Konrad-Zuse-Zentrum für Informationstechnik Berlin
Heilbronner Str. 10, W-1000 Berlin 31

Klaus Schöffel

# Ab initio Quantum Chemical Calculations with GAMESS–UK and GAUSSIAN90 Program Packages – A Comparison –

Technical Report TR 92–1 (January 1992)

Herausgegeben vom
Konrad-Zuse-Zentrum für Informationstechnik Berlin
Heilbronner Str. 10
1000 Berlin 31
Verantwortlich: Dr. Klaus André
Umschlagsatz und Druck: Rabe KG Buch-und Offsetdruck Berlin

ISSN 0933-789X

# Ab initio Quantum Chemical Calculations with GAMESS-UK and GAUSSIAN90 Program Packages - A Comparison -

Klaus Schöffel *
Konrad-Zuse-Zentrum für Informationstechnik Berlin,
Heilbronner Str. 10
D-1000 Berlin 31

**Abstract**

Two commercial available molecular electronic structure software packages GAUSSIAN90 and GAMESS-UK are compared. Basis for this comparison is a benchmark suite which is designed to highlight the typical range of calculations commonly performed by the *ab initio* computational chemist.

## 1 Introduction

The development of computer programs for electronic structure calculations of molecules with sufficiently general and comprehensive capabilities, so that they can be used for very accurate studies of small molecules or for less accurate studies of larger molecules, or for the determination of reaction paths ways or the calculation of electronic properties is by no means an easy task and certainly requires several man-years of work. These programs should be easy to use by non-experts. Over the last years several commercial software packages have become available, where the GAUSSIAN program system [1] became very popular and developed towards a standard for molecular electronic structure applications. The GAMESS-UK program package [2] is less widely used, mainly due to the fact that it has only very recently been commercialized. Previously it has been made available only to researchers within collaborative projects with the program author [3].

Since we think that GAMESS-UK is a system with a high functionality and applicability it is worth trying to benchmark this software package against the well known GAUSSIAN90.

*E-Mail: Schoeffel@sc.ZIB-Berlin.DE

The main features of the software packages will be briefly described, the benchmark suite will be presented and the final results discussed.

## 2 Program System GAMESS-UK

GAMESS-UK (Generalized Atomic and Molecular Electronic Struture System, developed in UK) is a general purpose *ab initio* molecular electronic structure program for performing SCF- and MCSCF-gradient calculations, together with a variety of techniques for post Hartree Fock calculations. On-going development of the code is carried out at Daresbury, with the program currently available on a wide range of machines, including the CRAY X-MP and Y-MP, IBM 3090-600E/VF, Amdahl VP, FPS-264 and M64/60, Convex C2, Cyber-205, Alliant FX8 and FX/2808, Intel iPSC/2 and iPSC/860, Meiko T800/20 + MK086 (i860), Apollo DN1000, SGI Power Series, IBM/RS6000, VAX, DEC S5000/200 Solbourne S4000 and SUN.

The program utilizes the Rys Polynomial or Rotation techniques to evaluate repulsion integrals over s,p,d and f type cartesian Gaussian orbitals. Open- and closed-shell SCF treatments are available within both the RHF and UHF framework, with convergence controls provided through a hybrid scheme of level shifters and the DIIS method. In addition generalized valence bond, complete active space SCF, and more general MCSCF calculations may be performed. The analytic energy gradient is available for each class of wavefunctions above mentioned. Gradients for s and p Gaussians are evaluated using the algorithm due to Schlegel, while gradients involving d and f Gaussians utilize the Rys Polynomial Method. Geometic optimization is performed using a quasi-Newton rank-2 update method, while transition state location is available through either a synchronous transit trust region or 'hill-walking' method. Force constants may be evaluated by numerical differentiation or analytically by using a coupled Hartree-Fock scheme.

Møller-Plesset perturbation calculations can be carried out up to the fourth order. 1st and 2nd derivatives are available analytically. Direct-MP2 calculations are possible for large systems.

*Ab initio* core potentials are provided in both semi-local and non-local formalism for valence-only molecular orbital treatments.

Conventional CI (using the table-driven selection algorithms within the framework of MR-DCI calculations), Direct-CI and Full-CI correlation treatments are available. The direct calculation of molecular valence ionization energies may be performed through Green's function techniques, using either the outer-valence Green's function (OVGF) or the two-particle-hole Tamm-Damcoff method (2ph-TDA).

A variety of wavefunction analysis methods are provided, including population analysis and distributed multipole analysis, localized orbitals, graphical analysis and calculation of 1-electron properties.

With conventional processing of two-electron integrals, GAMESS-UK is limited to a maximum of 255 basis functions . Direct-SCF calculations lift this restriction however - the program structure is open-ended in this mode, so that calculations up

to 1000 basis functions have been performed.

In addition to the functionality outlined above, the program now includes an utility module concerned with house-keeping, library file creation, file manipulation etc.

Areas of application for GAMESS-UK include:

- Theoretical studies of reaction surfaces. SCF and MCSCF calculations of equilibrium and transition state geometries, and the evaluation of force constants and vibrational frequencies at these stationary points.
- Energy assignments in the entire field of electronic spectroscopy. Calculation of spectroscopic properties of highly excited molecular and ionic states. Evaluation of transition moments.
- Graphical and numerical analysis of Hartree Fock and post-Hartree Fock wavefunctions.
- Generation of zero-order wavefunctions required in the more extensive treatments of dynamical correlation energy. Applicability to general systems with many electrons in this treatment of dynamical correlation energy. Benchmark treatments of correlation energy using Full-CI calculations.
- Theoretical mechanistic studies of chemisorption in heterogeneous catalysis.
- Treatment of bio-organic and related molecules through Direct-SCF calculations and electrostatic potential analysis.

# 3 Program System GAUSSIAN90

GAUSSIAN90 is a connected system of programs for performing semiempirical and *ab initio* molecular orbital calculations. It has been designed with the needs of the user in mind. Thus, all of the standard input is free-format and mnemonic.

GAUSSIAN90 provides an analysis of molecular potential energy surfaces to determine molecular structures and spectral properties of stable and unstable species. It provides a wide number of *ab initio* models, and some semiempirical models, using a user-friendly interface. The program is available for: Alliant, Ardent, Convex, CRAY X/Y-MP, IBM RS6000, IBM (MVS, VM), Multiflow and VAX.

The capabilities of the GAUSSIAN90 system include:

- Calculation of one- and two-electron integrals over s, p, d and f contracted Gaussian functions. The basis functions can either be cartesian Gaussians or pure angular momentum functions, and a variety of basis sets are stored in the program and can be requested by name. Integrals may be stored externally or recomputed as needed via the Direct SCF procedures.

- Self-consistent field calculations for restricted closed-shell (RHF), unrestricted open-shell (UHF), and open-shell restricted (ROHF) Hartree-Fock wavefunctions as multiconfigurational wavefunctions that fall within the Generalized Valence Bond-Perfect Pairing (GVB-PP) formalism.

- Evaluation of various one-electron properties of the Hartree-Fock, MP2 and CI energies and wavefunctions, including Mulliken population analysis, multipole moments and electrostatic fields.

- Semi-empirical calculations using the CNDO, INDO, MINDO/3, MNDO and AM1 model Hamiltonians.

- Automated geometry optimization to either minima or saddle points, numerical differentiation to produce force constants, polarizabilities and dipole derivatives and reaction path following.

- Correlation energy calculations using Møller-Plesset perturbation theory carried out to second (MP2), third (MP3) or fourth (MP4) order.

- Correlation energy calculations using configuration interaction (CI) or all double excitations (CID) or all single and double excitations (CISD).

- Correlation energy calculations using coupled cluster theory with double substitutions (QCISD).

- Analytic computation of the nuclear coordinate gradient of the RHF, UHF, ROHF, GVB-PP, MP2, RCID and RCISD.

- Computation of force constants (nuclear coordinate second derivatives), polarizabilities, hyperpolarizabilities, dipole derivatives and polarizability derivatives analytically for RHF and UHF energies, and numerically for RHF, UHF, MP2, CID and CISD energies.

- Harmonic vibrational analysis.

- Determination of intensities for vibrational transitions at the HF, MP2, and CI levels.

- Testing the SCF wave functions for stability under release of constraints.

- Excited States:
  Configuration Interaction with single excitations for energies and gradients. This allows excited states structures, adiabatic excitation energies, 0-0 transitions to be predicted and related properties such as oscillator strengths and electron densities to be computed.

- Direct SCF:
  Ground state vibrational frequencies and excited state properties are possible on the context of a direct SCF calculation.

- Semi-Direct MP2:
  This method optimizes the calculation procedure for the particular disk and memory configuration of the computer system.

- In-Core Methods:
  In-core methods for SCF energies, gradients and second derivatives and MP2 energies and gradients are available.

- QCI and CASSCF:
  Gradients as well as energies can be calculated at the Quadratic CI level. Associated properties, including polarizabilities, hyperpolarizabilities and electron densities, are also computed using the QCI level of correlation. The Complete Active Space SCF (CASSCF) method makes it possible to study systems with unusual electronic structures.

# 4 Benchmark-Suite

The following 13 benchmark calculations are designed to highlight the typical range of calculations commonly performed by the *ab initio* quantum chemist. They range from conventional single point Hartree-Fock calculations, over direct-SCF calculations, applicable to large molecules of biochemical interest, over effective-core-potential calculations of metallic clusters, to electron correlation studies using either Møller-Plesset perturbation theory or configuration interaction techniques. In the following each benchmark job is briefly described.

## 4.1 Hartree-Fock Calculations

Despite the fact that Hartree-Fock calculations are not the most demanding electronic structure calculations (computational effort scales as $N^4$ — or $N^3$ for very big molecules — , where $N$ is the number of basis functions), this type of calculation accounts for most of the cpu-time designated to todays computational molecular science applications, because it has developed to a tool which is routinely used by many researches from theoretical and experimental fields.
The iterative selfconsistent-field procedure is finished when the electronic energy has converged to within $10^{-9}$ hartree. The default available convergence aids and a main memory of 8 mega bytes is used.

### 4.1.1 Closed Shell SCF Calculation for Nitro-benzene (1)

Using a standard split valence 6-31G basis the applicability of the programs for medium sized molecules (91 basis functions) is tested.

### 4.1.2 Closed Shell SCF Calculation for Nickel-(0)-tetra-carbonyl (2)

This calculation is a slightly larger conventional SCF problem with 101 basis functions (3-21G basis set) with $T_d$ symmetry. For this transition metal complex a level shifter of 1.5 has been used.

### 4.1.3 Closed and Restricted Open Shell ECP Calculation for the $Na_7Mg^+$ Metal Cluster (3)

An double-$\zeta$ valence basis in connection with the Hay-Wadt effective-core-potential (ECP) is used. This test intends to explore the ECP capabilities of the programs. ECP calculations are particular important for the study of transition metal complexes, where they reduce the computational effort substantially.

### 4.1.4 Direct-SCF Calculation for Nitro-benzene (4)

Direct-SCF calculations avoid the external storage of the large two-electron integral file, by calculating these quantities during each iteration. The direct-SCF method is needed for the study of large molecules or for medium sized molecules with large basis sets.

## 4.2 Calculation including Static and Dynamic Electron Correlation

Complete active space selfconsistent field (CASSCF) calculations are employed, where static electron correlation plays an important role, as in the case of transition metal complexes and during breaking and forming of chemical bonds.

Møller-Plesset perturbation theory and configuration interaction techniques in the conventional table-driven or direct scheme are used to calculate, at least partial, the instantaneous electron correlation in a molecular system (dynamic electron correlation), where the former method is size consistent and the later isn't.

### 4.2.1 CASSCF-Calculation for $H_2CO$ (5)

This benchmark uses a double-$\zeta$ valence basis with polarization functions (42 basis functions) and correlates 10 electrons in 10 active orbitals using the CASSCF method.

### 4.2.2 CASSCF-Geometry Optimization for $H_2O$ (6)

Starting from a HF/DZP optimized geometry the internal coordinates are fully optimized within the $C_{2v}$ point group of the water molecule using a CASSCF wavefunction which correlates 10 electrons in 10 active orbitals and a double-$\zeta$ basis with polarization functions (26 basis functions).

### 4.2.3 Single and Multi-Reference CI-Calculation on $H_2CO$ (7)(8)

Using a triple-$\zeta$ valence basis with polarization functions (TZVP) a CI calculation is performed using all possible singly and doubly excited determinants out of a single reference and three references which gives 18513 configuration state functions for a single-reference case and 52252 for the multi-reference case.

### 4.2.4 Conventional-Multireference-Table CI (9)

The same molecule and basis set is used as in 4.2.3, but the table-driven selection algorithm is employed which allows the simultaneous determination of several roots. This technique is especially important when excited states of the same symmetry as the ground state are considered. According to an energy threshold criterion the 2447 most important symmetry adapted functions are selected by a perturbative technique out of 156144 functions. The correlation energy then is determined with an accuracy of $10^{-5}$ hartree.

### 4.2.5 MP2-Geometry Optimization for $H_2CO$ (10)

Starting from a HF/TZVP optimized geometry the internal coordinates are refined within the MP2 framework, where analytical MP2 gradients are used.

### 4.2.6 Direct-MP2 Calculation for $H_2CO$ (11)

As previously discussed, the direct techniques avoid the external storage of two-electron quantities. They are therefore well suited for the study of large systems, including correlation effects on the MP2 level. As an example the formaldehyd molecule is used with a TZVP basis.

## 4.3 Second Derivatives using HF or MP2 Wavefunctions

Second derivatives are particular important for the identification of stationary points on the potential energy surface and to predict vibrational frequencies. They can either be obtained numerically by a finite difference method of the analytical gradients or analytically via a coupled Hartree-Fock treatment.

### 4.3.1 Analytical Second Derivatives for $H_2CO$ (12)

Here, the second derivatives are calculated analytically using the coupled Hartree-Fock scheme and a TZVP basis set.

### 4.3.2 Second MP2 Derivatives for the $C_4$-molecule (13)

In this case the effect of electron correlation on the MP2 level on the second derivatives for the carbon $C_4$ cluster using the split valence 6-31G basis set with polarization functions is studied.

| No. | System | GAMESS-UK | GAUSSIAN90 |
|---|---|---|---|
| 1 | $\Phi$-$NO_2$ HF/6-31G (91) | 169.70 | 348.48 |
| 2 | $Ni(CO)_4$ HF/3-21G (101) | 121.46 | 387.42 |
| 3 | $Na_7Mg^+$ HF/ECPDZ (70) | 334.22 | — |
| 4 | $\Phi$-$NO_2$ Direct-HF/6-31G (91) | 1280.12 | 1034.22 |
| 5 | $H_2O$ CAS(10/10)OPT/DZP (26) | 548.10 | — |
| 6 | $H_2CO$ CAS(10/10)/DZP (52) | 188.80 | — |
| 7 | $H_2CO$ 1M-CI/TZVP (52) | 176.73 | 743.13 |
| 8 | $H_2CO$ 3M-CI/TZVP (52) | 302.71 | — |
| 9 | $H_2CO$ 4M-Table-CI/TZVP (52) | 341.12 | — |
| 10 | $H_2CO$ MP2-OPT/TZVP (52) | 803.71 | 1196.80 |
| 11 | $H_2CO$ 2ndHF/TZVP (52) | 342.28 | 768.40 |
| 12 | $C_4$ 2ndMP2/6-31G (60) | 1773.14 | 4603.13 |
| 13 | $H_2CO$ Direct-MP2/TZVP (52) | 192.21 | 132.93 |

Table 1: Timings for the benchmark jobs 1 - 13 performed with the GAMESS-UK and GAUSSIAN90 package on a single processor Convex C220. In the second column the molecular species is given with the number of basis functions in parenthesis. The third and fourth column gives the cpu-time for the two program packages in seconds. A dash (—) means that this pariticular benchmark could not be performed.

# 5 Results

Table 1 shows the timings for the benchmark jobs 1 - 13, obtained with the GAMESS-UK (G-UK) and GAUSSIAN90 (G-90) program packages using a single processor on a Convex C220. Not all benchmarks could be performed with G-90, the reasons are given below.

## 5.1 Hartree-Fock Calculations

For the nitro-benzene test case (1) G-UK is 2 times faster, where for the $T_d$ symmetry case (2) a factor of 3.2 is observed, which shows that for the conventional SCF approach G-UK clearly outperforms G-90, due to a well vectorized integral program part.

The calculation with the ECP (3) could not be performed with G-90 because the program aborts with floating point exception after having processed the user supplied basis set and ecp data. Using built in ECP and CEP-31G basis set the program aborts due to internal errors when the open shell calculation starts. The errors are reported to GAUSSIAN Inc.

Using the direct-SCF (4) approach G-90 is slightly faster than G-UK by a factor of 1.2, which shows that the direct-SCF part benefits from the code restructuring which took place when the GAUSSIAN88 was upgraded to GAUSSIAN90.

## 5.2 Electron-Correlation Calculations

Only two benchmark tests out of 6 could be managed with G-90 which is mainly due to limited local disk storage capacities and severe program limitations.

The $H_2O$ CASSCF geometry optimization (5) and $H_2CO$ single point calculation (6) could not be done with the given 300 Mbyte disk space. One reason for this intensive disk storage usage is that the CASSCF part of G-90 is not taking advantage of the spartial molecular symmetry, which may cause a tremendous increase in the use of computational resources. Another annoying thing is that the program generates 3 lines of output for each generated determinant which gives very easy thousands of redundant output lines. In order to get some indication on the rel. performance of the CASSCF part, the single point CAS(4/4)-cyclobutadiene G-90-handbook example was taken. For this particular case G-UK is about 1.8 times faster than G-90.

From the CI-calculations (7,8,9) only the single reference calculation (7) can be done with G-90 due to program limitations. G-UK clearly outperforms G-90 due to its highly vectorized direct-CI code by a factor of 4.2. Conventional Table-CI and multi-reference CI calculations are not possible using G-90.

The MP2-geometry optimization (10) shows that G-UK is about 1.5 times faster than G-90. Using the direct approach to calculate the MP2 energy (11) demonstrates again the good performance of the direct-SCF module in G-90. For this particular case G-90 is 1.4 times faster than G-UK.

### 5.2.1 Second Derivatives

For the analytic calculation of the second derivative using a HF-wavefunction (12) G-UK is 2.2 times faster than G-90. For the 2nd MP2 derivatives (13) a factor of 2.6 is observed, which is mainly due to the fact that the G-90 2nd MP2 derivatives are calculated numerically.

# 6 Summary and Conclusion

Two commercial available molecular electronic structure software packages have been compared according to a benchmark suite, which was designed to cover the typical range of calculations commonly performed by *ab initio* computational chemists.

The benchmark results show that in most of those cases where G-90 is applicable, G-UK clearly outperforms the commonly used G-90 with one important exception. Direct-SCF calculations, which do play an important role in modern electronic structure applications, perform slightly better with G-90.

The performance and applicability of G-90 for electron correlation studies is disappointing. CASSCF calculation are not manageable even when medium sized active spaces are applied. Single reference CI-calculations perform very badly, and multi-reference treatments are not possible. These are severe defficencies of the program package especially when systems are considered which are not properly described by a single determinant, as in the case of biradicaloid systems. Using the unrestricted CI scheme does not help in many cases, because the results may be drastically changed due to spin-contamination which has been demonstrated for many systems.

GAMESS-UK offers a larger functionality and better performance than the standard molecular electronic structure software package GAUSSIAN90. GAMESS-UK reflects modern trends in hardware for vector and parallel computers as well as in method development.

We believe that GAMESS-UK is indeed a very appealing alternative to the well-known GAUSSIAN90.

# References


[1] Gaussian 90, Revision I, M. J. Frisch, M. Head-Gordon, G. W. Trucks, J. B. Foresman, H. B. Schlegel, K. Raghavachari, M. Robb, J. S. Binkley, C. Gonzalez, D. J. Defrees, D. J. Fox, R. A. Whiteside, R. Seeger, C. F. Melius, J. Baker, R. L. Martin, L. R. Kahn, J. J. P. Stewart, S. Topiol, and J. A. Pople, Gaussian, Inc., Pittsburgh PA, 1990.

[2] M.F. Guest, GAMESS-UK User's Guide and Program Reference Manual, Daresbury 1990

[3] M.F. Guest, SERC Daresbury Laboratory, Warrington WA44AD