Perspective

Perspective on "Theory of self-consistent electron pairs. An iterative method for correlated many-electron wavefunctions"

Wilfried Meyer (1976) J Chem Phys 64: 2901

Hans-Joachim Werner

Institute for Theoretical Chemistry, University of Stuttgart, Pfaffenwaldring 55, D-70569 Stuttgart, Germany

Abstract. The title paper of Wilfried Meyer is a landmark for modern electron correlation theories and their efficient implementation. It described for the first time a matrix-formulated direct configuration interaction method. This approach eliminates all coupling coefficients involving external orbitals and the need for a "formula tape". Secondly, a full integral transformation is avoided, a prerequisite for integral-direct implementations. Third, the theory is formulated in a basis of nonorthogonal virtual orbitals (or atomic orbitals), which forms the basis of current local electron correlation treatments. Meyer's paper , which was written while I was a diploma student in his group, has strongly influenced my own work on multireference configuration interaction, coupled-cluster theory, and local electron correlation methods. The current paper reviews the history of the self-consistent electron pairs method and describes its relation to current electron correlation techniques.

Key words: Electron correlation – Configuration interaction – Direct configuration interaction – Self-consistent electron pair theories – Local correlation

The slow convergence of the dynamical correlation energy with respect to the number of Slater determinants or configuration state functions (CSFs) has been one of the major problems of quantum chemistry since the proposal of the configuration interaction (CI) method by Boys in 1950. It was realized early that at least the first-order interacting space of the Hartree–Fock (HF) wavefunctions, which is spanned by all single and double excitations from the HF reference function into virtual orbitals, is needed to recover a substantial part of the electron correlation energy. This led to the development of electron pair theories [1, 2] in the late 1960s and early 1970s. However, even this small subspace of the full CI space grows with the fourth power of the molecular size or the number of correlated electrons (assuming that the number of occupied and virtual orbitals grows linearly). Since, in principle, the Hamiltonian matrix must be constructed and diagonalized in the CSF basis, the applicability of the simplest and original version of the CI method was limited to very small systems. Early attempts to treat the correlation problem therefore introduced additional approximations, such as configuration selection by means of perturbational estimates of their importance. Alternatively, certain couplings were neglected, as for instance in the independent electron pair approximation (IEPA) [2] . By means of localizing the occupied orbitals and transforming the virtual orbitals to pair natural orbitals (PNOs), which are pairwise nonorthogonal, a dramatic reduction in the number of configurations could be achieved. In 1971 Meyer generalized the PNO-IEPA scheme to PNO-CI [1], which included all couplings between the different electron pairs. Moreover, he proposed the coupled-electron-pair approximation (CEPA) [1, 3], which accounts for the most important effects of higher excitations and is size-consistent. CEPA can be viewed as a good approximation to coupled-cluster theory. Using the PNO-CEPA method, Meyer, in 1973, was able to compute about 90% of the correlation energy for molecules such as H_2O and CH_4, and this was an important landmark in electronic structure theory. Due to the optimum convergence properties of the PNO orbital basis coupled with configuration selection, this could be achieved with only a few hundred configurations. The disadvantage of the PNO-CI/CEPA method is that different orbitals are needed for each electron pair, which limits the extension to larger systems. Furthermore, the perturbative determination of the PNOs and the subsequent orbital or configuration selection leads to small discontinuities on potential-energy surfaces. Despite the tremendous success of the PNO methods it was therefore desirable to develop methods which allowed the full space of single and double excitations to be included.

The first important step in this direction was made by Roos [4] in 1972 by proposing the direct CI method. He realized that the lowest eigenvectors and eigenvalues can

be computed iteratively without ever setting up the Hamiltonian matrix explicitly. Instead, it is sufficient to compute the vector $\boldsymbol{\sigma} = (\mathbf{H} - E)\mathbf{c}$, where $\mathbf{c}$ is a trial vector and $\boldsymbol{\sigma}$ a residual vector, from which an update to $\mathbf{c}$ can be obtained by first-order perturbation theory. The product $\mathbf{Hc}$ can be obtained directly from the two-electron integrals, the vector $\mathbf{c}$, and coupling coefficients, which depend on the formal structure of the CSFs. In the implementations by Roos and Siegbahn [5] symmetry-adapted orthonormal configuration state functions were used as N-electron bases. The integrals were fully transformed into the molecular orbital (MO) basis, and the construction of the $\boldsymbol{\sigma}$ vector was driven by the two-electron integrals. Depending on the interaction and integral type, about 140 different nonzero coupling coefficients were distinguished, resulting in a lot of logic and overhead.

In this situation, Meyer's paper represented significant progress in several aspects. Firstly, the logic for constructing the residual vector $\boldsymbol{\sigma}$ was not carried down any more to individual one- and two-electron integrals and configurations, but was reduced to deal with whole matrices of integrals and configuration coefficients. This is a generalization of what had already been proposed by Ahlrichs and Driessler [6] for the two-electron case. The integral matrices represent the half-transformed two-electron integrals with at least two occupied orbitals, while the coefficient matrices represent electron pair functions. The matrix formulation was achieved by partially giving up the normalization of the configurations: implicit in the theory is a different normalization of diagonal doubly external configurations, Φ_{ij}^{aa}, and off-diagonal configurations, Φ_{ij}^{ab} ($a \neq b$). This made it possible to eliminate all coupling coefficients depending on external (virtual) orbitals. In fact, this simplification follows most naturally by constructing the CSFs so that they can be considered as tensor components of the external orbital subspaces [7]. The equations for the residual vector were given in the form of matrix operations (mainly matrix multiplications), and there are only very few simple coupling coefficients depending on the spin-coupling (singlet or triplet) of the two external electrons. Thus, all complicated logic was removed, allowing an implementation of optimum efficiency. In fact, matrix multiplications are still the fastest operations one can perform on all kinds of computers today, ranging from personal computers with the Linux operating system or RISC workstations to vector supercomputers.

Secondly, Meyer showed in this paper that a full integral transformation can be avoided. The self-consistent electron pair (SCEP) method only requires transformed integrals with at least two occupied orbitals involved. The contributions of the remaining integrals with one or zero occupied orbitals can be treated by computing for each electron pair an "external" exchange operator that can be obtained directly in the atomic orbital (AO) basis, very much like the exchange contribution to the Fock matrix in the self-consistent-field (SCF) case. In fact, in the original SCEP formulation two such operators per pair were needed for the CI with single and double excitations (CISD) case, but it

was shown soon later [8] that a single operator per pair is sufficient. This is even true for full coupled-cluster with single and double excitations (CCSD) [9]. Avoiding the full integral transformation reduces the disc space for integral storage by a factor of 2 and the processing time for the transformation from $\mathcal{O}(N^5)$ to $\mathcal{O}(mN^4)$, where m is the number of correlated orbitals and N the number of basis functions. At first glance this seems most significant for highly accurate calculations with large basis sets on small molecules. At the time of Meyer's paper, most researchers considered this not to be a very important advantage, since the basis sets one was able to use were relatively small anyway. Moreover, computing the external exchange operators in the AO basis rather than in the MO basis from fully transformed integrals requires additional matrix transformations, and apart from this it increases the cost formally from about $n_{\text{pair}}(N - m)^4$ to about $n_{\text{pair}}N^4$. However, from the current point of view, avoiding the full transformation is essential for modern integral-direct correlation methods [10] that allow the use of many hundreds of basis functions. If in such methods screening of the two-electron integrals in the AO basis is performed, the scaling of the computational cost with molecular size can be reduced dramatically. Using local electron correlation treatments (see later) even linear cost scaling becomes possible.

The third new and, in retrospect, very important contribution in Meyer's paper was the proposal to compute not only the external exchange operators but also the complete residual directly in the nonorthogonal AO basis. He showed that the nonorthogonality only leads to quite minor complications in the equations for the residual; namely, some additional matrix multiplications with the overlap matrix are needed. The residual in the AO basis is transformed into the MO basis only for performing the perturbational update of the configuration coefficients. At the time when the paper was written, the AO formulation probably seemed to most people an unnecessary complication. From our present viewpoint, however, it was far ahead of the time: today we know that the AO formulation is essential for exploiting the local short-range character of electron correlation and makes possible linear cost scaling with molecular size even for coupled-cluster methods.

The fourth important contribution of the paper was to show how a Brueckner CI(D) is performed efficiently. In fact, even though the paper gives all matrix elements needed for a CISD, the iterative method described by Meyer was a Brueckner theory, since the single excitations were absorbed into the reference function. He proposed two different ways for absorbing the single excitations: either by a simple first-order update scheme of the orbitals with subsequent symmetrical orthonormalization or by performing a natural orbital iteration. He also introduced projectors to the external space, similar to those used nowadays in local correlation treatments. In the SCEP framework, the projectors were used to retain strong orthogonality between the external pair functions and the reference wavefunction once the occupied orbitals were modified in the Brueckner iteration. Meyer probably favored the Brueckner iteration mainly to avoid the construction of two external

exchange operators per pair in each iteration. As already mentioned, it was shown later that this is possible even with explicit inclusion of single excitations. Much later several groups showed that Brueckner theory has some important advantages, in particular in situations with symmetry-breaking.

The SCEP method was first implemented in 1976 for closed-shell reference functions by Dykstra and Meyer during a visit of Dykstra in Mainz and was first applied by Dykstra, Schaefer, and Meyer [11]. It was generalized in 1981 for multiconfiguration reference functions by Werner and Reinsch [8]. The latter method was the first implementation of an internally contracted multireference CI. The internal contraction scheme was proposed in another landmark paper of Meyer [12], which was written almost at the same time as the SCEP paper and is closely related to it. The internally contracted multireference SCEP method [8] used a similar matrix structure as Meyer's original SCEP, but was formulated in an orthogonal orbital basis and included all single and double excitations from the multiconfiguration SCF reference function. In 1988, a new implementation of the internally contracted multireference CI was completed [13, 14], which is still of unmatched efficiency. The internal contraction scheme is also widely used in multireference perturbation theory (MRPT2, CASPT2/3) [15–17], since the internally contracted configurations span exactly the first-order interaction space of a multiconfiguration reference wavefunction [12]. The SCEP theory was further generalized in 1984 by Pulay, Saebø, and Meyer [18] to the coupled-cluster doubles case [18] and later to closed and open-shell CCSD cases in our group [9, 19]. In Ref. [18] it was shown for the first time that the equations can be simplified by generating the configurations by spin-coupled pair excitation operators, $\hat{E}_{ai}\hat{E}_{bj}$. This leads not only to unnormalized but also to nonorthogonal configurations. Nevertheless, the number of matrix multiplications is reduced and for closed-shell reference functions the coupling coefficients are entirely eliminated (except for simple factors such as 2 in some terms).

As previously mentioned, the formulation of the theory with nonorthogonal correlation orbitals forms the basis for local electron correlation theories, as proposed by Pulay in 1983 [20] and first implemented for MP4(SDQ), CISD, and CEPA by Saebø and Pulay [21, 22]. A generalization for full CCSD was published in 1996 by Hampel and Werner [23]. In these methods the virtual space is spanned by a redundant set of projected AOs, which are orthogonalized on the occupied orbitals. Alternatively, AOs could be used directly as proposed by Meyer. The advantage of using AOs (or projected AOs) to span the virtual space is the fact that these are intrinsically localized. This allows the restriction of excitations from localized occupied orbitals to relatively small subspaces (domains) of virtual functions. Each pair is described by a different domain, but the size of each pair domain is independent of the molecular size. This immediately reduces the $\mathcal{O}(N^4)$ dependence of the number of configurations and variational parameters to $\mathcal{O}(N^2)$. Furthermore, by neglecting or approximating the very small contributions of distant electron pairs, linear

scaling, $\mathcal{O}(N)$, of the number of electron pairs and of the configurations is obtained, but nevertheless 98–99.5% of the canonical correlation energy is recovered. Linear scaling of the computational cost with molecular size has recently been achieved in our group for local MP2 [24], and this new program has made it possible to perform overnight LMP2 calculations for molecules with about 300 valence electrons and over 1500 basis functions on a single low-cost personal computer. A local coupled-cluster program with linear-scaling behavior is presently under development.

AO formulations are currently also being investigated by other groups with the aim of reducing the computational cost for large molecules. Ayala and Scuseria [25] have presented a linear-scaling MP2 method based on the Laplace transformation formalism of Häser and Almlöf [26]. More closely related to the SCEP formalism of Meyer is a recent tensor formulation of coupled-cluster theory by Head-Gordon et al. [27]. As far as the treatment of the virtual space is concerned, this is basically the same as what was described in Meyer's SCEP paper more than 20 years ago. (see also Ref. [7]); however, Head-Gordon et al. go a step further by also using a redundant set of nonorthogonal (projected) AOs to span the occupied space. This seems very wasteful at first glance, but as Head-Gordon et al. demonstrated for a number of test cases, a relatively small number of selected configurations is sufficient to recover most of the correlation energy [28]. In fact, in a very recent paper by Scuseria and Ayala [29] it is demonstrated for a related method that the number of configurations (or coupled-cluster amplitudes) needed to obtain a fixed fraction of the canonical correlation energy scales linearly with molecular size, as is also the case in local correlation methods with orthogonal localized orbitals.

During the last decade, linear-scaling density functional theory (DFT) methods have played a crucial role in extending the applicability of electronic structure methods to very large molecular systems; however, the problem remains that there is no possibility to systematically approach the exact result using DFT. Many attempts are presently being made to improve the available functionals, but this is mainly achieved by fitting DFT results to experimental data. DFT can therefore be viewed as a semiempirical theory. Accurate wavefunction-based local electron correlation methods are currently catching up with DFT in applicability to larger systems. It is my view that Meyer's 1976 paper forms the basis for these very exciting new developments. The SCEP theory was clearly far ahead of its time.

References

1. Meyer W (1971) Int J Quantum Chem Symp 5: 341
2. Kutzelnigg W (1977) In: Schaefer HF III (ed) Methods of electronic structure theory. Plenum, New York, pp 129–188, and references therein
3. Meyer W (1973) J Chem Phys 58: 1017
4. Roos BO (1972) Chem Phys Lett 15: 153
5. Roos BO, Siegbahn PEM (1977) In: Schaefer HF III (ed) Methods of electronic structure theory. Plenum, New York, pp 277–318

6. Ahlrichs R, Driessler F (1975) Theor Chim Acta 36: 275
7. Meyer W, Ahlrichs R, Dykstra CE (1984) In: Dykstra CE (ed) Advanced theories and computational approaches to the electronic structure of molecules. Reidel, Dordrecht, pp 19–38
8. Werner H-J, Reinsch EA (1982) J Chem Phys 76: 3144
9. Hampel C, Peterson KA, Werner H-J (1992) Chem Phys Lett 190: 1
10. Schütz M, Lindh R, Werner H-J (1999) Mol Phys 96: 719
11. Dykstra C, Schaefer HF, Meyer W (1976) J Chem Phys 65: 2740
12. Meyer W (1977) In: Schaefer HF III (ed) Methods of electronic structure theory. Plenum, New York, pp 413–446
13. Werner H-J, Knowles PJ (1988) J Chem Phys 89: 5803
14. Knowles PJ, Werner H-J (1988) Chem Phys Lett 145: 514
15. Roos BO, Linse P, Siegbahn PEM, Blomberg MRA (1982) Chem Phys 66: 197
16. Wolinski K, Pulay P (1989) J Chem Phys 90: 3647
17. Andersson K, Malmqvist P-Å, Roos BO, Sadlej AJ, Wolinski K (1990) J Phys Chem 94: 5483
18. Pulay P, Saebø S, Meyer W (1984) J Chem Phys 81: 1901
19. Knowles PJ, Hampel C, Werner H-J (1993) J Chem Phys 99: 5219
20. Pulay P (1983) Chem Phys Lett 100: 151
21. Saebø S, Pulay P (1987) J Chem Phys 86: 914
22. Saebø S, Pulay P (1988) J Chem Phys 88: 1884
23. Hampel C, Werner H-J (1996) J Chem Phys 104: 6286
24. Schütz M, Hetzer G, Werner H-J (1999) J Chem Phys (in press)
25. Ayala PY, Scuseria GE (1999) J Chem Phys 110: 3660
26. Häser M, Almlöf J (1992) J Chem Phys 96: 489
27. Head-Gordon M, Maslen PE, White CA (1998) J Chem Phys 108: 616
28. Maslen PE, Head-Gordon M (1998) J Chem Phys 109: 7093
29. Scuseria GE, Ayala PY (1999) J Chem Phys 111: 8330