

Molecular Shape

Peter Comba and Jan C. A. Boeyens

Abstract Molecular shape is recognized as an emergent property that complements the projection from four-dimensional space-time to tangent Euclidean space. Projection from hypercomplex algebra to real algebra necessitates the three-dimensional definition of concepts such as chirality, quantum uncertainty and probability density to compensate for errors of abstraction. The emergent alternative description of extranuclear charge density as spherical standing waves, optimized by a golden spiral, reveals atomic structure in line with the periodic table of the elements and underpinning the concepts of bond order, interatomic distance and stretching force constant, related to chemical interaction. The principles giving rise to molecular structure are shown to depend, like bond order, on the constructive interference of atomic wave fields, optimized by minimal adjustment to bond orders. The procedure is shown to be equivalent to the philosophy of molecular mechanics. Arguments based on the traditional interpretation of electronegativity, are presented to relate the parameters of strain-free bond lengths, dissociation energies and harmonic force constants, used in molecular mechanics, to quantum-mechanically defined ionization radii of atoms. Atomic electron densities and a bond-order function, both obtained by number-theory optimization, enable the direct calculation of interatomic distance, dissociation energy and stretching force constant for all pairwise interactions of any order. Torsional interaction determines the final shape of a molecule and presumably can only be understood as a four-dimensional effect.

Keywords: angular momentum, emergent property, molecular mechanics, molecular structure, torsional interaction

Corresponding author: Peter Comba
Anorganisch-Chemisches Institut, Universität Heidelberg,
e-mail: Comba, Peter <Peter.Comba@aci.uni-heidelberg.de>

Jan C. A. Boeyens
Unit for Advanced Scholarship, University of Pretoria, South Africa,
e-mail: jan.boeyens@up.ac.za

1 Introduction

The principal aim of chemical analysis is to develop a theoretical model of the interaction between atoms and molecules alike. Experimental work of the previous two centuries has resulted in a highly successful empirical account of chemical reactivity, but efforts to formulate a fundamental theory as a non-classical many-body problem have been less fruitful. By this approach chemical interaction is modelled in terms of probability-density distributions of independent electrons. Although the theory appears to work for one-particle problems, unforeseen effects emerge in the treatment of more complex systems [1]. In particular, the distribution of extranuclear electrons seems to obey an exclusion principle, not anticipated in the basic theory and there is no fundamental understanding of three-dimensional molecular shapes, as observed experimentally. The pivotal role of entropy, which controls the course of chemical reactions, is theoretically equally unexpected.

It is not unexpected that problems often occur in the fundamental analysis of emergent properties. Maybe the prudent response of the chemist should then be a critical re-examination of those assumptions that underpin the partially successful theory. In any theory there is a reductionist limit, beyond which there are no data to guide the recognition of more fundamental principles. In the theory of matter this limit occurs in the vacuum, or sub-aether [2], seen as the primeval form of matter, continuously spread across the endless void. On deformation of this featureless cosmos, ponderable matter emerges from the void as elementary distortions, which are perpetually dispersed, except in a closed system. We propose such a structure as the primary assumption in the theory of molecular shape and assume that persistent elementary units occur in the form of what will be called waves, for lack of better terminology. The elementary waves exhibit the first emergent properties of mass, charge and spin, which they possess in characteristic measure.

Matter in all forms can now be recognized as consisting of the three robust elementary forms known as proton, electron and neutrino. A fourth common form, known as neutron, only occurs in close combination with protons. It has a limited lifetime in free space where it decays into an electron, a proton and a neutrino.

The postulated elementary units should not be confused with waves as perceived in three-dimensional space, but rather as undulations in four-dimensional space-time. The mathematical description of these two types of swaying is fundamentally different. A steady state that results from four-dimensionally balanced forces, as a function of a scalar potential Φ , obeys an equation of the type

$$\frac{\partial^2 \Phi}{\partial x_0^2} + \frac{\partial^2 \Phi}{\partial x_1^2} + \frac{\partial^2 \Phi}{\partial x_2^2} + \frac{\partial^2 \Phi}{\partial x_3^2} = 0 \quad (1)$$

By assuming Φ in the form of a product function, each term in

$$\frac{1}{\Phi} \square^2 \Phi = \sum_{\mu=0}^3 \frac{1}{X_\mu} \frac{\partial^2 \Phi}{\partial x_\mu^2} = \sum k_\mu^2 = 0$$

defines a characteristic constant on condition that $\sum k_\mu^2 = 0$. The general solution of (1) is thereby defined as a four-dimensional hypercomplex function, known as a quaternion, in which all variables remain inextricably entangled.

A common approximation that reduces the equation into a three-dimensional wave equation assumes the separation of space and time coordinates, which is the basis of wave mechanics. For many purposes this is a good approximation in tangent Euclidean space, but it has no validity in curved four-dimensional space-time.

It is important to note that the property of spin is only defined in quaternion notation, which specifies a conserved quantity \mathbf{J} . It may be viewed as a four-dimensional symmetry operator, approximated by a three-dimensional angular-momentum operator \mathbf{L} and a one-dimensional spin, on separation of space and time variables. The approximation $\mathbf{J} = \mathbf{L} + \mathbf{S}$ implies that neither \mathbf{L} nor \mathbf{S} is a three-dimensional vector, both of them implying rotation in spherical mode [3]. The one-dimensional projections, \mathbf{L}_z and \mathbf{S}_z , in an applied magnetic field or in a molecular environment are vector quantities.

2 Space-like Correlations

It is appropriate to digress at this point into a discussion of the much-debated non-locality of quantum theory. In a nutshell it amounts to the observation that a wave-mechanical correlation, established at a given point, remains operational even as a correlated pair drifts apart indefinitely. The logical implication of this is instantaneous action at a distance; an idea much maligned over the ages, but well-defined as space-like interaction in the theory of special relativity. The interval between four-dimensional points in Minkowski space,

$$ds = \sqrt{dx_0^2 - dr^2},$$

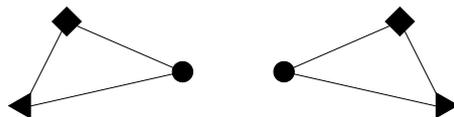
reduces to zero in the surface of the light cone and becomes complex on the outside. Whereas a stationary object within the light cone moves through time with constant space coordinates, the time coordinate remains constant for such an object in the space region. Irrespective of spatial separation any pair of points in four-dimensional space-time therefore remain in virtual contact and correlated.

Humans are conditioned to interpret the environment in three-, rather than four-dimensional detail and destined to experience events that appear normal in four dimensions as mysterious in three. Common examples include non-locality, the commutation properties and other mysterious features of wave-mechanical variables and the chirality of space.

3 Wave-mechanical Approximation

The non-classical mathematical description of the world follows equation (1), which in practice is solved by the separation of space and time variables. Although it is a good approximation it cannot render four-dimensional effects intelligible in three. The problem is highlighted by analogy with efforts to describe geometrical shapes in lower-dimensional space.

Mirror-related triangles in two dimensions clearly define a chiral pair, which appear achiral in three dimensions.



In the same way the chirality of a three-dimensional tetrahedron is resolved in four dimensions, which means that the three-dimensional chiral forms are identical when described four-dimensionally. Small wonder that all efforts to find a wave-mechanical difference between *laevo* and *dextro* enantiomers are inconclusive. The linear superposition principle, widely acclaimed as a distinctive property of quantum systems, is now recognized as no more than a partially successful device to mimic four-dimensional behaviour. This includes one of the pillars of chemical-bonding theory, known as the resonance principle.

Probably the most distinctive feature of quantum systems is the non-zero commutators of conjugate variables, said to represent a drastic departure from classical behaviour. In actual fact this is a standard feature of any \mathbb{H}^4 algebra, which only becomes problematic on trying to reformulate this in \mathbb{R}^3 . All of the foregoing is of decisive importance in a theory of molecular shape and creates a serious interpretational dilemma. The strictly three-dimensional empirical data simply cannot support a quasi four-dimensional theory, whereas many observed features cannot be accounted for classically. The point is that nuclear position, the decisive parameter, is a strictly classical particle property, but the interaction between atomic nuclei is wave-like. The traditional compromise to represent both nuclei and electrons by probability-density functions does not work.

4 Atomic Structure

The formation of molecules is driven by the interaction between the extranuclear electronic charge clouds that surround atomic nuclei. According to the wave model proposed here [4] such an electron cloud is conveniently considered as a spherical standing wave in the form of concentric annular shells. Each shell consists of a fixed number of electrons, proscribed by the quantum numbers l and s . All inner shells in a ground-state atom are considered closed with the maximum number of electrons. Except for some special cases such as the inert gases, the outermost valence shell

is, by definition, not a closed shell. However, given the sub-level degeneracy of $2l + 1$, with magnetic quantum numbers in the range $-l \leq m_l \leq l$, the valence shell remains spherical in all cases. The quantum number $m_l = 0$ defines a real wave function with spherical symmetry and any pair of complex functions corresponding to $\pm m_l$ also defines the same symmetry. The quantum numbers for any number of electrons in the valence shell can therefore always be assigned so as to define spherical symmetry. This stipulation is known as Hund's rule.

It is important to realise that all wave-mechanical predictions pertaining to atomic structure are only valid for the H atom. For this reason wave mechanics is only partially successful as a descriptor of the periodic table of the elements and hence, of the electronic configuration of many-electron atoms.

Simulation by number theory is the only known procedure that generates the detailed structure of the periodic table without further assumptions or *ad hoc* corrections. In its simplest form the simulation is based on the fact that any atomic nucleus consists of integral numbers of protons (Z) and neutrons (N), such that the ratio Z/N is a rational fraction. This ratio converges from unity to the golden ratio (τ) with increasing atomic number, and yields a distribution commensurate with the periodic table. The detailed structure of the periodic function is contained in the Farey sequence \mathcal{F}_4 of rational fractions and visualized in its Ford-circle mapping [5].

Noting that the periodic table derives from the extranuclear electronic configuration of atoms it would seem reasonable to assume that a number-theoretic simulation could reveal this distribution as well. In this instance we are dealing with the special distribution of matter around an active centre; the type of problem amenable to analysis by optimization in terms of logarithmic spirals. The only requirement is recognition of an appropriate convergence angle. Using the maximum valence-shell degeneracy of $2n - 1$ at the principal level n , a distribution that appears to replicate the radii of the Bohr model at $n^2 a_0$ for a variable convergence angle of $4\pi/(2n - 1)$, was indicated [6] and is shown here in Fig. 1.

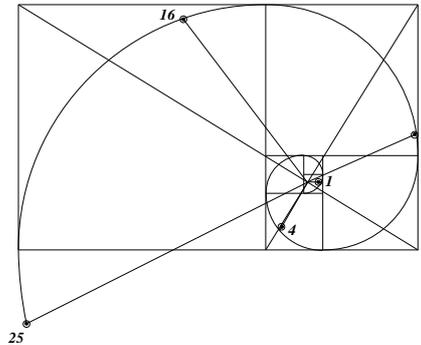


Fig. 1 Points generated in a golden rectangle by a Fibonacci spiral with variable convergence angle of $4\pi/(2n - 1)$. Numbers indicate distance to the spiral centre in units of a_0 .

However, in terms of the known periodic structure, the Bohr interpretation cannot apply. In order to generate the periodic table it is necessary to interpret the extremum condition as satisfied by the nodal surfaces of the spherical electron wave, as shown in Fig. 2. The detailed periodic structure, together with sub-shell degen-

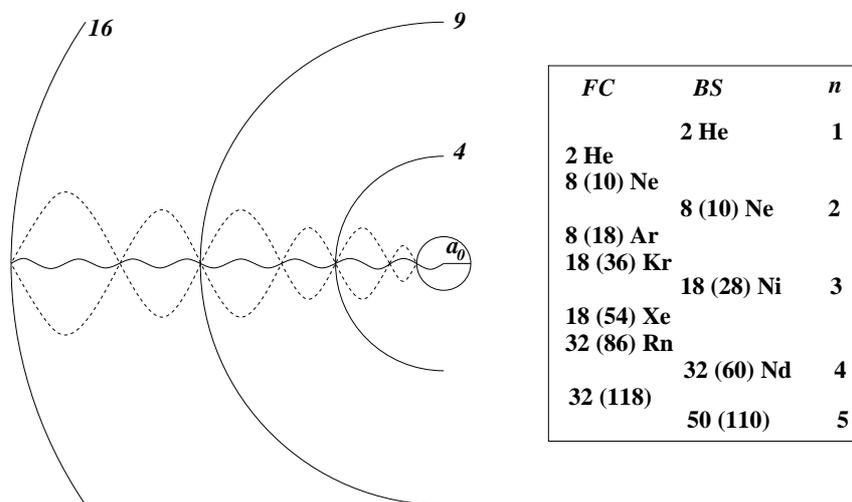


Fig. 2 Atomic shell structure as it emerges from electron-density optimization on a golden spiral. The variable convergence angle of $4\pi/(2n-1)$ manifests in the appearance of $2n-1$ additional cycles (*s, p, d, f*) in each interval between Bohr levels n and $n-1$, shown here as elementary ripples. In contrast to the Bohr-Schrödinger (BS) model, closed shells in the Ford-circle simulation (FC) invariably coincide with noble-gas configurations.

eracy of $2n-1$, corresponding to *s, p, d, f* spectroscopic states, emerges from the graphical representation. The model has been shown [4] to produce the electronic distribution on all atoms in quantitative detail.

It may be unexpected to find that number theory and traditional wave mechanics yield comparable reconstructions of extranuclear electronic configurations. However, both models are based on classical waves in three-dimensional space, appropriate for the understanding of atomic structure in tangent Euclidean space.

5 Molecular Structure

Towards an understanding of molecular shape it is important to enquire into the symmetry of an activated atom in a chemically crowded environment. Whether the quantum number m_l represents, as traditionally interpreted, a measure of directed orbital angular momentum, or an element of symmetry, is immaterial. An interatomic collision either redirects the orbital angular momentum of the extranuclear

electrons or distorts the symmetry of the charge clouds. The onset of covalent interaction is recognized in either the quenching of orbital angular momentum, or equivalently, optimization of the overall symmetry as specified by Laplace's equation. Either way this principle establishes a criterion in terms of which to predict the relative orientation of sub-molecular fragments that join up in a chemical reaction. Viewing the quantum number m_l as an orbital angular momentum vector along the polar direction (conventionally denoted by z) is the more convenient practice. It has the advantage that residual angular momentum is the recognized diagnostic of a magnetic moment that generates optical activity.

In the case of second-order covalent interaction the angular-momentum vectors line up antiparallel in a direction perpendicular to the axis of interaction. The energy which is required to decouple these vectors measures the steric rigidity of the arrangement, known as a barrier to rotation. In terms of the wave picture steric rigidity relates to the overlap mode of wave crests as shown for ethylene and ethyne in Fig. 3.

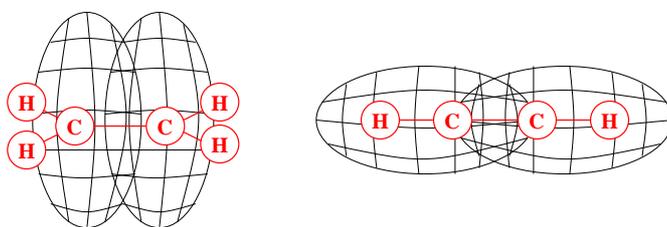


Fig. 3 Relative rotation about the axis of interaction destroys the interference pattern between the wave forms shown on the left, but not of those on the right.

The universally accepted model of the second-order dicarbon interaction, colloquially known as an ethylenic double bond, has developed from a misreading of a seminal paper which discussed the quantum theory of double bonds [7]. In this paper it is shown that by linear combination of the eigenfunctions

$$\psi(r, z) \frac{e^{\pm i\varphi}}{\sqrt{2\pi}},$$

which describe the angular momentum on a C atom, two eigenfunctions of different energy,

$$\psi(r, z) \frac{\cos \varphi}{\sqrt{\pi}} \quad \text{and} \quad \psi(r, z) \frac{\sin \varphi}{\sqrt{\pi}},$$

are produced. Interaction between two $-\text{CH}_2$ units is next analyzed by perturbation theory in terms of two derived functions of even and odd parity with respect to the plane $\varphi = 0$:

$$\psi_g(r, z, \varphi) \quad \text{and} \quad \psi_u(r, z, \varphi).$$

It was correctly pointed out that with these linear combinations a moment, as in O_2 , no longer exists around the C–C axis. ψ_u^2 is identified as the diatomic ground state which is shown to require a planar C_2H_4 arrangement, interpreted as the cause of the steric rigidity of ethylene.

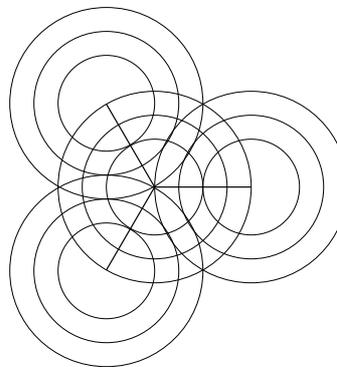
Derivative work, based on this analysis, erroneously assumes degeneracy of the linear combinations when defining a set of orthogonal real “orbitals”. There is no theoretical basis in the seminal paper to justify this assumption. However, what is clearly implied is that the orbital moment on each carbon atom is directed perpendicular to the molecular plane. The molecular angular momentum is therefore quenched vectorially only for this planar arrangement of the molecule. Torsional distortion which creates residual angular momentum therefore requires work, the true basis of a barrier to rotation.

It is instructive to note that the requirement of quenched angular momentum predicts the same tetrahedral geometry for methane [8] as the concept of symmetry optimization.

By exploiting these principles it becomes feasible to reconstruct the general topology of complicated molecules with known connectivity. Optimization of the topological shape to produce the geometrical details of molecular structure may be done by the methods of molecular mechanics.

During diatomic covalent interaction the spherical electronic waves on free atoms generate specific interference patterns, which define bond order, interatomic distance and a stretching force constant, characteristic of the interaction. These characteristic properties remain largely intact as the diatomic fragment becomes incorporated in larger molecular assemblies, the three-dimensional structure of which, in the first instance, depends on the interference between second-neighbour waves. It is almost axiomatic that the wave structure of the central atom in a planar arrangement will be elongated as shown for ethylene in Fig. 3.

Fig. 4 Schematic drawing of the interacting spherical electronic waves among four identical atoms in a plane.



The geometry of a four-atom molecule, as shown in Fig. 4, but with dissimilar atoms, is specified by three first-neighbour and three second-neighbour interatomic distances – a total of six independent parameters, not necessarily compatible with

constructive interference among all wave systems. In order to arrive at an arrangement that optimizes comprehensive constructive interference it may be necessary to make small adjustments to the six independent bond orders. The required adjustments will be functions of the relative stretching force constants. Each adjustment requires an amount of work,

$$w = \frac{1}{2}k_r\delta^2,$$

where δ defines a linear increase in interatomic distance. The optimized nuclear framework will in general not be planar, as suggested by the diagram.

The simple procedure, outlined here and extended over any number of connected atoms, with minimization of the total work required to produce the optimal structure, constitutes the philosophy of *molecular mechanics* (MM). In practice all secondary interactions are described in terms more familiar to structural chemists. In this way 1,3-interactions are formulated as deformation of characteristic *valence angles*, according to

$$w_\theta = \frac{1}{2}k_\theta(\Delta\theta)^2,$$

based on an *angle-bending* force constant. A 1,4-interaction is reduced to a *torsional* function that describes a periodic barrier to rotation and a so-called *non-bonded* interaction, which is also used to incorporate more remote interactions into the force field. Special parameters are added to deal with electrostatic interaction between polar regions and to maintain the planarity of special conjugated systems. Another refinement considers the interdependence of stretches and bends pertaining to common atoms.

Historically, molecular mechanics has developed from a purely empirical procedure to refine molecular trial structures by the minimization of steric energy as a function of nuclear coordinates [9, 10]. The trial structure is generated by assigning empirically idealized interatomic distances and valence angles according to the chemical connectivity pattern. Today, under certain conditions, it is not only the most efficient but also a very reliable procedure for the optimization of three-dimensional molecular structures, not only, but specifically also for large systems such as macromolecules, condensed phases and the analysis of conformational space [9, 11].

Although most of the parameters used in MM simulation are based on adequate theoretical concepts, the overall procedure to generate molecular structure remains essentially empirical. The recognized benchmark generally is an experimental structure and therefore not a free molecule but a selected rigid fragment from a bulk structure (*i.e.*, a crystal), and specific intermolecular interactions are generally ignored. It was noted that isotropic (averaged) intermolecular interactions are included in a force field based on crystal structures, and this mimics the situation in solution. Therefore, force fields based on experimental structures are not reproducing structures of isolated molecules and are slightly different from force fields derived from quantum-chemical structures, which are claimed to produce the structures of isolated molecules [9, 11, 12]. As intermolecular interactions are amongst the weakest forces in bulk phases, bond distances and valence angles may not be significantly affected by the environment. It is mainly the relatively weak torsional interactions

that differ significantly from their free-molecule counterparts. Since the torsional flexibility of (primarily carbon-carbon) single bonds is the basis of conformational variations, the shape of a molecule considerably depends on intermolecular interactions. It follows that MM is only able to correctly predict the shape if solvation, hydrogen bonding, long-range van der Waals interactions, electrostatics and polarization are accurately computed and if a full conformational analysis is performed.

5.1 Wave Mechanics

Quantum theory in general has problems similar to those discussed above for molecular mechanics. In addition, the computation of bulk phases and conformational analysis are computationally much more expensive and therefore virtually impossible. Matter, in its most elementary form, is nothing but a special configuration of four-dimensional space-time. The allowed shapes of material aggregates must therefore depend on the topology of space-time and any theory of molecular matter is inferred to incorporate some aspect of four-dimensional symmetry. However, the wave-mechanical model of matter, on assuming the separation of space and time coordinates, destroys the four-dimensional symmetry, as evidenced by the disappearance of the spin variable that links the shape of material aggregates to the topology of space-time.

In the form of a three-dimensional wave equation the theory defines a complex variable associated with a vector model of angular momentum. By the principle of symmetry optimization, *viz.* minimization of angular momentum, this variable may be used to predict the internal three-dimensional symmetry of assembling molecules [8]. Followed by MM optimization it might constitute a method of predicting classical molecular structure from first principles. However, the standard procedures of *quantum chemistry*, which rely on further separation of space variables, sacrifice the angular-momentum parameter, in order to eliminate the complex variable, and hence suppress the facility to predict internal molecular symmetry.

6 Molecular Mechanics

Among the techniques for structure optimization of chemical systems, molecular mechanics (MM) is by far the fastest and therefore, for large systems such as crystal lattices, polymers, proteins and solutions often the only useful method, especially, when a significant part of the conformational space and/or dynamics need to be included [9, 13]. In many areas, where accurate force fields have been carefully optimized, *e.g.* for carbohydrates and organic compounds in general [9, 14, 10], but also, *e.g.* for cobalt(III) hexamines [9, 15, 16, 17], the accuracy of the optimized bond distances is $< 0.01\text{\AA}$, and there also is good agreement between computed and experimental thermodynamic properties (relative strain energies) and vibrational fre-

quencies [9, 10, 18, 19], *i.e.*, the MM-derived parameters are nearly as accurate as the experimental data (*e.g.* X-ray crystallography) and in many cases better than those derived from quantum-chemical methods (QM).

Why then bother about much more expensive QM-based models? One reason is that MM may only lead to accurate results for molecules of the same type used for the optimization and validation of the force field, *i.e.* extrapolation is seen to be dangerous if not impossible [9]. This also extends to transition states and short-lived, unstable intermediates, and therefore to chemical reactivity. Since electrons are not considered explicitly in MM, electronic effects related to structural distortions, specific stabilities and spectroscopy cannot be modeled by MM. However, in all other areas, there is no good reason for not using a well optimized and validated MM model. Also, there are MM-based approaches to deal with most of the deficiencies listed above [9, 20, 21, 22, 23, 24, 25]. In the last decade, there have been a number of approaches, which have, based on simple rules [26], valence bond theory [27, 28, 29, 30] and ligand field theory [20, 21, 22, 23], allowed the simplification of the force-field optimization and validation procedure, and/or inclusion of electronic effects in MM models.

Therefore, the probably most serious disadvantage of MM compared to all other approaches in structural modeling is a seemingly missing theoretical basis, and the unspoken consensus is that, despite its successes, MM should eventually give way to more sophisticated QM-based models. It is primarily density functional theory (DFT), which in recent years, due to important developments in theory, hard- and software, has taken over some of the ground from MM. An interesting approach related to this observation is that the full potential of DFT is then realized when it is used to build-up a force field [31]. Another important aspect, of course, is that MM in general only produces optimized structures and minimized energies, *i.e.* no information about electronic ground and excited states.

6.1 Electronegativity

An alternative approach to QM, for the elucidation of the electronic basis of MM, is to reexamine the traditional model of covalent interaction mediated by electron pairs, based on the notion of chemical affinity. The assumption is that the interaction between a pair of atoms in a molecule only involves their valence electrons. If the two atoms are sufficiently alike, equally shared electron density between the atoms binds them together, if they are of different chemical nature, the sharing is unequal and the covalency reduced.

It has been a constant pursuit of chemists to classify the elements in terms of such an affinity factor, in order to predict details of their interaction. The discovery of the periodic table provided the first clue towards identification of such a chemical bonding parameter, which later became known as electronegativity. On plotting Lothar Meyer atomic volumes as a function of atomic number the elements are divided into

two kinds, according to the local slope of the connecting curve [32]. This observation led to the classification into electropositive and electronegative elements.

Electronegativity has recently been redefined [33] as the quantum potential of the atomic valence state, calculated from the ground-state energy of an electron, confined to the ionization sphere of radius r_0 :

$$E_g = \frac{h^2}{8mr_0^2}.$$

To stay in line with common practice it is convenient to assign a value of $\chi = \sqrt{E_g}$, with E_g in eV, for general use. Numerically this $\chi = 6.133/r_0$, with r_0 in Å units.

Since r_0 is characteristic for each atom, characteristic energies are predicted for atomic valence state electrons. It is the atomic equivalent of the Fermi energy of an electron at the surface of the Fermi sea in condensed phases, and in that sense represents the chemical potential of the valence electron for each atom. Electronegativity has been defined independently [34] in almost identical terms before. It is a function of only the electronic configuration of atoms and emerges naturally in the response of an atom to its environment. Alternatively, it is the tendency of an atom to interact with electrons and the fundamental property that quantifies chemical affinity and bond polarity.

It is instructive to examine the periodic variation of valence state electronegativities, as a function of atomic number. It separates into the same segments as the Lothar-Meyer curve and the qualitative trends are recognized as related to the known empirical trends of other electronegativity scales [33]. The slope of the curves at each atomic position represents a change in energy as a function of atomic number (*i.e.* number of electrons), and defines the chemical potential of the electrons, $dE/dn = -\mu$, at that point.

6.2 Simulation by Number Theory

In a previous analysis, based on the generalized covalency curve and empirically adjusted values of ionization radii, r_0 [8], it was shown how to obtain useful MM force-field parameters. An even simpler and more reliable method has now become available by recalculation of atomic ionization radii directly from numerical optimization of valence densities.

The recalculated ionization radii are essentially free-atom values and therefore no longer parametrically related to the general covalence curve. However, for homonuclear interactions of the same order, b , a common dimensionless interatomic distance d'_b is predicted, such that $d = d'_b r_0$. By considering the interaction as an interference between spherical standing waves, integer and half-integer bond-order parameters d'_b are readily optimized with a golden logarithmic spiral. These parameters vary between $d'_0 = 1$ and $d'_4 = \tau$ to allow the calculation of dissociation energy from a bond-order related to some power of the golden ratio [35]:

$$D_c = Kr_0^2 \tau^n.$$

K is a dimensional constant. For heteronuclear interaction,

$$D_c = Kr_0^3(1)\tau^n/r_0(2) \quad ; \quad r_0(1) > r_0(2).$$

The optimized quantized values of bond order correspond in an overwhelming majority of cases directly with the traditionally recognized single, double, triple and quadruple bonds that count electron pairs per interaction. Deformation of any diatomic system is resisted by the disturbance of an optimal arrangement, and measured as an harmonic stretching force constant. Its value depends on the differences $\Delta D'$ and $\Delta d'$, and the slope, s , of the line that represents the change between bond orders as a continuous function as shown in Fig. 2 of the Covalence paper [35]. Calculation of $\Delta D'$ is simplified by a special property of the golden ratio: $\tau^{n+1} - \tau^n = \tau^{n+2} \equiv \tau^+$. In the common units of Ncm⁻¹ or mdyne/Å, the force constant for a homonuclear stretch is defined as:

$$k_r = \frac{4.615 \tau^+ s}{(\Delta d' \cdot r_0)^2}.$$

6.2.1 Non-bonded and 1,3-Interaction

The tetrahedral environment of a covalently saturated carbon atom specifies the separation between 1,3-neighbours by simple trigonometry¹, with $d(\text{C-H})=1.12\text{\AA}$, and $d(\text{C-C})=1.54\text{\AA}$ as $d(\text{H}\cdots\text{H})=1.83\text{\AA}$, $d(\text{C}\cdots\text{C})=2.52\text{\AA}$, $d(\text{C}\cdots\text{H})=2.19\text{\AA}$.

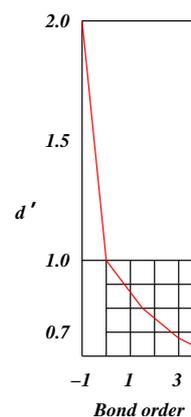
The H \cdots H distance is close to the $2r_0$ non-bonded limit, with an effective bond order of -1, as in Fig. 5. For C \cdots C, the approximate $d' = 2.52/1.78$ is interpreted to indicate an effective bond order of $-\frac{1}{2}$.

From these data and the slope of 10 in the non-bonded region, 1,3 stretching force constants are calculated and converted into angle-bending constants. In keeping with common MM practice other non-bonded contacts may be interpreted as van der Waals interaction, but no effort has been made so far to calculate an attractive component. Non-bonded distances of less than the van der Waals limit of $R_{vdW} = r_0(1) + r_0(2)$ amounts to repulsion against a force constant, as shown below. Both 1,3 and van der Waals interactions are simulated by the same function, the latter at somewhat lower order. To calculate k_r for the non-bonded C \cdots C interaction we consider a stretch from bond order $0 \rightarrow -\frac{1}{2}$ ($\Delta d' = 0.5$) for 1,3, and $0 \rightarrow 1$ ($\Delta d' = 1$) for van der Waals interaction. From table 3 [35] $n^+ = 10$. For H \cdots H and C \cdots H we consider a stretch from $0 \rightarrow 1$ ($\Delta d' = 1$). For 1,3 interactions $n^+ = 9$ and for van der Waals, with $R_{vdW} > R(1,3)$ we assume $n^+ = 10$. The results under these conditions are summarized by the following calculations:

1

$$c^2 = a^2 + b^2 - 2ab \cos C$$

Fig. 5 Variation of inter-atomic distance with bond order.



$$k_r(\text{C}\cdots\text{C}) = \frac{4.615 \times \tau^{10}}{(0.5 \times 1.78)^2} \times 10 = 0.47 \text{ Ncm}^{-1}$$

$$k_\theta = k_r \times 2.52/1.911$$

$$= 0.64 \text{ mdyne/\AA rad.}$$

$$R_{vdW} = 3.56 \text{ \AA} \quad ; \quad k_r = 0.12 \text{ Ncm}^{-1}$$

$$k_r(\text{H}\cdots\text{H}) = \frac{4.615 \times \tau^9}{(1 \times 0.98)^2} \times 10 = 0.63 \text{ Ncm}^{-1}$$

$$k_\theta = k_r \times 1.83/1.911$$

$$= 0.62 \text{ mdyne/\AA rad.}$$

$$R_{vdW} = 1.96 \text{ \AA} \quad ; \quad (\tau^{10})k_r = 0.40 \text{ Ncm}^{-1}$$

$$k_r(\text{C}\cdots\text{H}) = \frac{4.615 \times \tau^9}{1.78 \times 0.98} \times 10 = 0.35 \text{ Ncm}^{-1}$$

$$k_\theta = k_r \times 2.19/1.911$$

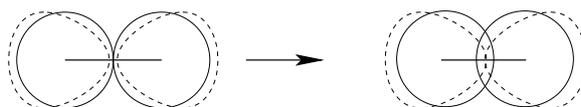
$$= 0.62 \text{ mdyne/\AA rad.}$$

$$R_{vdW} = 2.76 \text{ \AA} \quad ; \quad k_r = 0.22 \text{ Ncm}^{-1}$$

6.2.2 Interatomic distance free of strain

In principle these formulae could generate useful parameters for the simulation of a large variety of interactions by force-field methods. For simple diatomic interactions calculated values of interatomic distance for given bond order can be interpreted as the bond lengths free of strain (ideal bond distances), commonly used in MM. However, for small atoms, such as carbon, in sterically crowded environments this assumption needs adjustment.

The estimate of interatomic covalent distances proportional to $\tau^n r_0$ assumes spherical charge distributions and does not compensate for distortion by first-neighbour ligands. The effect of such distortion is illustrated schematically below.



Constructive interference between the distorted waves now occurs at a reduced interatomic distance compared to that of the undisturbed waves. Therefore, an interatomic distance free of strain in the correct molecular symmetry environment is reduced to

$$d_0 = d - \tau^{2n} r_0.$$

For first-order C–C interactions this results in

$$d_0 = 1.545 - 1.78\tau^{10} = 1.531\text{\AA}.$$

6.2.3 Covalent Interaction

The scheme outlined here has the potential to model all structural and thermodynamic effects, except for torsional flexibility, which depends on orbital angular momentum. Special parameters are needed to model these effects in MM, and these are not included here. Ionic contributions to covalent bonding have not been considered either but it should also be possible to do so with slight modifications of the model.

Rupture of a covalent bond occurs in stages in the number-theory-based model [35], by transformation into bonds of lower order, until only zero-order interactions (non-bonded) remain. At each stage, a significant rearrangement of the immediate chemical environment accompanies the lowering of bond order. An intimate relationship between all bonds in a molecule is therefore implied. It follows that intramolecular rearrangements, however drastic, do not involve the rupture of bonds. An example is the rearrangement, which involves end-on μ -peroxo-dicopper(II), side-on η^2 -peroxo-dicopper(II) and dioxodicopper(III) [36]. In such a process the molecule preserves its integrity and only experiences a concerted flow of valence electron density, whereby bond orders between all pairs of atoms are affected in a continuous process. The most visible effect could be a modification of the geometric arrangement of nuclei, naturally interpreted in terms of bonds, broken at some

points, and re-established elsewhere. This interpretation is certainly not right. Intramolecular cohesion should rather be seen as due to covalent interaction between all pairs of atoms, with non-zero bond orders restricted to neighboring pairs, and fading with increased separation.

In this sense covalency is a molecular property to be visualized as the interaction within a set of positively charged atom cores in a sea of valence electrons, spreading from local maxima around the nuclei through all of molecular space. The point-charge method models the interaction between all pairs of atoms in the equilibrium configuration. MM does the opposite: from the details of interatomic interactions it calculates the equilibrium configuration.

6.3 *MM Simulation*

The heart of MM is a force field [9], and the ultimate force field should be fully transferable between all types of molecules. However, progress towards comprehensive force fields, such as the Universal Force Field (UFF) and derivatives thereof [26, 37], is invariably accompanied by a large increase in the number of parameters. The elegant effort to reformulate MM in terms of valence bond concepts [27] has reduced the number of formal parameters at the expense of generic hybridization parameters. However, the model has not been shown to be applicable for transition metal complexes [30]. These are traditionally modeled by a points-on-a-sphere approach, where the angle function around the metal centers are replaced by non-bonded 1,3-interactions [9, 38, 39], in combination with minor corrections based on ligand-field theory [40] or, in the most general and advanced model for coordination compounds, with a ligand-field (angular overlap model, AOM) term included in the optimization routine [22, 23].

The complexity of problems addressed by molecular mechanics is such that multi-parameter modeling is almost unavoidable. The best to hope for is to find a parameter set, based on easily understood chemical concepts. The model outlined above is proposed in that spirit, although considerable refinement is required before it translates into an accurate and therefore useful tool. It is based on the classical concept of electronegativity, reinterpreted in terms of atomic ionization radii and the chemical potential of the valence state. The calculation of these parameters for non-hydrogen atoms does not involve empirical parameters or assumptions. It defines the valence state in terms of characteristic spheres to which a valence electron is confined at uniform charge density. Chemical bond formation occurs on the exchange of this valence charge density between atoms. The consequent polarization, when reduced to point-charge simulation in dimensionless units serves to describe all covalent dissociation energies as a function of interatomic distance. This function applies specifically to interactions free of strain (the ideal bond distances), which define parameters of fundamental importance in MM. Optimization of a quantized bond-order function allows direct calculation of diatomic dissociation energies and stretching force constants.

The bond-order function applies not only to integral and half-integer bond orders but also to interactions of zero and negative order, characteristic for all non-bonded interactions in a molecule. Based on these ideas it is in principle possible to define a force field, based on pair-wise interactions, which accounts for all structural and thermodynamic effects, apart from those related to orbital and spin angular momenta.

It is not the purpose of this paper to produce and present a new force field. We rather want to provide a theoretical basis for MM and therefore also to be able to efficiently produce generic force-field parameters. As it stands, one parameter (ionization radius) is needed to initiate the derivation of all other parameters to model all bond orders of any covalent interaction. It is therefore reassuring to note that the uniform valence density within a characteristic atomic sphere has the same symmetry as the 1s hydrogen electron. The first-order covalent interaction between any pair of atoms can therefore be modeled directly by the simple Heitler-London method for hydrogen to predict d , D and k_r [41]. The results are in agreement with those of the simpler number-theory simulation [35], which is therefore preferred for general use.

6.4 First Results

As a test for the proposed scheme, the parameters derived for C–C and C–H interactions by point-charge [8] and number-theory simulation were used as a force field for aliphatic hydrocarbons. Note that no optimization of the parameters was attempted, and the results are understood as a feasibility test that obviously needs further refinement. Table 1 summarizes our new parameters derived from number theory together with those from point-charge calculation.

Table 1 Force-field parameters for alkanes, as obtained by number theory. Corresponding values obtained by point-charge simulation [8] are shown in parentheses.

Parameter	C—C	C—H	H—H	Units
d_0	1.531 (1.51)	1.14 (1.07)		Å
k_b	5.12 (4.88)	4.39 (5.09)		Ncm ⁻¹
θ_0	106 (106)	110 (110)	108 (108)	deg.
$k_\theta(1,3)$	0.64 (0.6)	0.40 (0.6)	0.62 (0.6)	mdyne/Årad
d_{vdW}	3.56 (3.60)	2.76 (2.55)	1.96 (2.55)	Å
k_{vdW}	0.12 (0.6)	0.22 (0.6)	0.40 (0.19)	Ncm ⁻¹

Table 2 Experimental and computed C–C distances for substituted alkanes. (a) Experimental data taken from [9, 14]; (b) calculated from Hyperchem [43] using the MM+ force field [9, 43, 42] with $k(\text{C-C}) = 4.4 \text{ mdyne/\AA}$, $r_0(\text{C-C}) = 1.523 \text{ \AA}$; (c) this work.

Compound	C–C exp.(a) [\AA]	C–C cal.(b) [\AA]	C–C cal.(c) [\AA]	(a) – (c) [\AA]
$\text{CH}_3\text{-CH}_3$	1.532	1.531	1.531	0.001
$\text{CH}_3\text{CH}_2\text{-CH}_3$	1.534	1.534	1.533	0.001
$(\text{CH}_3)_2\text{CH-CH}_3$	1.535	1.537	1.534	0.001
$(\text{CH}_3)_3\text{C-CH}_3$	1.539	1.541	1.534	0.005
$(\text{CH}_3)_3\text{C-C(CH}_3)_3$	1.582	1.574	1.550	0.032
$((\text{CH}_3)_3\text{C})_2\text{-CH}_3(\text{C(CH}_3)_3)$	1.611	1.620	1.566	0.045

All of the values are in the expected range (see other force fields, such as MM3, Amber, Momec etc., *e.g.* in [10, 9, 13, 18, 26, 42]), but by no means are they refined and do not define an accurate force field. The major difference between the number-theory parameters and the alkane force field derived by point-charge simulation occurs in the strain-free C–C bond length with $d_0 = 1.53 \text{ \AA}$ compared to 1.51 \AA .

The optimized structural parameters of a series of aliphatic hydrocarbons, shown in Table 2, although less accurate than with a properly optimized force field, reflect the expected steric variations satisfactorily. Experimental and calculated values with an established empirical force field are shown for comparison [9, 14]. The differences between the sets of calculated bond distances demonstrate the expected similarity between the two force fields and confirm that the new approach may be used to generate a universal force field from first principles. It appears that, in the number-theory-based force field there is a small imbalance between the attractive bonding forces and the repulsive interaction, and some adjustment might be needed (see also discussion above and note that the non-bonded interactions have been modeled with the simplest possible approach). An obvious and important extension of the method described here, should it be used for the development of a general force field, is the calculation of “electronegativities” for metal ions, in order to also be able to parameterize metal-ligand interactions.

6.5 Discussion

The central idea behind number-theoretic simulation is that chemical interaction happens between reactants in their respective valence states. In the case of diatomic interaction the valence state is characterized by the ionization radii (electronegativities) of the atoms, and is described by simple formulae that relate bond order, interatomic distance and dissociation energy as functions of the golden ratio. It is important to note that the relationships are no more than good approximations for di-

atomic interaction in polyatomic molecules. For this reason it is perhaps premature to contemplate a comprehensive data set from which to generate the force field parameters for the simulation of any classical molecular structure in terms of pairwise interactions. It may even transpire that the predicted diatomic interactions at a many-ligand coordination site are not fully compatible. The simple formulae may require adjustment to be self-consistent. Harmonic force constants would be most sensitive to such modifications. In order to extend the number-theory approach rigorously to complex molecules it should be necessary to take the chemical environment of interacting atomic pairs into account. As an elementary example the interaction between two carbon atoms in ethane might be modelled as an interaction between two methyl groups in their molecular valence state. Such second-order corrections may be small and safely disregarded in general applications, but not completely ignored, as experienced in all efforts to construct universal force fields. The feasibility of deriving useful force fields based on number-theoretic diatomic interaction parameters has been demonstrated, but complications that arise on unrestricted extension of the approach are anticipated. The inability to recognize the principles that dictate molecular shape is seen as the most serious constraint. Preliminary ideas, which could serve as an initial guide in the development of algorithms that relate molecular shape to concepts in number theory, are explored next.

7 Molecular Conformation

Both molecular mechanics and wave mechanics are formulated to deal with the intricacies of molecular structure in three-dimensional tangent space. In many cases, where the procedure is clearly inadequate, only minor assumptions are apparently required to remedy small defects. Familiarity with such anomalies eventually conditions the chemist into accepting the *ad hoc* assumptions as fundamental concepts. The remarkable conviction of most chemists that optical activity only occurs as the collective property of chiral molecules in the bulk, is of this kind. It seems to avoid the absurd conclusion that the geometry of a chiral molecule could, by itself, cause optical rotation. Supposedly, it makes more sense to accept that a collection of molecules without S_n symmetry generates the helical motion of charge density that rotates the plane of polarized light. The wave-mechanical identification of molecular magnetic vectors that may interact with polarized light [8] relies on complex variables, but these are routinely eliminated by the separation of spatial variables in quantum-chemical analyses. It is obvious, therefore, that optical activity remains poorly understood. However, it is of more significance that certain aspects of optical activity cannot be rationalized in terms of the wave-mechanics of orbital angular momentum [6] and clearly depend on hypercomplex rotation of electronic charge. We identify this observation as one example that demonstrates the four-dimensional nature of molecular conformation.

7.1 Chirality

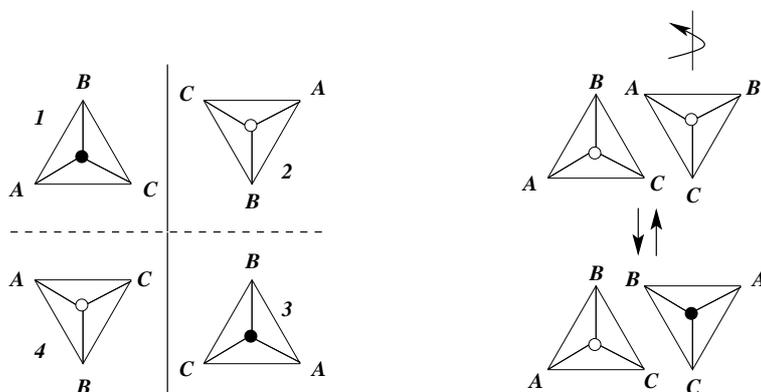


Fig. 6 In two-dimensional projection the identical pairs of rotated objects 1,4 and 2,3 appear to have opposite chiralities, whereas the enantiomeric pairs 1,2 and 3,4 appear to be identical. The racemization by 3D rotation within a 2D crystal is shown on the right.

We contend that molecular chirality appears as a four-dimensional symmetry which is incorrectly interpreted in three dimensions. The type of anticipated error is demonstrated by the way in which three-dimensional chirality is projected into two dimensions, as in Fig. 6. The two-dimensional chiral system is defined here in the plane that supports the triangular base of a three-dimensional chiral tetrahedron. The symmetry element, shown as a solid vertical line, represents an inversion (I) in three dimensions and a two-fold rotation (R) in two. The horizontal broken line represents a two-fold rotation in three dimensions, but a reflection (M) in two dimensions. To complete the argument, the three-dimensional reflection that operates diagonally, also appears as a two-dimensional reflection. Two-dimensional inversion is equivalent to rotation. In summary:

$$3DI \equiv 2DR$$

$$3DR \equiv 2DM$$

$$3DM \equiv 2DM$$

The two forms on the same side in the frame on the left have the same three-dimensional chirality but different absolute structures in two dimensions. Such a relationship would explain the variable sense of optical rotation, which depends on four-dimensional chirality, in a homochiral family of molecules. A mechanism for sterically unlikely rearrangements is illustrated by the racemization of a two-dimensional chiral crystal. The rearrangement is equivalent to a 2D reflection that

appears to be chemically and sterically forbidden within a crystal, but conveniently achieved by 3D rotation.

7.2 Torsional Interaction

The finer details of molecular shape depend to a large extent on the orientation of fluxional groups on the molecular surface. Whereas MM force fields are based on molecular structures observed in condensed phases, the torsion angles that fix the orientation of such groups cannot be specified. The same probably applies to all torsion angles in the molecular interior. The notorious difficulty of simulating the tertiary structures of large biomolecules in terms of MM pairwise interactions is due to this inability to model torsion angles.

In some carefully selected examples it has been shown how denatured proteins can spontaneously recover their natural folding pattern. The peptide torsion angles, which control the folding, return to the characteristic values of the native protein, known to be independent of chemical factors. Some long-range interaction appears to be at work.

We contend that the shape of large molecules in empty space is affected by the topology of the four-dimensional space-time manifold. Guided by the principle of cosmic self-similarity it is reasonable to assume that, like many spiral galaxies, extended molecules tend to curve like the surface of a golden spiral. It lies in an elliptic plane, which in four dimensions is the projective space, \mathbb{S}^3 , with a continuous group structure given by the quaternions, $\alpha + i\beta + j\gamma + k\delta$ with norm $\alpha^2 + \beta^2 + \gamma^2 + \delta^2 = 1$ [44]. On a local scale the shape of such a molecule is perceived to follow the surface of a large sphere, which is the projection of \mathbb{S}^3 into three dimensions. Long-chain molecules will then develop the same spiral structure as kudu horns while two-dimensionally connected macromolecules, such as graphene, will appear spherically distorted. In both cases the apparent torsion angles of π will more likely appear at a somewhat different value, such as $6/(5\tau^2)$ or $4\sqrt{\tau}$ [6]. Similarly directed torsional modification on an enantiomeric pair would impose different shapes on the molecules and destroy the apparent inversion symmetry that relates them. This minor difference in shape could be responsible for the mysterious homochirality in biological systems.

Acknowledgements The ideas outlined in this paper took shape during several visits of JCAB, sponsored by the Alexander von Humboldt Foundation, to the Institute of Inorganic Chemistry at the University of Heidelberg. Stimulating interaction with all members of the Institute is gratefully acknowledged.

References

1. Boeyens JCA (2011) in [45].
2. Eddington AS (1948) *The Nature of the Physical World*, University Press, Cambridge.
3. Batty-Pratt EB, Racey TJ (1980) *Int J Theor Phys* 19:437.
4. Boeyens JCA (2012) *Atomic Structure*. This volume.
5. Boeyens JCA, Comba P (2012) *Number theory*, This volume.
6. Boeyens JCA (2010) *Int J Mol Sci* 11:4267.
7. Hückel E (1931) *Z Phys* 70:204; English translation in [46].
8. Boeyens JCA (2008) *Chemistry from First Principles*. Springer.com, 2008.
9. Comba P, Hambley TW, Martin B (2009) *Molecular Modeling of Inorganic Compounds*. 3rd Completely Revised and Enlarged Edition. Wiley-VCH, Weinheim.
10. Burkert U, Allinger NL (1982) *Molecular Mechanics*, ACS Monograph 177.
11. Comba P, Kerscher M (2009) *Coord Chem Rev* 253: 564.
12. Remenyi R (2003) *Coord Chem Rev* 238-239:9.
13. Leach AR (1996), *Molecular modelling*, Longman, Edinburgh.
14. Allinger NL (1977). *J Am Chem Soc* 99:8127.
15. Snow MR, Buckingham DA, Marzilli DA, Sargeson AM (1969). *J Chem Soc Chem Comm.* 891.
16. Snow MR (1970). *J Am Chem Soc* 92:3610.
17. Buckingham DA, Maxwell IE, Sargeson AM, Snow MR (1970). *J Am Chem Soc* 92:3617.
18. Clark T (1985). *A Handbook of Computational Chemistry*, John Wiley & Sons, New York.
19. Niketic SR, Rasmussen K (1977). *The Consistent Force Field*, Springer-Verlag, Heidelberg.
20. Comba P, Zimmer M (1994) *Inorg Chem* 33:5368.
21. Comba P, Hambley TW, Ströhle M (1995) *Helv Chim Acta* 78:2042.
22. Burton VJ, Deeth RJ (1995) *J Chem Soc Chem Comm.* 573.
23. Deeth RJ (2006). *Eur J Inorg Chem* 2551.
— (2007). *Inorg Chem* 46:4492.
—, Paget VJ (1997). *J Chem Soc, Dalton Trans* 537.
—, Munslow IJ, Paget VJ (1997) in [47], p.77.
24. Jansen F, Norrby P-O (2003). *Theor Chem Acta* 109:1.
25. Norrby P-O (2001) in [48].
26. Rappé AK, Casewit CJ, Calwell KS, Goddard WA III, Skiff WM (1992). *J Am Chem Soc* 114:10024.
27. Root DM, Landis CR, Cleveland T (1993). *J Am Chem Soc* 115:4201.
28. Cleveland T, Landis CR (1996). *J Am Chem Soc* 118:6020.
29. Firman TK, Landis CR (2001). *J Am Chem Soc* 123:11728.
30. Landis CR, Firman TK, Cleveland T, Root DM (1997) in [47], p. 49.
31. Deeth RJ in [49].
32. Meyer L (1870) *Liebig's Annalen Supt* 7:354.
33. Boeyens JCA (2008) *Z Naturforsch* 63b:199.
34. Parr RG, Pearson RG (1983). *J Am Chem Soc* 105:7512.
35. Boeyens JCA (2012) *Covalent Interaction*, This volume.
36. Mirica LM, Ottenwaelde X, Stack TDP (2004). *Chem Rev* 104:1013.
37. Shi S, Yan L, Yang J, Fischer-Shaulsky T, Tatcher J (2003) *Comput Chem* 24:1059.
38. Hambley TW, Hawkins CJ, Palmer JA, Snow MR (1981). *Aust J Chem* 34:45.
39. Bernhardt PV, Comba P (1992) *Inorg Chem* 31:2638.
40. Comba P, Hambley TW, Ströhle M (1995). *Helv Chim Acta* 78:2042.
41. Boeyens JCA (1986). *S Afr J Chem* 82:361.
42. HyperChem™, Hypercube Inc.:1115 NW 4th Street, Gainesville, Florida 32601, USA.
43. Comba P, Hambley TW, Okon N (2001) in *MOMECLite*, a molecular modeling package for inorganic compounds; Demo version, Lauer & Okon Chemische Verfahrens- & Softwareentwicklung.
44. Stillwell J (1992) *Geometry of Surfaces*. Springer-Verlag, Heidelberg.

45. Putz MV (ed.) (2011) Quantum Frontiers of Atoms and Molecules, Nova, N.Y.
46. Hettema H (2000) Quantum Chemistry, World Scientific, Singapore.
47. Banci L, Comba P (eds) (2001) Molecular Modeling and Dynamics of Bioinorganic Systems. Kluwer, Dordrecht.
48. Cundari TR (ed.) Computational Organometallic Chemistry, Marcel Dekker Inc., New York, pp. 7.
49. Comba P (ed.) (2011) Modeling of Molecular Properties. Wiley-VCH.