Solving the Schrödinger equation of atoms and molecules: Chemical-formula theory, free-complement chemical-formula theory, and intermediate variational theory

Hiroshi Nakatsuji, Hiroyuki Nakashima, and Yusaku I. Kurokawa

Solving the Schrödinger equation of atoms and molecules: Chemical-formula theory, free-complement chemical-formula theory, and intermediate variational theory

Hiroshi Nakatsuji,[a] Hiroyuki Nakashima, and Yusaku I. Kurokawa

Quantum Chemistry Research Institute, Kyoto Technoscience Center 16, 14 Yoshida Kawaramachi, Sakyo-ku, Kyoto 606-8305, Japan

Chemistry is governed by the principle of quantum mechanics as expressed by the Schrödinger equation (SE) and Dirac equation (DE). The exact general theory for solving these fundamental equations is therefore a key for formulating accurately predictive theory in chemical science. The free-complement (FC) theory for solving the SE of atoms and molecules proposed by one of the authors is such a general theory. On the other hand, the working theory most widely used in chemistry is the chemical formula that refers to the molecular structural formula and chemical reaction formula, collectively. There, the central concepts are the local atomic concept, transferability, and from-atoms-to-molecule concept. Since the chemical formula is the most successful working theory in chemistry ever existed, we formulate our FC theory to have the structure reflecting the chemical formula. Our basic postulate is that as far as the SE is the principle of chemistry, its solutions for chemistry should have the structure that can be related to the chemical formulas. So, in this paper, we first formulate a theory that designs the wave function to reflect the structure of the chemical formula. We call this theory chemical formula theory (CFT). In the CFT, we place the valence ground and excited states of each atom at each position of the chemical formula of the molecule and let them interact using their free valences to form the ground and excited states of the molecule. The principle there is the variational principle so that the ground and excited states obtained satisfy the orthogonality and Hamiltonian-orthogonality relations. Then, we formulate the exact FC theory starting from the initial functions produced by the CFT. This FC theory is referred to as free-complement chemical-formula theory (FC-CFT), which is expected to describe efficiently the solution of the SE by the above reason. The FC-CFT wave function is modified from that of CFT. Since this modification is done by the exact SE, its analysis may give some insights to chemists that assist their chemistry. Thus, this theory would be not only exact but also conceptually useful. Furthermore, the intermediate theory between CFT and FC-CFT would also be useful. There, we use only integratable functions and apply the variational principle so that we refer to this theory as FC-CFT-variational (FC-CFT-V). It is an advanced theory of CFT. Since the variational method is straightforward and powerful, we can do extensive chemical studies in a reasonable accuracy. After finishing such studies, if we still need an exact level of solutions, we add the remaining functions of the FC-CFT and perform the exact calculations. Furthermore, when we deal with large and even giant molecules, the *inter*-exchange (*i*Exg) theory for the antisymmetry rule introduced previously leads to a large simplification. There, the inter-exchanges between distant electron pairs fade away so that only Coulombic interactions survive. Further in giant systems, even an electrostatic description becomes possible. Then, the FC-CFT for exactly solving the SE would behave essentially to order N for large and giant molecular systems, though the pre-factor should be very large and must be minimized. *Published by AIP Publishing.*
https://doi.org/10.1063/1.5040376

I. INTRODUCTION

Chemistry is governed by the principles of quantum mechanics,[1] the Schrödinger equation (SE),[2] and the Dirac equation (DE)[3] expressed commonly by

$$(H - E)\psi = 0, \tag{1}$$

where identical elementary particles involved must satisfy the fermion or boson exchange requirements. Electrons must satisfy the antisymmetry relation expressed by

$$P\psi = (-)^P \psi, \tag{2}$$

where P is a permutation operator. The molecular Hamiltonian of the SE is given by

a)Author to whom correspondence should be addressed: h.nakatsuji@qcri.or.jp

$$H = -\sum_i \frac{1}{2}\Delta_i - \sum_A \frac{1}{2}\Delta_A - \sum_{i,A} \frac{Z_A}{r_{iA}} + \sum_{i<j} \frac{1}{r_{ij}}$$
$$+ \sum_{A<B} \frac{Z_A Z_B}{R_{AB}}. \tag{3}$$

The theory presented here is applicable to both non-Born-Oppenheimer (BO) and BO cases,[4] and in the BO case, the last term is dealt with as a constant. In this paper, we assume the BO case.

If we can solve the SE exactly, elementary chemical phenomena should be predictable in high accuracy. This was a dream of many theoretical chemists. Heitler and London[5] applied quantum theory to the hydrogen molecule, and their theory became an origin of the valence bond (VB) theory.[6] Hylleraas tried to solve the SE of the helium atom considering the r_{12} coordinate explicitly as early as 1928, and highly accurate results were reported.[7] His studies became the origin of the so-called Hylleraas-CI.[7–10] Calais and Löwdin[11] reported a general method of calculating integrals including r_{12}. One of the authors considered a method of solving the Schrödinger equation projected on the density matrix space[12] with some interesting applications.[13,14] Explicitly correlated R12 and F12 methods have been actively developed and applied to many systems with high accuracy.[15–19] The explicitly correlated Gaussian method[20,21] utilized the mathematical merits of the special functions including the inter-electron coordinates r_{ij} explicitly and was used for the systems of basic importance. The full-CI extrapolation method gave highly accurate energies and potential curves of diatomic molecules without using r_{ij}.[22–24] Quantum Monte Carlo methods have a long history of calculating highly accurate wave functions including r_{ij} without using their difficult integrals and have been applied to many different subjects of chemistry.[25–27]

We studied the structure of the exact wave function in a series of papers[28–37] and developed the wave function theories that have the exact structure, namely, the theories that include the exact wave function within their variational space. Then, by applying the variational principle or the equivalent, we can get the exact wave function of the SE automatically. When the theory is given in an expanded form, we have to terminate at some order, which determines the accuracy of the result. When we solve the SE, we aim to obtain chemical accuracy, kcal/mol accuracy in the absolute value of the total energy.

It is easy to prove that by performing the recursive formula

$$\psi_{n+1} = [1 + C_n(H - E_n)]\psi_n \tag{4}$$

we can reach the exact solution of the SE, if the Hamiltonian H does not include diverging potentials. For harmonic oscillators, for instance, we can obtain their exact solutions in high accuracy. However, if the Hamiltonian includes the Coulombic potential like in Eq. (3), we have a divergence difficulty in solving Eq. (4) with the variational principle. To overcome this difficulty, the inverse SE[33] was first introduced and then the scaled SE defined by

$$g(H - E)\psi = 0 \tag{5}$$

was introduced.[34] Here, g is the scaling operator defined by

$$g = \sum_{i,A} r_{iA} + \sum_{i>j} r_{ij}, \tag{6}$$

which is always positive and so the scaled SE is equivalent to the original SE. The g operator was introduced to avoid the divergence due to the collisions between i and A and between i and j. When the electron i belongs to a special atom A, as in the local VB form, it can collide only with A and with j belonging to the same atom A so that the number of the terms in Eq. (6) reduces. This is a merit when a local description is possible. With the MO formalism, an electron delocalizes all over the molecule so that this merit does not arise. Then, it is easy to prove that by performing the modified recursive formula

$$\psi_{n+1} = [1 + C_n g(H - E_n)]\psi_n,$$
$$\psi^{exact} = \prod_n [1 + C_n g(H - E_n)] \cdot \psi_0, \tag{7}$$

starting from some initial function ψ_0, one gets the exact solution of the SE without meeting the divergence difficulty. When one performs several times the recursion formula (7) starting from an initial function ψ_0, the right-hand side becomes a sum of the analytical functions including C_n. If we collect all independent and non-diverging analytical functions from them as $\{\phi_I\}$, then the exact wave function is written as

$$\psi^{exact} = \sum_I c_I \phi_I. \tag{8}$$

In this equation, all independent analytical functions $\{\phi_I\}$ are given the independent coefficients $\{c_I\}$. So, Eq. (8) is more efficient than Eq. (7) for calculating the solution of the SE. This theory was initially referred to as free ICI [iterative complement (configuration) interaction] theory. But later, since it is not an iterative theory, we renamed it to the free complement (FC) theory. The functions $\{\phi_I\}$ are referred to as complement functions (cf's) since they are the elements of the complete solution of the SE. Thus, with the FC theory, the exact wave function is obtained by applying the FC theory to some initial function ψ_0, which is written as

$$\psi^{exact} = FC(\psi_0). \tag{9}$$

An important feature of the FC theory is that the exact wave function produced by the FC theory has the mathematical structure of the initial function ψ_0. Therefore, it is important to start from the initial function that is useful from a chemical point of view. We note here that an account of the FC theory and an earlier review were published[37] and that a similar formulation to solving the SE was also valid for solving the Dirac equation.[38]

In chemistry, the most important working theory lies in molecular structural formulas and chemical reaction formulas, which are collectively referred to in this paper as the chemical formula. Figure 1 illustrates some examples. The essence is the locality and transferability concept of atoms in molecules that has a long history from Dalton. Because of its basic importance in chemistry and biochemistry, we must learn some implications from the chemical formula.

Chemical Formula
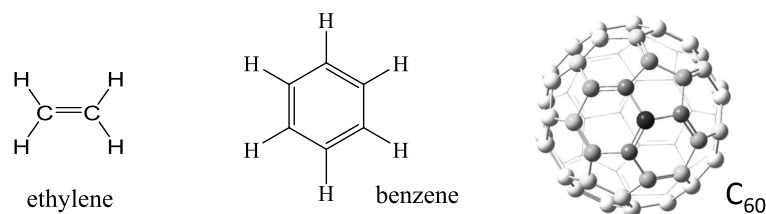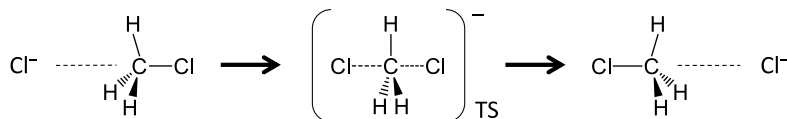
Molecular structural formula

FIG. 1. Example of the chemical formula.

Chemical reaction formula

In this paper, we will propose a general electronic structure theory that covers from approximate to exact levels in a unified way. We start from atoms, extend to molecules, and want to further expand to large and giant molecules. For this purpose, the first step will be to build an electronic structure theory that reflects faithfully the chemical formulas. We named this theory chemical formula theory (CFT). Then, we apply the FC theory to the wave functions produced by the CFT, which leads to the exact wave function that has the structure of the chemical formula. This theory is named free-complement chemical-formula theory (FC-CFT). We then investigate the possibility of the variational theory that lies in the intermediate of the CFT and the FC-CFT. Some applications of the theory presented in this paper will be given in Paper II[39] and the separate papers.

When systems become larger, some simplifications occur in their electronic structure. The largest reason lies in the fading away of the Fermion antisymmetry rule between distant electron pairs. The basis of this phenomenon has been described before by the *inter*-exchange (*i*Exg) theory.[40] For distant electron pairs, the exchange interactions fade away and the Coulomb interactions survive, resulting large simplifications in their electronic structures. In giant systems, these Coulombic interactions are further simplified to the electrostatic interactions even in the exact sense. The CFT for large and giant systems must describe these changes properly.

II. CHEMICAL FORMULA

Chemistry is the science dealing with atoms, molecules, and molecular systems including the biological ones. There, the most intuitive working theory is represented by molecular structural formulas and chemical reaction formulas which are collectively referred to as *chemical formulas*. Figure 1 shows some molecules and a reaction as examples of chemical formulas. The structural information given by the X-ray spectroscopy for biological systems is also included in the chemical formula. There, the basic concept is the locality and transferability of

atoms in molecules, while for bonds, some ambiguities and freedom are left, though they are also essentially local and transferable. The basic implication of the chemical formula is that molecules are formed from sparsely arranged atoms kept by bonds and that the electronic structure of each atomic region is essentially similar to that of the isolated atom. Hydrocarbons, for example, consist of carbons and hydrogens connected by the C—C and C—H bonds, but for the C—C bond, it may be single, double, triple, or even the median of them, and the nature of the CH bond is affected by them.

The locality in the chemical formula originates from the Coulombic potential included in the Hamiltonian given by Eq. (3). Near the nucleus, the electrons are captured in the exponential orbitals around the nuclei, like an electron of the hydrogen atom captured in the orbital $\psi = e^{-r}$. Therefore, the probability of finding an electron decays exponentially as r increases. The inter-electron repulsion term, the fourth term in the Hamiltonian, works to shield the effects of other nuclei so that the locality as expressed by "atoms in a molecule" is realized.

Chemists use chemical formulas in the molecular design and chemical synthesis, as if they had "seen" their molecules and reactions, and with adding their inspirations and knowledges accumulated through chemical studies, they finally synthesize successfully complex molecules and materials. Chemical formulas are the central tool for chemists to describe their ideas, i.e., their chemistry. When one refers to the modern text books of organic chemistry[41] and chemical biology,[42] one will find that almost all basic and central concepts are written using chemical formulas. Without chemical formulas, chemistry and biochemistry have no way of expressing their sciences. Thus, chemical formulas have long been the best chemical theory ever existed and used by chemists as their conceptual bases of the chemistry. Chemical formulas represent the essence of chemistry. Based on the ultimate importance of chemical formulas in the world of chemistry, this author postulates that as far as the SE is the principle of chemistry, its solutions for chemistry should have the structure that can be related to the chemical formulas.

Since the FC theory produces, as Eq. (9) implies, the exact wave function that has the structure similar to the initial wave function ψ_0, we want to formulate a general electronic structure theory that simulates faithfully the structure of the chemical formula and the origins thereof from a quantum chemical point of view. We refer to this theory as chemical formula theory (CFT).

III. CHEMICAL FORMULA THEORY AS A GENERAL ELECTRONIC STRUCTURE THEORY OF ATOMS AND MOLECULES

Since locality and transferability are central to the chemical formula, we use the local atomic concept from the beginning of the CFT. Another central implication of the chemical formula is that the molecule is formed from atoms whose three-dimensional arrangement is indicated in the formula. However, the chemical formula does not describe the chemical bonds and the interactions involved there in detail. Therefore, they must be described by the quantum-chemical formulations of CFT and the variational principle. In the bond formation process from atoms to molecule, not only the ground state but also the excited states of atoms are cooperatively transformed to the ground and excited states of the molecules. For example, when C—C and C—H bonds are formed from carbon and hydrogen, the ground state and many excited states of the carbon atom of $2s^2 2p^2$ and $2s2p^3$ configurations participate. The participation of the excited state from the $2s2p^3$ configuration is usually more important than that of the ground state of $2s^2 2p^2$ configuration. Therefore, CFT must describe the chemical changes from the ground and excited states of the constituent atoms to the ground and excited states of the molecule. This is an important feature of the CFT. The detailed information of the atomic energy levels is given in Moore's book, Atomic Energy Levels, and NIST Atomic Spectra Database including updated data.[43] They give basic information for building the CFT. For molecules, such levels of information are impossible because of the numerous numbers of molecules. For this reason, when we deal with large molecules, we still start from atoms and synthesize large molecules from individual atoms.

Let us consider some organic molecules. The ordinary σ-bonds that constitute their molecular framework would be described well with the local bonds represented by $(ab+ba)\alpha\beta$ and $(a + b)^2\alpha\beta$, which are VB and MO type, respectively. For the π-bonds, the VB and MO descriptions are different. The VB description is essentially local, like Kekulé and Dewar structures of the benzene molecule, but the MO description is delocalized over $6p_\pi$ orbitals. In the present theory, we choose the physically and chemically best description, or we can choose even both, if it is effective, since our theory is based on the non-orthogonal cf's. We further want to extend our theory to large and giant molecular systems. An example of a large system is the fullerene shown in Fig. 1. For the σ-framework, the local bond description as given above would be appropriate. For the π-framework, again the MO and VB descriptions exist, but the MO description may be easier to use. This 60 π description is still local, like 6 π of benzene. An interesting example of the giant system is a carbon

nanotube. For its σ-framework, it is a strong material, and from its π-framework, high electric conductivity along the tube is realized. The description of the σ-framework would be the same as above, but for the π-framework, a delocalized picture may become necessary. Thus, in the CFT, the descriptions of the bonds and the inter-atomic interactions are flexible. For large and giant systems, the CFT can utilize the simplicities shown in Sec. VI like the electrostatic model which can be exact, though it is used frequently in molecular dynamics (MD) simulations.[44,45] Finally, we note that in the FC calculations, the local description is computationally more feasible than the delocalized one.

If we define the exact wave function of a molecule in its Ith electronic state as ψ_I, then the ground and excited states of this molecule satisfy the following important relation:

$$\begin{aligned} \langle\psi_I|\psi_J\rangle &= \delta_{IJ}, \\ \langle\psi_I|H|\psi_J\rangle &= E_I\delta_{IJ}, \end{aligned} \tag{10}$$

where the Hamiltonian is given by Eq. (3) and E_I is the energy of the Ith electronic state of this molecule. Since we use the variational principle in CFT, the relation similar to Eq. (10) holds also for the CFT-level of the results. Since a motivation of CFT is to construct intuitively useful initial functions for the FC theory, it needs not to be highly accurate. Rather, it must be conceptually simple and useful. Accuracy can be improved later by applying the FC theory.

A. CFT for atoms

First, we consider the CFT for atoms. This phrase may sound strange because an atom is simply a dot in the chemical formula. In CFT, however, to this dot, all the electronic states of the atom must be associated. When we refer to the Atomic Energy Levels of Moore,[43] we find a detailed list of energy levels for each atom of the periodic table. Behind a "dot" of the chemical formula, plenty of atomic electronic structures are hidden. We want to describe these atomic electronic structures with CFT. Furthermore, for atoms, the explanation of CFT is simpler than that for molecules.

If we define the exact wave function of the electronic state I of each atom A by ψ_I^A, then they must satisfy the orthogonality and Hamiltonian-orthogonality relations as given by Eq. (10) with adding the superscript A on ψ_I, E_I, and H, where H^A is the atomic Hamiltonian of atom A given by

$$H^A = -\sum_i \frac{1}{2}\Delta_i - \sum_{i,A}\frac{Z_A}{r_{iA}} + \sum_{i<j}\frac{1}{r_{ij}}, \tag{11}$$

where the nucleus A is fixed at origin, which is a dot A of the chemical formula.

When only the ground state is of interest, one may limit the calculations to the ground state alone. Then, the ground-state wave function $\psi_G^{A'}$ may be expanded as $\psi_G^{A'} = \sum_i c_i\varphi_{i,G}^A$ (here, the prime indicates that it is tentative). The first complement function (cf) $\varphi_{0,G}^A$ is the best possible approximation to ψ_G^A, and other terms $\varphi_{i,G}^A$ ($i \neq 0$) are added to improve it. (We use the term "complement function," which was used in the FC theory, again here since the role of this function is essentially the same as before. It is abbreviated as cf, but note that it is different from CF.) In CFT, however, we want to describe both ground and excited states at the same time.

Equation (10) implies that when direct interactions between ground and excited states are included, both ground and excited states are improved efficiently. Suppose that there are M such excited states of the same symmetry within some energy range and that we have an approximate cf $\phi_{0,m}^A$ for each state of $m = 0 \ldots M$. Then, we make these approximate ground and excited states interact with each other as

$$\psi_I^{A,CFT} = \sum_{m=0}^M c_m^I \phi_{0,m}^A \quad (I = 0, \cdots, M) \quad (12)$$

and obtain, by applying the variational principle, the best possible expression of each state. The resultant states satisfy

$$\begin{aligned}\langle \psi_I^{A,CFT} | \psi_J^{A,CFT} \rangle &= \delta_{IJ}, \\ \langle \psi_I^{A,CFT} | H^A | \psi_J^{A,CFT} \rangle &= E_I^{A,CFT} \delta_{IJ}, \end{aligned} \quad (13)$$

which is an approximate relation corresponding to the exact relation given by Eq. (10).

The expression given by Eq. (12) corresponds to the "minimal" approximation for each state since only one cf $\phi_{0,m}^A$ was chosen for each electronic state m. Then, the first index of the subscript is 0. Since the minimal approximation is practically important, we recommend optimizing variationally not only the coefficients c_m but also the non-linear parameters like orbital exponents included in the cf $\phi_{0,m}^A$. If this minimal approximation is insufficient, we may introduce double, triple, ... approximations. But, the accuracy can be improved later by using the FC theory, which will be more efficient and systematic. The FC theory produces from $\phi_{0,m}^A$ (used as an initial function) the complete cf's $\{\phi_{i,m}^A\}$ $(i = 0, \ldots, L_m)$ for each mth electronic state of the atom A, as will be explained in Sec. IV A.

We next explain how to construct the function $\phi_{0,m}^A$. Suppose we are dealing with the beryllium atom, A = Be. It has four electrons, and the ground state is composed of two 1s and two 2s electrons as

$$\phi_{0,G}^{Be} = A[1s(1)1s(2)\alpha(1)\beta(2) \cdot 2s(3)2s(4)\alpha(3)\beta(4)], \quad (14)$$

where we use G for the ground state. Referring to NIST Atomic Spectra Database, we found that the next higher valence-like ^{1}S state over the ground state is the $2p^2$ state lying as high as at 76 190 cm^{-1}. This state is written similarly as

$$\begin{aligned}\phi_{0,2p^2}^{Be} = A[&1s(1)1s(2)\alpha\beta \cdot (2p_x(3)2p_x(4) + 2p_y(3)2p_y(4) \\ &+ 2p_z(3)2p_z(4))\alpha\beta], \end{aligned} \quad (15)$$

where the electron assignment on the spin function was abbreviated. Within this energy range, there are many Rydberg excited ^{1}S states at 54 677(3s), 65 245(4s), 69 322(5s), 71 321(6s), etc., in cm^{-1}. They are 2s → ns (n = 3, 4, 5, 6, ...) states and are written as

$$\phi_{0,2s\to ns}^{Be} = A[1s(1)1s(2)\alpha\beta \cdot (2s(3)ns(4) + ns(3)2s(4))\alpha\beta].$$
$$(16)$$

We truncate ns to 6s because higher Rydberg states scarcely interact with the valence electronic states. Therefore, the

number M is here 5. The atomic orbitals, ns and 2p, are given in a minimal Slater form by

$$\begin{aligned}1s &= \exp(-\alpha_{1s}r), \\ 2s &= (b_0 + r)\exp(-\alpha_{2s}r), \\ 2p_x &= x\exp(-\alpha_{2p}r), \quad 2p_y = y\exp(-\alpha_{2p}r), \\ &\quad 2p_z = z\exp(-\alpha_{2p}r), \\ 3s &= (b_0 + b_1 r + r^2)\exp(-\alpha_{3s}r), \\ 4s &= (b_0 + b_1 r + b_2 r^2 + r^3)\exp(-\alpha_{4s}r), \\ 5s &= (b_0 + b_1 r + b_2 r^2 + b_3 r^3 + r^4)\exp(-\alpha_{5s}r), \\ 6s &= (b_0 + b_1 r + b_2 r^2 + b_3 r^3 + b_4 r^4 + r^5)\exp(-\alpha_{6s}r).\end{aligned} \quad (17)$$

The coefficients b_i's and the exponents α's may be found in the literature or may be optimized by doing, for example, the atomic Hartree-Fock calculations.

Finally, our CFT wave function corresponding to Eq. (12) is given by

$$\begin{aligned}\psi_m^{Be,CFT} &= c_0 \phi_{0,G}^{Be} + c_1 \phi_{0,2p^2}^{Be} + c_2 \phi_{0,3s}^{Be} + c_3 \phi_{0,4s}^{Be} \\ &\quad + c_4 \phi_{0,5s}^{Be} + c_5 \phi_{0,6s}^{Be}.\end{aligned} \quad (18)$$

By applying the variational principle to this wave function, we obtain 6 solutions that satisfy Eq. (13). They correspond to the approximations to the ground state, 3s to 6s Rydberg states, and $2p^2$ valence excited states in energy order. A merit of CFT is that it gives the wave functions not only for the ground state but also for the excited states included in the energy range of the calculations. We can calculate therefrom the valence and Rydberg excitation energies and their wave functions. From them, we can calculate various properties of these states.

We have three notes here. First, in the above cf's, the inner core was written in a doubly occupied closed-shell form. Here, instead, we can introduce in-out correlation form, replacing 1s1s $\alpha\beta$ by (1s1s'+1s'1s) $\alpha\beta$, where 1s and 1s' have different exponents. The in-out form takes the in-out correlation effect into account and so can be advantageous. Second, in the above formulation, the Slater functions in Eq. (18) were minimal. Though the minimal approximation is a rather poor approximation in the modern quantum chemistry, the experiences from the applications of the FC theory show that the minimal Slater approximation is acceptable at least for small organic molecules: we could calculate chemical accuracy results with this approximation. Third, in the above formulation for atoms, the number of the atomic excited states having the same symmetry as the ground state, M, is usually small. A reason of including the ^{1}S($2p^2$) state lying as high as 76 190 cm^{-1} was that it is the lowest valence ^{1}S excited state. However, for molecules, this situation changes much.

Furthermore, a note is on the cationic and anionic states of atoms. As will become clear, for constructing the CFT of molecules, we need not only the neutral atomic states but also cationic and anionic states. For the cationic states, we have detailed energy levels in Moore's book,[43] but for the anionic states, only a few (stable) levels exist for atoms, which are available through the Internet.

B. CFT for molecules from CFT for atoms

For studying the molecular electronic structure with CFT, we need to know all lower electronic states of different symmetries for all atoms involved in the molecule under study. In molecules, the symmetry becomes lower so that the mixing of different atomic states is facilitated. Here again, Moore's and NIST tables of atomic energy level are very helpful for planning the CFT studies.

Let us now construct electronic structures of molecules with CFT. The procedure may be summarized as follows.

1. Put all constituent atoms at their positions of the chemical formula.
2. Examine the atomic cf's $\phi_{0,m}^{A}$ to the energy level of interest, E_{EL}, for all constituent atoms and prepare them. The non-linear parameters in the atomic cf's are assumed to be already optimized. The energy level, E_{EL}, is important since it decides the accuracy and the necessary amounts of computations. Moore's book[43] is very useful here as will be seen below for carbons.
3. This step is the "from atomic states to molecular states" step, which is central in the CFT. We produce molecular cf's from atomic cf's by making bonds with nearby atoms using free electrons of the ground & excited states of the constituent atoms. As will be given explicitly in the examples below, the two-center σ bond is expressed with the covalent term and the cationic-anionic ionic term. The local MO form is also useful near the equilibrium distance. When the system includes the π-framework, like benzene, the valence bond "resonance" form and/or the delocalized MO form are useful. The former has a local structure, but the latter does not. The paired electrons in the valence orbital of atoms may become a lone pair in a molecule. Different combinations of free valence electrons give different sets of bonds and lone pairs that correspond to different sets of molecular cf's. We designate the cf's of the molecule as $\phi_{0,I}^{\mathrm{Mol}}$. The suffix I represents a bonding form used in this cf, and the number of them is defined as M, similarly to the atomic case. The pre-suffix 0 represents that it is the initial function of the FC theory. This pre-suffix will be used later in the FC theory.
4. By a linear combination of all molecular cf's $\{\phi_{0,I}^{\mathrm{Mol}}\}$ prepared above, the CFT wave function of the target molecule Mol is described as

$$\psi_K^{\mathrm{Mol,CFT}} = \sum_I^M c_I^K \phi_{0,I}^{\mathrm{Mol}}. \tag{19}$$

5. Apply the variational principle to Eq. (19) and obtain CFT wave functions and energies of all the ground and excited states of the target molecule. After applying the variational principle, the suffix K denotes the ground and excited states of the molecule. They satisfy very important orthogonality and Hamiltonian-orthogonality relations like Eq. (10).

From the above procedure, some important features of CFT become evident. In CFT, both ground and excited states adjust themselves through their interactions, as expressed by Eq. (10). We understand from this equation that the stabilization energy relative to the atomic energy is distributed among the bond energies and that the electronic excitations occur with the weakening of the bonds involved. These concepts are useful for studying photo-chemical reactions.

Now let us apply the above procedures to the hydrocarbon molecule C_nH_m as an example. First, we put carbon and hydrogen atoms at the positions of the chemical formula. From the transferability, all n carbons are the same and the m hydrogens are the same. Second, we prepare the cf's of lower electronic states of the constituent atoms and investigate them. The hydrogen atom has a very discrete electronic structure. The ground state is the valence 1s state, and the degenerate 2s and 2p states are too high (82 259 cm^{-1}) and diffuse to form chemical bonds with other atoms. So, we need to consider solely the 1s state for hydrogen atoms. On the other hand, the carbon atom has many valence atomic states. They are, from lower with the energy in cm^{-1} in parentheses, the ^{3}P(0), ^{1}D(10193), and ^{1}S(21648) states of the $2s^22p^2$ configuration and the ^{5}S(33735), ^{3}D(64090), and ^{3}P(75256) states of the $2s2p^3$ configuration. There are no other valence states below the H(2s) state at 82 259 cm^{-1}.[43] The Rydberg states start from 60 333(^{3}P) to 61 982(^{1}P) cm^{-1}, which are 3s states, and the 3p Rydberg states exist from 68 856 to 73 976 cm^{-1}.[43] For the reactions and properties near the ground state of carbons, the ^{3}P, ^{1}D, and ^{1}S states of the $2s^22p^2$ configuration and the lowest ^{5}S state of the $2s2p^3$ configuration are important because they are relatively low (<34 000 cm^{-1}) in energy. Therefore, for ordinary chemistry of the ground state alone, the E_{EL} defined above may be around 34 000 cm^{-1}. The lowest ^{3}P($2s^22p^2$) state has two free valences and can form two bonds with the neighboring atoms, while the ^{5}S($2s2p^3$) state has four free valences and can form four bonds. Therefore, the ^{5}S($2s2p^3$) state can get very large stabilization by forming four bonds with the neighboring atoms, the total stabilization energy exceeding its high energy level (33 735 cm^{-1}) relative to the ground ^{3}P($2s^22p^2$) state. Therefore, this ^{5}S($2s2p^3$) state plays very important roles in carbon chemistry, which are sometimes more important than the roles of the ground ^{3}P state of the s^2p^2 configuration. The existence of these many interesting electronic states of carbons may be the origin of the variety of carbon chemistry and further the origin of life produced mainly by organic molecules and systems.

For the next step, we prepare the cf's for the 1s state of hydrogen as

$$\phi_{0,1s}^{\mathrm{H}} = 1s^{\mathrm{H}}\alpha \tag{20}$$

and for the ^{3}P($2s^22p^2$) and ^{5}S($2s2p^3$) states of carbon as

$$\phi_{0,^3\mathrm{P}(s^2p^2)}^{\mathrm{C}} = A[(1s^{\mathrm{C}})^2\alpha\beta \cdot (2s^{\mathrm{C}})^2\alpha\beta \cdot 2p_x^{\mathrm{C}}2p_y^{\mathrm{C}}\alpha\alpha] \tag{21}$$

and

$$\phi_{0,^5\mathrm{S}(sp^3)}^{\mathrm{C}} = A[(1s^{\mathrm{C}})^2\alpha\beta \cdot 2s^{\mathrm{C}}2p_x^{\mathrm{C}}2p_y^{\mathrm{C}}2p_z^{\mathrm{C}}\alpha\alpha\alpha\alpha], \tag{22}$$

respectively. Here, we omitted the electron numbers for brevity. For molecular calculations, the atomic state cf of the ^{5}S($2s2p^3$) state is often rewritten in the hybrid forms $\sigma_1^{\mathrm{C}}\sigma_2^{\mathrm{C}}2p_y^{\mathrm{C}}2p_z^{\mathrm{C}}$, $\sigma_1^{\mathrm{C}}\sigma_2^{\mathrm{C}}\sigma_3^{\mathrm{C}}p_z^{\mathrm{C}}$, and $\sigma_1^{\mathrm{C}}\sigma_2^{\mathrm{C}}\sigma_3^{\mathrm{C}}\sigma_4^{\mathrm{C}}$, depending on the sp, sp^2, and sp^3 hybrids, respectively. They are not strictly the atomic states, but useful for molecular calculations.

Now, in the third stage, we prepare molecular cf's from atomic cf's. This is the most important "from-atoms-to-molecule" step of CFT. In this step, we make bonds using free valence orbitals of atomic cf's. For the H_2 molecule, the cf of CFT is written as

$$\phi_{0,G1}^{H_2} = A[(1s^{H_1}(1)1s^{H_2}(2) + 1s^{H_2}(1)1s^{H_1}(2))\alpha(1)\beta(2)]$$
$$= A[(1s^{H_1}(1)1s^{H_2}(2))(\alpha(1)\beta(2) + \beta(1)\alpha(2))], \quad (23)$$

which is obtained by making a singlet coupling from the atomic cf $\phi_{0,1s}^{H}$ given by Eq. (20). This is the simplest Heitler-London function for the hydrogen molecule. Another different form of bond formation is possible, which is a local MO form given by

$$\phi_{0,G2}^{H_2} = A[(1s^{H_1} + 1s^{H_2})(1)(1s^{H_1} + 1s^{H_2})(2)\alpha(1)\beta(2)]. \quad (24)$$

This MO form has a well-known deficiency that it breaks off at the dissociation limit. So, we use this local MO form only when we study the H_2 molecule near its equilibrium geometry. There, we often use both cf's of the covalent VB and MO forms as our cf's of CFT: this corresponds to taking two-electron full CI into account. The so-called ionic contribution is included in the cf given by Eq. (24). We usually use the covalent VB and MO forms simultaneously for the local bond near the equilibrium distance, when we want to calculate the exact solution of the SE in chemical accuracy. We may consider the ionic terms separately. They are given by

$$\phi_{0,I1}^{H_2} = A[1s^{H_1}(1)1s^{H_1}(2)\alpha(1)\beta(2)],$$
$$\phi_{0,I2}^{H_2} = A[1s^{H_2}(1)1s^{H_2}(2)\alpha(1)\beta(2)]. \quad (25)$$

For the ground state of the CH molecule, whose symmetry is $^2\Pi$, the main cf is prepared from the ground $^3P(2s^2 2p^2)$ state of carbon $\phi_{0,^3P(s^2p^2)}^{C}$ given by Eq. (21) and the function $\phi_{0,1s}^{H}$ given by Eq. (20) as

$$\phi_{0,G}^{CH} = A[(1s^C)^2\alpha\beta \cdot (2s^C)^2\alpha\beta$$
$$\cdot (2p_x^C 1s^H + 1s^H 2p_x^C)\alpha\beta \cdot 2p_y^C\alpha] \quad (26)$$

by forming the CH σ bond with the singlet coupling of the $2p_x^C$ and $1s^H$ electrons. For the CH bond, the ionic terms $(2p_x^C)^2\alpha\beta$ and $(1s^H)^2\alpha\beta$ are also considered. Instead, the local MO bond form $(2p_x^C + 1s^H)^2\alpha\beta$ is convenient when we study only near the equilibrium distance. They give different independent cf's. The $^5S(2s2p^3)$ state does not contribute to this ground state by spin-symmetry, but the $^3P(75256)$ state of the $2s2p^3$ configuration contributes, but slightly because this state is much higher in energy than the $^3P(2s^2 2p^2)$ state. For general hydrocarbons, there are no symmetry restrictions. Therefore, both atomic states, $^3P(2s^2 2p^2)$ and $^5S(2s2p^3)$, given by Eqs. (21) and (22), respectively, play important roles, which will be discussed in the succeeding papers. The Rydberg 3s state of carbon $^3P(2s^2 2p3s)$ at 60 333 cm^{-1} shows some

role for the higher potential curves of the CH molecule around this energy level.

The C_2 molecule is an interesting example. It is a closed-shell molecule, and its bonding is still a subject of discussions.[24,46–48] Important contributions to the ground state come from the $\sigma_1^C \sigma_2^C 2p_y^C 2p_z^C$ expression of the $^5S(2s2p^3)$ state and from the $^3P(2s^2 2p^2)$ state given by Eq. (21). Before giving their expressions, let us define the spin part of the cf's used in this paper by

$$\alpha\beta\alpha\beta\cdots\alpha\beta\alpha\alpha\cdots\alpha, \quad (27)$$

as far as this spin function is possible for the chemistry under consideration. From the $\sigma_1^C \sigma_2^C 2p_y^C 2p_z^C$ expression of the $^5S(2s2p^3)$ state of the two carbons, C_1 and C_2, we obtain quadruple bonds

$$\phi_{0,G}^{CC} = A\{(1s^{C_1})^2(1s^{C_2})^2(\sigma_1^{C_1}\sigma_1^{C_2} + \sigma_1^{C_2}\sigma_1^{C_1})$$
$$\times (\sigma_2^{C_1}\sigma_2^{C_2} + \sigma_2^{C_2}\sigma_2^{C_1}) \cdot (p_y^{C_1}p_y^{C_2} + p_y^{C_2}p_y^{C_1})$$
$$\times (p_z^{C_1}p_z^{C_2} + p_z^{C_2}p_z^{C_1})\}, \quad (28)$$

where we have omitted the spin function of the form given by Eq. (27) and the electron numbers. We assumed the C_2 axis to be on x. From the $^3P(2s^2 2p^2)$ state of the two carbons, we obtain double π bonds

$$\phi_{0,G1}^{CC} = A\{(1s^{C_1})^2(1s^{C_2})^2(2s^{C_1})^2(2s^{C_2})^2(p_y^{C_1}p_y^{C_2} + p_y^{C_2}p_y^{C_1})$$
$$\times (p_z^{C_1}p_z^{C_2} + p_z^{C_2}p_z^{C_1})\} \quad (29)$$

and double σ and π bonds

$$\phi_{0,G2}^{CC} = A\{(1s^{C_1})^2(1s^{C_2})^2(2s^{C_1})^2(2s^{C_2})^2$$
$$\times [(p_x^{C_1}p_x^{C_2} + p_x^{C_2}p_x^{C_1})(p_y^{C_1}p_y^{C_2} + p_y^{C_2}p_y^{C_1})$$
$$+ (p_x^{C_1}p_x^{C_2} + p_x^{C_2}p_x^{C_1})(p_z^{C_1}p_z^{C_2} + p_z^{C_2}p_z^{C_1})]\}. \quad (30)$$

Equations (28)–(30) represent the most important cf's of the ground state of the C_2 molecule. The strength of the C—C bond of C_2 depends on the extent of the mixing of these two kinds of cf's. At infinite separation, the contribution of Eq. (28) becomes zero since the ground state of the carbon atom is the $^3P(2s^2 2p^2)$ state. The contribution of the cf given by Eq. (28) increases as the two carbons come closer. In the above, we omitted the cf's representing the contributions of the ionic terms, but now, it will be easily written down. When one studies only near the equilibrium geometry, the local bond forms are useful for both σ and π bonds. The detailed study will be reported in a separate paper.

The above examples were given only for some of the valence atomic states. There are many other states particularly for the states including carbon. The Rydberg atomic states lie at rather high energy levels, and their orbital exponents are diffuse in comparison with the valence states. Therefore, their contributions to the lower electronic states are limited.

Next is the fourth stage. We describe our CFT wave function by a linear combination of the molecular cf's, as given by Eq. (19). Note that the CFT is not a theory for the ground state alone, but it describes also the excited states of the target molecule. Usually, as the energy becomes higher, the accuracy decreases because we limit the energy range of the atomic cf's to be within some range.

The final fifth stage is the harvesting stage: the variational determination of the ground and excited states. The ground state is the most tightly bound state in which the stabilization energy relative to a sum of the atomic separate states is used to strengthen the chemical bonds within the molecule. The higher excited states consist of the chemical bonds whose sum of the bond energies is less than that of the ground state: some bonds are weaken in some way, which might be interested in the reactions of excited states.

The CFT is a general electronic structure theory for atoms and molecules that try to construct the electronic structures of the target system as naturally as possible like in the chemical formulas. Atomic and molecular CFT describes both ground and excited states lying in the energy range of interest. Molecular CFT wave functions are produced by making chemical bonds from the ground and excited states of the constituent atoms. For the reacting system, the states of interest along the reaction coordinate are calculated referring to the chemical reaction formula. These bond formation steps are probably the most important steps in CFT, and we use both VB and MO knowledge. The locality concept of VB is common to CFT and useful for reducing the number of the cf's. For most σ-framework, this would be valid. At the same time, if the delocalized nature is the essence of the bond, we must use it since by doing so the convergence to the exact solution of the SE will be accelerated. These CFT wave functions are used to study molecular structural properties and the potential energy surfaces along the reaction. The central concept of CFT is the locality concept of the electronic structure, its transferability, and from-atoms-to-molecule concept. The central principle of CFT is the variational principle so that the resultant wave functions satisfy the expressions, as given by Eq. (10). As in the chemical formula, the locality concept for bonds has still some ambiguity, which might be considered as an aspect of the diversity which is also an essential nature of chemistry and biology.

A motivation of CFT was to construct the wave function theory that mimics the chemical formula. Then, by applying the FC theory to the initial functions obtained with the CFT, we will be able to get the exact solution of the SE easier than starting from other initial functions. In the following section, we apply the FC theory to the initial functions obtained by the CFT. This level of theory is referred to as FC-CFT (free-complement chemical-formula theory).

IV. FREE COMPLEMENT THEORY APPLIED TO THE CFT WAVE FUCTION: FC-CFT

For the importance of the chemical formula in chemistry and biochemistry, the CFT would be a useful initial theory for solving the SE. For this purpose, we use the CFT as

$$\psi_{\text{exact}}^{\text{FC-CFT}} = FC(\psi_0^{\text{CFT}}), \tag{31}$$

where we use the FC theory as if it is an operator transforming the CFT function to the exact wave function. In this section, we describe in some detail about this operation. One important note is given initially. The CFT wave function is obtained by applying the variational principle to the cf's of CFT, as explained with Eq. (12) for atoms and with

Eq. (19) for molecules. We use them as the initial function of the FC theory. However, since the FC theory is linear on the cf's, the operation is done to the elementary cf's $\{\phi_{0,m}^A\}$ of Eq. (12) for atoms and $\{\phi_{0,I}^{\text{Mol}}\}$ of Eq. (19) for molecules. In other words, the operation of the FC theory is done to these elementary cf's, not to the optimized CFT wave function, because their coefficients are reoptimized again by the FC theory.

A. Atoms

For atoms, the cf's of the FC theory are produced using the recursion formula given by Eq. (7) with the scaling function given by Eq. (6). More simply, they can be obtained by multiplying the elements of the g operator to ϕ_0. Then, the cf's $\{\phi_{l,m}^A\}$ of atoms are written as

$$\{\phi_{l_m,m}^A\} = \phi_{0,m}^A, \; r_i\phi_{0,m}^A, \; r_{ij}\phi_{0,m}^A, \; r_j r_i\phi_{0,m}^A, \; r_k r_{ij}\phi_{0,m}^A,$$
$$r_{kl} r_{ij}\phi_{0,m}^A, \; \dots, \tag{32}$$

where $i, j, \dots$ denote the electrons and r_i is actually r_{iA} with A denoting the nucleus. We define L_m to be the number of the cf's produced from the initial cf $\phi_{0,m}^A$ and l_m to be the index running from 0 (initial cf) to L_m. With these definitions, the FC wave function for the Ith state of the atom A is written as

$$\psi_I^{A,\text{FC-CFT}} = \sum_m^M \sum_{l_m}^{L_m} c_{l_m,m}^I \phi_{l_m,m}^A \quad (I, m = 0, \dots, M). \tag{33}$$

The number of the cf's included in this expression, N_{CF}^A, becomes

$$N_{\text{CF}}^A = \sum_{m=0}^M L_m. \tag{34}$$

Thus, the SE of the atom A is solved for all the electronic states. In comparison with the CFT equation given by Eq. (12), the FC theory produced the cf's denoted by the suffix l_m, and with these functions, each atomic state becomes exact, interacting with other atomic states. Then, the results satisfy

$$\left\langle \psi_I^{A,\text{FC-CFT}} \middle| \psi_J^{A,\text{FC-CFT}} \right\rangle = \delta_{IJ},$$
$$\left\langle \psi_I^{A,\text{FC-CFT}} \middle| H \middle| \psi_J^{A,\text{FC-CFT}} \right\rangle = E_I^{A,\text{FC-CFT}} \delta_{IJ}, \tag{35}$$

which is essentially the same quality equation, as given by Eq. (10).

The above formulation shows the motivation of the CFT for the use as the initial functions of the FC theory. We compare Eqs. (12) and (13) for the CFT with Eqs. (33) and (35) for the FC-CFT. In both formulations, we use each atomic state as a basis for improving the set of the atomic states ($I, m = 0, \dots, M$). This originates from our experience of studying the ground and excited states with the SAC-CI (symmetry-adapted-cluster configuration-interaction) formalism.[49,50] Rather than using some arbitrary basis function set, it would be more efficient if we use both ground and excited states of the system as their interacting basis. Then, each state becomes better with the interactions among them and finally becomes exact, as the FC-CFT result. Looking at the chemical formula, we imagined the formulation to be as close as possible as the nature does and formulated the CFT and FC-CFT. The CFT was an initial stage of this formulation and the

FC-CFT is final: there are also interesting possibilities in the intermediate stage, as will be explained later.

For obtaining the chemical accuracy, it is necessary to consider to the second order, but we usually include r_i^3 and r_{ij}^3, but neglect the second-order terms of r_{ij} when electrons i and j belong to different shells. The term $r_{kl}r_{ij}$ is included only when all four electrons are different. However, depending on each atom, we modify this rule based on our experiences.

B. Molecules

The locality of the electronic structures of the CFT brings much simplicity to the calculations of the exact solution of the SE of molecules. First, we show the simplicity produced to the g-operator of the FC theory. As shown by Eq. (5), the g-operator was introduced to avoid the collisions between charged particles in a molecule that causes the divergence difficulty of the integrals in the variational solution of the SE. It was originally defined by Eq. (6). However, for the locality, the g-operator is much simplified. The electron belonging to atom A, i_A, has almost no chance to collide with the other nuclei B. Likewise, i_A and i_B do not have a chance to collide with each other. Taking these facts into account, the g-operator to be applied to the molecular initial functions of CFT is proposed as

$$g = \sum_{A,i_A} (r_{i_AA} + x_{i_AA} + y_{i_AA} + z_{i_AA}) + \sum_{A,i_A>j_A} r_{i_Aj_A}$$
$$+ \sum_{A,i_A,B}^{\text{A-B,bond}} r_{i_AB} + \sum_{i_A,j_B}^{\text{A-B,bond}} r_{i_Aj_B}, \qquad (36)$$

where the first two terms are the ordinary terms and the last two terms are sometimes added. Because molecules are asymmetric, we added x_{i_AA}, et cetera in addition to the spherical term r_{i_AA}. They work to polarize the orbital. The last two terms are exceptional since they are two-center r_{i_AB} and $r_{i_Aj_B}$ terms for the A-B bond. They might be unnecessary because they are essentially proportional to the AB bond distance and therefore might cause a redundancy. So, we consider them only when the hydrogen atom is involved in the bond since the bond distance is usually small.

However, when we use MO theory, the above simplifications are not realized: since all electrons are delocalized over all atoms in a molecule, we cannot assign an electron i to an atom A so that the number of the cf's increases. In Sec. III, we have considered the cases where the delocalized MO description might be better or easier than the local VB-type description. However, such cases occur usually for the system which has a large π-electron framework. Most σ-electron framework is still described by the local two electron bonds. We should choose the description as local as possible when we prepare the initial functions of the FC theory.

For molecules, the method of constructing the molecular cf's $\{\phi_{0,I}^{\text{Mol}}\}$ of the CFT was explained in Sec. III B, and the examples were given for some diatomic molecules. Here we use another example of the acetylene molecule, which is a linear closed shell molecule

$$H - C \equiv C - H.$$
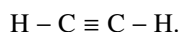
The principal part of the initial function for the ground state may be written as

$$\phi_{0,G}^{\text{Acetylene}} = \hat{A}\,[(1s^{C_1})^2(1s^{C_2})^2(\sigma(h_1C_1))^2(\sigma(C_1C_2))^2$$
$$\times (\pi_1(C_1C_2))^2(\pi_2(C_1C_2))^2(\sigma(C_2h_2))^2]. \qquad (37)$$

The first two are inner 1s cores, and the followings are the σ and π local bonds. Each of the local σ and π bonds in Eq. (37) may include covalent, ionic, and local MO forms, giving independent cf's. Next, we apply the g function given by Eq. (36). For chemical accuracy, we produce the molecular cf's mostly to order two, though some third-order terms are added when they are important. In addition to this standard general rule, we use the following rules to reduce the computational labors. Most of them are the results of the compromise between accuracy and costs.

1. When we study molecules in their equilibrium structure, the local bonds in the initial functions like the one given by Eq. (37) may be written in the local MO form: the reason is that when we use the sampling method of the local SE method,[51,52] the local MO form is twice faster than the covalent form. Then, each local bond is replaced, one by one order, with the covalent form and also with the ionic terms, if necessary.
2. When the functions r_{iA}, x_{iA}, and r_{ij} are applied to the electrons belonging to the local MO bond, the two-center local MO is changed to the covalent VB form to make it match with the local treatment.
3. For the 2s orbital, we add $\exp(-\alpha_{2S}r)$ at the first order when we use a nodeless 2s orbital in the initial function.
4. For the second order function $r_{i_Aj_A}r_{k_Al_A}$, we include only when all four suffixes are different. The function $r_{i_Aj_A}r_{j_Ak_A}$ may also be included, when it is thought not to be negligible.
5. For the two center terms for the bond between atoms A and B, like r_{i_AB} and $r_{i_Aj_B}$, originating from the last two terms of the g function given by Eq. (36), only the first order term is considered when A is the hydrogen atom.
6. Similarly, only up to the first order is considered for the inter-pair terms: between 1s and 2s pair, between 1s and bond pair, etc.
7. The diagonal terms are considered up to the third order, like r_{i_AA}, $r_{i_AA}^2$, $r_{i_AA}^3$ and $r_{i_Aj_A}$, $r_{i_Aj_A}^2$, $r_{i_Aj_A}^3$, when they are thought to be important.

When we have several initial cf's, they are grouped in order, according to the order of their "parent" initial function. For the first-order cf's, the g function is applied only to the first order. All cf's generated with these rules are again local functions.

The cf's thus produced by the FC theory for the molecule, Mol, may be written in a general form as

$$\phi_{l_I,I}^{\text{Mol}}(1,2,\ldots,N) = A[\{r_{ij}^{m_{ij}^{l_I,I}}\} \cdot u_1^{l_I,I}(1)u_2^{l_I,I}(2)\cdots u_N^{l_I,I}(N)] \qquad (38)$$

including the initial functions, where the ith electron function $u_i^{l_I,I}(i)$ is obtained from the initial functions by multiplying the functions like r_{iA}, x_{iA}, and/or r_{ij}. In Eq. (38), the suffix I of l_I is due to the fact that the number of the cf's produced from the

initial function $\phi_{0,I}$, which is defined as L_I, is generally different for each initial function I. Note that when $\phi_{l_I,I}$ includes the two-electron terms r_{ij}, the functions $u_i^{l_I,I}(i)$ and $u_j^{l_I,I}(j)$ are connected by r_{ij}, where usually the electrons i and j belong to the same atom because only in such a case the electrons i and j can collide. $m_{ij}^{l_I I}$ is 0, 1, 2, or 3 for the ij pair of the cf $l_I I$.

With these cf's, the FC wave function for the Kth state of the molecule is written as

$$\psi_K^{\text{Mol,FC-CFT}} = \sum_I^M \sum_{l_I}^{L_I} c_{l_I,I}^K \phi_{l_I,I}^{\text{Mol}} \quad (K,I = 0,\dots,M). \quad (39)$$

Thus, the SE of the molecule is solved for the ground and excited states. The number of the cf's included in this expression, $N_{\text{cf}}^{\text{Mol}}$, is given by

$$N_{\text{cf}}^{\text{Mol}} = \sum_{I=0}^M L_I. \quad (40)$$

In comparison with the CFT equation given by Eq. (19), the FC theory produced the functions denoted by the suffix l_I, and with these additional cf's, the ground and lower excited states become exact, interacting with the other states at the same time. Then, the results satisfy

$$\begin{aligned} \left\langle \psi_K^{\text{Mol,FC-CFT}} \middle| \psi_L^{\text{Mol,FC-CFT}} \right\rangle &= \delta_{IJ}, \\ \left\langle \psi_K^{\text{Mol,FC-CFT}} \middle| H \middle| \psi_L^{\text{Mol,FC-CFT}} \right\rangle &= E_K^{\text{Mol,FC-CFT}} \delta_{KL}, \end{aligned} \quad (41)$$

which is important showing the quality of the ground and excited states calculated.

Very often, the calculations are limited only to the ground state of a molecule. In such a case, the calculation can be simplified. The number of the cf's given by Eq. (40) can be reduced. There are two choices. One is to limit the calculations to the ground state from the beginning. Then, we calculate only the ground state, starting from the most probable initial function. Usually, we know from the experience that what atomic state is most important for molecular formation. For example, for carbon, we use the sp^3 configuration rather than the s^2p^2 configuration, though the former is the excited state, when the symmetry of the molecule permits this choice. Then, we construct the initial wave functions using the idea of the CFT and apply the FC theory to this initial function. Even in this case, we can get the exact wave function and the energy with the FC-CFT because the FC theory leads to the exact wave function from any initial function, if it has an overlap with the exact wave function, though we may need a larger number of cf's than the other cases we explain below. This is a simple way of performing the FC-CFT: the locality concept and other implications of the chemical formula are included in the initial wave function. Actually, for oxygen, for example, only the s^2p^4 configuration lies below 73 000 cm^{-1} so that this choice and the other choices may be the same.

The second choice is, for the carbon case, to choose the atomic states arising from both ground s^2p^2 and excited sp^3 configurations for making bonds with other atoms. Then, we construct all cf's at the CFT level and apply the FC theory to them to proceed the FC-CFT. The diagonalization at the CFT level is optional, if only the FC-CFT is the purpose, though we recommend doing so, because we can get some understanding

and information at the CFT level. If we need only the ground state, we may apply the FC theory only to those cf's that are important in the ground-state wave function at the CFT level to reduce the labor of the calculations.

Like the initial functions, the higher-order cf's of the FC-CFT also have local and transferable structures. Therefore, the final FC wave functions as the solutions of the SE have the local and transferable structures. This is an important feature of the present theory to obtain not only accurate but also understandable results. Then, the following quantities

$$\psi_{\text{FC-CFT}}^{\text{exact}} - \psi_0^{\text{CFT}}, \quad (42)$$

$$\Delta\Gamma_{\text{FC-CFT}}^{\text{exact}} = \Gamma_{\text{FC-CFT}}^{\text{exact}}(1,2,\cdots,N) - \Gamma_{\text{CFT}}(1,2,\cdots,N), \quad (43)$$

and

$$\Delta\rho_{\text{FC-CFT}}^{\text{exact}} = \rho_{\text{FC-CFT}}^{\text{exact}}(1) - \rho_{\text{CFT}}(1) \quad (44)$$

are of special interest for chemists, where Γ and ρ are the N-electron density matrix and electron density, respectively. These difference quantities are the modifications that the SE has given over the CFT quantities. The FC-CF theory is based on the non-orthogonal set of functions so that the formulations of Löwdin[53] would be useful for analyzing the quantities given by Eqs. (43) and (44). Since the SE is the governing principle of chemistry and biology, these quantities might give what we wanted to know. We want to learn from these quantities what are necessary to make quantum chemical calculations useful in the research of chemistry. The exact concept obtained with the FC-CFT would be useful since it is reliable. In the chemical formula, the descriptions of bonds and interactions are rather vague in comparison with the atomic concepts. Therefore, we want to learn from the results of the present theory about their exact nature and concepts for the system of interest or under the course of chemical reaction.

V. POSSIBILITY OF THE INTERMEDIATE THEORY: FC-CFT-V

The cf's $\{\phi_I\}$ produced by the FC-CF theory include both integratable and non-integratable functions. We group them into those integratable, which is called the i-set, and those non-integratable, which is called the n-set. The i-set plus n-set forms the complete set. The n-set cf's are characterized by the existence of the odd-powers of r_{ij} terms that makes them non-integratable. However, these cf's characterized by the r_{ij} terms may not be so important for chemical researches, though they work to lower the energy. For example, both MO and VB theories do not include r_{ij} terms, but they are very important for studying chemistry.

This suggests a possibility of the intermediate theory. We consider only the i-set cf's at some order of the FC theory and apply the variational principle to them. We refer to this theory as FC-CFT variational (FC-CFT-V), where "V" indicates variational. The accuracy of the FC-CFT-V can be improved by raising the order of the FC theory and by expanding the realm of the integral method. Because of its variational nature, we can study extensively the chemistry of the system of interest. Furthermore, when we start from the CFT, we generate the cf's not only for the ground states but also for the excited states. These ground and excited states satisfy the orthogonality and Hamiltonian orthogonality. This relation is important for the

study of the ground and excited states. The quantities defined by Eqs. (42)–(44) are useful to investigate the chemistry under study.

The FC-CFT-V is an advanced theory of the CFT. Since the FC theory always improves the theory, FC-CFT-V is always more accurate than the CFT itself. When integrations are possible for all cf's, the FC-CFT-V is just FC-CFT and highly accurate results are obtained. Such cases were already reported for small systems like the He atom[54–56] and H_2 molecule.[57] For H_2^+, highly accurate BO and non-BO results were reported.[58,59] In particular, highly accurate analytical potential curves of H_2^+ were derived from the non-BO wave function.[59]

Then, we propose the two-step approach. After studying the basic chemistry with the FC-CFT-V theory, we study only the most important subject by considering the effects of the n-set cf's by performing FC-CFT. Then, we can extend the accuracy to the exact level and we can predict and perform chemistry with confidence. For this purpose, we may use the local Schrödinger equation (LSE) method.[51,52] It is interesting to formulate the LSE method to be useful at this stage of the FC theory.

VI. POSSIBILITIES OF THE CFT AND FC-CFT FOR LARGE AND GIANT MOLECULES

When we deal with large and giant molecules, the local nature of the electronic structure of the CFT and the FC-CFT leads to much simplifications. This arises from the nature of the antisymmetry requirement for the electron pairs in molecules. The *inter*-exchange (*i*Exg) theory[40] tells us that as the distance of the electron pair increases, the antisymmetry rule between them fades away. As the system becomes larger, the number of the distant electron pairs in the molecule increases rapidly so that much simplifications result since antisymmetrizations are usually very time-consuming processes. Here, we discuss this important effect in large and giant molecules.

A. *i*Exg theory for antisymmetrization

Electrons are fermions, and therefore the electronic wave functions must satisfy the antisymmetry requirement, as expressed by Eq. (2). We have proposed an efficient antisymmetry theory, called Nk theory,[60] which is useful for correlated theory including r_{ij} explicitly. This algorithm is determinant-based, and its order is roughly N^{3-4}. Therefore, for large molecules, the antisymmetrization of electrons becomes a time-consuming step, and with the Nk theory alone, it is difficult to apply theories to very large molecules. To overcome this situation, we have proposed the *inter*-exchange (*i*Exg) theory of antisymmetrization[40] based on the locality of the molecular electronic structure. We briefly explain just the essence of the *i*Exg theory for applying it to large and giant molecules.

For explaining the *i*Exg theory, let us consider a diatomic molecule AB. The antisymmetrizer A_{AB} for the molecule AB can be rewritten as

$$A_{AB} = A_A A_B (1 + E_{AB}^{(1)} + E_{AB}^{(2)} + \cdots + E_{AB}^{(k)} + \cdots + E_{AB}^{(K)}), \quad (45)$$

where $E_{AB}^{(k)}$ is the k-electron exchange operator between A and B, K is the maximum number of electrons belonging to either A or B, and A_A and A_B are the antisymmetrizers for the electrons belonging to the atoms A and B, respectively. Note that each electron belongs to the orbital of an exponential form that belongs to either A or B. When we apply the one-electron exchange operator $E_{AB}^{(1)}$ to the molecular wave function ψ_{AB}, the magnitude of $E_{AB}^{(1)} \psi_{AB}$ becomes *exponentially smaller* than the magnitude of ψ_{AB} itself since by the exchange, the electron-nucleus distances r_{iA} and r_{jB} become longer by about the bond distance R. Then, as the distance R increases, the result of the exchange becomes smaller exponentially and finally becomes negligibly small, when the following Eq. (46) is satisfied. Such a situation can be measured by using the overlap T_{ab} between the squares of orbitals χ_a and χ_b to which the electron pair belongs,

$$T_{ab} = \left\langle \chi_a^2 \middle| \chi_b^2 \right\rangle \leq \lambda. \quad (46)$$

We use T_{ab} instead of S_{ab}^2 to avoid the case of $S_{ab} = 0$ by symmetry. For the pair of valence electrons of carbon atoms, λ is about 10^{-6} and the distance R_λ corresponds to the four bonds of hydrocarbons. For the k electron exchange $E_{AB}^{(k)}$, this decrease is k times faster than that of $E_{AB}^{(1)}$. In large molecules, there are many atomic pairs for which $E_{AB}^{(n)} \psi$ is negligibly small: then all terms of $k \geq n$ are negligible. Furthermore, the operation counts for $A_A A_B$ also reduce from $N_A! \times N_B!$ to $N_A! + N_B!$ and further to $N_A^3 + N_B^3$ with the use of the Nk theory. Thus, the labor of antisymmetrizations can be reduced much with a combinatorial use of the *i*Exg and Nk theories.

Now let us consider what the *i*Exg theory implies when we apply the FC theory to large molecules and even to giant molecular systems.

B. Simplification in large molecules

As molecules become larger, the number of electrons increases, but at the same time, the number of distant atomic pairs in the molecule also increases. With the use of the expression of the cf given by Eq. (38), the matrix elements between different cf's are written as

$$\langle \phi_I | H | \phi_J \rangle = \sum_{i,j=1}^{N} \left\langle u_i^I(1) \middle| -\frac{1}{2}\Delta_1 - \sum_A \frac{Z_A}{r_{A1}} \middle| u_j^J(1) \right\rangle D(I_i | J_j)$$
$$+ \sum_{\substack{i<j \\ k<l}}^{N} \left[\left\langle u_i^I(1) u_j^I(2) \middle| \frac{1}{r_{12}} \middle| u_k^J(1) u_l^J(2) \right\rangle \right.$$
$$\left. - \left\langle u_i^I(1) u_j^I(2) \middle| \frac{1}{r_{12}} \middle| u_l^J(1) u_k^J(2) \right\rangle \right] D(I_i I_j | J_k J_l),$$
$$(47)$$

where the minors $D(I_i | J_j)$ and $D(I_i I_j | J_k J_l)$ are defined as usual.[53] Here and below, the upper suffix l, I of $u_i^{l,I}$ is denoted by I alone for simplicity. In the above formula, we did not consider the effect of r_{ij} included in some cf's, but such cf's do not affect the following discussions because r_{ij} is actually $r_{i_A j_A}$, as explained before: r_{ij} connects electrons only within the same atom.

With this equation, let us consider the behaviors of the integrals when the orbitals involved are so distant that the exchange of the electrons between these orbitals is negligible as the iExg theory implies. For the one-electron part, when the orbitals u_i^I and u_j^J are far apart, the integral is negligibly small since each orbital is local and centered on a nucleus and decays exponentially from the center of the orbital, while the kinetic and nuclear attraction operators have little effect on these behaviors of the integrals. For the two-electron part, when all four orbitals, u_i^I, u_j^I, u_k^J, *and* u_l^J, are far apart to each other, the integral is negligible. When three of them are far apart, again the integral is negligible. When two of them, (i, k) and (j, l), in parentheses are close, but (i, l) and (j, k) are far distant, the first term remains but the second term is negligible. This case occurs when u_i^I and u_k^J belong to one atom and u_j^I and u_l^J belong to other different atom and the two atoms are far apart in the molecule. Similarly, when one is far apart though three are close, again both integrals are negligible. Both integrals survive only when all orbitals are close to each other. Similar behaviors are also seen for the overlap matrix elements involved in the co-factor matrices $D(I_i|J_j)$ and $D(I_iI_j|J_kJ_l)$. The above observations clearly imply that much simplification occurs when the molecule is very large: we can expect that roughly order-N behavior of the computational labor would result.

We apply the above argument to the interaction of the two atoms A and D that lie far apart from each other in a large molecule. We assume that both A and D have only one electron, and the distance between A and D is so far apart that the electron exchange between A and D does not occur in chemical accuracy. Then, the direct interaction term between A and D is written as

$$\langle \psi|H|\psi\rangle_{A,D} = \frac{Z^A Z^D}{R_{AD}} - \int \frac{Z^D}{r_{1_A D}} \rho^A(\mathbf{r}_{1_A}) d\mathbf{r}_{1_A}$$
$$- \int \frac{Z^A}{r_{1_D A}} \rho^D(\mathbf{r}_{1_D}) d\mathbf{r}_{1_D}$$
$$+ \int \rho^A(\mathbf{r}_{1_A}) \frac{1}{r_{1_A 2_D}} \rho^D(\mathbf{r}_{2_D}) d\mathbf{r}_{1_A} d\mathbf{r}_{2_D}, \quad (48)$$

where $\mathbf{r}_{1_A}$ represents the coordinate of an electron belonging to atom A and $\rho^A(\mathbf{r}_{1_A})$ is the electron density distribution associated with atom A,

$$\rho^A(\mathbf{r}_{1_A}) = \int \psi(\mathbf{r}_{1_A}\mathbf{r}_{2_B}\cdots\mathbf{r}_{N_G}) \psi(\mathbf{r}_{1_A}\mathbf{r}_{2_B}\cdots\mathbf{r}_{N_G}) d\mathbf{r}_{2_B}\cdots d\mathbf{r}_{N_G}.$$
$$(49)$$

The first term of Eq. (48) represents the nuclear repulsion, the second and third terms represent the attractions between the electron density of one atom with the nucleus of another atom, and the last term represents the repulsion between the electron density of the atoms A and D. Since the electron exchange between A and D is negligible, only the classical Coulomb interactions remain, and they sum up to the above simple formula.

We note that $\mathbf{r}_{1_A}$ in Eq. (48) denotes the coordinate of the electron that belongs to the atom A, but due to the overlap and antisymmetry principle, the electron delocalizes out of the atom A, but the delocalization limit is within the limiting

distance R_λ related to the parameter λ given by Eq. (46). To eliminate the overlapping effect between A and D, the AD distance necessary for Eq. (46) is roughly $2R_\lambda$. For a carbon pair, it is about 8 bond separation, ~24 a.u. (12.7 Å).

C. Further simplification in giant molecules

When we move to still larger systems, like giant molecular systems, further simplification occurs. There, the distance of the AD pair can be much larger than the above case. When the AD distance is so large that the widths of the orbitals a and d are negligible, then the width of the orbital d looks like a point when it is seen from the orbital a. In such a situation, the electron density distribution ρ^A looks like an electronic charge q^A on atom A, where

$$q^A = \int \rho^A(\mathbf{r}_{1_A}) d\mathbf{r}_{1_A} \quad (50)$$

and Eq. (48) is transformed to

$$\langle \psi|H|\psi\rangle_{A,D} = \frac{\delta^A \delta^D}{R_{AD}}, \quad (51)$$

where δ^A is the gross charge of atom A, $\delta^A = Z^A - q^A$.

Such simple electrostatic forms have been introduced in many molecular dynamics (MD) simulation programs,[44,45] though their gross charges are empirical and/or approximate. They were also used in the quantum mechanical models of giant molecular systems like the elongation method for polymers,[61] the divide and conquer method,[62,63] and the fragment molecular orbital method.[64] It is remarkable that this expression is valid even in the exact theory for solving the SE, which is the present FC theory, when it is combined with the iExg theory. Thus, the exact FC theory transforms its theoretical structure automatically to a lower-order N type theory for large molecules and even to the theory of giant molecular systems as the system size increases further. This transformation of the theoretical structure of the FC theory is ideal as a chemical theory. Intuitively, this is a natural transformation of the chemical theory that reflects the chemical formulas central to chemistry. The basis is the locality and the transferability of chemistry combined with the iExg theory.

We note that the quantum-mechanical theories for giant molecular systems, like the elongation method, divide-and-conquer method, and fragment MO method, cited above[61–64] are not the theories one obtains when the MO theory is applied to giant molecular systems, but these theories are the original products of the authors invented based on the MO theory. Furthermore, the correlation theories over Hartree-Fock are difficult to apply to giant molecular systems because integral transformations necessary for the MO-based correlation theories are too time-consuming for giant molecular systems.

The CFT for large and giant molecular systems must describe the above simplifications well. Though this is the subject of our separate paper, the FC-CFT based on such concepts must be designed to work to order N as the system size increases to giant systems. Fine coding will be necessary to fully utilize the simplifications in large and giant molecular systems.

VII. CONCLUDING REMARKS

We aimed in this paper to combine two major principles of chemical science, Schrödinger equation and chemical formulas, to produce useful predictive theory in chemistry. Chemical formulas are used daily by chemists and biochemists as working concepts and the sources of their ideas. Therefore, if the wave function of the exact theory can have the structure of the chemical formula, it must be very useful for chemists. The FC theory has a merit that it produces the exact wave function that has the structure assumed to the initial function used at the beginning. Therefore, in this paper, we first proposed the CFT that gives the wave function whose structure reflects that of the chemical formula. Then, starting from CFT, we can produce, by FC-CFT, the exact wave function that has the structure similar to the chemical formula by using the FC theory. The difference of the results of the CFT and the FC-CFT is due to the exact SE itself. So, by analyzing the difference, the chemists may get deep understanding that assists their researches. Thus, the CFT is designed from the beginning to be used by the FC theory as its initial functions so that it affects the descriptions of the exact wave function.

In CFT, our basic idea is that the molecular electronic states are produced from atomic electronic structures. The detailed knowledge of the atomic electronic structures is summarized in Moore's book of Atomic Energy Levels,[43] which provides us useful basis. The basic ideas on the natures of the interactions between atoms are included in chemical formulas. It is not decisive, but suggestive and must be described well by the methods of quantum chemistry. We use all real electronic states, ground and excited states, and ionic and anion states of atoms as if they are the basis functions for constructing the molecular ground and excited states. Atoms react with the nearby ones using free valences with forming bonds for stabilization. The governing principle is the variational principle. We use both VB and MO concepts, but the basic and dominant local structures in chemical formulas are close to the VB concept. The delocalized π-framework may be described well with the MO concept. The resultant electronic states produced by CFT satisfy the orthogonality and the Hamiltonian orthogonality between the different states.

This result of CFT is elevated to the exact level by the FC theory, which is referred to as FC-CFT. When we study both ground and excited states of the molecule with CFT, all of them are made exact by the FC theory. The CFT is designed to be used by the FC theory as its initial function so that this connection is smooth and natural: the important relation, the orthogonality and the Hamiltonian orthogonality expressed by Eq. (10), is satisfied at all levels of CFT, FC-CFT, and FC-CFT-V, and only the first indices l_m and l_I, respectively, for atoms and molecules are used to represent the cf's produces by the FC theory.

The cf's of the FC-CFT consist of the integratable and non-integratable functions, the latter characterized by the inclusion of the r_{ij} terms, which may not be important for chemical studies, as seen from the MO and VB theories which also do not include r_{ij} terms. So, even if we use only the integratable functions, the chemistry of the system under study would be described reasonably well. This is the FC-CFT-V

theory and is an advanced variational theory of the CFT. For the efficiency of the variational method, we can study chemistry extensively. Afterwards, if it is necessary to go to the exact SE level, we add the non-integratable cf's, perform FC-CFT for both ground and excited states, if necessary, and obtain the exact-level of the results. The chemical formula concepts are valid in all levels of the theory.

The examinations of the locality-based concepts for the possibility toward very large and even giant molecular systems gave some light for the future. Because of the locality-based structure of the present theory, the *inter*-exchange (*i*Exg) theory of antisymmetrization introduced previously could produce large simplifications in the treatment of large systems. There, the exchanges between distant electron pairs fade away and Coulombic interactions become dominant. For still larger giant systems, even the electrostatic descriptions become possible. Then, we have to develop the CFT for such systems considering the simplifications investigated in this paper. The code of the FC-CFT for exactly solving the SE must be designed to work to essentially order N for large and giant molecular systems.

In applications of the present theory, the automated intelligence (AI) technology would be useful to simplify the inputs, to connect different interfaces of the FC theory, and to clarify the implications of the calculated results, given in chemical formula language, obtained by the various levels of the present theories. Furthermore, the feedbacks from the experimentalists about the merits and the faults of our theory will be useful for further development of our methodology. Such collaborations may be a key for the future progress of the theory and algorithms.

ACKNOWLEDGMENTS

[1] P. A. M. Dirac, Proc. R. Soc. London, Ser. A **123**, 714 (1929).
[2] E. Schrödinger, Ann. Phys. **384**, 361 (1926); **384**, 489 (1926); **384**, 734 (1926); **385**, 437 (1926); **386**, 109 (1926); Phys. Rev. **28**, 1049 (1926).
[3] P. A. M. Dirac, Proc. R. Soc. London, Ser. A **117**, 610 (1928).
[4] M. Born and R. Oppeheimer, Ann. Phys. **389**, 457 (1927).
[5] W. Heitler and F. London, Z. Phys. **44**, 455 (1927).
[6] S. S. Shaik and P. C. Hiberty, *A Chemist's Guide to Valence Bond Theory* (John Wiley & Sons, New Jersey, 2008).
[7] E. A. Hylleraas, Z. Phys. **48**, 469 (1928); **54**, 347 (1929); **65**, 209 (1930); **71**, 739 (1931).
[8] J. S. Sims and S. Hagstrom, Phys. Rev. A **4**, 908 (1971).
[9] G. W. F. Drake and Z.-C. Yan, Chem. Phys. Lett. **229**, 486 (1994).
[10] M. B. Ruiz, J. Math. Chem. **46**, 1322 (2009).
[11] J.-L. Calais and P.-O. Löwdin, J. Mol. Spectrosc. **8**, 203 (1962).
[12] H. Nakatsuji, Phys. Rev. A **14**, 41 (1976).
[13] H. Nakatsuji and K. Yasuda, Phys. Rev. Lett. **76**, 1039 (1996).
[14] M. Nakata, H. Nakatsuji, M. Ehara, M. Fukuda, K. Nakata, and K. Fujisawa, J. Chem. Phys. **114**, 8282 (2001).

[15] W. Kutzelnigg, Theor. Chim. Acta **68**, 445 (1985); W. Kutzelnigg and W. Klopper, J. Chem. Phys. **94**, 1985 (1991).

[16] J. Noga, W. Kutzelnigg, and W. Klopper, Chem. Phys. Lett. **199**, 497 (1992).

[17] J. Rychlewski, *Explicitly Correlated Wave Functions in Chemistry and Physics—Theory and Applications* (Kluwer Academic, Dordrecht, 2003).

[18] S. Ten-No, Chem. Phys. Lett. **398**, 56 (2004).

[19] W. Klopper, F. R. Manby, S. Ten-No, and E. F. Valeev, Int. Rev. Phys. Chem. **25**, 427 (2006).

[20] J. Mitroy, S. Bubin, W. Horiuchi, Y. Suzuki, L. Adamowicz, W. Cencek, K. Szalewicz, J. Komasa, D. Blume, and K. Varga, Rev. Mod. Phys. **85**, 693 (2013).

[21] S. Bubin, M. Pavanello, W. Tung, K. L. Sharkey, and L. Adamowicz, Chem. Rev. **113**, 36 (2013).

[22] L. Bytautas and K. Ruedenberg, J. Chem. Phys. **122**, 154110 (2005).

[23] L. Bytautas and K. Ruedenberg, J. Chem. Phys. **124**, 174304 (2006).

[24] J. S. Boschen, D. Theis, K. Ruedenberg, and T. L. Windus, Theor. Chem. Acc. **133**, 1425 (2014).

[25] B. L. Hammond, W. A. Lester, Jr., and P. J. Reynolds, *Monte Carlo Methods in Ab Initio Quantum Chemistry* (World Scientific, Singapore, 1994).

[26] M. P. Nightingale and C. J. Umrigar, *Quantum Monte Carlo Methods in Physics and Chemistry* (Springer, 1998).

[27] J. B. Anderson, *Quantum Monte Carlo: Origins, Development, Applications* (Oxford University Press, 2007).

[28] H. Nakatsuji, J. Chem. Phys. **113**, 2949 (2000).

[29] H. Nakatsuji and E. R. Davidson, J. Chem. Phys. **115**, 2000 (2001).

[30] H. Nakatsuji, J. Chem. Phys. **115**, 2465 (2001).

[31] H. Nakatsuji, J. Chem. Phys. **116**, 1811 (2002).

[32] H. Nakatsuji and M. Ehara, J. Chem. Phys. **117**, 9 (2002).

[33] H. Nakatsuji, Phys. Rev. **65**, 052122 (2002).

[34] H. Nakatsuji, Phys. Rev. Lett. **93**, 030403 (2004).

[35] H. Nakatsuji, Phys. Rev. A **72**, 062110 (2005).

[36] H. Nakatsuji and M. Ehara, J. Chem. Phys. **122**, 194108 (2005).

[37] H. Nakatsuji, Acc. Chem. Res. **45**, 1480 (2012).

[38] H. Nakatsuji and H. Nakashima, Phys. Rev. Lett. **95**, 050407 (2005).

[39] H. Nakatsuji, H. Nakashima, and Y. I. Kurokawa, J. Chem. Phys. **149**, 114106 (2018).

[40] H. Nakatsuji and H. Nakashima, J. Chem. Phys. **142**, 194101 (2015).

[41] P. Y. Bruice, *Organic Chemistry: Global Edition* (Pearson, 2016).

[42] D. L. Nelson and M. M. Cox, *Lehninger Principles of Biochemistry*, 7th ed. (W. H. Freeman & Co, 2017).

[43] C. E. Moore, "Atomic energy levels," Report No. NSRDS-NBS 35, Vol. I–III, 1971; for updated data, see A. Kramida, Yu. Ralchenko, J. Reader, and NIST ASD Team, NIST Atomic Spectra Database (ver. 5.5.6) (National Institute of Standards and Technology, Gaithersburg, MD, 2018), available online at https://physics.nist.gov/asd.

[44] M. Karplus, Annu. Rev. Biophys. Biomol. Struct. **35**, 1 (2006).

[45] R. Salomon-Ferrer, D. A. Case, and R. C. Walker, Wiley Interdiscip. Rev.: Comput. Mol. Sci. **3**, 198–210 (2013).

[46] S. Shaik, D. Danovich, W. Wu, P. Su, H. S. Rzepa, and P. C. Hiberty, Nat. Chem. **4**, 195 (2012).

[47] D. Wilian Oliveira de Sousa and M. A. Chaer Nascimento, J. Chem. Theory Comput. **12**, 2234 (2016).

[48] M. Hermann and G. Frenking, Chem. Eur. J. **22**, 4100 (2016).

[49] H. Nakatsuji, Chem. Phys. Lett. **59**, 362 (1978); **67**, 329 (1979); **67**, 334 (1979).

[50] H. Nakatsuji, "SAC-CI method: Theoretical aspects and some recent topics," in *Computational Chemistry—Reviews of Current Trends*, edited by J. Leszczynski (World Scientific, Singapore, 1997), Vol. 2, pp. 62–124.

[51] H. Nakatsuji, H. Nakashima, Y. I. Kurokawa, and A. Ishikawa, Phys. Rev. Lett. **99**, 240402 (2007).

[52] H. Nakatsuji and H. Nakashima, J. Chem. Phys. **142**, 084117 (2015).

[53] P.-O. Löwdin, Phys. Rev. **97**, 1474 (1955).

[54] H. Nakashima and H. Nakatsuji, J. Chem. Phys. **127**, 224104 (2007).

[55] Y. I. Kurokawa, H. Nakashima, and H. Nakatsuji, Phys. Chem. Chem. Phys. **10**, 4486 (2008).

[56] H. Nakashima and H. Nakatsuji, Phys. Rev. Lett. **101**, 240406 (2008).

[57] Y. I. Kurokawa, H. Nakashima, and H. Nakatsuji, Phys. Rev. A **72**, 062502 (2005).

[58] Y. Hijikata, H. Nakashima, and H. Nakatsuji, J. Chem. Phys. **130**, 024102 (2009).

[59] H. Nakashima and H. Nakatsuji, J. Chem. Phys. **139**, 074105 (2013).

[60] H. Nakashima and H. Nakatsuji, J. Chem. Phys. **139**, 044112 (2013).

[61] A. Imamura and Y. Aoki, Int. J. Quantum Chem. **32**, 137 (1987); Y. Aoki and F. L. Gu, Phys. Chem. Chem. Phys. **14**, 7640 (2012).

[62] W. Yang, J. Mol. Struct.: THEOCHEM **255**, 461 (1992); W. Yang and T. Lee, J. Chem. Phys. **103**, 5674 (1995).

[63] M. Kobayashi, Y. Imamura, and H. Nakai, J. Chem. Phys. **127**, 074103 (2007); T. Yoshikawa, M. Kobayashi, A. Fujii, and H. Nakai, J. Phys. Chem. B **117**, 5565 (2013).

[64] K. Kitaura, E. Ikeo, T. Asada, T. Nakano, and M. Uebayasi, Chem. Phys. Lett. **313**, 701 (1999); D. G. Fedorov, N. Asada, I. Nakanishi, and K. Kitaura, Acc. Chem. Res. **47**, 2846 (2014).