

分子軌道法計算（その7）

- Gaussian 03による金属錯体の計算 -

和佐田（筒井）祐子

. はじめに

千変万化する物質の変化を予想する手だてではないか。これは、万物が流転すると人類が認識して以来の願望です。20世紀になって、量子力学が発展した結果、物質の性質や変化は、電子の運動の現れであることがあきらかになり、物質の問題は、電子の運動を記述したSchrödinger方程式を解く応用数学にすぎないとまで言われました。

しかし、同時に、数学は水素原子のような限られた系を除くと正確には解けない問題であることを示しています。この結果、電子状態の近似解を与えるさまざまな方法が開発されました。現在、コンピュータの計算能力を最大限に利用した方法として、Hartree-Fock方程式を基礎とする非経験的分子軌道法とKohn-Sham方程式を解く密度汎関数法が広く利用されています。

Gaussianシリーズは、物質を扱うあらゆる分野の人々が、物質に対するイメージそのままにこれらの方法を利用できるように発展してきました。Gaussianでは物質の構造を迅速に決定し、スペクトルデータなどの実測値と直接比較できる量を解をもとにプログラム内で計算して出力します。それゆえ、現在では、初心者から専門家までの幅広い利用者層に愛用される計算プログラムとして全世界的に利用されています。

ここでは、Gaussianシリーズを初めて利用する人と、これまでGaussianシリーズを利用してきたがGaussian 03は初めてという人を想定しています。また、今回のバージョンアップで特に身近になったと思われる開殻系の計算に注目し、常磁性金属イオンを含む化合物 錯体 の計算を例にして話をすすめてゆきます。方法論や利用方法の基本は以前のバージョンと何らの変化もないので、詳しくは以前のセンターニュースなど⁽¹⁾⁽⁸⁾を参照してください。

. Gaussian 03の新機能

Gaussianシリーズは、バージョンを重ねるごとにプログラムが改良されるだけでなく、可能な計算方法を増やしてきました。もともと分子軌道法のプログラムであったGaussianに密度汎関数法がGaussian 92/DFTとして加わり、Gaussian 98からは、分子力学が組み込まれ、総合化学プログラムの感が強くなりました。Gaussian 03での主要な新機能については、以下の項目が挙げられます。分子軌道法プログラムとしての新機能には、

対象になる物質の拡張。これまでの有限な大きさの原子、分子やクラスターだけではなく、周期構造を有する高分子、表面、結晶、多量体に対象を広げました。

電子励起状態の計算方法の多様化。Restricted Active Space Multiconfiguration Self Consistent Field (RASSCF) 法が加わり、Complete Active Space Multiconfiguration Self Consistent Field (CASSCF) よりも配置関数を絞り込んだ励起状態の計算が可能になったうえ、CASSCF やRASSCF波動関数を用いてNatural Bond Orbital (NBO) による結合状態の解析が可能になりました。また、内殻励起状態に有用な、京都大学の中辻らによるSymmetry Adapted Cluster/Configuration Interaction (SAC-CI) 法が付け加わりました。

重原子の相対論効果。金や白金などの重原子の化学で重要な相対論効果は、従来、内殻電子を有効内殻ポテンシャル (effective core potential, ECP) で近似して評価していましたが、Douglas-Kroll-Hessによるスカラー相対論ハミルトニアンを用いた全電子計算が可能になりました。post-SCF計算や構造最適化、振動解析が可能です。

この他にも、分子軌道法の草分けで、現在も軌道相互作用の理論に活発に利用されている拡張Hückel法もパラメータが整備されました。また、電子相関の効果を基底関数の拡張や方法論の展開により外挿して少ない計算労力で正確なエネルギー値を計算するG3法が組込まれました。

分子軌道法以外の新機能としては、分子動力学による定量的な反応機構の追跡が可能になりました。また、Gaussianでは測定値と直接比較できる物理量を算出できます。NMRのスピン スピンカップリング、赤外線吸収スペクトルの非調和項、調和及び非調和の振動 回転カップリング、紫外、可視領域での円偏光二色性、磁化率、分極率、超分極率のほか、入射光に活性が依存する共鳴Ramanスペクトルの活性などが計算できます。

このような新機能が付け加わったことで、有機分子などでの計算によく利用されていたGaussianが、金属元素を含む化合物や固体などの集合体の物性をも対象とする、物質科学の総合計算プログラムとしてさらに発展してきたといえます。

・ Gaussian 03によるSCFの収束の向上

他の非経験的分子軌道法計算プログラムと比較しても、Gaussianシリーズは以前からSCFエネルギー計算の収束が速いことが特徴でした。これは、他のプログラムに比べて初期値がよいこと、PulayのDirect Inversion in the Iterative Subspace extrapolation method (DIIS) 法をはじめとする多様な収束アルゴリズムが用意されていることによります。初期値や収束アルゴリズムの詳細については、それぞれ、User's ReferenceのGuessやSCFの項目とその参考文献を参照してください。しかし、ラジカルや遷移金属錯体などの開殻系については、基底状態に近い励起状態が存在するなどの理由で必ずしも収束状況はよくなく、発散するならまだしも、収束はしたものの、不適切な高い電子状態が得られることもたびたびでした。実際筆者らの $[\text{Cu}(\text{H}_2\text{O})]^{2+}$ のエネルギー計算は、Gaussian 94以前のプログラムでは普通の方法では収束しませんでした。Gaussian 98で通常用いられる拡張Huckel法による波動関数を初期値とするDIIS法では、SCFエネルギーが上昇しながら収束した結果、本来ならばCu上に存在しなければならないスピンの、二個の酸素の上に半分ずつ存在するなど、誤った電子状態を与えました。しかし、Gaussian 03では、初期値に配位結合などの共有結合よりも弱い相互作用を簡便に記述できるHarrisによる汎関数を用いることで、特に金属イオンを含む化合物などでは初期値が飛躍的に向上して収束しやすくなりました。また、

DIIS法に対しても大幅な改良が加えられました。閉殻有機分子などのSCFサイクル20回以内よりは時間がかかりますが、 $[\text{Cu}(\text{H}_2\text{O})]^{2+}$ では、73回で正しい電子状態にほぼ一様に収束します。また、Gaussian 98のB3LYPにおいてDIIS法ではSCFサイクル160回で収束した $[\text{CuPt}(\text{H}_2\text{O})_4(\text{NH}_3)_2(\text{C}_6\text{H}_5\text{N})]^{2+}$ も73回で収束しました。また、計算時間もGaussian 98では、9時間37分であったのに対し、Gaussian 03では4時間35分(経過時間比)であり、収束の速さが計算時間の短縮に結びついています。ただし、後述するquadratically convergent SCFを必要とする $[\text{Ti}(\text{NCH})]^{2+}$ は収束しませんでした。収束しにくい化合物が何でも収束しやすくなったわけではないのですが、以前よりも開殻系を計算しやすくなったといえます。

・ Gaussian 03による並列計算

Gaussian 98でも並列化による計算時間の短縮がかなり進んでいました。通常よく利用されるHartree-Fock法や密度汎関数法(DFT)励起状態の計算に使用されるCASSCF法やsingle-excitation CI(CIS)法がかなり効率よく並列化されていました。また、センターのVPP5000(vpp)や研究室でよく利用されているPCクラスターマシンでは、手軽なpost-SCF計算として利用されているMøller-Plessetの二次摂動法(MP2)もまたLindaにより効率よく並列化されていました。Gaussian 03では、GP7000(gpcs)のような、Lindaを利用できないコンピュータでもMP2の並列計算が可能になったり、これまで並列化されていなかった計算が並列化されたりするなど、並列計算による高速化が一段と進んでいます。

ここでは、実際の研究で行われる程度の規模の並列計算を、以前解説した計算例⁽⁷⁾を交えながら、センターのグリッドコンピュータHP2500(ngrd)で実行した結果を示します。ngrdではCPU時間が正確に測定できないので、経過時間で計算時間を評価しました。また、各計算について計算時間が長いプログラム単位(Link)について、実際の計算時間と計算速度比((並列計算速度)/(1CPUでの計算速度))を表に示しました。計算速度比がCPU数に近いほど並列化効率が良いといえます。

(1) 有機リン化合物 $\text{C}_{19}\text{H}_{34}\text{NO}_4\text{P}$ に対するB3LYP/6-31G(d)による振動解析計算、 C_1 対称性、59原子、s, p, d軌道を含む447軌道⁽²⁾

振動解析計算は、分子振動をそのまま利用する赤外線吸収やRaman散乱などの振動スペクトルの帰属のみならず、平衡定数や反応速度といった、化学反応の方向を判断する熱力学的パラメータの算出、最適化された構造が適切なエネルギー極小点に存在するか、正しい遷移状態を特定したのかの判断など、理論的な手法により構造や反応経路を議論するのに必須です。このように重要な振動解析計算ですが、大きな分子や開殻系では、膨大なメモリーと計算時間が要求されます。近年、小分子を対象として発達してきたHartree-Fock法と、結晶や金属錯体を対象として発達してきた密度汎関数法を融合させて、実験結果をよく再現できるように開発されたB3LYPが6-31G(d)基底関数との組み合わせで中程度の大きさの分子系の電子状態計算によく利用されています。59原子447軌道の C_1 対称性の有機分子のB3LYP/6-31(d)による振動解析計算に対する1, 2,

4, 8 及び 16 CPU 並列化の効果について表 1 に示しました。

エネルギー計算を行う L502 と各原子の座標にかかる力の計算を行う, L703 の並列化効率は 16 CPU まできわめてよいのに対し, もっとも時間が長い L1110 や L1002 などの力の定数の計算では 8 CPU 以上で計算効率が下がっています。このため, 全体時間では, CPU 数を増やせば速くなるものの, 16 CPU 計算は理想的な並列化で期待される 9 CPU 分の働きしかありません。

なお, 通常の構造最適化ではこれらの時間のかかるオーバーレイは利用されないため, L502 と L703 による高い並列化効率が 16 CPU 程度まで利用できることとなります。また, vpp の Gaussian 98/Linda 並列では, 2 CPU で 11 時間半ほど, 8 CPU で 4 時間弱と予想されるので⁽⁹⁾, グリッドコンピューティングによる並列化計算はかなり速いといえます。

表 1 B3LYP/6-31G (d), C₁₉H₃₄NO₄P での計算時間

CPU 数	1	2	4	8	16
実時間					
L502	0:57:58	0:27:59	0:14:28	0:07:54	0:04:13
L1110	8:18:20	4:08:53	2:30:20	1:22:14	0:56:05
L1002	17:37:03	8:58:06	5:36:31	3:10:51	2:08:25
L703	4:25:31	2:14:51	1:07:07	0:35:31	0:17:33
全体	31:21:21	15:51:13	9:29:22	5:17:16	3:27:00
並列処理で実際に利用された平均 CPU 数 (計算速度比)					
L502	1.0	2.1	4.0	7.3	13.7
L1110	1.0	2.0	3.3	6.1	8.9
L1002	1.0	2.0	3.1	5.5	8.2
L703	1.0	2.0	4.0	7.5	15.1
全体	1.0	2.0	3.3	5.9	9.1

SCF は DIIS 法で 15 回収束し, Gaussian 98 よりも 2 回少ないが同じエネルギー値に収束した。

メモリー使用量は 800 MB, ディスク使用量は 1307 MB であった。

(2) UHF 法による [F α (NCH)]²⁺ の構造最適化, SCF=QC, C_i 対称性, 22 原子, s, p, d 軌道を含む 265 軌道

エネルギーの収束が悪い遷移金属錯体などの開殻分子では, quadratically convergent SCF は, しばしば最後の手段として用いられます。この方法は, 通常利用される DIIS 法よりも収束しやすいのですが, 時間がかかること, あまりにも悪い初期値を使用すると最小ではない解に強制的に収束させてしまう危険があることなどから, 利用前に他の収束法での収束状況が発散方向にないことを確認するなどの慎重さが必要です。[F α (NCH)]²⁺ の非制限 Hartree-Fock (UHF) 法レベルでの構造最適化計算で最も時間がかかる L508 に対する 1, 2, 4 及び 8 CPU 並列化の効果につい

て表 2 に示しました。

SCFの回数がGaussian 98の場合と同様にCPU数ごとに大きく異なるので、CPU数の増加が単純に高速化に結びつかないのですが、計算時間をSCF回数について平均した結果について比較すると、4 CPUで2.3倍程度しか計算速度が向上しなかったGaussian 98/Lindaよりも並列化効率がよく、8 CPU程度までの並列化が良好です。このため、同じCPU数でVPPと比較すると、1 CPUでは2.2倍の時間がかかると予想されるのに対し、4 CPUでは1.3倍程度、8 CPUでは1.1倍程度と期待され、vpp程度の高速計算が期待できます。

表 2 UHF (SCF=QC) による $[F\alpha(NCH)]^{\dagger}$ の計算時間

CPU数	1	2	4	8
全体の時間	4:52:54	3:52:51	1:49:17	1:03:11
SCF回数	39	61	56	54
L508の実時間	4:48:45	3:50:34	1:48:05	1:03:11
L508の1回あたり時間	0:07:24	0:03:47	0:01:56	0:01:10
並列処理で実際に利用された平均CPU数 (計算速度比)	1.0	2.0	3.8	6.3

Gaussian98と同じエネルギー値に収束した。

メモリー使用量は200MB、ディスク使用量は47MBであった。

(3) MP 2 による $[T(NCH)]^{\dagger}$ の構造最適化, C_2 対称性, 22原子, s, p, d軌道を含む265軌道^(4,7)

Hartree-Fock法では十分な評価がなされていない電子相関を簡便に扱える方法として、MP 2 法はよく利用されています。DFT計算が普及する以前には、有機金属錯体等でMP 2 による構造最適化が必要ともされてきました。B3LYPでは過大評価しやすい弱い相互作用の研究などでよく利用されています。

Gaussian 98では、最も時間を要するL906がvppのLindaで並列化されていましたが、gpcsでは並列化が行われていませんでした。Gaussian 03では2 CPUまでの並列化が良好で、4 CPUではvpp Gaussian 98/Lindaの2.5倍 (実時間比) よりも効率が悪くなります。

表3 UMP2による $[\text{T}(\text{NCH})]^{2+}$ の計算時間

CPU数	1	2	4
実時間			
L502	0:14:33	0:07:35	0:04:01
L906	2:53:36	1:43:59	1:19:46
L1002	0:19:01	0:09:45	0:05:03
L703	0:03:49	0:01:55	0:01:00
全体	3:31:37	2:03:38	1:30:08
並列処理で実際に利用された平均CPU数 (計算速度比)			
L502	1.0	1.9	3.6
L906	1.0	1.7	2.2
L1002	1.0	2.0	3.8
L703	1.0	2.0	3.8
全体	1.0	1.7	2.3

Gaussian98と同じエネルギー値に収束した。

メモリー使用量200MB, ディスク使用量16115MBであった。

(4) MP2による $[\text{CrC}(\text{H}_2\text{O})]^{2+}$ の振動解析計算, C_2 対称性, 20原子, s, p, d軌道を含む182軌道

手軽なpost-SCF計算のMP2がB3LYPほど利用されない大きな理由のひとつとして, 振動解析計算に膨大な計算機資源を必要とするので, 分子振動を計算できないことが挙げられます。しかし, 名古屋大学情報連携基盤センターの豊富な計算機資源を利用すれば, MP2振動解析計算は20原子程度の分子でも可能です。

Gaussian 98ではLindaでも並列化されていませんでしたが, Gaussian 03では効率が悪いながらも並列化が行われています。2 CPU並列に比べると4 CPUでは2割程度, 5時間の計算時間の短縮がみられ, 並列化は必要であると考えられます。

表4 UMP2による $[\text{CrCl}(\text{H}_2\text{O})_5]^{2+}$ の計算時間

CPU数	1	2	4
実時間			
L906	4:52:00	2:33:34	1:27:13
L811	10:08:05	6:59:39	5:34:46
L1112	22:46:20	16:44:03	14:16:41
全体	39:41:53	27:26:51	22:13:15
並列処理で実際に利用された平均CPU数 (計算速度比)			
L906	1.0	1.9	3.3
L811	1.0	1.4	1.8
L1112	1.0	1.4	1.6
全体	1.0	1.4	1.8

メモリー使用量は1200MB, ディスク使用量は22951MBであった

(5) CISによる $[\text{N}(\text{CH}_3\text{CN})_5]^{2+}$ の第50励起状態までのエネルギー計算, C_1 対称性, 36原子, s, p, d軌道を含む396軌道⁽⁵⁾

紫外線や可視光の吸収スペクトルや光化学反応などの電子励起状態を研究する第一段階として, CIS法は広く利用されています。計算時間が長いオーバーレイは, 積分の変換を行い一電子励起CI行列を作成するL804, この行列の対角化を行うL914です。しかし, L804である程度並列化されているものの, L914が並列化されていないために効率はよくなく, 励起状態の構造最適化などの計算時間が長い場合には有用という程度です。

表5 CISによる $[\text{N}(\text{CH}_3\text{CN})_5]^{2+}$ の第50励起状態までの計算時間

CPU数	1	2	4
実時間			
L804	10:42:13	6:31:24	4:57:18
L914	2:38:53	2:31:30	2:50:04
全体	13:25:39	9:05:31	7:48:55
並列処理で実際に利用された平均CPU数 (計算速度比)			
L804	1.0	1.6	2.2
L914	1.0	1.0	0.9
全体	1.0	1.5	1.7

メモリー使用量は800MB, ディスク使用量は50671MBであった。

(6) 2電子4軌道のCASSCFによる $[T(NCH)]^{\ddagger+}$ の第一励起状態のエネルギー計算, D_{5h} 対称性, 22原子, s, p, d軌道を含む265軌道^(5,7)

CASSCFは単一の配置関数では表現できない基底状態や励起状態の電子状態を研究するのに最もよく利用されています。しかし, active spaceの電子や軌道が多すぎると計算が大変になるので active spaceを選別する必要があります, 計算結果が利用者の化学的直感に依存すること, また, 収束が悪いのも難点です。計算規模を縮小できるようにしたのがRASSCFです。Gaussian 98でもgpcsのLindaを利用しない計算での並列化の効率が良好でした。ngrdのGaussian 03の8 CPU並列化効率率はgpcsと同程度です。また, Gaussian 98と異なり, CPU数が異なってもSCFの回数が増えないので, 並列化の効果がCPU数とよく相関しています。また, gpcs上でGaussian 98とGaussian 03で同じ計算を行ったところ, それぞれ65回と33回で収束し, SCF 1回あたりの時間は10分16秒と10分21秒でそれほど差がなかったことから, 安定した速やかな収束により計算が速くなったと結論されます。ngrdのSCF 1回あたりの時間はgpcsと比較して0.15倍ときわめて高速になっています。

表6 CASSCF (2, 4), state averageによる $[T(NCH)]^{\ddagger+}$ の第一励起状態の計算時間

Link	1	2	4	8
実時間				
L510	1:38:23	0:49:19	0:25:51	0:14:41
全体	1:38:35	0:49:32	0:26:04	0:14:54
並列処理で実際に利用された平均CPU数 (計算速度比)				
L510	1.0	2.0	3.8	6.7

Gaussian98と同じエネルギー値に35回で収束した。

メモリー使用量は90MB, ディスク使用量は53MBであった。

・ センターでのGaussian 03の利用法

Gaussian 03はセンターのvpp, gpcs及びngrd1, ngrd2 (以下ngrdとする) にインストールされています。ただし, 2004年4月現在, vppでは並列化効率がよくない計算についてベクトル計算機の高速度性を利用したい場合など, 非並列での利用のみとなります。先の例ではMP 2での振動解析計算が該当します。このため, 並列化による計算速度の向上を期待したい場合には, gpcsとngrdではGaussian 03を利用できますが, vppでは以前と同じくGaussian 98を用いてください。

(1) 入力ファイルの作成

Gaussianの入力データには, Gaussian 94や98と同じデータが利用できます。詳細な解説については以前のセンターニュースを参照してください。例えば, $[Fe(H_2O)]^{\ddagger+}$ の構造最適化では図1に示すデータを用意する必要があります。

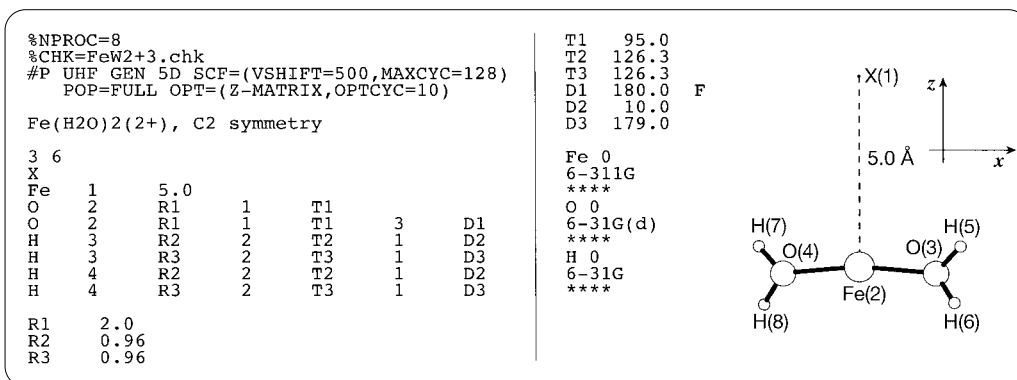


図1 $[\text{Fe}(\text{H}_2\text{O})_2]^{2+}$ のUHFによる構造最適化計算の入力データ

一行目の%NPROCは並列計算に用いるCPU数を指定します。ここでは8 CPUを利用しています。非並列計算ではこの行は必要ありません。また、vppやPCクラスタの入力ファイルではCPU数の指定が%NPROCLとなっていますが、流用する場合にはLindaを使用しないのでLを除いてください。二行目の%CHKでは、チェックポイントファイルと呼ばれる、基底関数などの計算条件や最適化構造や分子軌道などの計算結果を格納するファイルを指定します。これは省略できますが、チェックポイントファイルを保存しておくこと、計算結果を利用してつぎの計算を行うときなどに再入力する手間が省けること、SCF計算や構造最適化計算が収束しなかった場合に収束最終回から再実行できるなどの利点があります。このような%で始まる命令はLink 0 コマンドと呼ばれています。#で始まる行はルートカードと呼ばれ、電子状態の計算方法など行いたい計算を指定します。例では、電子状態の計算法はUHFです。GENは基底関数を原子ごとに指定するキーワードで、基底関数の情報は構造の情報の下に空白行をおいてから入力します。5Dのオプションは、d軌道を5軌道で展開することを意味します。⁽²⁾ SCFのオプションのVSHIFTは初期空分子軌道のエネルギーを指定した値（ここでは500 mHartree）だけ上昇させ、SCF手続きの途中で被占軌道と混ざりにくくします。接近した電子励起状態がある場合には有効な方法です。収束の悪い錯体の計算ではMAXCYCオプションでSCFサイクルを64回以上に設定します。また、構造最適化のキーワードとオプションOPT=Z-MATRIXは、Zマトリックス指定された内部座標について構造最適化を行うことを示しています。また、電子状態の判断の材料になる電荷密度や分子軌道をすべて出力するためのPOP=FULLを指定します。開殻系の計算で重要なことは#の後ろにPをつけることです。Pは詳細モードで出力を行うことを意味します。これでSCFの各サイクルごとのエネルギー値が出力されることになり、収束や発散の判断の材料になります。Pなしで収束しない場合には何も出力されないまま時間ばかりが経ち、あるときConvergence failureのメッセージとともにエネルギーの値のみが出力されることになります。

空白行でルートカードを終了し、つぎにタイトルカードになります。計算についての情報、データを作成した日付など、後々のデータ整理に必要な情報を入れ、空白行で終了とします。

つぎに計算対象となる物質の電荷、スピン状態、構造についての情報を入力します。最初の36は物質全体の電荷及びスピン多重度で、+ 3価、六重項、すなわち スピンが スピンよりも

5個多いことを意味します。つぎに構造を指定します。Zマトリックスで指定する場合には、図のように、構造最適化したい内部座標変数名を含むZマトリックスと内部座標の初期値データの順に指定します。なお、Zマトリックス指定の詳細については、参考文献を参照してください⁽¹⁾。X線構造解析の結果得られた座標を利用したい場合には、デカルト座標による指定も可能です。この場合には、上と同じ構造を例とすると、

Fe	0.000	0.000	0.000
O	1.992	0.000	0.174
O	-1.992	0.000	0.174
H	2.625	0.134	-0.535
H	2.491	0.014	0.994
H	-2.625	-0.134	-0.535
H	-2.491	-0.014	0.994

のように指定し、OPT=Z-MATRIXの代わりに単にOPTと指定します。また、X線構造解析の結果には水素がなかったり、原子がdisorderを起こしたりすることがありますが、これらの原子の座標だけを、Zマトリックスにより指定することも可能です。例えば、銅イオンに配位した水分子の水素の位置をZマトリックスで指定します。

Cu	0	0.123	-0.023	2.543	
					: 10原子の座標
O	0	0.248	0.056	4.697	
					:
H	0	4.455	1.712	-2.695	
H	12	0.979	1 108.53	43 180.0	
					:

注意しなければならないのは、混合座標指定を用いる場合には、デカルト座標で指定されている部分では元素記号の直後に「0」が必要なことです。また、Zマトリックス指定は最初の3原子以外のどこでも可能ですが、内部座標で参照する原子の座標がすべて定義された後に指定します。上の例では、Zマトリックスで指定する水素原子は、12、1、43番の原子を参照しているため、44番以後に指定する必要があります。この場合にもOPT=Z-Matrixの代わりにOPTを指定します。空白行を一行入れて座標指定は終了とします。

(2) スクリプトファイルの作成

研究室のワークステーションなどでは、ここでg03コマンドを実行すればプログラムが入力ファイルにしたがって動作します。しかし、共同利用のセンターで長時間にわたって大量のメモリーやディスクを利用しながら計算する場合には、ジョブ管理システムNQSにジョブを投入して実行します。NQSを利用するには、コマンドを記述したスクリプトファイルが必要です。Gaussian 03はGaussian 98と同様なスクリプトファイルで実行できます。ここではgpcsを例に説明します。

```
# @$-q m -eo -o g03.err -lT 12:00:00 -lp 8
cd bench
g03 few2.dat few2.out
```

図2 スクリプトファイルの例

一行目ではmキューに12時間のジョブを8 CPU並列で実行し、システムからのメッセージやエラーをg03.errに出力することをNQSに知らせています。10時間以上 (ngrdのgzでは100時間以上) の計算時間が予想される場合には、キューの制限値を越えない程度に少し長めの時間を指定する必要があります。また、この時間は、各CPUで利用した時間の合計値であり、10時間の指定で2 CPUをフル稼働すると5時間しか計算しないことにも注意が必要です。なお、非並列ジョブでは-lp以下は必要ありません。二行目は、入出力ファイルをディレクトリbenchに作成するために、このディレクトリに移動しています。ログインディレクトリの直下に入力ファイルを作成した場合にはこの行は必要ありません。三行目ではg03コマンドを実行しています。ここでは、入力ファイルをfew2.dat、出力ファイルをfew2.outとしています。研究室のワークステーション等のg03と異なるのは、入力ファイルの前に「<」、出力ファイルの前に「>」をつけず、単に空白で区切るようにすることです。グリッドコンピュータの場合には、使用するキューをgyまたはgzにする他は同じです。vppの場合にはxとします。

(3) ジョブの実行と確認

入力ファイルとスクリプトファイルが用意できたらコマンドラインからNQSにジョブを投入します。スクリプトファイル名はg03.cmdとします。

```
gpcs% qsub g03.cmd
Request 27249.gpcs submitted to queue: m.
```

投入されたジョブは、他のジョブで混んでいる場合など必ずしもすぐには実行されません。メッセージは単にジョブをNQSがmキューに受け付け、IDを27249.gpcsとしたことを伝えているだけで実行したとは伝えていないのです。ジョブの状況はqstatコマンドで確認します。

```

gpcs% qstat m
m@gpcs;  type=BATCH;  [ ENABLED, RUNNING ];  pri=16
  0 exit;   2 run;   0 queued;   0 wait;   0 hold;   0 arrive;

      REQUEST NAME          REQUEST ID          USER PRI          STATE PGRP
1:      g98.cmd             27234.gpcs-g       w49108a 31        RUNNING 4702
2:      g03.cmd             27249.gpcs-g       w49108a 31        RUNNING 19755

```

STATEがRUNNINGの場合は実行されています。また、qstat -aとすると、すべてのキューについての混雑状況がわかります。

実行しているジョブは、出力ファイルの内容をmoreコマンドなどで確認し、望んだとおりの計算が行われているか否かを確認します。ジョブを停止する必要がある場合には、ジョブのREQUEST IDをqstatコマンドで確認します。前述のジョブではIDが27249ですので、

```

gpcs% qdel -k 27249
Request 27249.gpcs is running, and has been signalled.

```

とします。

. Gaussian 03の新機能の利用法

入力データの作成方法はGaussian 98とかわらないとはいうものの、Gaussian 03で新たに付け加わった計算プログラムを利用したい場合には、入力データをどのように作成したらよいか途方にくれるところです。Gaussian 03のプログラム群があるg03ディレクトリの下のtestsディレクトリの下comディレクトリに、入力例があります。gpcsやngrdでは、/opt/local/g03/g03/tests/comにあり、いずれの機械でも同じ内容です。695個の入力ファイルから目的とする入力例を探すには適当なキーワードについてgrepコマンドで検索します。例えば、相対論計算を行いたい場合には、「relativ」などのキーワードで検索します。

```

gpcs% cd /opt/local/g03/g03/tests/com
gpcs% grep 'relativ' *.com
test566.com:GaF B3LYP, scalar relativistic
test566.com:GaF B3LYP, scalar relativistic, point nuclei

```

test566.comが目的の計算例であることがわかります。このように、入力ファイルのタイトル行にキーワードが含まれていれば、検索されてきます。SAC-CIなどの略称でなければ、ルートカードのオプションであっても小文字で検索します。不安な場合には-iオプションをつけてgrep -i 'relativ' *.comとし、大文字小文字の区別をなくしてください。検索されたファイルをコピーして自分の課題で実行できることを確認してから編集して利用します。

また、Gaussian社のホームページにオンラインマニュアルが用意されていますので、参照して

ルートカードの内容を確認しながら利用するのがよいでしょう。GaussianのホームページのURLは

```
http://www.gaussian.com/
```

であり、Tech SupportをクリックするとキーワードやIOPなどの項目が選べるようになります。

. Gaussian 03でのチェックポイントファイルの利用

複雑な分子構造やようやく収束した分子軌道，長時間の計算の後に得られた力の定数など，チェックポイントファイルには膨大な情報が格納されています。研究室のコンピュータで長い時間をかけて最適化した構造などをもとに，センターの計算機資源をふんだんに利用して力の定数を計算し，この力の定数を利用して研究室のコンピュータで効率よく再度構造最適化を行うといった計算をする場合には，チェックポイントファイルをコンピュータ間でやりとりして計算を進めるのが効率的です。しかし，バイナリファイルのチェックポイントファイルは，センターのgpcsとngrdの間などの異なったシステムにそのまま移すと読めなくなってしまうので，あらかじめformchkコマンドでテキストファイルに変換します。例えば，gpcsやグリッドコンピュータでは以下のように利用します。

```
%gpcs /opt/local/g03/g03/formchk FeW2+3.chk
Read checkpoint file FeW2+3.chk
Write formatted file FeW2+3.fchk
Rotating derivatives, DoTrsp=T IDiff= 2 LEDeriv=      407 LFDPrp=
0 LDFDPr=      0.
```

FeW2+3.chkはテキストデータのFeW2+3.fchkに変換されます。このfchkファイルを計算を行いたいコンピュータに転送し，バイナリのチェックポイントファイルに戻します。

```
%gpcs /opt/local/g03/g03/unfchk FeW2+3.fchk
Read formatted file FeW2+3.fchk
Write checkpoint file FeW2+3.chk
```

この例では，チェックポイントファイルとしてFeW2+3.chkが作成されるので，これを計算に利用します。この変換により新しいチェックポイントファイルは元よりも小さくなります。このファイルには，ZマトリックスやECPの情報が含まれていませんので，これらの情報を必要とする場合には，入力ファイルにZマトリックスやECPを再入力する必要があります。なお，このようなデータの劣化がないchkmoveコマンドはGaussian 03からはサポートされなくなりました。

一方，同じシステムのGaussian 98またはそれ以前バージョンのチェックポイントファイルは，このようなデータの劣化を起こすことなくGaussian 03で利用できます。古いバージョンのファイルGauss98.chkを変換したい場合には，

とすると、Gauss98.chkがGaussian 03のチェックポイントファイル形式に書き換えられます。

．おわりに

計算による方法では、実験による方法と異なり、既存の物質だけではなく、反応の遷移状態や未確認の中間体、不安定な励起状態といった化学変化のキーポイントになる状態や、思考実験としてのモデル化合物といった実在しない物質の電子状態をも対象として物質の変化を議論できるという利点があります。とくに、実測データをパラメータとして利用しない非経験的分子軌道法は、測定実験可能な物質と不可能な物質を扱えるので、実験を繰り返すことなく物性や反応性を予測して物質を設計できるのではという期待も高まります。しかし、測定条件を誤った測定値は実験事実として信頼してはならないのと同様に計算結果についても正しい解であるか否かについては慎重な検討が必要です。例えば、電子状態が正しいか否かの検討は、低い励起状態が存在する開殻系の計算では常につきまといます。⁽¹⁰⁾ Gaussianのアルゴリズムや計算機の性能の向上により計算対象が多様化すればするほど、さらに、物質科学的に高度にとぎすまされた直観力が要求されるのです。このような困難を克服し、さまざまな物質の性質を電子状態に基づいて解析し、自然現象の本質をつかむ日が来ることを願ってやみません。

謝 辞

グリッドコンピューティングにあたり、実行時のトラブルの処理等についてさまざまな便宜を図っていただきました、名古屋大学情報連携基盤センターの津田知子先生と高橋一郎氏に深く感謝いたします。また、VPP800/63のGaussian 98による計算処理データの作成にあたり、多大な便宜を図っていただきました京都大学学術情報メディアセンターの小澤義明氏に深く感謝いたします。

参考文献

- (1) 山辺信一，和佐田裕昭，筒井祐子：名古屋大学大型計算機センターニュース，Vol. 29, 19 (1998)
- (2) 和佐田裕昭，筒井祐子，山辺信一：名古屋大学大型計算機センターニュース，Vol. 29, 88 (1998)
- (3) 筒井祐子，山辺信一，和佐田裕昭：名古屋大学大型計算機センターニュース，Vol. 29, 197 (1998)
- (4) 山辺信一，和佐田裕昭，筒井祐子：名古屋大学大型計算機センターニュース，Vol. 29, 277 (1998)
- (5) 和佐田裕昭，筒井祐子，山辺信一：名古屋大学大型計算機センターニュース，Vol. 30, 16 (1998)

- (6) 和佐田 (筒井) 祐子 , 山辺信一 , 和佐田裕昭 : 名古屋大学大型計算機センターニュース , Vol. 31, 195 (2000)
- (7) 和佐田 (筒井) 祐子 : 名古屋大学大型計算機センターニュース , Vol. 32, 192 (2001)
- (8) 古賀伸明編 : 第 5 版実験化学講座12計算化学 , 丸善 (2004)
- (9) 京都大学学術情報メディアセンターの旧スーパーコンピュータシステムVPP800の計算時間の0.83倍としてVPP5000の計算時間を見積もった値。以下のVPP5000の計算時間は , この見積りを利用しています。VPP800による計算時間は以下のURLの「計算結果比較」を参照して下さい。
- http://www.kudpc.kyoto-u.ac.jp/service/Application/Gaussian98/g98_report.html
- (10) T. Bally, W. T. Borden, “Reviews in Computational Chemistry”, edited by K. B. Lipkowitz and D. B. Boyd, Wiley-VCH, New York, vol. 13, 1 (1999)

(わさだ (つつい) ゆうこ : 名古屋工業大学大学院工学研究科)