

N体問題高速化アルゴリズムの検討

濱口信行

(株)日立製作所 ソフトウェア開発本部

N体問題では、演算量が N^2 に比例するため N の大きさに応じて適切なアルゴリズム、プログラミングも異なってくる。
本稿では、幾つかの N に対してアルゴリズムの検討を行い、テストケースに於いて一千万体問題で、1タイムステップ当たり、3秒で実行出来た経過を報告する。

Examination of accelerative algorithm in N-body problem

Nobuyuki Hamaguchi

Software Development Center , Hitachi Ltd.

In N-body problem, the while sum of operations is directly proportional to square of N , and suitable algorithm and programming differ with size of N .

IN this paper, I report progress that examination of algorithm for some sized N and ten millions body problem in test case is performed in three seconds per one time step simulation.

1. はじめに

N体問題では、その1タイムステップ当たりの演算量は、Nの二乗に比例し、およそその演算量E (FLOP)と Nの間には

$$E = 30 \times N^2 \quad \text{の関係があります。}^{1)}$$

そこで、1タイムステップ当たり3秒以内で実行出来る事を、'現実的に実行可能'とする基準を定め、各種計算機で実行可能なNを出来る限り大きくするアルゴリズムを検討しました。

まず、シミュレーション手法を粒子と粒子の相互作用を解く粒子-粒子(PP)法²⁾に固定し、汎用機、ワークステーション、スーパーコンピュータで実行可能なNを測定により求めました。

N=1000程度なら、計算機の機種によらず実行可能。N=10000程度なら S-3800 クラスのスーパーコンピュータで実行可能。

N=100000となると、S-3800 クラスのスーパーコンピュータでは実行不可能となる事を確認し、シミュレーション手法の見直しを行なった。

ポテンシャル ϕ を導入し、 ϕ による場から、粒子に働く力を求める粒子-格子(PM)法³⁾に変更して、適切なアルゴリズムを検討し、S-3800/180 クラスのスーパーコンピュータでN=10000000 (一千万体問題)が実行可能となりました。

本報告では、この最終結果にいたるまでの、検討内容と実測結果を示します。

2. P-P (PARTICLE-PARTICLE) 法での測定結果と解析

P-P法は、各粒子-粒子間に働く力を求める方法です。ある粒子(i)を考え、その質量を M_i としたとき、この粒子に作用する力 \vec{F}_i は

$$\vec{F}_i = G * M_i \sum_{\substack{j \\ j \neq i}}^N M_j / R_{ij}^3 * \vec{R}_{ij} \quad (R_{ij} = |\vec{R}_i - \vec{R}_j|) \quad G: \text{万有引力定数} \quad \vec{R}: \text{粒子の位置ベクトル}$$

と表わせます。

この式に忠実にコーディングを行ない、N体問題をシミュレートするのに要する時間を測定しました。測定方法は、100タイムステップ実行しそのタイムステップあたりに要する平均所要時間を1タイムステップの実行時間としています。

これ以降の記述では、1タイムステップの実行時間を実行時間と記述します。

2.1 N=1000 での実測結果

汎用機、ワークステーションでの実行時間は表1のようになります。

表1 汎用機、ワークステーションでの実行時間(秒)

CPU	M-680H	M-880H	3050RX (60MHz)
実行時間(秒)	2.28	1.07	1.25

この測定結果から判ります様に、このサイズのN体問題では、使用する計算機や、アルゴリズムは特に意識する必要がないと結論できます。

2.2 N=10000 での実測結果

S-3800/180 での実行時間は、2.56秒となり、実行可能となりました。

この値は、数GFLOPsの計算機では微妙な値(3秒以内で実行出来ないスーパーコンピュータも出てくる可能性がある)ですので、この手法で実行するときのプログラミングについて、若干の検討を行ないました。

粒子に作用する力の計算では、平方根計算と除算がその所要時間の大半を占めていますので、高速な平方根の逆数を算出するルーチンを作成することが有効です。また、粒子に作用する力の計算は行列とベクトルの乗算となりますので、アンローリング等のプログラミングの工夫により、高速化出来ます。

S-3800/180 では、高速な平方根の逆数ルーチンにより、実行時間は1.80秒、アンローリング等のプログラミングで0.76秒となりました。この事からシミュレーション手法を変更しないと言う前提条件のもとでは、数GFLOPsクラスのスーパーコンピュータでは、 $N=100000$ のオーダーのNまでが、実行可能な大きと言えます。

2.3 N=100000での実測結果

このサイズでは、定義式に忠実なコーディングで、その並列化特性の確認と、粒子群をグルーピング化する方式の2ケースの測定を実施しました。その測定結果を表2、表3に示しました。

表2 P-P法の並列化効果(S-3800)

CPU台数	1 CPU	2 CPU	3 CPU	4 CPU
実行時間(秒)	252.76	127.04	86.40	67.64
台数効果	1	1.99	2.93	3.74

表3 グルーピングP-P法の実測結果(S-3800/180)

グルーピング数	100	1000
実行時間(秒)	2.65	0.47

表2の結果より、この計算は、並列化効果が、非常に高い事が判ります。但し、分散メモリの場合、PE数を増やすと、その通信量も考慮する必要があります。表3の結果では、グルーピングP-P法は、効果のある事が判りますが、そのグルーピング方法や、グルーピング数の設定などにより詳細な検討が必要です。ただし、P-P法でも、工夫しだい、数GFLOPs クラスのスーパーコンピュータでも、 $N=100000$ のオーダーのNまでは、実行可能な大きさになる可能性がある事が判りました。

3. P-M (PARTICLE-MESH) 法での測定結果と解析

粒子に働く力を求めるのに、重力ポテンシャル ϕ を導入し、その質量密度 ρ に対し、ポアソン方程式

$$\Delta \phi = G \rho \quad G: \text{万有引力定数}$$

を解いて、その ϕ から $\vec{F} = M * (-\nabla \phi)$ より粒子に働く力を算出する方法があります。この計算のさい、 ρ 、 ϕ を求める為、空間に、格子(グリッドまたはメッシュ)を作ります。この手法は、粒子間の相互作用を ϕ を通じて求める方法と言えます。

この方式で、プログラミングする場合、留意すべき処理としては、

- (1) 粒子の質量と位置より、格子点上の、質量密度 ρ 分布を求める処理。
- (2) ポアソン方程式の求解処理
- (3) 格子点上に働く力から、各粒子に働く力を求める処理。

があります。

今回は、粒子の質量を同じ値とし、初期位置を一様乱数により定め、初期速度は0として、Nの値を変更していった場合の結果、解析事項を示します。処理時間の測定方法はP-P法と同じで、演算は全て行なう(定数値、同一値に対して演算は削減しない)様になっています。

3.1 $N \leq 10000$ の場合

P-P法では $N=10000$ の場合、汎用機、ワークステーションでは実行不可能であったので、P-M法では実行可能か否かを調査した。その結果下記の様になりました。

メッシュサイズを $11 \times 11 \times 11$ に固定し、M-880Hでの実行時間 E (nsec) を最小二乗法で N の関数として求めると

$$P-P法 \quad E=1052 \times N^2 \quad P-M法 \quad E=6169 \times N + 139635341$$

となり、 $N=10000$ ではP-M法では0.20秒と実行可能であることが判りました。

$N=1000$ での測定値(表4)より推定すると $N=10000$ のオーダーでは、汎用機、ワークステーションでも実行可能と言えます。

表4 汎用機、ワークステーションでの $N=1000$ での実測結果

CPU	M-880H	M-680H	3050RX (60MHz)
実行時間(秒)	1.06	2.29	2.03

また、このオーダーの N に対してはメモリに余裕がありますので、ポアソン方程式の求解処理では、ガウスの消去法を使用出来ます。これは、各タイムステップで左辺の行列が同じ形となり、LU分解が1度で可能な為です。

3.2 $N=100000$ の場合

S-3800/180でメッシュサイズ $11 \times 11 \times 11$ で実行時間は0.28秒となり実行可能となりました。ただし、メッシュサイズを大きくしようとするとポアソン方程式の求解で直接法は使用しにくくなります。容易に判りますように、ガウスの消去法で解く場合、各辺のメッシュサイズを m 倍すると、必要なメモリ量は m^6 倍になります。かなり複雑なコーディングをすれば、メモリの制限も除く事は可能ですが、更に大きな N をターゲットにしていたので、数GFLOPsのスーパコンピュータでも実行可能という確認をとるにとどめました。

3.3 $N=1000000$ の場合

ポアソン方程式の求解としては、反復法とFFTの使用を検討しましたが、任意のメッシュ数が選択出来ると言うことで、今回は反復法を試みました。FFTでも、2から9基底の3-D FFTを準備すれば、実用的には充分なので、今後の課題としています。

反復法としては、SOR法、加速係数1.9、収束判定値 10^{-6} での実測結果を表5に示しました。(直接法、グルーピングP-P法との比較の為それらの実測結果も記載)

表5 種々の方法による実行時間 (S-3800/180)

解法	メッシュサイズ/グルーピング数	実行時間(秒)	平均反復回数/タイムステップ
P-P法	10000	4.72	-----
P-P法	1000	25.60	-----
直接法	$11 \times 11 \times 11$	2.73	-----
SOR	$11 \times 11 \times 11$	2.63	113
SOR	$51 \times 51 \times 51$	11.45	126

この結果より、数GFLOPsの計算機では、実行可能な N の大きさは、制限付きで 10^6 が限度と考えられました。また、この処理では、大半の時間がポアソン方程式の求解に要していると思われましたので、前処理付き反復法であるSCG法²⁾の導入を検討しました。結果的には、SOR法と余り差がなく $N=10000000$ での測定の過程で判明した事ですが意外な見落としが原因となっています。

3.4 N=10000000 の場合

前処理付き反復法としてSCG法を使用してSOR法との比較した結果を表6、実行時間の比率の高い部分を検出するために実施した(メッシュサイズ 11×11×11に固定)結果を表7に記述しました。

表6 SCG法とSOR法(加速係数1.9)での比較。(S-3800/180)

メッシュサイズ	SOR法		SCG法	
	実行時間(秒)	反復回数 *	実行時間(秒)	反復回数 *
16×16×16	25.77	118	25.86	42
32×32×32	27.02	112	26.52	66
64×64×64	27.63	79	26.67	71
128×128×128	31.03	74	29.05	73

* 平均反復回数/タイムステップ 収束判定値 10^{-6}

表7 メッシュサイズ固定での実行時間比較 (S-3800/180)

反復解法	加速係数	実行時間(秒)	反復回数
直接法	----	27.43	----
SOR	1.9	25.75	113
SOR	1.8	25.66	56

この2つの測定結果より、N=10000000でのN体問題で所要時間のかかる処理は、以下の2つでこれらの処理の性能は計算機の構成(特性)に大きく依存してきます。

(1) 粒子の質量、位置より、格子点上の質量密度 ρ 分布を求める処理

ベクトル化不可、並列化不可

ランダムメモリアクセス

(2) 格子点上に働く力より各粒子に働く力を求める処理

ベクトル化可、並列化可

ランダムメモリアクセス

(1)、(2)の処理は共にランダムメモリアクセスの多、並列実行する場合((1)は対策をこうじて並列化した場合)、キャッシュに性能が依存するワークステーションでは性能が大きく低下し、分散メモリの場合は通信の増加により性能が低下します。

(1)の並列化、ベクトル化では、ワーク領域の大きさは、ベクトル化では、ある程度の性能が得られるループ長に依存し、並列化では、PE数に依存します。

改善後の、実験結果を表8に示します。

表8 SCG法とSOR法(加速係数1.9)での比較。(S-3800/180) (改善後)

メッシュサイズ	SOR法		SCG法	
	実行時間(秒)	反復回数	実行時間(秒)	反復回数
16×16×16	1.36	118	1.24	42
32×32×32	1.44	112	1.27	66
64×64×64	2.23	79	1.75	71
128×128×128	3.25	74	2.56	73

以上から、 $N=10000000$ のオーダーのN体問題では、P-M法により、数GFLOPs クラスのスーパーコンピュータで実行可能となりました。
 ちなみに、反復解法については、表9に示した様にSOR法はその反復回数(実行時間)が加速パラメータに大きく依存しますので、SCG法の方が有利と考えられます。

表9 SOR法における加速係数の影響 (S-3800/180)

加速係数	平均反復回数/タイムステップ	実行時間(秒)
0.1	1744	19.59
0.3	890	10.52
0.5	623	7.53
0.7	474	6.04
0.9	373	5.35
1.0	331	4.97
1.1	295	4.55
1.2	261	4.28
1.3	229	3.84
1.4	199	3.52
1.5	170	3.09
1.6	141	2.72
1.7	110	2.45
1.8	76	2.11
1.9	79	2.11

メッシュサイズ 64×64×64 に固定

4. まとめ

N体問題において、幾つかのサイズのNによる実験を行なう事により、アルゴリズム、プログラミングに関する多くの知見を得る事が出来、簡単なテストケースでは、数GFLOPs クラスのスーパーコンピュータ S-3800/180 を使用して、一千万体問題で1タイムステップ当たり3秒以内(現実的に実行可能)でシミュレーションを実行する事が出来た。

まだスタートしたばかりで、簡単なケースから実行していったが、今後はより一般的な問題、それにおおじて、FFT等の使用や、解法の検討を行ない、最終的には適切な計算機構成の検討もふくめて 10^9 のN体問題を実行可能にする手法に挑戦していくつもりでいます。

参考文献 :

- (1) 杉本太郎 : 手作りスーパーコンピュータへの挑戦 BLUE BACKS (1992)
- (2) 島崎真昭 : 計算機科学/ソフトウェア技術講座9 スーパーコンピュータとプログラミング 共立出版 (1989)
- (3) 岡田利男, 川田重男 共著 : シミュレーション物理学3 力学, 近代科学社 (1991)