

Interdisciplinary R&D

12

Theoretical Chemistry: An Overview on Modern Trends

Chandra N. Patra

Analytical Chemistry Division, Bhabha Atomic Research Centre (BARC), Trombay-400085, INDIA

$$\begin{aligned}
 H\psi &= E\psi \\
 \vec{F} &= m \frac{d^2\vec{x}}{dt^2} \\
 \Omega &\leq \Omega_0
 \end{aligned}$$

ABSTRACT

Theoretical and Computational Chemistry forms an integral component of research in chemistry and is now also emerging, through its interface with allied disciplines as a major contributor to frontier areas of interdisciplinary research. With the availability of powerful computational resources, it is no longer a dream to predict the structure and dynamics of complex systems, and to design novel molecules and materials with desired properties for various applications. A major challenge one faces in this endeavor arises from the need to use different tools for different length and time scales inherent in the description of materials and phenomena. The current review is aimed at integrating all the domains; The microscopic domain, where the electronic structure obtainable through the solution of Schrodinger equation of quantum mechanics is relevant, while in the intermediate mesoscopic length scale, classical equations of motion, statistical mechanical description and atomistic simulation are commonly used and in the larger macroscopic length scale, continuum mechanics might suffice.

KEYWORDS: *Interdisciplinary research, Novel molecules, Microscopic domain, Atomistic simulation*

Introduction

The school chemistry book starts with the concept 'Chemistry: An experimental Science', this has taken a new direction with the advent of Computational Chemistry, where theoretical and experimental research go hand-in-hand. This not only impacted the subject of chemistry in public domain but also emboldened the chemists to work along the seamless boundaries of molecules to materials. In recent years, the development of fundamental concepts and mathematical foundations has increased tremendously without any ambiguity, especially with the advancement of computational techniques along with state-of-the-art computational tools.

The genesis of 'Theoretical Chemistry', started nearly 400 years ago, when in the seventeenth century, Johannes Kepler [1] speculated on the symmetry of snowflakes and also on the close packing of spherical objects. The symmetrical arrangement of closely packed structures in the late nineteenth century led to many theories of crystallography and solid-state inorganic chemistry. John Dalton [2] represented compounds as aggregations of circular atoms, and Johann Josef Loschmidt [3] created diagrams based on circles using two-dimensional analogues. August Wilhelm von Hofmann is credited [4] with the first physical molecular model that was essentially topological. Joseph Le Bel [5] and Jacobus Henricus van't Hoff [6] introduced the concept of stereochemistry, with Van't Hoff showing tetrahedral molecules representing the three-dimensional properties of carbon. John Desmond Bernal gave [7] the first model of liquid water. It is now over 30 years since the first computer simulation of a liquid was carried out [8] at the Los Alamos National Laboratories, using the most powerful MANIAC computer.

*Author for Correspondence: Chandra N. Patra
E-mail: chandra@barc.gov.in

Electronic Structure Theory Describes the Motions of the Electrons and Produces Energy Surfaces

Hartree-Fock Theory & Beyond

The shapes and geometries of molecules, their different energy levels and wavefunctions, as well as the interactions of these states with electromagnetic fields lie within the realm of structure theory. In the Born-Oppenheimer model of molecular structure, it is assumed that the electrons move so quickly that they can adjust their motions instantaneously with respect to any movements of the heavier and slower moving atomic nuclei. This assumption motivates us to view the electrons moving in electronic wave functions within the molecule's atomic framework. These electronic functions are found by solving a Schrödinger equation [9].

$$H_k \Psi_k = E_k \Psi_k$$

thus, depend on the locations $\{Q_i\}$ at which the nuclei are sitting i.e. the E_k and Ψ_k are parametric functions of the coordinates of the nuclei and this dependence of electronic energies on the positions of the atomic centers cause them to be referred to as electronic energy surfaces.

The electronic energies $E_k(Q)$ allow one to determine the geometries and relative energies of various isomers that a molecule can assume by finding those geometries $\{Q_i\}$ at which the energy surface E_k has minima, $\partial E_k / \partial Q_i = 0$ with all directions having positive curvature (as monitored by the Hessian matrix, $H_{ij} = \partial^2 E_k / \partial Q_i \partial Q_j$, with no negative eigen values) [10]. Such geometries describe stable isomers. Also produced in electronic structure simulations are the electronic wave functions $\{\Psi_k\}$ and energies $\{E_k\}$ of each of the electronic states. The separation in energies can be used to make predictions about the spectroscopy of the system. Not only can electronic

wave functions tell us about the average values of all physical properties for any particular state (i.e., Ψ_k above), but they also allow us to tell how a specific “perturbation” (e.g., an electric field in the Stark effect, a magnetic field in the Zeeman effect, light's electromagnetic fields in spectroscopy) can alter the specific state of interest. The late Professor John Pople, made developments leading to the suite of Gaussian computer codes [11] that now constitute the most widely used electronic structure computer programs. For his contributions, he shared the 1998 Nobel Prize in Chemistry.

The full N -electron Schrödinger equation governing the movement of the electrons in a molecule is

$$\left(-\frac{\hbar^2}{2m_e} \sum_i \nabla_i^2 - \sum_a \sum_I \frac{Z_a e^2}{r_{Ia}} + \sum_{ij} \frac{e^2}{r_{ij}} \right) \Psi = E \Psi$$

However, by approximating the full electron-electron Coulomb potential $\sum_{ij} \frac{e^2}{r_{ij}}$ by a sum of terms, each depending on the coordinates of only one electron $\sum_j V(r_j)$, one arrives at N separate Schrödinger equations:

$$\left(-\frac{\hbar^2}{2m_e} \nabla_i^2 - \sum_a \frac{Z_a e^2}{r_{Ia}} + \frac{e^2}{r_{ij}} \right) \phi_i = \varepsilon_i \phi_i$$

one for each of the N so-called orbitals ϕ_i whose energies ε_i are called orbital energies. It turns out that much of the effort going on in the electronic structure area of theoretical chemistry has to do [12] with how one can find the “best” effective potential V_{eff} ; that is, the V_{eff} , which depends only on the coordinates r of one electron, that can best approximate the true pairwise additive Coulomb potential experienced by an electron due to the other electrons. The approximation in different forms along with the symmetries led to Hartree-Fock (HF) and other high-level theories for accurate description of solution of Schrödinger equations. The CI, [13] MCSCF, [14] MPPT/MBPT, [15] and CC [16] methods move beyond the single-configuration picture by adding to the wave function more configurations whose amplitudes they each determine in their own way. This can lead to a very large number of CSFs in the correlated wave function, and, as a result, a need for extraordinary computer resources.

Density Functional Theory (DFT)

Here one solves a set of orbital-level equations

$$\left(-\frac{\hbar^2}{2m_e} \nabla_i^2 - \sum_a \frac{Z_a e^2}{r_{Ia}} + e^2 \int d\mathbf{r}' \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + U(r) \right) \phi_i = \varepsilon_i \phi_i$$

in which the orbitals ϕ_i “feel” potentials due to the nuclear centers (having charges Z_a), Coulombic interaction with the total electron density $\rho(r')$ and a so-called exchange-correlation potential denoted $U(r')$. The particular electronic state for which the calculation is being performed is specified by forming a corresponding density $\rho(r')$. Before going further in describing how DFT calculations [17] are carried out, let us examine the origins underlying this theory.

The so-called Hohenberg-Kohn theorem [18] states that the ground-state electron density $\rho(r)$ describing an N -electron system uniquely determines the potential $V(r)$ in the Hamiltonian

$$H = -\sum_j \left(-\frac{\hbar^2}{2m_e} \nabla_j^2 + V(r_j) + \frac{e^2}{2} \sum_{j,k} r_{j,k} \right)$$

and, because H determines the ground-state energy and wave function of the system, the ground-state density $\rho(r)$ determines the ground-state properties of the system. It is also likely that extensions of DFT to excited states (many workers are actively pursuing this) be placed on more solid ground and

applicable to molecular systems. Because the computational effort required in DFT approaches are smaller than that for conventional (SCF, MCSCF, CI, etc.) methods, DFT offers great promise and is expanding its applicability beyond atoms and molecules to materials and biomolecules. Useful implementations of DFT using plane wave basis sets and pseudopotentials replacing the core electrons as employed in Car-Parrinello molecular dynamics (CPMD) and Vienna *Ab-initio* Simulation Package (VASP) codes find wide applications in material science and solid states.

Statistical Mechanics

Classical Theories

The goal of statistical mechanics is to interpret and predict the properties of macroscopic systems in terms of their microscopic counterparts [19]. It provides the basis for understanding numerous natural phenomena and for design and optimization of chemical processes [20]. The importance of statistical mechanics in many different branches of basic and applied chemistry has long been realized, although its tunability to specific structural and dynamical problems has become feasible only in recent times [21]. The obvious reason for such a systematic development is its versatile applicability and suitability to cover the entire spectrum of physicochemical problems, thereby providing an inherent proximity to real experimental data [22]. As a corollary, a number of approximate and semiempirical methods [23] emerge which use ingenious combinations of basic concepts from statistical mechanics.

The quanta of sophistication in statistical mechanical methods are naturally driven by the specificity of problems used to represent various physicochemical systems. So far a number of more rigorous theoretical methods have been devised, which are based on molecular simulations, liquid-state theories, self-consistent field theory, and classical density functional theory. Thus, efficient lattice-Boltzmann methods have been developed for predicting the structure and dynamics of charged colloidal systems, integral equation theory has been established for equilibrium phase behavior of fluids involving virtually any system of practical interest, density functional theory in various versions has been formulated [24] to study the conformational behavior of polymer solutions at interfaces. With the rapid increase of computational capabilities, molecular simulation and *ab initio* quantum mechanics started providing the major impetus in statistical mechanics. However, simulation itself has a long way to go for the actual realization of the fundamental concepts in chemistry, not only because significant progress has yet to be made about the strategies for modeling multiple length and timescale phenomena but, more importantly, interpretation of simulation data, much like experimental results, often requires theoretical tools for analysis and representation. Thus, the present and the next decade is going to have an amalgamation between the analytical and simulation methodologies for predicting the molecular constituents of a system with “tailored” properties, commonly used in practical applications including control of gene expression, synthesis of biomacromolecules, and fabrication of nanomaterials [28].

Basic Concepts and New Developments

Statistical mechanical description of classical systems [1] involves the concept of an ensemble, which is an arbitrarily large collection of imaginary systems, all of which are characterized by the same macroscopic parameters, but have different sets of coordinates and momenta of the particles. The system dealing with an interface is an open system in general, and is characterized by fixed values of volume V , temperature

T , and chemical potential μ with the ensemble designated as a grand canonical ensemble. Theoretical description of classical fluids at an interface starts with the description of the single particle density $\rho(r)$, of the fluid, conveniently expressed as [3]

$$\rho^{(1)}(\mathbf{r}) = \rho(\mathbf{r}) = \langle \hat{\rho}(\mathbf{r}) \rangle = \left\langle \sum_{i=1}^N \delta(\mathbf{r} - \mathbf{r}_i) \right\rangle$$

The determination of this fundamental quantity has led to a large number of methodologies, which can be broadly classified, into four categories. The first one is based on the integral equation theory (IET), which has been found to be quite successful in both homogeneous and inhomogeneous forms of description. The second one belongs to methods based on density functional theory (DFT), which has been applied to simple and complex fluids in its various forms of description, viz. perturbative and non-perturbative methods like weighted density approaches (WDA). The hybrid methods that include both the density functional and the integral equation theory constitute the third method, which has been applied quite recently to study the fluid mixtures and the electric double layer. The fourth one involves the Monte Carlo (MC) and molecular dynamics (MD) simulations, which provide direct physical insights into different important aspects of the structure of fluids.

Alongside the developments of versatile theoretical tools, the past two decades have seen a phenomenal growth in computer experiments [25] mainly because they provide essentially exact, quasi-experimental data on well-defined models, hence theoretical results can be tested unambiguously in a manner that is generally impossible with data obtained in experiments on real liquids. It is also possible to obtain information on quantities of theoretical importance that are not readily measurable in the laboratory. In a conventional MD simulation, a system of N particles is placed within a cell of fixed volume. A set of velocities is also assigned, usually drawn from a Maxwell-Boltzmann distribution corresponding to the temperature of interest. It was several years before a successful attempt was made to solve the equations of motion for a set of Lennard-Jones particles [26].

The subsequent trajectories of the particles are then calculated by integration of the classical equations of motion

$$m\ddot{\mathbf{r}}_i = -\nabla V_N(\mathbf{r}^N).$$

The particles are assumed to interact through some prescribed force law and the bulk of the computational labor is concerned with the calculation at every step of the forces acting on each particle. The static and dynamic properties of the system are then obtained as time averages over the dynamical history of the system

$$A_{obs} = \langle A \rangle_t = \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t dt A[\Gamma(t)]$$

where Γ represents a particular point in phase space. Apart from the choice of initial conditions, a molecular dynamics simulation is, in principle, entirely deterministic in nature. By contrast, as the name suggests, a probabilistic element is an essential part of any Monte Carlo computation. In a classical Monte Carlo simulation, a system of N particles interacting through some known potential is again assigned a set of arbitrarily chosen initial coordinates; a sequence of configurations of the particles is then generated $[\Gamma(t) \rightarrow \Gamma(t+1)]$ by successive random movements with the probability density $\rho_{ens}(\Gamma)$ of the ensemble. Any reasonable initial distribution should lead to the same average value.

$$A_{obs} = \langle A \rangle_{ens} = \langle A | \rho_{ens} \rangle = \sum_{\Gamma} A(\Gamma) \rho_{ens}(\Gamma)$$

with the ergodicity holding good for the specified system of interest. A number of hybrid methods are also in rolling, viz. the self-consistent density functional approach (SCDFA) [27] from DFT and IET and the Monte Carlo density functional theory (MCDFT) [28] from MC and DFT. Whereas the former is based on the calculation of density functional quantities from IET and mostly applied to simple fluids including the ionic systems, the later requires single chain simulation with enumeration of configurations and finds applications in complex polymeric fluids.

Concluding Remarks and Future Directions

The goal of materials design is the optimization of specific properties such as high strength and low density together with a number of other critical aspects including manufacturing cost and environmental acceptability. Computational materials design help in addressing all these aspects. The present decade is passing through an evolution of methodologies [29] for designing new materials through multiscale materials modeling [30]. At the quantum scale, ground state (and sometimes the excited state) energies as well as other properties (e.g., molecular geometry, vibrational and NMR spectroscopic data, multipolar moments, etc.) are calculated. Currently, ab initio molecular orbital theory [31] provides the most accurate prediction of molecular properties. The results of quantum mechanical calculations are often used in the design of molecular force fields providing a connection to the next scale, that of atomistic simulations. Using statistical mechanics, the results of atomistic or molecular scale calculations can then be applied to describe behavior at the mesoscopic and macroscopic scale (e.g., process or bulk properties). Mesoscale computations describe behavior and properties of systems that reflect the molecular composition of materials. Some current predictive approaches include linear statistical modeling, fractal models, renormalization models, lattice-Boltzmann approaches, wavelets, homogenization solutions of partial differential equations, self-consistent mean field theory, dynamic mean-field density functional methods, and dissipative particle dynamics. Application of statistical mechanical theories in physicochemical problems has opened a new vista in understanding microscopic structural and dynamic properties and linking the same to macroscopic phenomena occurring in real time. Many developments like SCF, IET, DFT, MD and MC as well as combination of these methods like SCDFA and MCDFT have strengthened the field. The practical value of all these developments is reflected not only by its generality but also by its versatility for solving complex problems. This review gives some specific examples of recent developments hovering around these methods. The applicability of the present methods demands the scope for new methodologies as well as versatile numerical algorithms.

The contents discussed in the present review are illustrative rather than exhaustive. The methods described above could be reliably applied for studying a number of current topics like the DNA salt binding, the molecular self-assembly, the solvation dynamics, the wetting transition. The organization of nanostructures within self-assembled templates has attracted a great deal of attention in recent years as this method can possibly be used in developing hybrid composite materials [32]. Molecular simulations corroborate the fact that by tethering oligomers to specific locations on nanoparticle surfaces could facilitate the self-assembly of nanoparticles into specific structures which will eventually be

dependent upon the complex geometry and topology of the tethered nanoparticles [33].

Future applications of the methods discussed above depend on continuing progress in the representation of real systems and, more importantly, on their clever implementations. Further, efficient algorithms and faster data processing are the need of the hour and quite a good number of works on artificial intelligence and machine-learning techniques are worth mentioning. [34,35] Although much current work in the literature concerns relatively simple models, much of it in the coming years will depend upon the development of more realistic force fields for more complex systems. Important advances are already emerging in applications of these methods to material fabrication, environmental protection, biomolecular engineering, and nanotechnology for various high-end applications in space, atomic energy, and defence equipment. Modeling and simulation in general and theoretical chemistry in particular will play an extremely crucial role in all these areas of developments for mankind in the foreseeable future!

Acknowledgments

The author is greatly indebted to Swapan K. Ghosh for inspiration and numerous fruitful discussions. He gratefully acknowledges Prof. Arun Yethiraj for helpful guidance. He is also benefited from insightful discussions with L.B. Bhuiyan, A.L. Benavides, Alok Samanta, T. Bandyopadhyay, N. Choudhury, and B. Modak, and other members of Theoretical Chemistry Section. He sincerely thanks Dr. A.K. Tyagi for his kind interest and constant encouragement.

References

- [1] Kepler Johannes, *The Six-Cornered Snowflake*, 1611; translated by L. L. Whyte, 1966 (Oxford Univ. Press).
- [2] Dalton, John A *New System of Chemical Philosophy* (Henderson & Spalding, London, UK, 1808).
- [3] Loschmidt J, *Chemische Studien* (Vienna, Austria-Hungary: Carl Gerold's Sohn, 1861).
- [4] August Wilhelm von Hofmann, *Extracts from Proceedings of the Royal Institution*, 4, 401-430 (1865).
- [5] Le Bel, J.-A. , *Bulletin de la Société Chimique de Paris* (in French). 22: 337 (1874).
- [6] *La Chimie dans l' Espace*, Bazendijk: Rotterdam, 1875
- [7] Bernal, J. D., *Proceedings of the Royal Society of London*. 113 (763): 117 -160 (1926).
- [8] Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H. and Teller, E., *J. Chem. Phys.*, 21, 1087 (1953).
- [9] Schrödinger, E. *Physical Review*. 28 (6): 1049-70 (1926).
- [10] P Pulay, *Molecular Physics* 17 (2), 197-204 (1969).
- [11] M. J. Frisch et al, *Gaussian*, Gaussian inc., USA.
- [12] P.W. Atkins and Ronald Friedman, *Molecular Quantum Mechanics*, (Oxford University Press, 2005).
- [13] Sherrill, C. David; Schaefer III, Henry F. ,Löwdin, Per-Olov (ed.). *Advances in Quantum Chemistry*. Vol. 34. San Diego: Academic Press. pp. 143-269. (1999).
- [14] Cramer, Christopher J. (2002). *Essentials of Computational Chemistry*. Chichester: John Wiley & Sons, Ltd. pp. 191-232.
- [15] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry* (MacMillan, New York, NY, 1982).
- [16] Mukherjee, D, *Chem. Phys. Lett.* 274 (5-6): 561-566 (1997).
- [17] R.G. Parr and W. Young, *Density-Functional Theory of Atoms and Molecules* (Oxford University Press, 1995).
- [18] P. Hohenberg and W. Kohn *Phys. Rev.* 136, B864 (1964).
- [19] D. Chandler, *Introduction to Modern Statistical Mechanics* (Oxford University Press, New York, 1987).
- [20] *Molecular Modeling and Theory in Chemical Engineering*, edited by A. Chakraborty, (Academic Press, San Diego, 2001).
- [21] *Fundamentals of Inhomogeneous Fluids*, edited by D. Henderson (Dekker, New York, 1992).
- [22] K. Efimenko, M. Rackaitis, E. Manias, A. Vaziri, L. Mahadevan, and J. Genzer, *Nature Materials* 4, 293 (2005).
- [23] A. Yethiraj, *Adv. Chem. Phys.* 121, 89 (2002).
- [24] *Chemical Applications of Density Functional Theory*, edited by B.B. Laird, R.B. Ross, and T. Ziegler (American Chemical Society, Washington DC, 1996).
- [25] M.P. Allen and D.J. Tildesley, *Computer Simulation of Liquids* (Clarendon Press, Oxford, 1991).
- [26] A. Rahman, *Phys. Rev.* 136, A405 (1964).
- [27] C.N. Patra and S.K. Ghosh, *J. Chem Phys.* 117, 8938 (2002).
- [28] A. Yethiraj and C.E. Woodward, *J. Chem. Phys.* 102, 5499 (1995).
- [29] Y. Wang, L. Herron, P. Tiwary, *Proc. Natl. Acad. Sci. USA* 119 (32), e2203656119 (2022).
- [30] *Multiscale Modelling and Simulation*, edited by S. Attinger, and P. Koumoutsakos (Springer, New York, 2004).
- [31] W.J. Hehre, L. Radom, P.V. Schleyer, and J. Pople, *Ab Initio Molecular Orbital Theory* (Wiley-Interscience, New York, 1980).
- [32] Sujit K. Ghosh, et al *Nature Communications* 15, 1278 (2024)
- [33] M.J. Mitchell et al, *Nature Communications* 14, 75 (2023)
- [34] S Menon, S Adhikari, *J MondalbioRxiv*, 01. 15.575555 (2024).
- [35] D Wang, P Tiwary *Biophysical Journal* 123 (3), 429a (2024).