計算科学技術特論A

古典分子動力学計算法の高速化

2019年7月4日

名古屋大学大学院工学研究科 附属計算科学連携教育研究センター 吉井範行







- 多数の原子や分子が集まってできた分子集団系
- ・原子、分子、さらには分子団の力を計算
- ・運動方程式を数値積分し、運動を追跡
- 得られた原子、分子、もっと大きな塊の軌跡から 熱力学、統計力学的性質を抽出

MD計算の例





MD 計算の 流れ

- 初期条件の設定
- ・分子間カの計算
- 運動方程式の数値積分
 時間刻みΔt
- 初期状態の平衡化
- 平衡化できたか?
- ・軌跡、物理量の出力
- 十分に統計はとれたか?
- ・最終座標・速度の出力







•<u>原子数</u>

100~100,000,000 個 (長距離相互作用のある場合) 対象系の規模や物理現象による。

・基本セルの一辺の長さ

数から数十nm

·<u>周期境界条件</u>

基本セル(立方体、直方体、平行六面体etc) イメージセルを基本セルの周囲に配置。

→表面効果の除去

3次元のみでなく、2,1次元方向にイメージセルを配置 2次元:界面、平面膜、2枚の平板にはさまれた状態 1次元:細孔、棒状分子

・<u>ポテンシャルカット</u>

Lennard-Jones相互作用 Ewald法の実空間部分 通常、原子直径の3倍(3o)程度。



相互作用(力)



- ・ 分子間相互作用(力) 2体力の和 $V_{\text{intermolecule}}(r^N) = \sum_{i>j} V(r_{ij})$ Lennard-Jones(LJ)相互作用 $V_{\text{LJ}}(r) = 4\varepsilon \left\{ \left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^6 \right\}$ クーロン相互作用 $V_{\text{Coulomb}}(r) = \frac{q_1q_2}{r}$
- 分子内相互作用(力)
 分子内の原子を剛体的に取り扱わず、フレキシブルに運動できるとする。(フレキシブルモデル)
 分子内ポテンシャルを伸縮、変角、torsionなどで表現する。

$$V_{\text{intramolecule}} = V_{\text{st}} + V_{\text{be}} + V_{\text{tor}}$$

$$= \sum k^{\text{st}} (r - r_0)^2 + \sum k^{\text{be}} (\theta - \theta_0)^2 + \sum_n \frac{1}{2} k_n^{\text{tor}} \{1 + \cos(n\phi - \phi_0)\}$$

分子間相互作用の精度がMD計算の質を決める

分子モデルと運動方程式





MD計算の例



対象	系·現象	対象	系·現象
低分子液体、溶液	水、水溶液、 溶媒和、化学反応	タンパク質	折りたたみ、酵素、イオンチャンネル、阻害剤、タンパク質 複合体
固体	相転移、フォノン	DNA、RNA	
不均一系、界面	結晶成長、吸着、トライ ボロジー、蒸発、凝縮、 透過、分配、界面増強ス ペクトル、界面張力	生体膜	脂質膜、ラフト、抗生物質、薬 物透過、吸収
両親媒性分子水溶液 ミセル、単分子膜、二 分子膜、バイセル	球状ミセル、棒状ミセル、 L膜、LB膜、ベシクル	ウイルス	カプシドタンパク質
化学反応	溶液内化学反応、 電解反応、 自由エネルギー曲面	ガラス	ガラス化、スローダイナミクス
超臨界流体	分子系、混合系、クラス ターの生成	金属、半導体	電極、液体金属
高分子	分離膜、電解質膜、複雑 系、スローダイナミクス	非平衡マクロ動力学	対流、ずり、核生成、流体力 学

MD計算の歴史



発表年	手法	対象	発表者
1953	MC	初めてのMC、剛体球	Metropolis et al.
1957	MC	球状分子	Wood and Parker
1957	MD	初めてのMD、剛体球	Alder and Wainwright
1964	MD	球状分子	Rahman
1969	MC	水	Barker and Watts
1971	MD	水(回転の運動方程式)	Rahman and Stillinger
1971	MC	イオン	Woodcock
1977	MD	拘束動力学	Rychaert et al.
1977	MD	タンパク質分子	McCammon et al.
1980	MD	圧力一定	Andersen
1980	MD	圧カテンソル一定	Parrinello and Rahman
1981	MD	経路積分、量子系	Chandler and Wolynes
1984	MD	温度一定	能勢
1985	MD	第一原理MD	Car and Parrinello
1986	MD	溶液内化学反応	Hynes et al.
1987	MD, MC	FMM	Greengard
1988	MD, MC	汎用ポテンシャル	Jorgensen et al.
1991	MD	非断熱量子動力学	Rossky et al.

MDの現状



汎用ポテンシャルセットの開発
 対象系の分子のポテンシャルが簡単に構築できる

・粗視化モデル
 大規模系が取扱い可能に

ポテンシャル名	開発者
AMBER	Kollman
CHARMM	Karplus
OPLS	Jorgensen
GROMOS	Berendsen

プログラム名	開発者
AMBER	Kollman
CHARMM	Karplus
GROMACS	Hess
TINKER	Ponder
NAMD	Schulten
DESMOND	Shaw
LAMMPS	Plimpton
MODYLAS	Okazaki
GENESIS	Sugita
MRABLE	Ikeguchi

MD計算の方向性



- 大規模化(原子数増)
 高並列コンピュータ+高並列用ソフトウェア
 高結度化(アンサンブル増)
- 高精度化(アンサンブル増)
 高並列コンピュータで並列同時実行
 +汎用ソフトウェア
- 長時間化(ステップ数増)
 専用計算機+専用ソフトウェア



ANTON by Wikipedia







大規模化·高精度化

<u>ユーザー</u>

- 問題設定
- 必要な計算資源はどの程度か
- ソフトはあるか
- プレ・ポスト処理できるか

初級の開発者

- 機能拡張を自分で行う
- ソフトはソース書き換え可能か(ソース公開。技術的にも)

上級の開発者

- 問題規模
- アルゴリズムの選択
- データ構造
- 非並列のしっかりしたプログラム作成
- 並列化(MPI, thread...)
- 開発期間

内容



MD計算の要素 各要素の高速化、並列化

MD計算とは

- MD計算の流れ、取り扱う原子数と周期境界条件、相互作用(力)、分子モデルと運動方程式
- ・ MD計算の例、歴史、現状

非並列のMD計算

- 短距離相互作用(bookkeeping, cell index法)
- ・ 長距離相互作用(クーロン相互作用) Ewald, PME法
- 効率化のための方法 拘束動力学 (Δtを大きくする) Multiple time step (Δtを大きくする) On the flyによる最適化 (収束計算をなくす)

高速化のために、キャッシュの有効利用、キャッシュミス

MD計算の並列化

• 高速多重極展開法(FMM)

データ構造 メタデータ法、通信量の最小化、データ局在化による計算効率の向上

ベンチマーク MODYLAS

非並列のMD計算



- シングルコアで高速化を進める MD計算手法の選択 短距離力の計算(bookkeeping, cell index) 長距離力の計算(Ewald, PME, FMM) 拘束動力学の利用 Multiple time stepの導入 on the fly計算の利用
- 高速化のためのエ夫



短距離相互作用 bookkeeping法 1

- ある分子からの距離がr。よりも短い相手分子の番号すべてをリストアップ。
- リストの作成は何ステップかに1回毎.
 固体のような流動性のないものなら更新不要
- r_cよりもΔrだけ長い距離のものまで余分にリストに載せる.
- リストを用いて相互作用を計算する際には、リ ストを複数のノードやスレッドに割り振って並列 計算させるループ分割法が容易に実装可能。





L	Ν	N'	L/r _l	N'/N
9σ	600	150	2.6	0.25
12σ	1500	150	3.4	0.10
15σ	3000	150	4.3	0.05
18σ	5000	150	5.1	0.03



短距離相互作用 bookkeeping法

k=0

do i=1,n-1 do j=i+1,n i, j間の距離がrl以下 k=k+1 list(1,k)=i リスト作成 list(j,k)=j リスト作成 enddo enddo

do i=1,k リストをもとに相互作用計算 enddo

MPI並列なら 相互作用の計算結果をノード間で通信 スレッド並列 通信不要

- 数ステップに一度 相互作用対のリス ト(list)を更新する
- 毎ステップリストを 用いて相互作用計 算する。

短距離相互作用 Cell index法



- 帳簿を作成, 更新する際のN(N-1)/2通りの距離計算が大きな負荷.大き なメモリを要する.
- L/r_iが4以上の場合には、一辺の長さがr_iよりも大きくなる範囲内で基本 セルそのものをx,y,zそれぞれの方向に4分割以上の分割を行う。
- 注目している分子が属しているセルとそれに直接接しているセルあわせ て27個の分割セル内にある分子との間だけでbookkeeping法を行う
- 分割セルの一辺の長さをよりも短くする場合には、第二近接のセルまで 考慮に入れる
- Cell index法についても作成したリストを複数のノードやスレッドに割り 振って相互作用の並列計算を行えばよい.





長距離相互作用 クーロン相互作用

- エワルド法 O(N²) あるいは
 Particle Mesh Ewald法 O(NlogN)
 →小、中規模系向け
- ・ 高速多重極展開法(Fast Multipole Method: FMM) O(N)
 →大規模系向け





Ewald法の要点

- 減衰の遅いクーロン相互作用を、減衰の速い関数と遅い関数に分ける。
- ・遅い関数はフーリエ変換することにより逆格子ベクトルを用いて評価。
- ・速い関数は実空間で評価。

$$V_N(\mathbf{r}^N) = \frac{1}{2} \sum_{\mathbf{n}} \sum_{i} \sum_{j}' \frac{1}{4\pi\varepsilon_0} \frac{Q_i Q_j}{|\mathbf{r}_i - \mathbf{r}_j + L\mathbf{n}|}$$

- (減衰の遅い1/r) = (減衰の速い関数) $\frac{1}{r} = \frac{\operatorname{erf}(\alpha r)}{r} + \frac{\operatorname{erfc}(\alpha r)}{r}$
- 減衰の速い関数 ··· この関数形で評価 $\frac{\operatorname{erfc}(\alpha r)}{r}$
- 減衰の遅い関数 ··· フーリエ変換したもので $\frac{\operatorname{erf}(\alpha r)}{r}$ 評価





長距離相互作用 Ewald法





Particle mesh Ewald(PME)法

逆格子ベクトルの項は

$$V_N^{(2)} = \frac{1}{2\varepsilon_0 L^3} \sum_{\boldsymbol{G}\neq 0} \frac{\exp\left(-\frac{|\boldsymbol{G}|^2}{4\alpha^2}\right)}{|\boldsymbol{G}|^2} \sum_i Q_i \exp(i\boldsymbol{G}\cdot\boldsymbol{r}_i) \sum_j Q_j \exp(-i\boldsymbol{G}\cdot\boldsymbol{r}_j)$$
$$= \frac{1}{2\varepsilon_0 L^3} \sum_{\boldsymbol{G}\neq 0} H(\boldsymbol{G}) S(\boldsymbol{G}) S(-\boldsymbol{G})$$



$$S(\boldsymbol{G}) = \sum_{i} Q_{i} \exp(i\boldsymbol{G} \cdot \boldsymbol{r}_{i})$$

S(G)を毎ステップ計算しなければならない。 S(G)を格子点uにおける補間を用いて表現すると

$$S(\boldsymbol{G}) = \sum_{\boldsymbol{u}} \hat{S}(\boldsymbol{u}) \exp(i\boldsymbol{G} \cdot \boldsymbol{u})$$

となる。ここでS(u)はS(G)のuにおける重み因子。これにより $V_N^{(2)} = \frac{1}{2\varepsilon_0 L^3} \sum_{u} \hat{H} * \hat{S}(u) \hat{S}(u)$

S(G)の計算に補間法を利用(格子点で評価) 高速フーリエ変換(FFT)利用 O(NlogN)





Δtを大きくする

- 拘束動力学 (⇔剛体回転子)
- Multiple time step

収束計算をなくす

• On the fly による最適化

拘束動力学



・拘束条件つきの運動方程式

拘束条件

$$g = (\mathbf{r}_1 - \mathbf{r}_2)^2 - d_{12}^2 = 0$$

運動方程式

$$m_{1}\ddot{r}_{1}=F_{1}+\lambda\frac{\partial g}{\partial r_{1}}$$

ラグラジアン

$$L = \sum_{i} \frac{1}{2} m_{i} \dot{r}_{i}^{2} - V(r^{N})$$

オイラー・ラグランジュ方程式
 $\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{r}_{1}} \right) - \frac{\partial L}{\partial r_{1}} = 0$
運動方程式
 $m_{1} \ddot{r}_{1} = F_{1}$



Multiple time step



同じMD計算のなかで変化の速いものと遅いものがある場合には、速い変化のほうに対応した小さな∆tを選択しなければならない。

例

United atomのブチルアルコールのフレキシブルモデル



 Hに対する運動方程式とほかの重い原子に対する運動方 程式を区別して、異なったムtで解くことができれば非常に 都合のよい(部分的にムtを大きくできる)ことになる。この 考え方に基づいた方法がMultiple time step法である。 Multiple time step



- 時間発展演算子法により、重い(遅い)運動と軽い(速い)運動とを分離して数値積分する。
- 時間発展演算子法

リウビル演算子 L

$$iL = iL_r + iL_p = v\frac{\partial}{\partial r} + \frac{F}{m}\frac{\partial}{\partial v}$$

時間発展演算子

$$e^{i \, \mathsf{L} \Delta t} \qquad q_i(\Delta t) = e^{i \, \mathsf{L} \Delta t} \, q_i(0)$$

異なる自由度など2つの部分に分離できると

$$e^{i\mathbf{L}\Delta t} = e^{i\mathbf{L}_p\Delta t/2}e^{i\mathbf{L}_r\Delta t}e^{i\mathbf{L}_p\Delta t/2} + O(\Delta t^3)$$
$$r_i(\Delta t) = e^{i\mathbf{L}_p\Delta t/2}e^{i\mathbf{L}_r\Delta t}e^{i\mathbf{L}_p\Delta t/2}r_i(0)$$

Multiple time step



 時間発展演算子法により、重い(遅い)運動と 軽い(速い)運動とを分離して数値積分する。

リウビル演算子 L

$$iL = iL_r + iL_{heavy} + iL_{light} = v\frac{\partial}{\partial r} + \frac{F_{heavy}}{m_{heavy}}\frac{\partial}{\partial v_{heavy}} + \frac{F_{light}}{m_{light}}\frac{\partial}{\partial v_{light}}$$

時間発展演算子

$$e^{i \, \mathbf{L} \Delta t} \qquad q_i(\Delta t) = e^{i \, \mathbf{L} \Delta t} \, q_i(0)$$

$$e^{i\mathbb{L}\Delta t} \approx e^{i\mathbb{L}_{heavy}\Delta t/2} e^{i\left(\mathbb{L}_{r}+\mathbb{L}_{light}\right)\Delta t} e^{i\mathbb{L}_{heavy}\Delta t/2}$$
$$\approx e^{i\mathbb{L}_{heavy}\Delta t/2} \left\{ e^{i\mathbb{L}_{light}\Delta t\delta t/2} e^{i\left(\mathbb{L}_{r}\right)\delta t} e^{i\mathbb{L}_{light}\delta t/2} \right\}^{n} e^{i\mathbb{L}_{heavy}\Delta t/2}$$

On the fly による最適化



- 第一原理MD(Car-Parrinello法)
- 電荷揺らぎモデル(TIP4P-FQ、SPC-FQモデル)
- 分極モデル(点双極子モデル)
- 最適値を得るために繰り返し計算を必要する変数を含む
- 繰り返し計算を行わず、最適値近傍を動的に揺らがせる近似解で満足できる場合
- その変数に仮想的な慣性量、運動エネルギーを与えて運動方程式を解く。
- 繰り返し計算がなくなるため高速化が可能。
- 運動エネルギーの相当する分だけ最適値からずれて、誤差が生じる。

分極モデル

$$U = U_0 + U_{\text{pol}}$$

 $U_{\text{pol}} = -\sum_i E_i \cdot \mu_i + \frac{1}{2\alpha} \sum_i \mu_i \cdot \mu_i$
 $E_i = T_{ij} \cdot \mu_i$
 $\mu_i = \alpha E_i$

拡張系のラグラジアン 粒子と双極子モーメントの自由度がある。 $L(\mathbf{r}^{N}, \boldsymbol{\mu}^{N}) = \frac{1}{2} \sum_{i=1}^{N} m \dot{r}_{i}^{2} + \frac{1}{2} \sum_{i=1}^{N} M \dot{\mu}_{i}^{2} - U$

双極子モーメントの運動方程式
$$M\dot{\mu}_{i} = \frac{\partial}{\partial\mu} L(\mathbf{r}^{N}, \boldsymbol{\mu}^{N}) = -\frac{\boldsymbol{\mu}_{i}}{\alpha} + \boldsymbol{E}_{i}$$

高速化のために



遅い演算(除算と組み込み関数)

- ・ 除算は他の四則演算より遅い。
- 組み込み関数も時間のかかる演算である。
- ホットスポット内ではできるだけこれらの演算を減らすように工夫 する。



高速化のために



遅い演算(組み込み関数) 組み込み関数のfortranコード化 環境によって速くなる場合とそうでない 場合がある。両者の比較が必要。

- $e^{x} \times e^{y} \longrightarrow e^{x+y}$
- $\log_a x + \log_a y \longrightarrow \log_a xy$
- $\sin\theta \times \cos\theta \longrightarrow 0.5\sin2\theta$
- $If(sqrt(x^*x+y^*y) < R) \longrightarrow If(x^*x+y^*y < R^2)$





高速化のために



分子シミュレーションへの適用

• Lennard-Jones相互作用

$$V = \left(\frac{1}{r}\right)^{12} - \left(\frac{1}{r}\right)^{6}$$

 $r^{6}=r^{2}*r^{2}*r^{2}$ $r^{12}=r^{6}*r^{6}$ $V=1.d0/r^{12}-1.d0/r^{6}$ $r^{6}=r^{2}*r^{2}*r^{2}$ $r^{12}=r^{6}*r^{6}$ $V=(1.d0-1.d0*r^{6})/r^{12}$ J相互作用への適用

キャッシュの有効利用



CPUの必要なデータに高速アクセス

- ① データはメモリに格納
- ② 必要に応じて隣接するデータと共にキャッシュ上に
- ③ 演算で必要となるとレジスタにコピー
- ④ 演算された結果は再びキャッシュに上書き
- ⑤ キャッシュ上で用いられなくなるとメモリ上に上書きコピーされる.



キャッシュミス



- データ:レジスターキャッシューメモリ
- ・レジスタ (キャッシュ:高速アクセス
- キャッシュ ← メモリ: ← 低速アクセス
- 演算を行う時にデータがキャッシュ上にない ときはメモリまで取りに行く。キャッシュミス。
- データがキャッシュに乗っているように工夫することが重要。

キャッシュミスを防ぐ



- ・1次元配列の変数:引数の順にキャッシュに格納
- ・1次元配列の変数へのアクセス→配列の引数について連続的にアクセスするように心がける。
- 2次元以上の配列の変数: Fortranでは配列左側の引数の順にキャッシュに格納。
 (C言語は右側から)
- 2次元以上の配列の変数へのアクセス→Fortran
 では配列の左側の引数について連続的にアクセスするように心がける。A(I,j)

(C言語は右側)

MD計算の並列化



- ループ分割:計算負荷の大きなdo ループを並列化
 ⇒並列化が容易
 ⇒全ノード間で通信(MPI)
 ⇒MPIの高並列になると効率が悪い
- ・領域分割:基本セルを複数の領域に分割
 ⇒並列化が大変。
 ⇒通信は隣接領域間のみ。
 ⇒超並列でも効率がよい。
 データ構造を大幅に変更
 ほぼ最初から書き直し

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	16

データ構造



原子のデータ構造

- ・ 原子座標、速度、カ、セグメントデータ → メタデータ化
- プロセスに属するサブセルを割り当てて固定 (図では4×4のサブセルを1つのプロセスに割り当て)
- サブセル内の原子のデータをプロセスに局在化させる。グローバルに持たない→省メモリ。
- データ領域に空きを用意 → サブセル内の原子数の増減に対応。各サブセル内の原子数に合わせて要素数が変化。要素数情報もあわせて保持。
- メタデータ配列のまま演算も通信も行う。サブセルを表す3次元配列を持っているので、取り扱いが容易
- バッファリングやindirect accessがない。
- ・ サブセル単位のメタデータがキャッシュに載るようにサブセル内の原子数を調整。
- 原子の運動に伴って原子の帰属するサブセルを更新。











自セルと相互作用する原子情報 を集める。

- 通信回数と通信量の最小化
- 隣接セルと通信して1次元(z 軸)方向にデータを結合し棒 状のデータを作る。
- 棒状データをy軸方向に通信 し結合させて、平面状データ にする。
- x軸方向に平面状データを通信し結合させて、立方体状の データにする。
- Kースパコンの多重同時通信 (各軸ごとに+、一方向)





キャッシュの利用効率の向上

- L1キャッシューL2キャッシューメモリ
- L2-メモリ間、L1-L2間のデータ通信の最小化 最重要ホットスポット
- 2体間相互作用計算
- FMMのM2L計算

2体間相互作用計算

3重ループ構造
 do 5つのサブセルを指定するインデックス
 do 注目する1つのサブセル内の粒子
 do 隣接する5つのサブセル内の原子
 相互作用計算
 enddo



enddo

enddo

- これにより、5つのサブセルのデータが効率的に再利用される。
- ただし5つのサブセルの座標データをL1キャッシュに載せておくことが重要





静電相互作用

- ・電気を帯びた系:水、電解質溶液、タンパク質、脂質などの両親媒性分子、イオン液体、、、
- 周期境界条件 カットオフ

Ewald法 O(N²)

Particle Mesh Ewald法 O(NlogN)

Darden, et al. (1993), Essman, et al. (1995)

Multilevel summation法 O(N)

D. J. Hardy, et al. (2015)

高速多重極展開法(FMM)O(N)

Greengard, Rokhlin (1987)



原子とサブセルとの相互作用.

GreengardらによるFMMの定式化

球面調和関数

$$Y_n^m(\theta,\phi) = \sqrt{\frac{(n-m)!}{(n+m)!}} P_n^m(\cos\theta) e^{im\phi}$$
 (ρ,α,β)

D

$$\frac{1}{r'} = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \rho^n Y_n^{-m} (\alpha, \beta) \frac{Y_n^m (\theta, \phi)}{r^{n+1}}$$

$$\phi(P) = \frac{q_i}{r'} = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} q_i \rho^n Y_n^{-m}(\alpha, \beta) \frac{Y_n^m(\theta, \phi)}{r^{n+1}}$$

多極子モーメント
$$M_n^m = \sum_i q_i \rho_i^n Y_n^{-m}(\alpha_i, \beta_i)$$

FMMにおける変換

• P2M

$$M_n^m = \sum_i q_i \rho_i^n Y_n^{-m} (\alpha_i, \beta_i)$$

• M2M

$$M_{j}^{\prime k} = \sum_{n=0}^{j} \sum_{m=-n}^{n} \frac{M_{j-n}^{k-m} i^{|k|-|m|-|k-m|} A_{n}^{m} A_{j-n}^{k-m} \rho^{n} Y_{n}^{-m} (\alpha, \beta)}{A_{j}^{k}} \qquad A_{n}^{m} = \frac{(-1)^{n}}{\sqrt{(n-m)(n+m)}}$$

• M2L

$$L_{j}^{k} = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \frac{M_{n}^{m} i^{|k-m|-|k|-|m|} A_{n}^{m} A_{j}^{k} Y_{n+j}^{m-k} (\alpha, \beta)}{(-1)^{n} A_{n+j}^{m-k} \rho^{n+j+1}}$$

• L2L

$$L_{j}^{k} = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \frac{L_{n}^{m} i^{|m|-|m-k|-|k|} A_{n-j}^{m-k} A_{j}^{k} Y_{n-j}^{m-k} (\alpha, \beta) \rho^{n-j}}{(-1)^{n+j} A_{n}^{m}}$$

• L2P

$$\phi_{i} = \frac{q_{i}}{4\pi\varepsilon_{0}} \sum_{j=0}^{\infty} \sum_{k=-j}^{j} L_{j}^{k} Y_{j}^{k} \left(\theta_{i}, \phi_{i}\right) r_{i}^{j}$$
$$f_{i} = -q_{i} \sum_{j=0}^{n_{\max}} \sum_{k=-j}^{j} L_{j}^{k} \nabla_{i} \left\{ r_{i}^{n} Y_{j}^{k} \left(\theta_{i}, \phi_{i}\right) \right\}$$
$$\nabla_{i} \left\{ r_{i}^{j} Y_{j}^{k} \left(\theta_{i}, \phi_{i}\right) \right\} \Longrightarrow r_{i}^{j-1} Y_{j-1}^{k\pm 1} \left(\theta_{i}, \phi_{i}\right)$$

Solid harmonics化

Legendre陪関数

$$P_{\ell}^{m}(\cos\theta) = (-\sin\theta)^{m} \frac{d^{m}}{(d\cos\theta)^{m}} P_{\ell}(\cos\theta)$$
$$= \frac{1}{2^{\ell}\ell!} (-\sin\theta)^{m} \frac{d^{\ell+m}}{(d\cos\theta)^{\ell+m}} (\cos^{2}\theta - 1)^{\ell}$$

Regular solid harmonics

$$R_{\ell}^{m}(r,\theta,\phi) = \frac{r^{\ell}}{(\ell+m)!} P_{\ell}^{m}(\cos\theta) e^{im\phi}$$

Singular solid harmonics $S_{\ell}^{m}(r,\theta,\phi) = (-1)^{\ell+m} \frac{(\ell-m)!}{r^{\ell+1}} P_{\ell}^{m}(\cos\theta) e^{im\phi}$

$$\frac{1}{\|\boldsymbol{r}_{S} + \boldsymbol{r}_{R}\|} = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} S_{\ell}^{-m} (\boldsymbol{r}_{S}) R_{\ell}^{m} (\boldsymbol{r}_{R})$$

FMMにおける変換

• P2M

.

$$M_{\ell}^{m} = \sum_{i} q_{i} R_{\ell}^{m} \left(\mathbf{r}_{i} \right)$$

• M2M $M_{\ell}^{\prime m} = \sum_{\lambda=0}^{n_{\max}} \sum_{\mu=-\lambda}^{\lambda} M_{\ell-\lambda}^{m-\mu} R_{\lambda}^{\mu}$

M2L
$$L_{\ell}^{m} = \sum_{\lambda=0}^{n_{\max}} \sum_{\mu=-\lambda}^{\lambda} M_{\lambda}^{\mu} S_{\ell+\lambda}^{-(m+\mu)}$$

L2L

$$L_{\ell}^{m} = \sum_{\lambda=\ell}^{n_{\max}} \sum_{\mu=-\lambda}^{\lambda} L_{\lambda}^{\mu} R_{\lambda-\ell}^{\mu-m}$$
L2P

$$\Phi = \frac{1}{4\pi\varepsilon_0} \sum_{\ell=0}^{n_{\max}} \sum_{m=-\ell}^{\ell} L_{\ell}^m M_{\ell}^m$$

$$\boldsymbol{f}_{i} = -\boldsymbol{q}_{i} \sum_{\ell=0}^{n_{\max}} \sum_{m=-\ell}^{\ell} L_{\ell}^{m} \begin{pmatrix} \frac{1}{2} \left(R_{\ell-1}^{m+1} - R_{\ell-1}^{m-1} \right) \\ -\frac{1}{2} \left(R_{\ell-1}^{m+1} + R_{\ell-1}^{m-1} \right) \\ R_{\ell-1}^{m} \end{pmatrix}$$

実部、虚部の対称性の利用(実数化)

Regular, singular solid harmonicsや多極子モーメント、
 局所展開係数の実部、虚部の間にある対称性を利用してこれらの要素数を減らし、変換の演算量の削減を図る。

$$L_\ell^{m} = \sum_{\lambda=\ell}^{n_{ ext{max}}} \sum_{\mu=-\lambda}^{\lambda} L_\lambda^\mu R_{\lambda-\ell}^{\mu-m}$$

実部、虚部の間にある対称性

 $R_n^{-m} = (-1)^m R_n^{*m}$

$$S_n^m \hspace{0.2cm} L_n^m \hspace{0.2cm} M_n^m$$
も同様の式が成り立つ

mの非負部のみ取り扱えばよい $-n \le m \le n$ $0 \le m \le n$

要素数
$$(n_{\max} + 1)^2$$
 =25 $(n_{\max} = 4)$
 $(n_{\max} + 1)(n_{\max} + 2)$
 2 =15

実部、虚部の対称性の利用(実数化)



$$\begin{split} \boldsymbol{\sharp} \boldsymbol{\mathfrak{S}} \boldsymbol{\mathfrak{W}} \boldsymbol{\mathfrak{L}} & L_{\ell}^{m} = \sum_{\lambda=\ell}^{n_{\max}} \sum_{\mu=-\lambda}^{\lambda} L_{\lambda}^{\mu} R_{\lambda-\ell}^{\mu-m} \\ & = \sum_{\lambda=\ell}^{n_{\max}} \left[L_{\lambda}^{0} R_{\lambda-\ell}^{-m} + \sum_{\mu=1}^{\lambda} L_{\lambda}^{\mu} R_{\lambda-\ell}^{\mu-m} + \sum_{\mu=-\lambda}^{-1} L_{\lambda}^{\mu} R_{\lambda-\ell}^{\mu-m} \right] \\ & = \sum_{\lambda=\ell}^{n_{\max}} \left[L_{\lambda}^{0} R_{\lambda-\ell}^{-m} + \sum_{\mu=1}^{\lambda} L_{\lambda}^{\mu} R_{\lambda-\ell}^{\mu-m} + \sum_{\mu=1}^{\lambda} L_{\lambda}^{\mu} R_{\lambda-\ell}^{\mu-m} \right] \\ & = \sum_{\lambda=\ell}^{n_{\max}} \left[L_{\lambda}^{0} R_{\lambda-\ell}^{-m} + \sum_{\mu=1}^{\lambda} L_{\lambda}^{\mu} R_{\lambda-\ell}^{\mu-m} + \sum_{\mu=1}^{\lambda} L_{\lambda}^{\mu} R_{\lambda-\ell}^{\mu-m} \right] \\ & = \sum_{\lambda=\ell}^{n_{\max}} \left[\left(-1 \right)^{m} L_{\ell}^{0} \Re R_{\lambda-\ell}^{m} + \sum_{\mu=1}^{\lambda} \left\{ \Re L_{\lambda}^{\mu} \left(\Re R_{\lambda-\ell}^{\mu-m} + \left(-1 \right)^{m} \Re R_{\lambda-\ell}^{\mu+m} \right) + \Im L_{\lambda}^{\mu} \left(\Re R_{\lambda-\ell}^{\mu-m} + \left(-1 \right)^{m+1} \Re R_{\lambda-\ell}^{\mu+m} \right) \right\} \right] \\ & + i \sum_{\lambda=\ell}^{n_{\max}} \left[\left(-1 \right)^{m+1} L_{\ell}^{0} \Im R_{\lambda-\ell}^{m} + \sum_{\mu=1}^{\lambda} \left\{ \Re L_{\lambda}^{\mu} \left(\Im R_{\lambda-\ell}^{\mu-m} + \left(-1 \right)^{m+1} \Im R_{\lambda-\ell}^{\mu+m} \right) + \Im L_{\lambda}^{\mu} \left(\Re R_{\lambda-\ell}^{\mu-m} + \left(-1 \right)^{m+1} \Re R_{\lambda-\ell}^{\mu+m} \right) \right\} \right] \end{split}$$

直交座標の漸化式の利用

- 従来は、P2M,L2Pにおいて、原子座標として直交座標を球 面座標に変換し、球面調和関数を計算していた。
- ここでは、直交座標から直接solid harmonicsを求める漸化 式を導入し、座標変換を消去する。

rでスケールされたLegendre陪関数

 $\tilde{P}_{\ell}^{m}(r,z) = r^{\ell} (r\sin\theta)^{-m} P_{\ell}^{m}(\cos\theta)$

 $\tilde{P}_{\ell}^{m}(r,z)(x+iy)^{m}=r^{\ell}P_{\ell}^{m}(\cos\theta)e^{im\phi}$

Regular solid harmonics

$$R_{\ell}^{m}(x,y,z) = \frac{1}{(\ell+m)!} \tilde{P}_{\ell}^{m}(r,z)(x+iy)^{m}$$

Singular solid harmonics

$$S_{\ell}^{-m}(x, y, z) = \frac{(\ell - m)!}{r^{2\ell + 1}} \tilde{P}_{\ell}^{m}(r, z) (x + iy)^{m}$$

m ≧ 0についての漸化式 $\tilde{P}_{\ell}^{m}(r,z) = \begin{cases}
(-1)^{m}(2m-1)!! & (\ell = m) \\
(2\ell - 1)z\tilde{P}_{\ell-1}^{m}(r,z) & (\ell = m+1) \\
\frac{2\ell - 1}{\ell - m}z\tilde{P}_{\ell-1}^{m}(r,z) - \frac{\ell + m - 1}{\ell - m}r^{2}\tilde{P}_{\ell-2}^{m}(r,z) & (\ell \ge m+2)
\end{cases}$

K. Nitadori, Particle mesh multipole method: An efficient solver for gravitational/electrostatic forces based on multipole method and fast convolution over a uniform mesh.

多重ループの1重ループ化

- M2M,M2L,L2Lにおいて、solid harmonicsの添え字についての多重 ループが存在する。
- 1重ループ化しループ長を伸ばすことにより、SIMDなどのマシン性 能を引き出す。

$$L_{\ell}^{m} = \sum_{\lambda=0}^{n_{\max}} \sum_{\mu=-\lambda}^{\lambda} M_{\lambda}^{\mu} S_{\ell+\lambda}^{-(m+\mu)}$$

2重ループ



最内側ループ長 15 → 225

1重ループ

Benchmark test

- 水系(19530原子)
- 1000MDステップ実行
- 周期境界条件
- NVEアンサンブル
- 実行環境 京コンピュータ 8ノード(8MPI, 1Thread) MODYLAS使用
- FMMの展開次数 n_{max}=4 球面調和関数版 改良版



球面調和関数使用時、FMMは全体の65%。 M2Lだけで60%を占める。 P2MやL2Pで4%ほど。 FMMの中ではM2L,P2M,L2Pが99%を占めている。

YN7

計算条件、全体に対するFMMの割合を示す。 YOSHII Noriyuki, 2017/05/11 YN7

対称性を利用した実数化(ループ2重) → ループの1重化 → 漸化式利用による座標変換の削除









YN3

計算精度

変更前後における計算結果の比較(1step)

MODYLASの熱力学量に関する標準出力,

	Hamiltonian	potential-E
M2L(single loop)	-3.901749600398E-16	-4.449136174563E-16
M2L(double loop)	-3.901749600398E-16	-4.449136174563E-16
original MODYLAS	-3.901749600398E-16	-4.449136174563E-16
	kinetic-E	temperature
M2L(single loop)	5.473865741657E-17	2.030212809202E+02
M2L(double loop)	5 //738657/1657F-17	2 030212809202F+02
	J.4/J00J/410J/L-1/	2.0302120032022.02

- ・12ケター致。
- 正しい静電相互作用計算を行えている。

計算が正しく行えていることを示す。 YOSHII Noriyuki, 2017/05/11 YN3

FMM

- FMMを、Greengard & Rokhlinに従う球面調和関数を基底としたものから、Solid harmonicsへと変更した。
- 実部と虚部の対称性を利用による演算量の削減を図った。
- これに伴いP2M、M2M、M2L、L2L、L2P変換の演算量が減少。特にM2Lが2.4倍速くなった。
- 直交座標から直接solid harmonicsを評価できる 漸化式を利用することにより、座標変換を消去した。これによりP2M,L2Pが3.5倍速くなった。
- FMM全体で2.4倍高速化することに成功した。







MD calculation



NPT ensemble P=1atm、T=310. 15 K SHAKE/RATTLE/ROLL CHARMM22 with CMAP LJ 12 Å cut off Coulomb FMM

Empty capsid

VP1、VP2, VP3	, VP4 60 each
Sphingosine	60
H ₂ O	1、884、218
Na ⁺	5, 310
K ⁺	136
CI [_]	5、206
	6, 480, 236 atoms
PDB 1HXS	
pH 7.4	
positive charges	4、860
negative charges	5, 100

 $\Delta t = 0.5 \text{ fs} / 2 \text{ fs} / 4 \text{ fs}$ 200 ns



Local pressure from virial theorem



Origin of the negative pressure



Local pressure from virial theorem 240 excess negative charges at pH=7.4

j≠i

$$P_{\alpha} = \frac{N_{\alpha}kT}{V_{\alpha}} + \frac{1}{3V_{\alpha}} \sum_{i}^{N_{\alpha}} \sum_{j \neq i}^{N} \mathbf{r}_{ij} \mathbf{F}_{ij}$$

$$Q(\mathbf{r}_{i}) = \frac{1}{3V_{\alpha}} \sum_{j \neq i}^{N} \mathbf{r}_{ij}$$

$$P(\mathbf{r}_{i}) = q_{i}Q(\mathbf{r}_{i})$$

Exchange of water



Stroboscopic picture of a water molecule crossing the capsid.



1. Water exchange is in thermal equilibrium and is very rapid.

 \rightarrow Chemical potential of water molecules is the same between the inside and outside of the capsid.

2. The capsid doesn't allow the ions to pass through.

 $\rightarrow\,$ The capsid works as a semipermeable membrane.

プロジェクトに育てられたMODYLAS

- NAREGIプロジェクト 自前で汎用プログラムを作ろう!(2004年秋) (自前コードがないと、開発した手法のテスト、最先端の手法の導入 といった方法論に関する 研究ができない。) 開発方針:多少遅くてもよい、高精度の計算ができなければならない。 まずは自分たちの研究用として利用。 数ノードで並列化MPI+SMP(富士通SR8000)、数万原子系に対応。 ⇒ループ分割、簡単。
 - スパコンを占有した研究のための実証計算。
 - 分子研SR11000(2005年正月)
 - ⇒両親媒性分子集合体の熱力学的安定性についての
 - 自由エネルギー計算。
 - ウイルス系(1000万原子)への挑戦
 - ⇒高速多重極展開法(FMM)実装(2007年春)。

プロジェクトに育てられたMODYLAS

- Nano Grand Challenge プロジェクト 領域分割化。
 並列化を考慮した変数の取り直し。プログラムの書き直し。
 高並列化へと歩み出す。
 開発方針:領域分割化し、高並列に対応。
- 京プロジェクト 変数のコピーレス化、変数を通信用、演算用に分けない。 TOFUによる隣接通信の活用。
 SIMDの活用、オンキャッシュ化 開発方針:1000万原子系を1MDステップ10ms以下に。(実際には 5msにまで高速化)
 ソフト公開、ソースコード公開。
 ミニアプリ化。
- ポスト京 プロジェクトの要求、マシンの仕様に合わせて最適化してきた。 将来のユーザーのニーズ、マシンのアーキテクチャーを見据えな がら、時代に合ったソフトウェア開発を進めなければならない。

今日のまとめ



MD計算の要素 各要素の高速化、並列化

MD計算とは

- MD計算の流れ、取り扱う原子数と周期境界条件、相互作用(力)、分子モデルと運動方程式
- ・ MD計算の例、歴史、現状

非並列のMD計算

- 短距離相互作用(bookkeeping, cell index法)
- ・ 長距離相互作用(クーロン相互作用) Ewald, PME法
- 効率化のための方法 拘束動力学 (Δtを大きくする) Multiple time step (Δtを大きくする) On the flyによる最適化 (収束計算をなくす)

高速化のために、キャッシュの有効利用、キャッシュミス

MD計算の並列化

• 高速多重極展開法(FMM)

データ構造 メタデータ法、通信量の最小化、データ局在化による計算効率の向上

ベンチマーク MODYLAS

診考文献



MD全般

- D. Frenkel, B. Smit, "Understanding Molecular Simulation," Academic Press, San Diego (1996).
- ・ 上田 顕,「コンピュータシミュレーション」,朝倉書店(1990).
- M. P. Allen, D. J. Tildesley, "Computer Simulation of Liquids," Oxford Science, Oxford (1987).
- J. P. Hansen, I. R. McDonald, "Theory of Simple Liquids," Academic Press, London (1986).
- 岡崎 進, 吉井 範行, コンピュータ・シミュレーションの基礎(第2版), 化学同人(2011). 力学
- H. Goldstein, 「ゴールドスタイン新版古典力学(上), (下)」(瀬川富士, 矢野 忠, 江沢康生 訳), 吉岡書店(1983).

数値積分

- G. J. Martyna, M. E. Tuckerman, D. J. Tobias, M. L. Klein, Mol. Phys., 87, 1117(1996). Particle mesh Ewald法
- T. Darden, D. York, L. Pedersen, J. Chem. Phys., 98, 10089(1993).

FMM

- L. Greengard, V. Rokhlin, J. Comput. Phys., 73, 325(1987). SHAKE
- G. Ciccotti, J. P. Ryckaert, Comp. Phys. Rep., 4, 345(1986).
- P. Gonnet, J. Comput. Phys. 220, 740(2007).

MODYLAS

- Y. Andoh, et al., J. Chem. Theory Comput. 9, 3201 (2013)
- Y. Andoh, et al., J.Chem. Phys. 141, 165101 (2014).