental parameters like atomization energies for a group of molecules. A widely used example of Pople basis sets is that denoted as 6-31G(d)

that has been worked out for all atoms of the periodic table. This basis set and its extensions (such as 6-311++G(3df,2p); generally the longer the notation, the better the basis set) work very well for the determination of equilibrium geometries of common organic molecules.

A different principle guides the selection of the optimization target for some other basis sets: they are intended to reproduce the results of very accurate ab initio calculations (mostly performed for atoms) or are derived from those. The correlation-consistent basis sets developed by Dunning and coworkers (Dunning, Jr. 1989) (the basic example being the cc-pVDZ set) and the ANO (atomic natural orbital) basis sets first proposed by Almlöf and Taylor (1987) belong to this class. Each basis set type has a number of extensions, the largest basis sets allowing the estimation of the HF energy in the infinite basis set limit, which is the accurate solution of the HF equations. This way the error caused by the fourth approximation in Table 1 can be recovered. If the number of orbitals in a basis set is K, the SCF procedure generates K MOs whose orbital energies are generally different. If there are N electrons in a molecule, then in the ground (lowest-energy) state of the molecule the *electron configuration* consists of the $N/2$ lowest-energy MOs filled with two opposite-spin electrons (if N is odd, $N=2M+1$, a single electron occupies the $M+1$-th MO). The MOs that are filled by electrons are called occupied MOs; the rest constitute the ensemble of virtual MOs. In the calculation of the total electronic energy of the molecule, only the occupied MOs count.

Even if a *de facto* complete basis set is used, the exact solution of the HF equations is not the exact solution of the molecule's electronic Schrödinger equation, because the effects of electron correlation are not included in it. The exact energy eigenvalue is lower than that of the HF energy, by the amount that is called *correlation energy*. This is the error caused by the third (mean field) approximation in Table 20.1.

The computational (and intellectual) effort to correct the lack of electron correlation in the HF method is much larger than getting the HF energy and wave function. There are numerous methods for estimating the correlation energy, which are able to recover the correlation energy to various levels. The highest-level methods can approach the accurate solution of the full nonrelativistic electronic Schrödinger equation very closely.

The principal fact that constitutes the foundation of these methods is that the set of all electron configurations (the assignment of electrons to different MOs, including now those that are unoccupied in the ground state) constitute a basis set for the given molecule that allows the calculation of the correlation energy. The configurations are generally derived from the ground-state configuration as a reference, by "moving" one, two, three etc. electrons from an occupied to a virtual MO. Knowing the configurations, the molecule's wave function is expressed as the linear combination of configurations and the energy is calculated using the *exact nonrelativistic Hamilton operator*. The best wave function will be the one that provides the lowest energy, which is found by varying the linear combination coefficients. This method is called *configuration interaction* (CI). When *all*

possible configurations arising from a HF calculation are used in the expansion, the method is the full configuration interaction (FCI) method (which is not necessarily the accurate solution, because the basis set for which such a calculation is possible is generally not close to complete). For large basis sets the number of configurations can be huge (billions even for molecules with a few nonhydrogen atoms) and in practice is reduced according to different principles. The simplest principle is that only configurations coming from the HF ground-state configuration by single, double etc. *excitations* (in which at most one, two etc. electrons are simultaneously "moved" from occupied to virtual orbitals) are included in the expansion, constituting the CI-singles (CIS), CI-singles and doubles (CISD) etc. methods. Except for the CIS method, which provides the same energy as the HF method (but is good to estimate the energy of excited states), the CI methods are considered to produce very accurate energies, accounting for as much as 99% of the correlation energy. The price, however, is high: with a reasonably large basis set, CISD calculations can be performed only for molecules containing not more than 10 or 20 non-hydrogen atoms.

Another approach to the calculation of the correlation energy is perturbation theory. The most commonly used version is *Møller-Plessett perturbation theory*, in which the "unperturbed" problem is the HF description of the system under consideration and the perturbation is the difference between the accurate and the HF Hamiltonian. Using perturbation theory actually a number of selected excited configurations are introduced into the wave function. The "price", the computational time and the memory and/or disk requirement increases quickly with the size of the system and the order of perturbation; in routine calculations higher than fourth-order perturbation theory is rarely applied. The goodness of the results rarely converges with the increase of the order of perturbation. The second-order version, MP2 theory accounts for approximately 60 % of the correlation energy. The third-order version, MP3, often yields worse results than MP2. MP4 generally produces very good results, but the computation time is comparable to the CI calculations. Care should be taken when applying MP theory: it is deemed to fail for systems where there are low-lying excited states (because the perturbational correction, which is expected to be small, contains the difference of the excitation energy in the denominator that is low in this case).

The other class of perturbational type theories is the *coupled cluster* (CC) approach, which is currently considered to be the most promising approach, producing benchmark results for other approximate methods. In the CC methods a special principle is used to systematically increase the number of configurations to be included in the wave function. The state-of-the art energy calculations are performed using the CC singles-doubles version supplemented by various approximate treatments of the triple excitations denoted as, for example, CCSD(T). The widely used QCISD(T) (Quadratic CI...) method, originally developed as a CI procedure, proved to be a truncated version of CCSD(T).

The HF method, by its nature, is not able to describe certain situations, where the problem is that the Slater determinant contains various arrangements of

electrons among atoms with equal weight, this way producing a wave function that is physically incorrect. What is missed in such cases is the so-called *nondynamical* or *chemical correlation*. Textbook examples for such are the decomposition of the H_2 molecule (the HF wave function at large H-H distances yields an H^+ and an H^- ion instead of two atoms) and the ozone molecule. The right way of describing the electronic structure of this kind of molecules is inclusion of several configurations from the beginning. Selecting the most important electron configurations like in the valence bond method of early quantum chemistry requires care and is hard to automatize (in fact, the configurations should be selected by hand). There is, however, a version, the *complete active space (multiconfiguration) self-consistent field* (CAS [MC]-SCF) method, for which smart automatic algorithms have been worked out. The basic idea is that one selects the set of occupied and virtual orbitals explored preciously in a HF calculation as the ones between which electrons are allowed to "move" (generating new configurations). This is called the active space. The number of included occupied orbitals will determine the number of electrons (say *m)* that can be moved. In the CAS-SCF version of the MC-SCF method all possible configurations that can be generated within the active space are included in the expansion (which is nothing but a full CI calculation within the active space). Then, the coefficients and the atomic orbitals in the MOs *and* the coefficients in the CI expansion are simultaneously optimized. One often used notation for the method includes the number of orbitals in the active space (*n*) and the number of electrons (*m*) like CAS(*m,n)* or [*n,m*]-CASSCF. The number of configurations rises quickly with the size of the active space and the number of electrons; routine geometry optimizations are hard to do for larger than CAS(14,14) combinations. What is the key, however, is the kind of orbitals in the active space. It is desirable to include the antibonding pair of each included bonding MO (for example, a σ(C-C) MO together with the corresponding σ^*(C-C) type MO). The selection of MOs included in the active space is critical; with the wrong orbitals completely incorrect results can be obtained (they can correspond to some exotic excited state). An additional complication is that the method does not cover the correlation energy involving the electrons outside of the active space. This can be remedied by various perturbational schemes that in routine calculations rarely go beyond the second order. Even less routine are CI calculations starting from MC-SCF wave functions as references (MR-CI). Codes capable of doing such calculations should not be used as black boxes. The use of the CAS-SCF method and its extensions is recommended only for advanced users.

As mentioned in Section 20.2, density functional theory is a conceptually different approach: it is intended to handle electron correlation from the beginning. The difficulty is that the way of doing so is not known and only approximate, more-or-less ad hoc formulas are used. However, these formulas and their parameterization have been worked out with amazing intuition so that the resulting methods work very well, especially for the determination of molecular geometries – at dramatically lower computational cost than ab initio methods

producing the same accuracy. The relative energies are less accurate, especially for potential barriers of reactions. Most DFT methods also fail when the role of dispersion-type interactions are important in determining molecular geometries. There is an intense search for methods to cover dispersive interactions, mostly in empirical ways.

Even with various levels of sophistication, ab initio and first principles methods remain approximate. As was shown above, electronic structure theory is based on some fundamental approximations. The inaccuracies caused by the two lowest-level approximations, the expansion of MOs in AO basis sets and the mean field approximation can be corrected by using a large enough basis set and a high enough treatment of electron correlation. This allows one, in principle, to obtain the accurate solution. In real calculations, however, there are limitations: for large molecules a high-level treatment of electron correlation can not be performed with currently available computers. Calculations using the most accurate methods are out of reach already for moderate-sized molecules containing as few as 10 or 20 non-hydrogen atoms. In order to utilize the available computing power, one needs to select a method to be used. A so-called model chemistry (Hehre et al 1986) is specified by the electronic structure method used (HF, a DFT parameterization or some ab initio level of treating electron correlation) and the basis set used for the MO expansion. The combination of the method and the basis set is also referred to as "level of theory". There is a systematic change in the accuracy of various levels of theory.

The Pople diagram in Figure 20.5 shows the basic principles. The two important properties that control the accuracy of a calculation are the basis set and the level of handling electron correlation. Basis sets of increasing size (or completeness) are placed on the x axis, and the methods of higher and higher level of electron correlation treatment are along the y axis (horizontal in the figure), with the HF method being at the origin. A point on the diagram corresponds to a level of theory. The qualitative change of energy is shown along the depth axis (vertical in the figure). The goal is to get the accurate solution of the nonrelativistic Schrödinger equation, the complete basis limit of the perfect treatment of electron correlation, which is far in the direction of the bottom right corner of the diagram. With the increase of the basis set along the x axis (without any treatment of electron correlation) the energy becomes lower and lower, sooner or later converging to the HF limit. If one uses a relatively low level of electron correlation treatment (e.g. MP2), the tendency is the same but the energies are consistently lower than the corresponding HF energies, etc.

When chemical properties of a series of molecules are needed, the level of theory (the model chemistry) should be selected so that the smallest and largest molecule in the set could be treated with it. The results will not be consistent if for each molecule one uses the highest level of theory allowed by the available computational resources. Somewhat different attitude is more efficient when thermodynamic properties of a set of molecules are needed. In this case the main

issue is not that the qualitative and quantitative features be obtained with comparable accuracy; instead, we need accurate data.
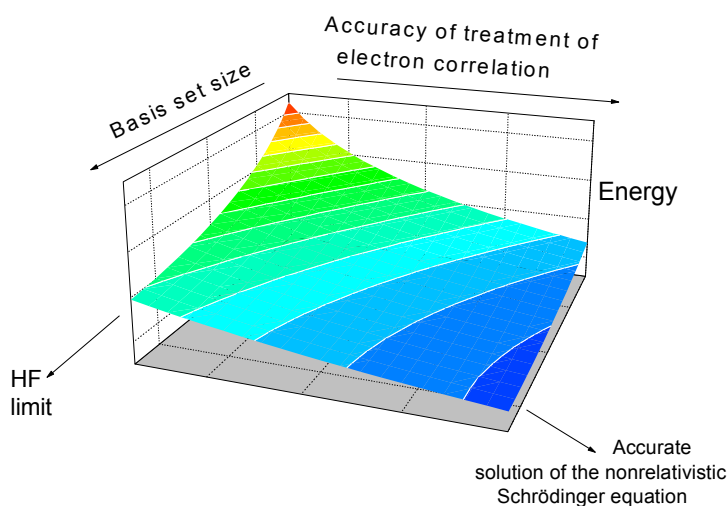

Figure 20.5 The change of the energy calculated for a system by solving the electronic Schrödinger equation with ab initio methods, with the increase of the basis set and of the level of treatment of electron correlation (a modified version of the original Pople diagram)

In such a case it is better to use a high-level theory for the small molecules in the set, and test less accurate methods against experimental data and/or the high-level theory as benchmarks. This enables one to select the method that proves to provide acceptably accurate results for the handling of the larger molecules in the set. Often one can observe some compensation of errors so that a relatively low level of theory produces better agreement with the experiment than a somewhat higher level. Care should be taken in such cases, because it is hard to predict where the errors start not to compensate each other. It is important to note that when thermodynamic cycles are formed, only energies calculated with the same level of theory should be used for each compound in the cycle. If in a set the same molecule appears in different thermodynamic cycles described at different levels of theory, it has to be re-calculated at each level used for the other members in the respective cycle.

Acknowledging that electronic structure methods are approximate, but there are well defined regularities in the details of the calculations, a number of composite electronic structure methods have been worked out. One can generate his/her own

method that is optimized for the selected set of molecules. There are several "empirical" observations underlying the development of composite methods.

The first is that the goodness of the molecular geometry converges much faster than the goodness of energy. When a molecule's geometry is optimized, and one starts from scratch, an efficient strategy is that the first optimization is performed at a relatively low theoretical level. If one optimizes a guessed molecular structure for example successively with the HF, MP2, CCSD and CCSD(T) methods, each optimization starting from the result of the previous lower level, the geometry change in the HF→MP2 step will generally be relatively large, then much smaller in the MP2→CCSD step and generally negligible in the last. In addition, the error introduced by calculating the energy not at the optimum geometry corresponding to the given level of theory but at that from a lower level is relatively small.

The next observation is that the vibrational frequencies needed for statistical thermodynamics converge even faster. In addition, the frequencies with a given method differ systematically from the experimental ones (for example, the HF/6-31G* method in average overestimates the frequencies by a factor of about 1.1). This enables one to obtain good frequency sets by using a relatively low-level theory and employing "empirical" scaling factors.

The third observation is that, once the basis set is large enough, the improvement of the energy when one step from one basis set to a higher one is approximately the same with a lower and a higher level of electron correlation treatment.

Applying such observations one can simulate a high-level calculation from the results of lower-level calculations. A classical composite or multilevel scheme is the Gaussian-2 method of Pople and coworkers (Curtiss et al. 1991). In this method, the geometry is first optimized with the HF/6-31G(d) method, and vibrational frequencies are also calculated at this level. Then the geometry is re-optimized with the MP2/6-31G(d) method, and further calculations are performed at this level. The basis set is smaller and smaller as one goes from MP2 to MP4 and QCISD(T). (In the Pople diagram this corresponds to starting at a point relatively close to the origin on the correlation treatment axis but far from the origin along the basis set axis, at a large basis set, and stepping simultaneously one to the right and one toward smaller basis sets, always stretching the available computational power to approximately the same level.) Then the additive basis-set corrections are stepwise collected and added to the QCISD(T)/6-311G(d,p) energy and with an empirical correction that depends on the number of electrons, a simulated QCISD(T)/6-311++G(3df,2p) energy can be obtained. Thermodynamic properties (originally, atomization energies) are calculated from these energies and the scaled HF/6-31G* frequencies of the compounds in the thermodynamic cycle.

Along these lines many composite methods have been developed. DFT has also been included because it produces MP2 quality geometries at essentially HF price. It is important to note, however, that these methods are optimized for a set of molecules in a statistical sense, and for different members of the set the accuracy can be different. For example, halogen-containing molecules often form an island

of large experiment-theory deviations. The more similar the molecules in a set to be studied are, the larger is the possibility to work out a project-oriented composite method. The condition for this is that the set should contain several members for which reference data are available for testing.

For the interested reader, here we list some textbooks of different levels that can be used to understand the physical, mathematical and computational details used on electronic structure methods.

1. Introduction to Computational Chemistry (Jensen 2006): a comprehensive textbook on up-to-date methodology. It provides basic information not only on electronic structure methods, but also on statistical mechanics. Applications are also listed.

2. Molecular Electronic-Structure Theory (Helgaker et al. 2000, 2013): another excellent comprehensive advanced textbook covering the foundations and mathematical details of electronic structure theory, including computational aspects

3. Quantum Chemistry: Fundamentals to Applications (Veszprémi and Fehér 1999): a readable and didactic introduction to the basics of electronic structure theory; numerous applications are presented that help the reader to get hands-on experience.

4. Ab Initio Molecular Orbital Theory (Hehre et al. 1986): a textbook written for a chemist user, covering the foundations and a lot of applications at the level of the 1980s. The philosophy of the Pople school (work out methods that can provide accurate results approaching experiments, using empirical corrections if needed) can be well understood.

5. Exploring Chemistry with Electronic Structure Methods (Foresman and Frisch 1996): intended to be a "manual" to the Gaussian suite of programs, this book contains a lot of applications that help the reader to learn how to do certain calculations. There are a number of practical tricks and caveats where a slightly incorrect calculation provides vastly incorrect results.

6. Approximate Molecular Orbital Theory (Pople and Beveridge 1970): written at the time when ab initio quantum chemistry was not routinely used, the book is an excellent presentation of the foundations of theory and the philosophy of introducing approximations and the use of empirical parameters.

7. Quantenchemie. Ein Lehrgang (Zülicke 1973, Zülicke 1985): another excellent textbook, written in German and also available in Russian, covering the quantum mechanical foundations and the mathematical details of quantum chemical methods.

8. The Molecular Orbital Method and the Reactivity of Organic Molecules (Basilevsky 1969; in Russian): the first part of this book provides an excellent brief introduction to Hartree-Fock theory. Although methods for accurate thermodynamic calculations are not included, the second part of the book covers a comprehensive overview and introduction to the application the Hückel method to chemical problems that is instructive for a beginner.

9. Theorems, proofs, and derivations in quantum chemistry (Mayer 2003): a treatise written for advanced users, focusing on the physical and mathematical foundations and details many of which are just touched in regular textbooks. The book is self-contained: proofs of all theorems are provided, satisfies even the "purist's" expectations.

10. Quantum Theory of Molecules and Solids (Slater 1963, 4 volumes): a comprehensive set of monographs covering all details of earlier quantum chemical methods.

11. Essentials of Computational Chemistry, Theories and Models (Cramer 2004): a textbook providing details extending from molecular mechanics, molecular orbital theory and density functional theory to methods to calculate thermodynamic properties and some of its technical caveats, as well as spectroscopic properties of gaseous species. Additionally it provides information about implicit and explicit models for calculation in condensed phases. Corrections for errors found in the book are reported on the internet (Cramer 2013).

12. Quantum Chemistry (Levine 2007): this excellent textbook provides a solid basis for understanding physical and mathematical aspects of quantum mechanics and molecular electronic structure theory with clear explanation and hints about pitfalls. Examples with calculation results are used to explain the methods and limits of their applicability.

20.4 Additivity schemes

A basic principle of chemistry is that the properties of functional groups are "almost" transferable from one compound to the other. While the transferability of chemical properties is hard to quantify, investigation of interrelationship of measured physical chemical properties of related compounds like heats of formation, yielded quantitative rules that allow the calculation of properties of unknown compounds from data on related ones. Transferability of properties means that the thermodynamic quantities of a compound can be obtained by summing the contributions from its constituent subunits. The crudest approximation, namely, that properties of atoms are transferable (which is not bad, for example, for paramagnetic susceptibilities) does not work in thermodynamics. At the next level, bonds, additivity of thermochemical properties works better (Pitzer 1940, Platt 1947, Janz 1955, Greenshields and Rossini 1958, Janz 1958, Somayajulu and Zwolinski 1966), especially for well-defined classes of compounds, for example, hydrocarbons, but it is not accurate when used for a wide range of substances. It is the next level of complexity, that of properly selected groups of atoms (related, but not necessarily identical to functional groups in chemistry), for which thermochemical properties prove to be transferable. The scheme proposed by Benson and Buss (1958) became the most

widely applicable (see also (Laidler 1956, Allen 1959) and (Cox and Pilcher 1970) about the equivalence of the three approaches). This method is widely referred to as "group additivity" scheme (maybe "group contribution additivity" would express more the idea beyond it).

In the terminology of Benson (Benson 1968), a group is "a polyvalent atom (ligancy$\geq$2) in a molecule together with all of its ligands", denoted as X-$(A)_i(B)_j(C)_k(D)_l$, where X is the central atom attached to i atoms of the sort A, j atoms of sort B, and so on. The prototype groups in hydrocarbons are $C-(C)(H)_3$, $C-(C)_2(H)_2$, $C-(C)_3(H)$, and $C-(C)_4$, also denoted as P, S, T, Q, respectively, for primary, secondary etc. The heat of formation (or entropy, or heat capacity at a given temperature) of a molecule will be the sum of the contributions of these groups. For example, the heat of formation group values (GV), also referred to as group additivity value (GAV), for the $C-(C)(H)_3$ and $C-(C)_2(H)_2$ groups are -41.8 kJ mol^{-1} and -20.9 kJ mol^{-1}, respectively (Cohen and Benson 1993). The estimated heat of formation of ethane (two P groups) then is

$$2 \times GV(C-(C)(H)_3) = 2 \times (-41.8) = -83.6 \text{ kJ mol}^{-1}$$

and propane (two P and a S group) one gets

$$2 \times GV(C-(C)(H)_3) + GV(C-(C)_2(H)_2) = 2 \times (-41.8) - 20.9 = 104.5 \text{ kJ mol}^{-1}.$$

Better estimates can be obtained if one takes into account the next neighbor atoms also. Such units are called components. The evaluation of component contributions requires a large, very accurate experimental database, and due to the lack of such, only a few component evaluations are available. As a result, component additivity values are not routinely used in additive estimates. Instead, the type of atoms are distinguished, for example for carbon the hybrids sp^3 (C), sp^2 (C_d), sp (C_t), aromatic sp^2 (C_b), and fused ring aromatic (C_{bf}) and allenic, =C= (C_a) (Figure 20.6). In addition, ring strain and nonbonded interactions are also taken into account, such as the repulsive 1,4 or *gauche* interaction of two methyl groups, and the *cis* and *ortho* interactions. Attempts were made to evaluate other non-covalent interactions, for example hydrogen-bonding interactions between two OH groups in vicinal diols and related compounds (Cohen 1996). However, not many went into common practice, partly due to the lack of accurate experimental data, and partly because of the intrinsic non-additivity of the contributions of such extended
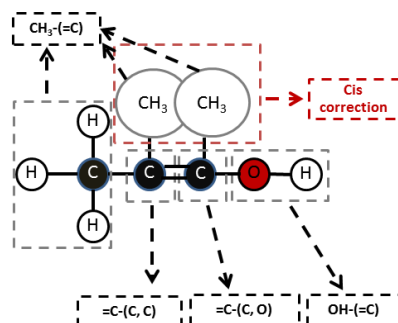
Figure 20.6 Decomposition of an unsaturated alcohol with a branched hydrocarbon chain into Benson-type groups

groups. The same holds for the group values of heteroatoms such as sulfur, phosphorus, halogens or silicon and corrections for their non-covalent interactions. It is important that care should be paid to the proper consideration of symmetry numbers when entropy contributions are calculated. GVs and interaction values are listed in several papers and books (Benson 1968, Cohen and Benson 1993, Poling et al. 2001).

The procedure of estimating the thermochemical properties of a compound is then:

1. write down the structural formula of the compound;

2. analyze the structure, find the groups constituting it and count the occurrences of each type (an example is shown in Figure 20.6);

3. take into account the possible nonbonding interactions and count them by type;

4. look up the group values in a compilation, multiply them by the number the group or interaction occurs in the compound, and sum them all.

This is an algorithm that is easy to program and several computer codes have been developed for the purpose (Stein et al. 1991, Ritter 1991, Ritter and Bozzelli 1991, Muller et al. 1995, CHETAH 1998, Blurock et al. 2012, RMG webpage 2012) (for the parameters used in CHETAH see Poling et al. (2001)). Some useful remarks on the applicability of the first three codes can be found in (Burcat 2009).

An interesting development is the evaluation of GVs for groups corresponding to transition structures (Cohen 1982, Sumathi et al. 2001a, Sumathi et al. 2001b, Sumathi et al. 2002) that can be used in the approximate calculation of reaction rate coefficients.

It should be noted that the accuracy of GVs depends not only on how accurately the additivity assumption is fulfilled but also on the reliability of the experimental database. Inconsistencies in the database yield ambiguous GVs. The most reasonable way of reducing the influence of data inconsistency is the use of active tables such as the Active Thermochemical Tables (see Section 20.5) developed at Argonne National Laboratory (Ruscic et al. 2005, Goos et al. 2013).

When this approach is used, it is especially obvious that every time the thermochemical database or the set of groups or interactions is extended, the whole set of GVs has to be re-optimized, instead of keeping the previous values fixed and optimizing only the newly added ones.

A great help is provided by ab initio calculations of thermochemical properties. Theory can provide "replacements" for data unavailable experimentally, that can be used instead of measured ones. When possible, it is desirable to use very accurate high-quality ab initio data. When large molecules such as polyaromatic hydrocarbons are considered, however, the currently available strategy is that one selects the appropriate model chemistry. If there is a wide enough range of compounds calculated using electronic structure methods, and yet ambiguities arise in assigning GVs, one should also consider not only the possibility that the database is inconsistent but also that additivity may not be perfect.

Benson type GVs, for standard heats of formation, entropy and heat capacity contributions have been tabulated for groups from which one can build alkanes, alkenes, alkynes, aromatics, cycloalkanes and cycloalkenes, polycyclic aromatic hydrocarbons, as well as oxygenated, nitrogenous, halogenated, organosulfur, organophosphorus, organoboron, and organometallic compounds, and for free radicals. There are many papers in the literature listing GVs derived later with added values for groups not included in the earlier compilations. When starting a calculation using the additivity principle, it is important that data from a selected compilation not be mixed with those from other tables because of the danger of reducing the accuracy of the estimation. A caveat: one should watch for errors within the same table (several inconsistent data have been detected).

Three remarks on the use of additivity schemes:

1. The additivity of properties beyond the atomic masses is actually not guaranteed by any law of Nature and is generally not strictly fulfilled. Bader and coworkers (Wiberg et al 1987, Bader and Bayles 2000) using the topological definition of an atom in a molecule based on the analysis of the electron density (Bader 1990), showed that transferability of groups is generally only apparent because changes of properties of one group are often compensated by the opposite changes of the neighboring group.

2. Use of additivity methods to estimate thermodynamic properties seems to be obsolete in light of obtainable ab initio data of equal or better quality. However, there are cases (for example, mechanism generation in combustion chemistry) when data on such a large number of compounds is needed that ab initio calculations would be too slow to get the desired data, especially for larger molecules, so that the speed of estimation offered by additivity-based techniques is a must.

3. The common philosophy used in deriving group values is that one starts with the simplest compounds containing the selected group, and expands the GV evaluation stepwise. Experimental or ab initio data for large molecules are rarely used when GVs are derived. As a result, "exporting" the GVs from small molecules to large ones does not necessarily yield accurate results. Application of

the active table paradigm can help to get consistent GVs when extending the data set to include larger molecules, but that requires some solid data for the latter. Unfortunately, such numbers are not easily obtainable, because the accuracy of both the experimental and theoretical methods is reduced when the size of molecule increases.

20.5 Databases

Thermochemical data has been measured since in the second half of the 19th century. Extensive work performed by Thomsen and by Berthelot yielded numerous enthalpies of reaction (especially of combustion). Probably the first two databases on thermochemical data are their two monographs (Thomsen 1886, Berthelot 1879) in which the authors summarized their experimental data. The data were organized based on the principles and theorems they worked out. Notably, many of the data they listed are so accurate that their difference from the currently accepted values is less than a percent. Since then, numerous compilations have appeared that contain data on a few thousand compounds. A non-exhaustive overview of them up to 1988 was given by Gurvich (1988). Looking up data for a specific, not very trivial compound is quite cumbersome if the books are not to hand. In the following, we list a few of these printed compilations and present a few electronic databases. It seems to be important to emphasize that internet search for thermodynamic data, i.e. "the quick and easy method," is not recommended. According to our experience, data obtained from such sources should be treated with caution as their quality and reliability is often dubious and difficult to verify due to inadequate referencing. .

Thermodynamic data is available in various formats. Virtually, there has not yet been a widely accepted standard on how to archive thermodynamic data. In early sources, enthalpies of formation, heat capacities and entropies were listed at a single or a few temperatures (mostly at 298.15 K). To represent the temperature dependence of thermodynamic data, polynomials are generally used. Widely used are the NASA polynomials (the earlier (McBride and Gordon 1967, Gordon and McBride 1971) 7- and the more recent (McBride et al. 2002) 9-parameter version). The basic formula determines the heat capacity C^0_p as a fourth-order polynomial of the absolute temperature in the 7-parameter version, enhanced by a first- and second-order term of the inverse temperature in the nine-parameter version. Also popular are the Shomate polynomials (Shomate 1954), representing C^0_p as a third-order polynomial in T plus a T^{-2} term. The polynomials allow the interpolation between measured or calculated data points. The validity of the polynomials is restricted to certain temperature ranges. In some cases, the polynomials produce a discontinuity when switching from one temperature-range to another; it is advisable to check whether the data extracted from a database suffer from this error. Nowadays the data are stored in machine-readable format. One standard is

the NASA format (Gordon and McBride (1971) also adopted by the widely used CHEMKIN simulation code (Kee et al. 1996). More recently, two IUPAC projects were devoted to the development of a standard for storage and communication of thermodynamic data utilizing the Extensible Markup Language (XML) (ThermoML IUPAC Standard). This standard is recommended by many peer-reviewed journals, the data being made publicly available through the ThermoML Archive (ThermoML Web Archive 2013) (note that there is no reference to this database on the homepages of the journals).

The most reliable thermodynamic properties are those established by the Committee on Data for Science and Technology (CODATA) (Cox, Wagman and Medvedev 1989) for some key chemical substances. Use of these recommended, internally consistent values is encouraged in the analysis of thermodynamic measurements, data reduction, and preparation of other thermodynamic tables. As an extension, an IUPAC committee critically evaluated thermochemical properties of selected radicals (CH, CH_2(triplet), CH_2(singlet), CH_3, CH_2OH, CH_3O, CH_3CO, C_2H_5O, $C_6H_5CH_2$, OH, and NH_2) (Ruscic et al. 2005; extension is in progress).

Evaluated thermochemical data of thousands of substances can be found in different volumes of the Journal of Physical and Chemical Reference Data. Older reprints or monographs are available for free download (JPCRD and monographs webpage). Thermodynamic properties of a large number of hydrocarbons can be found in the book of Stull, Westrum and Sinke (Stull et al. 1969) and in the compilations published by the US National Bureau of Standards (Rossini et al. 1952, 1953, Domalski 1972).

Older, but excellent, thermochemical data tabulated for organic and inorganic substances and ions, are available in Gurvich's comprehensive thermochemical compendium (Gurvich et al. 1989, 1991, 1994, 1997). Note that the data for reactive species like radicals and biradicals in the Gurvich tables are outdated.

The compilation by Pedley (Pedley 1994, Pedley et al. 1986) contains critically evaluated enthalpies of formation for more than 3000 organic compounds. Evaluated data on organic compounds are also listed in another publication of the Thermodynamic Research Center (Marsh et al. 1988). About ten years ago TRC joined NIST, and the database (including much more than thermochemical data) is commercially available as the NIST-TRC Web Thermo Tables (WTT) (NIST TRC Web Thermo Tables). Another commercial database with thermochemical data is from Design Institute for Physical Properties (DIPPR 801 Database 2013).

The NIST-JANAF Thermochemical Tables (Chase et al. 1985, Chase 1998) have been great sources of data and papers on both inorganic and organic molecules for a long time. It is wise to read the introduction of these volumes where it is explained how the data were evaluated, and what the uncertainties or problems with the different techniques are. (Note that the older volumes contain numerous errors). Newer issues (e.g. Dorofeeva et al. 2001) include well documented quantum chemical data, too. The JANAF database also is available online (NIST Standard Reference Database 13).

The NIST Chemistry Webbook (NIST Chemistry Webbook) contains thermochemical data for more than 7000 organic and small inorganic compounds, virtually independently from the JANAF tables. The listed data are obtained from the literature and are generally not subject to critical evaluation. It is advisable to check the cited literature source and check whether the values were taken correctly from the reference and whether corrections or updates are available (a relatively productive way is to look up papers that cite the original publication).

The NASA Glenn thermodynamic database (McBride et al. 2002) covers about 2000 species. Carefully evaluated temperature-dependent data are provided in the form of 9-term NASA polynomials. They can be accessed online using the ThermoBuild feature of the NASA Glenn chemical equilibrium computer program (NASA Thermo Build webpage). The sources of data are referenced. Although it was intended to be continuously updated, the database seems to not contain data more recent than 2002.

Thermodynamic data on ions is listed in the JANAF (e.g. Chase 1998, Chase et al. 1985) and the Gurvich tables (Gurvich 1988, Gurvich et al. 1988, 1989, 1991, 1994, 1997) and the NASA Glenn database (NASA Thermo Build webpage 2012). A wealth of data on ions is listed in the more than 800-page long compendium of Lias et al. (Lias et al. 1988). Concerning thermodynamics of ions, there are two conventions, depending on how a gas-phase electron is handled. The *thermal electron convention* (TEC), mostly used by thermodynamicists, defines the electron as a *standard chemical element* and treats its thermochemistry accordingly. In other words, at 298 K it has its own enthalpy, namely, the integrated heat capacity corresponding to an ideal Boltzmann gas, i.e. 6.197 kJ/mol. The *ion convention* (IC) is mostly used by the mass spectrometry community and defines the electron as a *sub-atomic particle*, having no thermodynamic properties. The heat of formation values obtained by the different conventions can be converted to each other: for positive ions (as well as for the formation of an electron) the IC heat of formation is 6.197 kJ/mol less than in the TEC scheme, while for negative ions the opposite holds. The ion convention is used for example in (McBride et al. 2002, NASA Thermo Build webpage) and (Lias et al. 1988); the thermal electron convention in the JANAF tables (e.g. Chase et al. 1985, Chase 1998 and Gurvich´s work (Gurvich 1988, Gurvich et al. 1989-1997).

One database containing systematically measured and calculated thermochemical data is the NIST Computational Chemistry Comparison and Benchmark Database (NIST CCCBD 2013). It lists experimental and calculated data for species with well-established enthalpies of formation. Covered are mostly species containing C, H, O, and N with less than 6 non-hydrogen atoms, with more details than in the NIST Chemistry Webbook. Valuable are the comparisons of data from experiments and from quantum chemical calculations obtained with numerous semi-empirical (AM1, PM3), density-functional and composite ab initio methods (G1, G2MP2, G2, G3, G3B3, CBS-Q).

The other extensive database that also focuses on quantum chemical data is the Extended Third Millennium Thermochemical Database (Goos, Burcat and Ruscic 2013). It is a collection of evaluated thermochemical data of gaseous compounds, some liquids and solid, and contains data for pure elements, metals, inorganic and organic compounds and reactive species like radicals and ions. In addition, it includes all inert gases and a limited number of compounds of 4th to 6th row elements such as Br, I, Cr, Mn, Fe, Co, Ni, Cu, Hg, Mo, Pb, Pd, Pt, Sb, Sn, W, Zn, Zr. For easy use in old and modern kinetic modeling and computational fluid dynamics software, temperature-dependent thermochemical data are provided in 7- or 9-term NASA polynomial format. The thermodynamic properties of about one half of the included species are also calculated with the G3B3 method (Baboul et al. 1999) (making this database the largest collection of data obtained with this method). G3B3 is a composite quantum chemical method producing molecular and thermochemical properties that compare very well with experimental results, the accuracy (with a 95% confidence limit) being around ± 10 kJ/mol or better. The database is critically evaluated by the authors. They provide not only referenced literature values (calculated and measured ones) but also results of their own calculations. In the database the accuracy of the data is shown in detail, together with data used to calculate the partition functions. The deviations of the fitted temperature dependence of thermochemical properties in terms of NASA polynomials from the "accurate" data are also provided. The database is updated and enhanced on a regular basis, as well as on user requests.

This database contains more consistent thermochemical properties than other collections, because (for a part of the included species) a novel paradigm, the active table approach is used to check the congruency of data on different species. The active table approach (Ruscic et al. 2004, Ruscic et al. 2005, Active Thermochemical Table Webpage 2013) addresses fitting a whole set of data simultaneously, in contrast to the usual "sequential" approach. In the latter, data for a few basic compounds are nailed down, and these serve as fixed values when new and new species are added. The danger here is that one can get different values for the same property when the calculation is based on different pathways. The active table approach, instead, considers the data for the whole set of species (a *thermochemical network*) simultaneously. The network (essentially a set of interlocking Haber-Born cycles) contains redundant information that makes it easy to pick data that are not consistent with the rest. Statistical analysis including simultaneous error propagation then produces thermochemical data usually with smaller error bars as compared to that derived for the same quantity from a set of measurements aimed at the determination of only the property in question. More importantly, the dataset obtained will be consistent, which is essential in, for example, combustion modeling when a multitude of compounds are involved in complex chemical mechanisms whose enthalpy production/consumption influences the kinetics itself. The approach has already provided valuable information by reducing the error bars on some key compounds (Ruscic et al. 2004). Even more impressive is that the approach helped to correct heats of

formation of some radicals that were thought to be well established, such as OH (Ruscic et al. 2002), HO_2 (Ruscic et al. 2006), C_6H_5 (Stevens, Ruscic and Baer 2010) or NCN (Goos, Sickfeld et al. 2013). Future application of the active table approach will certainly have a major role in making our knowledge on thermochemical data more reliable and consistent.

20.6 Summary

Knowledge of accurate thermochemical data is essential to the modeling and design of reacting systems. Measurement of enthalpies of formation, the key quantity, is limited, not only because of experimental restrictions, but also because of the extraordinary number of compounds involved, for example, in combustion systems. The role of non-experimental methods is essential in the collection of data. Electronic structure methods (quantum chemistry) offer various levels of precision, more accurate calculations being more expensive (in terms of computation time, memory and disk-space requirements). High-level calculations can match the accuracy of measurements but are limited to relatively small molecules. Larger molecules can be handled by relatively less expensive but at the same time less accurate methods. When using such levels of theory, calibration to experimental data on compounds related to those of interest is desirable. The multitude of species for which thermochemical data are needed is hard to handle even with such lower-level methods. The observation that thermochemical properties of functional groups and other molecular constituents are transferable from molecule to molecule allows inexpensive estimation of thermodynamic data. Various group additivity schemes have been developed with group values for different groups, often depending on the environment. The very high speed of these calculations allows quick derivation of thermochemical data, but for relatively large molecules, calibration against experimental or more easily available quantum chemical data is desirable.

The data available in the literature have been collected in databases, the more valuable being those involving expert evaluation. Especially useful is the active table approach that involves simultaneous fitting of data on several species using a thermochemical network connecting all species of interest including the relationships among their thermochemical data. Evaluated data are available for a few thousand species. For molecules not included in the tables, group additivity schemes can be used. If very high accuracy is needed or the goodness of the group additivity data is dubious, and the experimental determination of the thermochemical parameters is not possible, application of electronic structure methods can be a solution.

References

Active Thermochemical Table Webpage http://atct.anl.gov/ (accessed 14 January 2013)

Allen TL (1959) Bond Energies and the Interactions between Next‑Nearest Neighbors. I. Saturated Hydrocarbons, Diamond, Sulfanes, S8, and Organic Sulfur Compounds. J Chem Phys 31:1039-1049

Almlöf J, Taylor PR (1987) General contraction of Gaussian basis sets I Atomic natural orbitals for first‑ and second‑row atoms. J Chem Phys 84:4070-4078

Atkins P, De Paula J (2005) Elements of Physical Chemistry, Oxford University Press

Ayala PA, Schlegel HB (1998) Identification and treatment of internal rotation in normal mode vibrational analysis. J Chem Phys 108:2314-2325.

Baboul AG, Curtiss LA, Redfern PC, et al (1999) Gaussian-3 theory using density functional geometries and zero-point energies, J Chem Phys 110:7650-7657

Bader RFW (1990) Atoms in Molecules - A quantum theory. Oxford University Press, New York

Bader RFW, Bayles D (2000) Properties of Atoms in Molecules: Group Additivity. J Phys Chem A 104: 5579-5589.

Basilevsky MV (1969) The Molecular Orbital Method and the Reactivity of Organic Molecules Khimia Publishers, Moscow (in Russian)

Benson SW (1968) Thermochemical kinetics. John Wiley and Sons, New York.

Benson SW, Buss JH (1958) Additivity Rules for the Estimation of Molecular Properties. Thermodynamic Properties. J Chem Phys 29: 546-572.

Berthelot M (1879) Essai de mécanique chimique fondée sur la thermochimie, Dunod, Paris

Blurock ES, Warth V, Grandmougin X, et al F (2012) JTHERGAS: Thermodynamic estimation from 2D graphical representations of molecules, Energy 43:161-171

Burcat A (2009) Ab Initio Calculations of Carbon-Containing Species and Comparison with Group Additivity Results. Part II. C4 Species. J Chem Eng Data 54:1829–1835

Chase MW Jr. (Editor) (1998) NIST-JANAF Thermochemical Tables, 4th ed. J Phys Chem Ref Data, Monograph 9, 1-1951

Chase MW, Davies CA, Downey JR Jr., et al (1985) JANAF Thermochemical Tables, 3rd ed.; J Phys Chem Ref Data 14, Supplement 1

CHETAH Version 7.2: The ASTM Computer Program for Chemical Thermodynamic and Energy Release Evaluation (NIST Special Database 16), 4th ed (1998) URL http: //www.astm.org.

Cohen N (1982) The use of transition-state theory to extrapolate rate coefficients for reactions of oh with alkanes. Int J Chem Kinet 14:1339–1362

Cohen N (1991) Are reaction rate coefficients additive? Revised transition state theory calculations for OH + alkane reactions. Int J Chem Kin 23:397–417

Cohen N (1996) Revised Group Additivity Values for Enthalpies of Formation (at 298 K) of Carbon-Hydrogen and Carbon-Hydrogen-Oxygen Compounds. J Phys Chem Ref Data 25:1411-1481

Cohen N, Benson SW (1993) Estimation of heats of formation of organic-compounds by additivity methods. Chem Rev 93: 2419-2438

Cox JD; Pilcher G (1970) Thermochemistry of Organic and Organometallic Compounds. Academic Press, New York (Chapter 7)

Cox JD, Wagman DD, Medvedev VA, CODATA Key Values for Thermodynamics, Hemisphere Publishing Corporation, New York (1989); available at http://www.codata.org/resources/databases/key1.html (accessed Dec. 27, 2012)

Cramer CH J (2004) Essentials of Computational Chemistry, Theories and Models, 2nd Edition, Wiley.

Cramer CH (2013) Essentials of Computational Chemistry, Theories and Models, Errata on http://pollux.chem.umn.edu/Book.html (accessed 3 Jan 2013)

Curtiss LA, Raghavachari K, Trucks GW, et al (1991) Gaussian2 theory for molecular energies of first and second row compounds. J Chem Phys 94:7221-7230.

DIPPR 801 Database (2013) http://www.aiche.org/dippr/events-products/801-database

Domalski ES (1972) Selected Values of Heats of Combustion and Heats of Formation of Organic Compounds Containing the Elements C,H,N,O,P and S. J Phys Chem Ref Data 1:221-277

Dorofeeva OV, Novikov VP, Neumann DB (2001) NIST-JANAF Thermochemical Tables. I. Ten Organic Molecules Related to Atmospheric Chemistry. J Phys Chem Ref Data 30:475-514

Dunning TH, Jr (1989) Gaussian basis sets for use in correlated molecular calculations I The atoms boron through neon and hydrogen. J Chem Phys 90:1007-1023

Ellingson BA, Lynch VA, Mielke SL, et al (2006) Statistical thermodynamics of bond torsional modes: Tests of separable, almost-separable, and improved Pitzer–Gwinn approximations. J Chem Phys 125: 084305.

EMSL Basis Set Exchange Library https://bse.pnl.gov/bse/portal (Accessed 8 January 2013), Schuchardt KL, Didier BT, Elsethagen T, Sun L, Gurumoorthi V, Chase J, Li J and Windus TL (2007) Basis Set Exchange: A Community Database for Computational Sciences. J Chem Inf Model 47:1045-1052

Foresman JB, Frisch AE (1996) Exploring Chemistry with Electronic Structure Methods. 2nd edn. Gaussian Inc.

Goos E, Burcat A, Ruscic B (2013a) Ideal Gas Thermochemical Database with Updates from Active Thermochemical Tables. Available via ftp://ftp.technion.ac.il/pub/supported/aetdd/thermodynamics mirrored at http://garfield.chem.elte.hu/burcat/burcat.html and also available from http://www.dlr.de/vt/en/.
Last print version (outdated by now): Burcat A, Ruscic B (2005) Third millennium ideal gas and condensed phase thermochemical database for combustion with updates from active chemical tables, Technical Report TAE 960 and ANL-05/20, Technion-IIT, Haifa, and Argonne National Laboratory, Argonne, Illinois

Goos E, Sickfeld Ch, Mauß F, et al (2013b) Prompt NO formation in flames: The influence of NCN thermochemistry. Proc Comb Inst 34:657-666

Gordon S, McBride BJ (1971) Computer Program for Calculation of Complex Chemical Equilibrium Compositions, Rocket Performance, Incident and Reflected Shocks and Chapman-Jouguet Detonations, NASA Report SP-273

Greenshields JB; Rossini FD (1958) Molecular Structure and Properties of Hydrocarbons and Related Compounds. J Phys Chem 62: 271-280

Gurvich L (1988) Reference books and data banks on the thermodynamic properties of individual substances. Pure & Appl Chem 61:1027-1031

Gurvich LV, Veyts IV, Alcock AB, Thermodynamic Properties of Individual Substances, 4th Ed.,Vol. 1, Hemisphere Publishing Co. (1989) Vol. 2 (1991), Vol. 3 (1994), Vols. 4–6 (1997)

Hehre WJ, Ditchfield R, Pople JA (1971) Self-Consistent Molecular orbital Methods XII Further Extensions of Gaussian-Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules. J Chem Phys 56:2257-2261

Hehre WJ, Radom L, Schleyer PvR, and Pople JA (1986) Ab Initio Molecular Orbital Theory. John Wiley & Sons, New York

Helgaker T, Jørgensen P, Olsen J (2000) Molecular Electronic-Structure Theory. Wiley, 1st Ed, (2013) Paperback reprint of 1st edition, ,2nd revised Edition Paperback, expected May 2013.

Janz GJ (1955) The estimation of thermodynamic properties for organic compounds and chemical reactions. Quart Rev Chem Soc 9: 229-254. DOI: 10.1039/QR9550900229

Janz, GJ (1958) Estimation of Thermodynamic Properties of Organic Compounds; Academic Press: New York

Jensen F (2006) Introduction to Computational Chemistry. 2nd edn, Wiley

JPCRD and monographs reprints are available for free from http://www.nist.gov/srd/reprints.cfm or http://www.nist.gov/srd/monogr.cfm

Kee RJ, Rupley FM, Meeks E, et al (1996) CHEMKIN-III: A FORTRAN Chemical Kinetics Package for the Analysis of Gasphase Chemical and Plasma Kinetics, Sandia Report UC-405, SAND96-8216

Laidler KJ (1956) A System of molecular thermochemistry for organic gases and liquids. Can J Chem 34:626-648

Lias SG, Bartmess JE, Liebman JF, et al (1988) Gas-Phase Ion and Neutral Thermochemistry. J Phys Chem Ref Data 17, Supplement No. 1.

Levine IN (2007) Quantum Chemistry. 6th International edition. Alpha Books, Pearson

Marsh KN, Das A, Frenkel M, et al (1988) TRC Thermodynamic Tables, Non-Hydrocarbons, Vols. I-VIII and Hydrocarbons, Vols. I-XII, Thermodynamics Research Center Texas A&M University, College Station, Texas; TRC (1997) Selected Values of Properties of Chemical Compounds, Thermodynamics Research Center, Texas A&M University, College Station, Texas

Mayer I (2003) Simple theorems, proofs, and derivations in quantum chemistry. Kluwer Academic/Plenum Publishers, New York

Marsh KN (2001) Calorimetry, in: Moore JH, Spencer ND (2001) Encyclopedia of Chemical Physics and Physical Chemistry, Volume 2, Institute of Physics Publishing, Bristol and Philadelphia

McBride BJ, Gordon, S (1967) FORTRAN IV Program for Calculation of Thermodynamic Data, NASA TN-D 4097

McBride BJ, Zehe MJ, and Gordon S (2002) NASA Glenn Coefficients for Calculating Thermodynamic Properties of Individual Species. NASA report TP-2002-211556

Muller C, Michel V, Scacchi G, et al (1995) THERGAS – A computer program for the evaluation of thermochemical data of molecules and free-radicals in the gas phase. Journal de Chimie Physique et de Physico-Chimie Biologiques 92:1154-78

NASA Thermo Build http://www.grc.nasa.gov/WWW/CEAWeb/ceaThermoBuild.htm (accessed 29 Dec 2012)

NIST Chemistry WebBook, NIST Standard Reference Database Number 69, http://webbook.nist.gov/chemistry (Accessed 30 Dec 2012)

NIST CCCBD, Computational Chemistry Comparison and Benchmark Database, NIST Standard Reference Database Number 101, Release 15b, August 2011, Editor: Russell D. Johnson III, http://cccbdb.nist.gov/

NIST Standard Reference Database 13: NIST-JANAF Thermochemical Tables, http://kinetics.nist.gov/janaf/ (accessed 28 Dec 2012)

NIST-TRC Web Thermo Tables (WTT) http://trc.nist.gov/tde.html (accessed 30 Dec 2012)

Pedley JB (1994) Thermochemical data and structures of organic compounds. Vol. 1, (TRC Data Series) Thermodynamics Research Center, College Station, TX, USA

Pedley JB, Naylor RD, Rylance SP (1986) Thermochemical Data of Organic Compounds, 2nd Edn. Chapman and Hall, London

Piszczatowski K; Lach G; Przybytek M, et al (2009) Theoretical Determination of the Dissociation Energy of Molecular Hydrogen. J Chem Theor Comp 11:3039-3048

Pitzer KS (1940) The Vibration Frequencies and Thermodynamic Functions of Long Chain Hydrocarbons J Chem Phys 8: 711-720

Pitzer KS, Gwinn WD (1942) Energy Levels and Thermodynamic Functions for Molecules with Internal Rotation: I Rigid Frame with Attached Tops. J Chem Phys 10:428-440

Platt JR (1947) Influence of Neighbor Bonds on Additive Bond Properties in Paraffins J Chem Phys15:419-420

Platt JR (1952) Prediction of Isomeric Differences in Paraffin Properties J Phys Chem 56:328-336

Poling BE, Prausnitz JM, O´Connell JP (2001) The Properties of Gases and Liquids, 5th edition, McGraw Hill, New York

Polyansky OL, Császár AG, Shirin SV, et al (2003) High-Accuracy ab Initio Rotation-Vibration Transitions for Water, Science 299, 539-542

Pople JA, Beveridge DL (1970) Approximate Molecular Orbital Theory. McGraw Hill, New York

Ritter ER (1991) THERM: a computer code for estimating thermodynamic properties for species important to combustion and reaction modeling. J Chemical Information and Computer Sciences 31:400-408

Ritter ER, Bozzelli JW (1991) Therm: thermodynamic property estimation for gas phase radicals and molecules. Int J Chem Kin 23:767-778

RMG – Reaction Mechanism Generator, http://greengroup.github.com/RMG-Py/ (accessed 27 Dec 2012; the link connects to the Python version, but there is a Java version, too)

Rossini FD, Pitzer KS, Arnett RL, et al (1953) Selected values of physical and thermodynamic properties of hydrocarbons and related compounds, American Petroleum Institute, Carnegie Press, Pittsburgh, PA

Rossini FD, Wagman DD, Evans WH, et al (1952) Selected Values of chemical thermodynamic properties, Circular of the National Bureau of Standards 500, US Department of Commerce, National Bureau of Standards

Ruscic B, Boggs JE, Burcat A et al. (2005) IUPAC Critical Evaluation of Thermochemical Properties of Selected Radicals: Part I, J Phys Chem Ref Data 34:573-656

Ruscic B, Pinzon RE, Morton ML et al (2004) Introduction to Active Thermochemical Tables: Several "Key" Enthalpies of Formation Revisited. J Phys Chem A 108:9979-9997

Ruscic B, Pinzon RE, Morton ML, et al (2006) Active Thermochemical Tables: Accurate Enthalpy of Formation of Hydroperoxyl Radical, HO_2, J Phys Chem A 108:6592-6601

Ruscic B, Pinzon RE, Laszewski Gv et al (2005) Active Thermochemical Tables: Thermochemistry for the 21st Century. J Phys Conf Ser 16:561-570

Ruscic B, Wagner AF, Harding LB et al. (2002) On the Enthalpy of Formation of Hydroxyl Radical and Gas-Phase Bond Dissociation Energies of Water and Hydroxyl. J Phys Chem A 106:2727-2747

Schuchardt KL, Didier BT, Elsethagen T, et al (2007) Basis Set Exchange: A Community Database for Computational Sciences. J Chem Inf Model 47(3):1045-1052, doi:101021/ci600510j

Shomate CH (1954) A Method for Evaluating and Correlating Thermodynamic Data J Phys Chem 58:368-372

Slater JC (1963-1974). Quantum Theory of Molecules and Solids, Vols. 1-4: Electronic Structure of Molecules. McGraw-Hill, New York

Somayajulu GR; Zwolinski BJ (1966) Generalized treatment of alkanes. Trans Faraday Soc 62: 2327-2340

Stein SE, Rukkers JM, Brown RL (1991) NIST Standard Reference Database 25: NIST Structures and Properties Database and Estimation Program; National Institute of Standards and Technology: Gaithersburg, MD, 1991; see also http://webbook.nist.gov/chemistry/grp-add/S-and-P.html#ref5 (accessed 29 Dec 2012)

Stevens WR, Ruscic B, and Baer T (2010) The Heats of Formation of C_6H_5, $C_6H_5^+$, and C_6H_5NO by TPEPICO and Active Thermochemical Tables Analysis. J Phys Chem A 114:13134-13145.

Stull DR, Westrum EF, Sinke GC (1969) The Chemical Thermodynamics of Organic Compounds. Wiley, New York

Sumathi R, Carstensen HH,. Green WH Jr. (2001) Reaction Rate Prediction via Group Additivity Part 1: H Abstraction from Alkanes by H and CH_3. J Phys Chem A 105:6910-6925

Sumathi R, Carstensen HH, GreenWH Jr. (2001) Reaction Rate Prediction via Group Additivity Part 2: H-Abstraction from Alkenes, Alkynes, Alcohols, Aldehydes, and Acids by H Atoms. J Phys Chem A 105:8969-8984

Sumathi R, Carstensen HH, GreenWH Jr. (2002) Reaction Rate Predictions Via Group Additivity. Part 3: Effect of Substituents with CH_2 as the Mediator. J Phys Chem A 106:5474-5489

ThermoData Engine (TDE), NIST TDE 103a - Pure Compounds, NIST TDE 103b - Pure Compounds, Binary Mixtures, Ternary Mixtures and Chemical Reactions

ThermoML IUPAC Standard, http://www.iupac.org/namespaces/ThermoML/index.html (accessed 27 Dec 2012) is an XML based standard for thermodynamic data communication that was initially developed within IUPAC Project 2002-055-3-024, later extended within the IUPAC project 2007-039-1-024. See also: Frenkel M et al. (2006) Pure Appl Chem 78:541-612, Pure Appl Chem (2011) 83:1937-1967

ThermoML Web Archive, http://trc.nist.gov/ThermoML.html (accessed 14 January 2013)

Thomsen J (1886) Thermochemische Untersuchungen Vol. IV, J. A. Barth, Leipzig

Truhlar DG (2008) Molecular Modeling of Complex Chemical Systems J Am Chem Soc 130, 16824–16827, Glossary, http://pubs.acs.org/JACSbeta/JVI/issue3.html and http://pubs.acs.org/JACSbeta/jvi/glossary.html

Vansteenkiste P, Van Neck D, Van Speybroeck V, et al (2006) An extended hindered-rotor model with incorporation of Coriolis and vibrational-rotational coupling for calculating partition functions and derived quantities. J Chem Phys 124:044314; additions and corrections (Publisher's Note) J Chem Phys 125: 049902

Veszprémi T, Fehér M (1999) Quantum Chemistry: Fundamentals to Applications. Kluwer Academic/Plenum Publishers, New York

Wiberg KB, Bader RFW, Lau CDH (1987) Theoretical Analysis of Hydrocarbon Properties. 2. Additivity of Group Properties and the Origin of Strain Energy. J Am Chem Soc 109, 1001-1012.

Zülicke L (1973) Quantenchemie. Ein Lehrgang. Band 1: Grundlagen und allgemeine Methoden. VEB Deutscher Verlag der Wissenschaften, Berlin (in German; a Russian edition is available)

Zülicke L (1985) Quantenchemie - Ein Lehrgang. Band 2, Atombau, Chemische Bindung und molekulare Wechselwirkungen, Hüthig Verlag, Heidelberg (in German; a Russian edition is available)