

1Pa035

大規模ab initio分子軌道計算における2電子積分超行列に関する研究

(NTT物性基礎研・ATR適応研)

○寺前裕之、下川信祐、大田原一成

Hartree-Fock分子軌道法は1970年代頃から色々な分子の電子状態計算に適用されるようになったが、より現実的な大きさの分子への適用を考えた場合に分子のサイズが大きくなると急激に増加する二電子積分をどのように取り扱うかが大きな問題となる。

Gaussian70やHondo5のような初期のプログラムでは、二電子積分を4つの添字と共にファイルに書き込み、SCF過程における繰り返し計算ではそのファイルを読み込んで利用するようにプログラムされていた。外部記憶に書き出された二電子積分が添字 r, s, t, u と共に読み出されて一つの二電子積分についてFockの行列要素に対して6種類の寄与分が計算される。

RaffenetiはこのFockの行列に対する二電子積分の寄与を計算するのに、より効率の良い方法としてPK超行列として知られる方法を提案した。以下ではこれをPK法と呼び、用いなかった場合についてNOPK法と呼ぶことにする。

$$I_{rstu} = \langle rs | tu \rangle - \frac{1}{4} \langle rt | su \rangle - \frac{1}{4} \langle ru | st \rangle$$

と線形結合を取り直せば、

$$F_{rs} \cdots P_{tu} I_{rstu}$$

$$F_{tu} \cdots P_{rs} I_{rstu}$$

と演算が大幅に簡略化できることになり、線形結合を取り直す前と後で二電子積分の数が変わらず、線形結合の取り直しに要する時間が十分に短いなら、演算が6通りから2通りに減少するため、かなりの高速化が見込まれる。PK法は実際にもうまく働いたので、Hondoには当初から組み込まれており、またGaussianにも組み込まれている。ただし現在ではdirect SCF法を用いて二電子積分を補助記憶にはストアしないのが主流であるため、どのような計算にもPK法が有効であるかは良く調べられていない。

一方、近年のコンピューター本体や周辺機器の性能の向上と価格の下落により、パーソナルコンピュータークラスターを使用し、Hartree-Fock分子軌道計算においても上述の二電子積分ならびにFockの行列要素 F_{rs} の計算に並列処理を利用する事が可能になってきた。従来では考えられなかった大容量のメモリーや高速なCPUを多数用いる事ができるために、旧来の常識的な考えが通用しない、すなわちパラダイムシフトが起こる場合がある。例えば、我々が以前に報告したように、並列処理を行うことにより二電子積分は各ローカルのハードディスクに分割されて保存されるので全体としての容量が増加し、また入出力も分割されたファイルに対して行われるのでこれも並列入出力となり負荷が大幅に軽減される。またオペレーティングシステムによりファイルの入出力はメモリー上にバッファーされるので、極端な場合にはファイルの入出力は各1回の書き込みと読み込みのみ行われて、残りの読み込みは実際には行われず

にメモリー上に保持されたまま行われるので、非常に高速な処理となる。この場合にはプログラムの書き直しを行う必要もなく簡便に高速化が達成できることとなる。従って、我々が行っている分子軌道計算では、direct SCF法よりも通常SCF法を用いた方が高速に処理できる場合がほとんどとなっている。通常SCF法では上で述べたように、二電子積分の取り扱いが重要であり、従ってPK法が大規模な計算においても果たして速度面で有利であるかどうかを調べることは重要である。

我々は近年、Hartree-Fock分子軌道法を用いた分子ダイナミクスの計算を行っており、通常一つの分子に対して1000-3000点程度のHartree-Fock分子軌道計算を行う。そのため一点あたりではわずかな差であっても最終的には大きな差となりうる。従って、本研究では我々は、中規模のPCクラスター環境を使用した並列処理計算において217-274基底関数までのHartree-Fock分子軌道計算を行い、PK法とNOPK法の経過時間およびCPU使用時間の比較を行ったので報告する。

8筐体で8CPUからなるPCクラスターを用いた並列計算を行った。OSにはRedhat Linuxを使用した。分子軌道計算にはGAMESSを使用した。並列化のライブラリはGAMESS付属の物をそのまま使用し、並列処理自体もGAMESSの元のコードのままとした。問題となる二電子積分の入出力の部分は並列化がなされている。速度の測定に用いた分子は、実際に市販されているベンゾジアゼピン(BZP)系およびチエノジアゼピン(TZP)系の向精神薬で、フルトプラゼパム、トリアゾラム、クロチアゼパム、エチゾラム、フルタゾラム、ロラゼパムの各分子である。基底関数には3-21G基底を使用した。計算時間の計測はSCF計算およびエネルギー勾配の一点計算を各10回を行い、実時間が最も速い値を採用した。

表にPK法とNOPK法とで計測した経過時間をクロチアゼパムについて示す。1CPUでの計算時はNOPK法がPK法に比べて実時間で0.35程度となっており、NOPK法が速く処理が終わることがわかる。さらにCPU使用時間とSystem処理時間の和を比べても、NOPK法が有利となっている。経過時間については、いずれの分子の場合にも、適当なノード数となってメモリー上に二電子積分が展開されると、NOPK法に比べてPK法の方がより経過時間が短くなり、結果的にPK法とNOPK法の差がなくなることがわかる。クロチアゼパムでは8CPU時に差がほぼ消失している。Fultazolamのみは8CPU時でもNOPK法がPK法に比べて0.59倍であり、差が解消されなかった。ただし、計算の並列度が上がるにつれ、両方で差が解消していく傾向があるのは興味深い。

表 通常フォーマット(NOPK)と PK フォーマットによる計算時間の比較

Molecule	NO PK				PK			NO PK/PK Ratio	
	N	CPU	SYS	Wall	CPU	SYS	Wall	CPU+SYS	Wall
clotiazepam	1	124.98	21.23	259.14	113.34	51.18	750.53	0.89	0.35
	2	64.69	8.63	94.58	59.71	26.06	320.55	0.85	0.30
	4	34.38	4.62	50.58	32.04	10.64	100.82	0.91	0.50
	8	19.20	2.72	32.50	18.10	4.84	33.74	0.96	0.96