

Multiconfigurational second order perturbation theory applied to the calculation of electronic spectra of conjugated systems

Björn O. Roos*, Luis Serrano-Andrés and Manuela Merchán

Department of Theoretical Chemistry, Chemical Center, P.O.B. 124, S-221 00 Lund, Sweden;
Departamento de Química Física Universitat de València, Spain

Abstract

A newly proposed method for the calculation of complex electronic structures in molecules is applied to the calculation of excited states. The method is a two step procedure: The Complete Active Space (CAS) SCF method is used to calculate the molecular orbitals and a reference function. This step thus takes care of the basic interactions including the coupling of nearly degenerate configurations, which is very common for excited states. Dynamic electron correlation effects are added in the second step, where second order perturbation theory is used with the CASSCF wave function constituting the reference function. The procedure has been successfully applied to a number of electronic spectra, yielding results of much higher accuracy than has been possible to obtain with traditional CI based methods. The lectures will give examples from studies of benzene, the azabenzenes, short polyenes, and three hetero-cyclic pentadienes. In several of these studies have the theoretical results provided a new understanding of the spectra and suggested a number of new assignments.

1. INTRODUCTION

Accurate calculations of the electronic spectra of conjugated systems has remained a challenge for *ab initio* quantum chemistry since the first calculations were performed on the benzene molecule in the late 60ies. As an example of these difficulties let us take a closer look at pyrrole. In spite of numerous efforts since the first calculations were performed and the first UV spectrum was measured, the interpretation of the electronic spectra of this molecule is still not complete. The difficulties are caused by the appearance of a number of strongly allowed Rydberg series, which overlap two broad and diffuse bands usually assumed to arise from π - π^* valence type excitations. The maxima of the two intense bands are located in the regions around 6.0 and 7.5 eV, respectively. Simpler molecular orbital (MO) based arguments would assign the first feature to a 1B_2 valence excited state and the second to a ${}^1A_1^+$ state of the same type. These arguments also predict the existence of a weakly allowed transition, ${}^1A_1^-$, close to the more intense 1B_2 band.

Theoretically, the early CNDO calculations by Del Bene and Jaffé¹ support this interpretation. In contrast, the *ab initio* studies performed so far give a completely different picture. The most comprehensive treatment is probably the Symmetry-Adapted-Cluster (SAC) CI calculations performed by Nakatsuji *et al.*². They find the 1B_2 state at energies more than one eV above the maximum of the first intense peak and locate the ${}^1A_1^-$ transition at energies above 7 eV in all three molecules. The intense ${}^1A_1^+$ state is not found at all. Other theoretical studies have obtained similar results³⁻⁵. On the other hand, these theoretical studies have been able to calculate the location of a number of Rydberg states with reasonably good accuracy. Several investigators

have from these results drawn the false conclusion, that the first major feature in the pyrrole spectrum is composed entirely of bands belonging to the R(3p) Rydberg series.

This situation is more common than exception. Only very large multi-reference (MR) CI can account in a quantitative way for the large changes in electron correlation that are often brought about by an electronic excitation. An example of the difficulties is provided by a recent MRCI study of the valence excited states in the pyrimidine molecule⁶. The problems are most severe for valence excited states, especially those which show large ionic contributions to the electronic structure. It is in these cases necessary to account for the considerable effects of the dynamic polarization of the σ -electrons, implying that a calculation must include a major part of the electron correlation in the σ shells. This might be difficult in systems with many electrons, due to the limitations inherent in the CI methods. Calculations of Rydberg excited states seem on the other hand to be much more straightforward and rather accurate excitation energies are normally obtained, once an appropriate basis set is used.

For some time, we have attempted to gain insight into the correlation effects on spectral and structural features of conjugated systems through *ab initio* quantum chemical calculations on the electronic spectra of small and medium size (up to 20 first row atoms) molecules. Recently, we reported on the π - π^* singly excited states of benzene⁷ as an illustration of a novel approach that has the aim of being accurate to better than 0.5 eV. We have also reported results for the π - π^* and n - π^* singly excited states of the azabenzenes⁸, for valence and Rydberg excited singlet and triplet states in the short polyenes ethene, butadiene, and hexatriene⁹, and for the hetero-cyclic pentadienes: cyclopentadiene, pyrrole and furan¹⁰. In all these cases were the excitation energies computed with an accuracy higher than the proposed aim. The largest error among all the valence excited states in benzene occurred for the ${}^3E_{1u}$ state and was 0.26 eV.

To compute the excitation energies from first principles, electron correlation has to be treated in a balanced way and all states of the same symmetry, that are close in energy, have to be considered simultaneously. The most important correlation effects are described by configuration state functions (CSFs) mixing the π -electrons among the π -orbitals. However, to obtain quantitative results dynamic correlation effects have also to be considered. These effects are dominated by the dynamic polarization of the σ -electrons, which is described by configurations involving simultaneous σ - σ^* and π - π^* excitations, and can differ substantially from state to state. Typically, large σ -polarization effects are observed for states of strong ionic character.

The approach used here is capable of handling these problem and correctly account for the different correlation contributions to the excitation energies as has been shown in a number of applications⁷⁻¹⁰. The method is based on the Complete Active Space (CAS) SCF approximation¹¹, which has proven to be particularly suited to deal with situations where the electronic structure varies strongly, for example, in the vicinity of transition states or in excitation processes. The CASSCF wave functions will comprise all strongly interacting electronic configurations, but will not to a large extent account for the dynamic electron correlation effects. They are added in a subsequent step, where the CASSCF wave function serves as the reference function in a second order perturbation calculation of the correlation energy (CASPT2)^{12,13}. The validity of a second order perturbation calculation of the differential correlation effects on excitation energies has been demonstrated⁷⁻¹⁰. The same approach has also been shown to yield accurate geometries, binding energies, and other properties of molecular systems both in their ground and excited states.

2. METHOD AND COMPUTATIONAL STRATEGY

2A. Basis sets

The choice of an appropriate basis set is especially intricate in studies of a large number of electronic states of a molecule. The basis set then has to be adequate to describe not only the relatively compact ground state, but also more diffuse valence excited states and Rydberg states. It also has to be flexible enough to describe the

important correlation effects, especially the dynamic σ -electron polarization. We have performed a number of tests of different basis sets and have concluded that it is necessary to include at least one set of d-type functions on first row atoms and in addition a p-type function on the hydrogen in order to account quantitatively for these effects. Generally contracted basis sets of the atomic natural orbital (ANO) type^{14,15}, which are obtained from C,N,O(14s,9p,4d)/H(8s,4p) primitive sets have been used in all applications made so far. These basis sets are constructed to optimally treat correlation and polarization effects and should be large enough to describe the electronic structure of the valence excited states with the desired accuracy. The smallest contraction used is C,N,O[4s3p1d]/H[2s1p]. When the aim is to describe both valence excited states and Rydberg states, the original basis sets have to be supplemented with diffuse functions. In the study of the linear polyenes⁹ these diffuse functions were centred on the different carbon atoms. As a result, a large number of Rydberg orbitals were created, which turned out to make it difficult to locate the valence excited states in the CASSCF calculations. We therefore in the study of the pentadienes instead added diffuse functions in the average charge centroid for the 2A_2 and 2B_1 states of the cation. With such basis sets it turned out to be much easier to separate valence and Rydberg excited states. The same approach is successfully used in on-going studies of the spectra of benzene, phenol¹⁶, and cytosine¹⁷.

2B. CASSCF and CASSI calculations

The first step in the computational procedure has the aim of determining the molecular orbitals (MOs). Ideally each electronic state should be described with its own optimized orbitals, a procedure which implies independent CASSCF calculations for each excited state. Such calculations frequently lead to convergence problems, however, due to the proximity of excited states of the same symmetry. The experience gained so far shows that enough accurate MOs can actually be obtained from average CASSCF calculations, where the averaging is over all desired states of a given symmetry. The small errors introduced in the MOs are corrected in the subsequent perturbation calculation. Sometimes it is necessary to include more states in the average CASSCF calculation than are actually to be studied. In the recent study of furan eight states of 1A_1 symmetry had to be used, although only four were of interest. The reason was that the second valence excited state of this symmetry appeared as state number eight at the CASSCF level, but as the fourth state in the final CASPT2 calculation. The four extra states were of quasi Rydberg character (higher Rydberg states not properly described by the basis set). In order to obtain meaningful results it is consequently necessary to know the general structure of the desired excited states, such that they can be identified in the CASSCF calculation. It is likely that similar problems, related to the difficulty in defining proper reference states, has prevented the identification of some valence excited states in previous MRCI studies.

The choice of active orbitals for the CASSCF calculations has to be done with great care. Again it is necessary to have an idea of the character of the excited states. The pyrrole molecule¹⁰ can serve as an illustrative example: It was decided to include in the treatment all valence excited states (π - π^*) plus excitations from the highest occupied π -orbitals to the 3s, 3p and 3d Rydberg orbitals. Consequently the valence π -orbitals and the Rydberg orbitals constitutes a minimal active space. Using the symmetry (C_{2v}) of the molecules we can label this active space as (4253) corresponding to the four symmetries a_1 , b_2 , b_1 , and a_2 , where b_1 and a_2 represent π -orbitals. But the active σ -orbitals (of Rydberg type) will not be occupied in states of A_1 and B_2 symmetries (π - π^* states). For these symmetries it is then sufficient to use the active space (0053), which includes the valence π -orbitals plus the Rydberg orbitals $3p\pi(b_1)$, $3d\pi(b_1)$, and $3d\pi(a_2)$. For the same reasons the active space can be reduced to (4232) for states of symmetries B_1 and A_2 , now excluding the Rydberg orbitals of π symmetry. It is sometimes necessary to extend these active spaces somewhat in order to avoid the appearance of intruder states in the CASPT2 calculations. If excitations out of σ -orbitals are considered, for example lone pairs, they have of course also to be included in the active space.

The MOs obtained in the different CASSCF calculations are in general not the same. Also, CASSCF wave functions, which have been optimized for different states of the same symmetry, are not orthogonal to each other. The CASSCF state interaction (CASSI) method¹⁸ has been developed to compute transition properties from non-orthogonal state functions and is used to compute the oscillator strengths. In the formula for the oscillator strength we used the energy differences corrected for by the second order perturbation method (PT2F). This approach for the oscillator strength is by now well documented in a number of applications and has in a recent study of the pyrimidine molecule been shown to give results similar to those obtained with the multi-reference CI method⁶.

2C. The CASPT2 method

The CASPT2 method^{12,13} computes the first order wave function and the second order energy in the full CI space without any further approximation with a CASSCF wave function constituting the reference function. The zeroth order Hamiltonian is defined as a Fock type one electron operator and is constructed such that Möller-Plesset type perturbation theory is obtained in the closed shell single determinant case. Two different formulations of the zeroth order Hamiltonian are possible: one which utilizes only the diagonal part of the Fock matrix (called PT2D) and one, which includes also the non-diagonal elements (PT2F). The first choice is computationally simpler and leads in most cases to results not very different from PT2F, as illustrated, for example, in the study of the electronic spectrum of the benzene molecule⁵. It should be emphasized, however, that it is only the non-diagonal approach, which is invariant to rotations of the molecular orbitals. The full approach must therefore be used in cases, where such invariance is important, for example, in calculations of potential surfaces. It was shown in the calculations on the polyenes⁹ and the pentadienes¹⁰, that the difference between PT2D and PT2F can be substantial for excited states of the Rydberg type, while it is small for valence excited states. PT2F is therefore the only recommended procedure, even if it is somewhat more time-consuming than PT2D.

The CASPT2 equations are formulated exclusively in terms of one-, two- and three-body densities and is therefore independent of the actual size of the reference function. The limiting factor is not the number of correlated electrons, but the number of active orbitals, which determines the size of the density matrices. The current implementation of the CASPT2 method allows a maximum of 14 active orbitals. A larger number is rarely needed in any application and could anyway not be handled by the CASSCF program, except in cases with very few active electrons (or holes).

The computational procedure outlined above is included in the quantum chemistry software MOLCAS-2¹⁹, which can be obtained by contacting the author of the present contribution.

3. THE ELECTRONIC SPECTRUM OF CYCLOPENTADIENE, PYRROLE, AND FURAN

As an example we present in the table the results from a recent study of three molecules. This example is especially intriguing, since earlier attempts, both experimentally and theoretically to assign the spectra have been incomplete or in error. Some of the assignments made in the table are made here for the first time and are to some extent tentative. However, the overall agreement with experiments gives general support to the present results. The valence excited states 1B_2 and ${}^1A_1^+$ have only been firmly assigned before for CP and the location of the ${}^1A_1^-$ state in the three molecules is here made for the first time. A more detailed discussion of the assignments can be found in ref. 10. With the present assignments the agreement with experiment is surprisingly good. The deviations for the valence excited states are all smaller than 0.1 eV, while for the Rydberg states they lie between -0.06 and + 0.17 eV.

Table. Calculated and experimental excitation energies and oscillator strengths in cyclopentadiene (CP), pyrrole, and furan

State	CP		Excitation energy (eV)		Furan		Osc. strength		
	Calc.	exp.	Calc.	exp.	Calc.	exp.	CP	Pyrrole	Furan
Singlet states									
¹ A ₂ (1a ₂ - 3s)	5.65	5.68	5.08	5.22	5.92	5.94		forbidden	
¹ B ₂ (1a ₂ - 3pπ)	6.25	6.31	5.78	5.86	6.48	6.48	.0248	.0399	.0471
¹ A ₂ (1a ₂ - 3pσ)	6.30	6.26	5.83	-	6.59	6.61		forbidden	
¹ B ₁ (1a ₂ - 3pσ)	6.24	6.25	5.85	5.88	6.46	6.48	.0306	.0239	.0309
¹ A ₁ (valence)	6.31	≈6.2	5.92	5.88	6.16	-	.0003	.0195	.0015
¹ B ₂ (valence)	5.27	5.26	6.00	5.98	6.04	6.06	.1475	.1253	.1543
¹ B ₁ (2b ₁ - 3s)	7.95	8.03	5.97	5.98	7.21	7.38	.0252	.0006	.0192
¹ B ₁ (1a ₂ - 3dδ)	6.87	-	6.40	6.43	7.15	-	.0003	.0117	.0000
¹ A ₂ (1a ₂ - 3dδ)	6.85	-	6.42	-	7.00	-		forbidden	
¹ A ₂ (1a ₂ - 3dσ)	6.91	-	6.51	-	7.22	-		forbidden	
¹ B ₂ (1a ₂ - 3dπ)	6.86	6.80	6.53	6.50	7.13	-	.0101	.0011	.0074
¹ A ₁ (1a ₂ - 3dπ)	6.93	7.05	6.54	6.50	7.31	7.28	.0208	.0010	.0003
¹ B ₁ (2b ₁ - 3pσ)	-	-	6.62	6.78	-	-	-	.0161	-
¹ A ₁ (2b ₁ - 3pπ)	-	-	6.65	6.78	-	-	-	.0127	-
¹ A ₂ (2b ₁ - 3pσ)	-	-	6.77	-	-	-		forbidden	
¹ A ₂ (2b ₁ - 3dδ)	-	-	7.31	-	-	-		forbidden	
¹ B ₁ (2b ₁ - 3dδ)	-	-	7.32	7.26	-	-	-	.00005	-
¹ A ₁ (2b ₁ - 3dπ)	-	-	7.36	7.43	-	-	-	.0047	-
¹ B ₁ (2b ₁ - 3dσ)	-	-	7.39	7.43	-	-	-	.0025	-
¹ B ₂ (2b ₁ - 3dπ)	-	-	7.43	7.43	-	-	-	.0005	-
¹ A ₁ (valence)	7.89	≈7.9	7.46	7.54	7.74	7.82	.4415	.3261	.4159
Triplet states									
³ B ₂ (valence)	3.15	3.10	4.27	4.21	3.99	4.02			
³ A ₂ (1a ₂ - 3s)	5.63	-	5.04	-	5.86	-			
³ A ₁ (valence)	4.90	>4.7	5.16	5.10	5.15	5.22			
³ B ₁ (1a ₂ - 3pσ)	6.25	-	5.82	-	6.42	-			

Acknowledgements

The research reported in this communication has been supported by a grant from the Swedish Natural Science Research Council (NFR), by IBM Sweden under a joint study contract, by the Cooperación Científica y Técnica of Spain, and by project OP90-0042 of the Secretaría de Estado de Universidades e Investigación.

REFERENCES

1. J. Del Bene and H. H. Jaffé, *J. Chem. Phys.* **48**, 4050 (1968).
2. H. Nakatsuji, O. Kitao, and T. Yonezawa, *J. Chem. Phys.* **83**, 723 (1985).
3. K. Tanaka, T. Nomura, T. Noro, H. Tatewaki, T. Takada, H. Kashiwagi, F. Sasaki, and K. Ohno, *J. Chem. Phys.* **67**, 5738 (1977).
4. W. Butscher and K. H. Thunemann, *Chem. Phys. Letters* **57**, 224 (1978).
5. D. C. Rawlings and E. R. Davidson, *Chem. Phys. Letters* **98**, 424 (1983);
D. C. Rawlings, E. R. Davidson, and M. Gouterman,
Int. J. Quantum Chem. **26**, 237 (1984)

6. P.-Å. Malmqvist, B. O. Roos, M. Fülcher, and A. Rendell, *Chem. Phys.* **162**, 359 (1992).
7. B. O. Roos, K. Andersson, and M. P. Fülcher, *Chem. Phys. Lett.* **192**, 5 (1992).
8. M. P. Fülcher, K. Andersson and B. O. Roos, *J. Phys. Chem.*, (in press).
9. L. Serrano-Andrés, M. Merchán, I. Nebot-Gil, R. Lindh, and B. O. Roos, *J. Chem. Phys.* (submitted).
10. L. Serrano-Andrés, M. Merchán, I. Nebot-Gil, M. Fülcher, and B. O. Roos, *J. Am. Chem. Soc.* (submitted).
11. For a review of the CASSCF method see: B. O. Roos in: *Ab Initio Methods in Quantum Chem.-II* (ed. K.P. Lawley), J. Wiley & Sons Ltd, 399 (1987).
12. K. Andersson, P.-Å. Malmqvist, B. O. Roos, A. J. Sadlej, and K. Wolinski, *J. Phys. Chem.* **94**, 5483 (1990).
13. K. Andersson, P.-Å. Malmqvist, and B. O. Roos *J. Chem. Phys.* **96**, 1218 (1992).
14. J. Almlöf and P. R. Taylor, *J. Chem. Phys.* **86**, 4070 (1987).
15. P.-O. Widmark, P.-Å. Malmqvist, and B. O. Roos, *Theor. Chim. Acta* **77**, 291 (1990).
16. J. Lorentzon, P.-Å. Malmqvist, and B. O. Roos, *to be published*.
17. M. Fülcher and B. O. Roos, *to be published*.
18. P.-Å. Malmqvist, *Int. J. Quantum Chem.* **30**, 479 (1986);
P.-Å. Malmqvist and B. O. Roos, *Chem. Phys. Letters* **155**, 189 (1989).
19. K. Andersson, M. P. Fülcher, R. Lindh, P.-Å. Malmqvist, J. Olsen, B. O. Roos, A. J. Sadlej, and P.-O. Widmark,
MOLCAS Version 2, User's Guide, University of Lund, Sweden (1991).
The program can be obtained in versions for IBM VM/XA and AIX (for RS/6000 workstations) by contacting B.O. Roos).