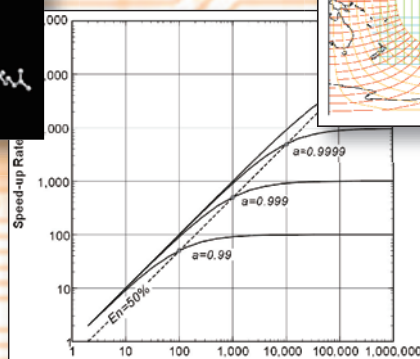
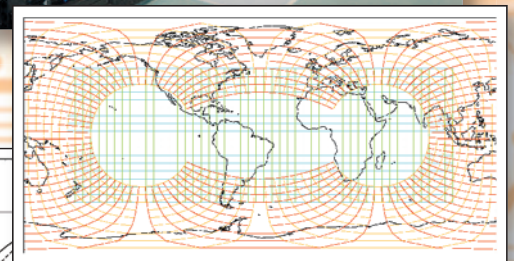
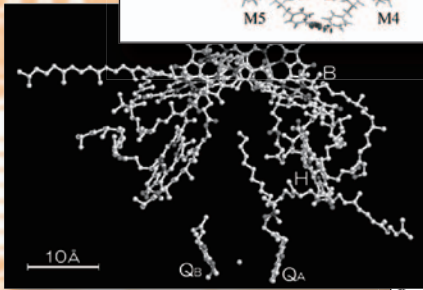
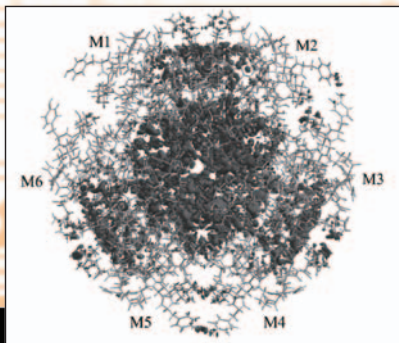


スーパーコンピューティング ニュース

Vol.11 No. Special Issue 1 2009. 2

特集：ペタスケール・シミュレーションの信頼性



東京大学情報基盤センター
(スーパーコンピューティング部門)

目 次

特集：ペタスケール・シミュレーションの信頼性

特集号「ペタスケール・シミュレーションの信頼性」発刊にあたって	1
中島研吾（東京大学情報基盤センター）	
気象・気候変動予測シームレスシミュレーション	5
高橋桂子（独）海洋研究開発機構地球シミュレータセンター	
地震波伝播と強震動の大規模並列 FDM シミュレーション	35
古村孝志（東京大学大学院情報学環総合防災情報研究センター／地震研究所地震 火山災害部門）	
密度汎関数法によるタンパク質全電子シミュレーション	65
佐藤文俊（東京大学生産技術研究所）	
有限要素法に基づく第一原理分子動力学法について	99
土田英二（産業技術総合研究所計算科学研究部門）	
粒子法シミュレーションの大規模化と高速化	123
越塚誠一（東京大学大学院工学系研究科）	

特集号「ペタスケール・シミュレーションの信頼性」

発刊にあたって

中島 研吾

東京大学情報基盤センター

1. はじめに

東京大学情報基盤センタースーパーコンピューティング部門では、隔月で広報誌「スーパーコンピューティングニュース」を発行しているが、この他にも2006年度から年一回の頻度で特集号を発行しており、いずれも好評を博している。

- 2006年度：スーパーコンピューターの拓く世界 ―異分野の研究に触れる―
- 2007年度：コーディングしてみよう！ ―スパコンプログラミングを極める I―

3回目となる2008年度は「ペタスケール・シミュレーションの信頼性」と題して各分野の最先端の研究を実施されている皆様に、様々な視点から大規模シミュレーションの「信頼性」についてご執筆を頂くことを企画した。2006年度は、大規模シミュレーションとそれによって開拓される新しい科学、2007年度はそれを支える基盤技術に着目した。2008年度は、そのような基盤技術のプログラムへの適用方法も含めて、支配方程式・モデリングと現象との整合性、数値的な安定性など様々な「信頼性」に着目した。

2. 背景, 方針

計算機シミュレーションを駆使した「計算科学」が、「理論」、「実験」に続く「第三の科学」と呼ばれるようになって久しい。長い間、計算科学は理論や実験の結果の検証や補間が主たる役割であったが、近年、計算機のハードウェア、ソフトウェア技術の飛躍的な発展と理論面での進展、実験・観測技術の向上、様々なデータの蓄積も相俟って、真の科学としての一步を歩み出そうとしている。

特に日本において、2002年より稼動を開始した「地球シミュレータ」（ピーク性能40TFLOPS）が果たした役割は大きい。地球温暖化、台風の進路予測、地震波動伝播のシミュレーション等、高解像度の大規模モデルを使用して飛躍的に高精度なシミュレーションが可能となった他、様々な現象が複合した大規模マルチフィジックス連成モデルによって複雑な現象のメカニズムの解明も可能となり、「新しい科学」の創成に大きく貢献している。

2012年に完成予定の理化学研究所「次世代スーパーコンピューター」（ピーク性能10PFLOPS）は、バイオテクノロジーを中心とした更に広範囲の分野での成果が期待されてい

る。当情報基盤センターでも 2008 年 6 月より「T2K オープンスパコン」（ピーク性能 140 TFLOPS）が稼動を開始しており，科学技術分野のみならず，大規模テキスト処理，経済物理シミュレーションなど様々な分野での活用が始まっている。

サンディア国立研究所の「ASCI Red」が初めてピーク性能 1 TFLOPS を超えたのは 1996 年末のことであった。それから 10 年あまり，2008 年「Road Runner（ロスアラモス国立研究所）」によって人類はついに「PFLOPS（ペタフロップス）」時代を迎えたが，解決すべき問題はまだまだ多い。これまで，たとえば大規模データの処理，可視化あるいは数十万単位のプロセッサコア間の通信などの問題については様々なところで議論が進められており，研究開発において着実な進歩が見られる。本特集号では，「ペタスケール・シミュレーションの信頼性」と題して，各分野の最先端の研究者の方々に，様々な視点から大規模シミュレーションの「信頼性」についてご執筆を頂いた。

執筆者の方々には以下のような項目を念頭においてご執筆いただくようお願いした：

- 対象となる問題の支配方程式，モデリング手法，適用するアルゴリズム，プログラミング手法，並列化までの一連の流れが分野外の読者にもある程度理解できるよう，特に以下の点に着目して簡単に解説する。
 - 「密行列固有値計算」，「疎行列ベクトル積」等，利用する数値計算プロセス
 - 並列化，大規模化の戦略，手法
 - 利用しているライブラリ，外部プログラム等
 - 複数のアプローチが考えられる場合には，各手法の長所・短所，現在の手法を使用している理由，メリット
 - 大規模シミュレーションにおける問題点，その解決法，あるいは将来更に大規模になることで生じる可能性のある問題点とその解決法。

- シミュレーションの精度，信頼性を保証するための様々なアプローチについて解説する。
 - 計算結果の解釈の方法
 - プログラムの正当性，精度を検証するための方法
 - ・ 適正なメッシュ数，粒子数などの決定方法，基準となる考え方
 - ・ 伝統的に実施されてきた方法，現在実施されている方法
 - ・ 今後実施される可能性のある方法
 - ・ 実験・観測が困難な現象・条件における検証方法
 - ・ 実験・観測結果の検証や補間ではない真の「予測科学」への展望
 - ・ 問題点，モデリングの限界
 - 安定性，精度，効率のトレードオフ
 - ・ 数値的な不安定性と対応
 - いわゆる「マルチフィジックス」の導入

- ・ 有効性
 - ・ 現状, 問題点
 - ・ 将来展望
- ハードウェア, コンパイラ, ライブラリ等, 基盤技術分野の研究者, 技術者に求めることについても提言を行う。
- T2K オープンスパコン, 情報基盤センターへの期待についてもコメントする。

3. 概要

本特集号では, 以下の 5 編の記事から構成されている (掲載順) :

- ① 気象・気候変動予測シームレスシミュレーション (高橋桂子 (海洋研究開発機構))
- ② 地震波伝播と強震動の大規模並列 FDM シミュレーション (古村孝志 (東京大学情報学環))
- ③ 密度汎関数法によるタンパク質全電子シミュレーション (佐藤文俊 (東京大学生産技術研究所))
- ④ 有限要素法に基づく第一原理分子動力学法について (土田英二 (産業技術総合研究所))
- ⑤ 粒子法シミュレーションの大規模化と高速化 (越塚誠一 (東京大学工学系研究科))

①, ②, は地球科学分野における事例であり, 「地球シミュレータ」を駆使した大規模計算の他, ②では「T2K オープンスパコン (東大)」との比較についても触れられている。共に, 「格子系」のモデルを使用し, 「計算⇒観測結果との比較⇒モデル改良」というサイクルであり, ①では, 大気海洋結合モデル, 局所的に詳細なモデルを扱うことによって計算結果の観測結果に対する精度が大幅に改善することが示されている。②の扱う範囲は地下空間であり, 大気・海洋と比較して不明な点が多い。異常震域のシミュレーションのように, 地下構造の不均質性を予測しながら計算モデルを構築する必要がある場合もある。

③, ④は共に分子の電子状態を求めるための第一原理シミュレーションであるが, ③では密度汎関数法, ④では「格子系」のモデルの一つである有限要素法が使用されており, それぞれ, 支配方程式, モデル化, 数値アルゴリズムについて詳細に解説されている。有限要素法は連続体力学向けのシミュレーション向けの手法として広く知られているが, 適応格子, 多重格子法など, 連続体力学向けの有限要素法で一般的に使用されている手法がそのまま活用されているのは非常に興味深い。

⑤は, 工学系のシミュレーションであるが, 特に非線形現象の取り扱いに優れている「粒子法」について解説いただいた。差分法, 有限要素法などの「格子法」との比較, 基本的アルゴリズム, 様々な計算事例の他, 物理ベースコンピュータグラフィックスへの適用の他, 大規模化・高速化については, 昨今注目されている GPU (Graphics Processing Unit) に

も触れられている。最後に、シミュレーション結果の信頼性を定量化する方法である Verification and Validation (V&V) について、実際アメリカ (AIAA, ASME) で策定されたガイドラインが簡単に紹介されている。

4. 結言

本特集号は、大規模シミュレーションに取り組んで行こうとする様々な分野の研究者、技術者、学生など幅広い層の読者に役立つものと期待される。シミュレーションプログラムを自ら開発して行こうという人にとってはもちろんであるが、既存プログラムや商用コードを利用してシミュレーションを実施する研究者、技術者にとっても大いに役立つ一冊となることを祈念する。

最後に、多忙な中、本特集号の記事をご執筆いただいた 5 名の著者の皆様（高橋桂子，古村孝志，佐藤文俊，土田英二，越塚誠一（敬称略，掲載順））には篤く御礼申し上げます。

気象・気候変動予測シームレスシミュレーション

高橋桂子

(独)海洋研究開発機構地球シミュレータセンター

1. はじめに

私たちになじみの深いお天気は、さまざまな時間、空間スケールの現象が複雑に影響し合っ
て形成されている。通常、3日から5日程度の時間の長さを中心となる天気の前測は、数時間
の数10km程度の狭い範囲での降雨や、それより狭い範囲で急激な現象の変化を伴う数時間単
位の集中降雨などの異なるスケール現象をも前測の対象となり、前測に必要な気象モデルは、
これらのそれぞれのスケールの現象を再現、前測可能であるとともに、より広い範囲で、より
長い時間にわたる現象をも、その相互影響を考慮できるモデルである必要がある。

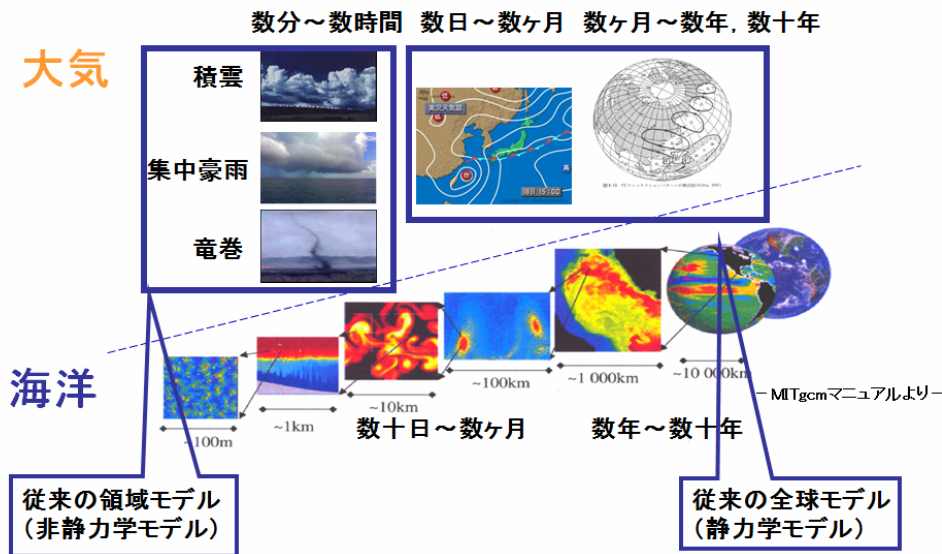
一方、温暖化現象に代表される気候変化は、さまざまな分野において今後も多くの関心とそ
の対処法が考えられる課題であるが、温暖化の条件下での気象前測、またより長い時間の前測
である季節の変動前測や、さらにはエルニーニョやインド洋ダイポール現象などの変動がどの
ように気象へ影響を与えるかについては、未解決の課題が多い。第4次IPCCレポート(気候
変動に関する政府間パネル(Intergovernmental Panel on Climate Change: IPCC)レポート)の
報告以降、気象現象前測と気候変動前測との現象の“連続性”あるいは“階層性”を考える必要があ
るのではないかと、という関心が高まっている[1]。この課題への回答は、温暖化が進んだ状況下
で、私たちの生活に直接関係がある気象現象においてどのような影響がおよぼされるのかにつ
いての具体的な回答を与えることに繋がる。本来、これらの課題への解答や情報が必要である
ことは、多くの人々が認めるところである。

これまでの気象前測に用いられてきた前測モデルと、全球を対象とするような長い時間、お
よび大きな空間を対象とする気候変動前測モデルは、対象とするスケールが異なることから、
もともと用いている前測モデルが異なる。特に、気象を扱う際には、空間の水平スケールと鉛
直スケールの差がないために、大気の大気対流を直接扱うことが求められるが、気候変動を扱うモ
デルは、鉛直方向の運動方程式で、重力と気圧傾度力が釣り合っていると近似した静水圧近似
が用いられ、水平スケールが鉛直スケールより十分大きい現象については、非常によく近似を
することが知られている。よって、気象と気候変動前測に用いられてきたモデルは、現象の時
間、空間スケールにより異なるモデルを用いてきた(図1)。

しかしながら、たとえば、赤道上の雲の分布が地球全体の気候変動へ影響を与える可能性が
示唆されていること、あるいは、雲の種類が地球の放射過程を介して全地球の熱的な循環や気
象の変化、気候変動に影響を与えている可能性があることなど、雲に代表される空間単位が数
100m、時間スケールとしては数時間の現象と、雲の生成過程や消滅過程の時空間単位よりは
るかに長く大きい気候変動のスケールの現象との相互影響を同時に扱う必要性が強く示唆され
てきている。

上記のように気象前測と気候変動前測において異なるモデルを用いてきた理由のひとつは、
計算機資源の問題である。2002年3月に稼働を開始した地球シミュレータは、その目的のひ

大気と海洋の現象と時間・空間スケール



第1図: 大気と海洋の様々な時間, 空間スケールの模式図。

大気と海洋では、時間と空間スケールの関係が、ほぼ線形の関係にあることが知られている。たとえば、大気現象を例にとると、積雲の成り立ちや集中豪雨、竜巻などは、発生から収束までが数分から数時間の時間スケールであり、局所的な空間スケールである。それに対し、低気圧や高気圧の一生はそれより長い数日から10日程度で数1000kmの空間スケールである。これらの時間、空間のスケールの大きさと現象は、海洋においても成り立ち、数1000kmスケールでの現象を大気では総観現象といい、水平スケールが鉛直スケールより卓越することから、より小さな、鉛直スケールを無視できないスケール（積雲などの鉛直対流が顕著な現象など）で異なるモデルを用いてきた。

特に、従来の気候変動予測で用いられてきた水平数100kmの解像度を約10km程度とし、問題規模を約1000倍大きくして、例えば台風の構造をより詳細にとらえることができるようにということのひとつの目標として掲げた。稼働当初において、地球シミュレータ以前では領域に対してであった水平解像度10kmでシミュレーションが、全地球を対象にすることが可能となり、大気や海洋の詳細構造が捉えられることが示された[2]。しかしながら、全地球10kmの水平解像度での長期のシミュレーションは、高速化の工夫をした上に地球シミュレータをもってしても実現はできていない。加えて、本来ならば、上記にもふれたように、気象と雲が解像できるほど、つまり、雲の生成単位である数10mから数100mでの解像度で、かつ長期の予測シミュレーションを可能とすることが求められるような、現在大きな関心を集めている気象現象と気候変動予測を結びつけるためには、さらに大規模な計算機資源と大規模シミュレーションを可能とする高速化手法と高精度な計算手法、および高解像度でなければ表せない諸気象現象を再現できる気象と気候モデルが必要となる。

地球シミュレータセンターでは、地球シミュレータ稼働当初から、上記に述べてきたような気象と気候変動予測を結びつけるための、複数の時間、空間スケールを一度に扱えるようなシミュレーションモデルは、どのようなモデルが必要、かつ最適であるかについての検討を進め、地球シミュレータを最大限に活用した場合、モデルはどこまで現象を再現、予測できるかについての研究開発を推進してきた。本稿では、この目的を果たすために開発を進めてきた大気海

洋結合モデル: Multi-Scale Simulator for the Geoenvironment (MSSG)の特徴と、超大規模なシミュレーションを可能とした計算性能最適化の手法、および MSSG 研究開発の一環において、地球シミュレータ上において実施した、台風等の進路の予測精度向上を目指したパイロットランとしてのシミュレーションの成果、現時点における研究開発の成果の一端を紹介し、今後の課題と展開について紹介する。

2. 大気海洋結合モデル MSSG の概要と特徴

2. 1 支配方程式系と離散化

大気海洋結合モデル: Multi-Scale Simulator for the Geoenvironment (MSSG)は、大気の流れは、静力学平衡を仮定せずに、完全圧縮型の 3 次元 Navier-Stokes 方程式を何の仮定をすることもなく解くことである。本大気モデルを MSSG-A とよび(A は, atmosphere の意味の略記), MSSG-A の流れは、以下の時間発展方程式系で表現することができる。

$$\frac{\partial \rho'}{\partial t} + \frac{1}{G^{\frac{1}{2}} a \cos \varphi} \frac{\partial (G^{\frac{1}{2}} G^{13} \rho u)}{\partial \lambda} + \frac{1}{G^{\frac{1}{2}} a \cos \varphi} \frac{\partial (G^{\frac{1}{2}} G^{23} \cos \varphi \rho v)}{\partial \varphi} + \frac{1}{G^{\frac{1}{2}}} \frac{\partial (\rho w^*)}{\partial z^*} = 0, \quad (1)$$

$$\frac{\partial \rho u}{\partial t} + \frac{1}{G^{\frac{1}{2}} a \cos \varphi} \frac{\partial (G^{\frac{1}{2}} G^{13} P')}{\partial \lambda} = -\nabla \cdot (\rho u \mathbf{v}) + 2f_r \rho v - 2f_\varphi \rho w + \frac{\rho v u \tan \varphi}{a} - \frac{\rho w u}{a} + F_\lambda, \quad (2)$$

$$\frac{\partial \rho v}{\partial t} + \frac{1}{G^{\frac{1}{2}} a} \frac{\partial (G^{\frac{1}{2}} G^{23} P')}{\partial \varphi} = -\nabla \cdot (\rho v \mathbf{v}) + 2f_\lambda \rho w - 2f_r \rho u - \frac{\rho u u \tan \varphi}{a} - \frac{\rho w v}{a} + F_\varphi, \quad (3)$$

$$\frac{\partial \rho w}{\partial t} + \frac{1}{G^{\frac{1}{2}}} \frac{\partial P'}{\partial z^*} + \rho' \mathbf{g} = -\nabla \cdot (\rho w \mathbf{v}) + 2f_\varphi \rho u - 2f_\lambda \rho v + \frac{\rho u u}{a} + \frac{\rho v v}{a} + F_r, \quad (4)$$

$$\frac{\partial P'}{\partial t} + \nabla \cdot (P \mathbf{v}) + (\gamma - 1) P \nabla \cdot \mathbf{v} = (\gamma - 1) \nabla \cdot (\kappa \nabla T) + (\gamma - 1) \phi, \quad (5)$$

$$P = \rho R T, \quad (6)$$

$$\rho w^* = \frac{1}{G^{\frac{1}{2}}} \left(G^{\frac{1}{2}} G^{13} \rho u + G^{\frac{1}{2}} G^{23} \rho v + \rho w \right). \quad (7)$$

方程式系は、連続の式 (1), 独立変数 ρu , ρv , ρw を求めるための静力学平衡の仮定を用いない 3 次元 Navie-Stokes 方程式 (2), (3), (4), 熱力学方程式 (5), 状態方程式 (6) から成っている。ここで, ρ' , P はそれぞれ $\rho' = \rho - \bar{\rho}$, $P' = P - \bar{P}$ で定義され, $\bar{\rho}$ はと \bar{P} は, それぞれ標準大気密度と気圧を表す。f, μ , κ , γ は, それぞれコリオリ力, 粘性係数, 拡散係数, 乾燥大気の定圧比熱と定積比熱の比を表す。また, F は, ρu , ρv , ρw で表される大気の流れによる移流など断熱過程での時間変化項と雲や放射などの非断熱過程による時間変化項による外力項, G は球面座標系で考える際のメトリック項, λ は緯度, φ は経度を表す。

MSSG-A と同様に海洋モデル(MSSG-O とよぶ, O は ocean の略記)についても、下記の方

式系で流れを表現する。大気との違いは、気体の変化を完全圧縮性としているのに対し、海洋は物性の違いからブジネスク近似((9)式)を採用している点である。方程式系を以下に示す。

$$\frac{\partial c}{\partial t} = -\mathbf{v} \text{grad} c + F_c \quad (8)$$

$$\frac{\partial T}{\partial t} = -\mathbf{v} \text{grad} T + F_T \quad (9)$$

$$0 = \nabla \cdot \mathbf{v} = \left(\frac{1}{r \cos \varphi} \frac{\partial u}{\partial \lambda} + \frac{1}{r \cos \varphi} \frac{\partial (\cos \varphi v)}{\partial \varphi} + \frac{1}{r^2} \frac{\partial (r^2 w)}{\partial r} \right) \quad (10)$$

$$\frac{\partial u}{\partial t} = -\mathbf{v} \text{grad} u + 2f_r v - 2f_\varphi w + \frac{vu \tan \varphi}{r} - \frac{wu}{r} - \frac{1}{\rho_0 r \cos \varphi} \frac{\partial P'}{\partial \lambda} + F_\lambda \quad (11)$$

$$\frac{\partial v}{\partial t} = -\mathbf{v} \text{grad} v + 2f_\lambda w - 2f_r u - \frac{uv \tan \varphi}{r} - \frac{wv}{r} - \frac{1}{\rho_0 r} \frac{\partial P'}{\partial \varphi} = +F_\varphi \quad (12)$$

$$\frac{\partial w}{\partial t} = -\mathbf{v} \text{grad} w + 2f_\varphi u - 2f_\lambda v + \frac{uw}{r} + \frac{vw}{r} - \frac{1}{\rho_0} \frac{\partial P'}{\partial r} - \frac{\rho'}{\rho_0} \mathbf{g} + F_r \quad (13)$$

$$\rho = \rho(T, c, P_0) \quad (14)$$

$$\frac{d}{dr} P_0 = -\rho_0 g(r) \quad (15)$$

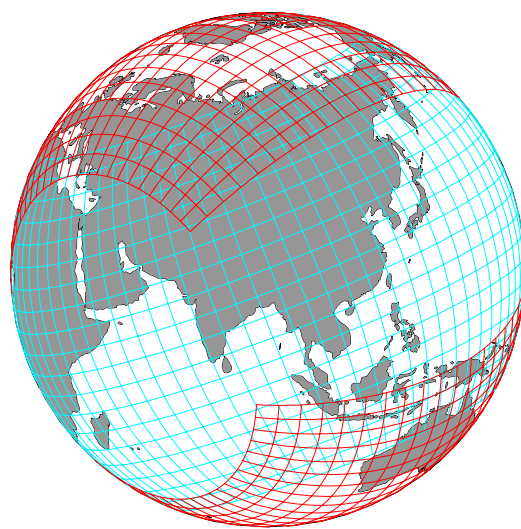
本海洋モデルにおいても静力学平衡の近似は用いることなく、3次元非静力学モデルとなっている点の一つの特徴である。また、方程式から明らかなように、解くべき変数もほぼ大気モデルで解く変数とほとんど変わりはない。(14)式であらわされる密度のみが、海洋を特徴づける塩分 C を含む式となっており、UNESCO の式と呼ばれる経験式として与えられる。

MSSG-A および MSSG-O の乱流モデルとしては、Lilly and Smagorinsky タイプの乱流モデルを採用し、サブグリッドスケールの混合過程のモデルとして採用している。また、MSSG-A においては、雲のモデルとして雲微物理モデルを採用し、放射モデルは短波と長波の影響を考慮した単純な透過モデルを現時点では採用している。本稿では、採用したこのモデルに関する詳細な記述は割愛するので、下記の表と MSSG-A について採用された各スキームやモデルについては参考文献[3, 4, 5, 6, 7, 8, 9, 12]を、MSSG-O については参考文献[10, 11, 12, 13]を参照されたい。

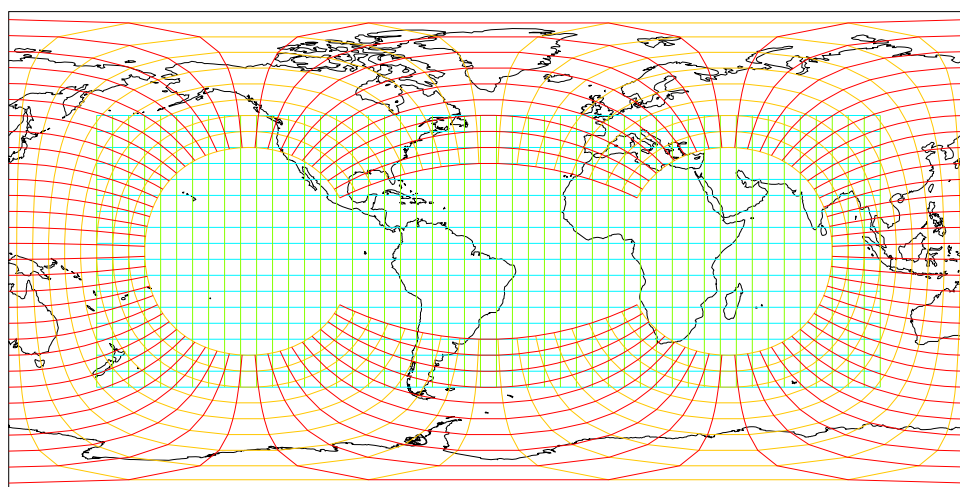
2. 2 座標系と離散化

2. 1 において示した MSSG-A と MSSG-O の時間発展方程式系を離散化し、全球を表す全きゅう座標系上の各格子点において、時間発展方程式を各時刻における予報変数の値を求めることによって、時系列的な変化を求めることができる。

まず、3次元空間上の離散化について概説する。全球の3次元空間を扱う際の、まず水平方向の空間は、MSSG-A と MSSG-O のいずれにおいても、Yin-Yang グリッド[14]上で離散化している。Yin-Yang グリッドは、2次元の2つのパネルをちょうど野球のボールのように、球面に張り付けた構造をした重合格子である。図2に、その略図と展開図を示す。Yin-Yang グリッドは、従来の緯度経度座標系の一部（北緯45度と南緯45度、経度0度から270度で囲ま



(a)



(b)

第2図: Yin-Yang グリッドの模式図。

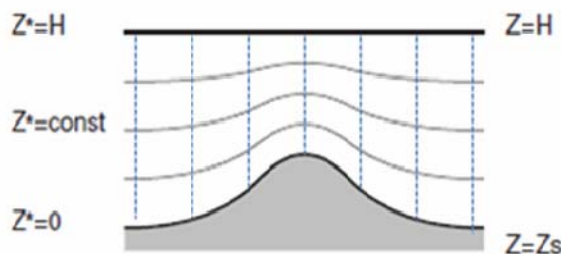
(a)は、Yin-Yang グリッドを地球上に張り付けた模式図であり、(b)は展開図である。グリッドが重複するりょういきについては、それぞれの領域内におけるグリッド上において計算を行い、各パネルの境界に相当するインターフェイス上のグリッドにおいてのみ、他方のグリッドから補間して値を決定する。重複領域のグリッドの値を求める手法は、多くの手法が提案されているが、MSSGでは、インターフェイス上のグリッドにおいてのみのパネル間の情報交換を行う手法を採用している。

れた領域)をひとつのパネルとし(N系とよぶ、図2(a)で水色で示した座標系)と、緯度と経度方向に90度回転させたもうひとつのパネル(図2(a)でピンク色で示した座標系、E系とよぶ)を張り合わせて構成する。

Yin-Yang グリッドは、基本的には緯度経度座標系であるので、例えば、高精度の計算手法や領域モデルとしての境界条件の設定など、これまでに蓄積されてきた多くの緯度経度座標系における知見をそのまま活用することができる。また、緯度経度座標系では、北極や南極に極点として特異点が存在し、その周辺ではグリッドが集中するため CFL 条件が厳しくなる。しかしながら、Yin-Yang グリッドは極点としての特異点が存在せず、全球ほぼ一樣な大きさのグリッドから構成できることから、緯度経度座標系に比べて CFL 条件を緩和できることも特長である。

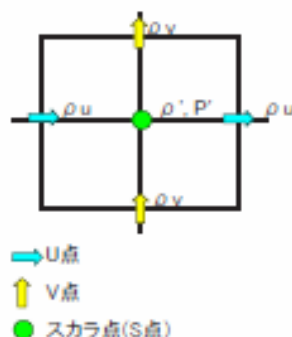
その反面、グリッドが重複して定義される重合領域におけるグリッド上の計算と、それらの計算を他のパネル上のグリッドの値とどのように整合性を保つか、特に保存量の扱いには注意が必要であることも付け加えておく。

3次元空間へ拡張し、3次元全球に対応する3次元格子を構成するには、上記の2次元Yin-Yangグリッドを鉛直方向に、定義する総数並べたグリッド系で3次元を構成する。このため、鉛直報告には、予報変数が定義されるグリッド点が、鉛直方向には直線的に並ぶことになる。特に、山岳などの地表面上の地形を扱うためにZ*座標系、MSSG-Oでは海底地形の起伏を表すために、z座標系を採用している。Z*座標系は、 $Z^* = H(z-z_s)/H-z_s$ で表現される。ここで、zは物理的な標高であり、Z*はモデルの鉛直座標上での高度となる。Z_sは地表面の物理的な標高であり、Hはモデルで定義する上端の物理的な標高である。球面水平方向には、Yin-Yangグリッド系が、図3の実線にそって配置され、鉛直方向には点線上にグリッドが配置されることになる。水平方向のYin-Yangグリッド上での予報変数の離散化は、スタガード格子の一種であるArakawa-Cグリッドを用いて離散化している(図4)。



第3図: 鉛直方向のグリッドの配置の模式図。

Z*座標系を用いることによって、鉛直方向のグリッドは山岳などの地形に沿って定義される。地形に沿った流れが効率よく表現可能な特長がある半面、水物質や温度などのスカラートレーサー値は、矩形の鉛直グリッドに比較して、地形に沿って移流され易くなる特徴をもつことが知られている。また、急峻な地形の時には、不適切なグリッド配置となるので、注意が必要である。ベクトルとスカラー量はともに、同じグリッド点上で計算している。

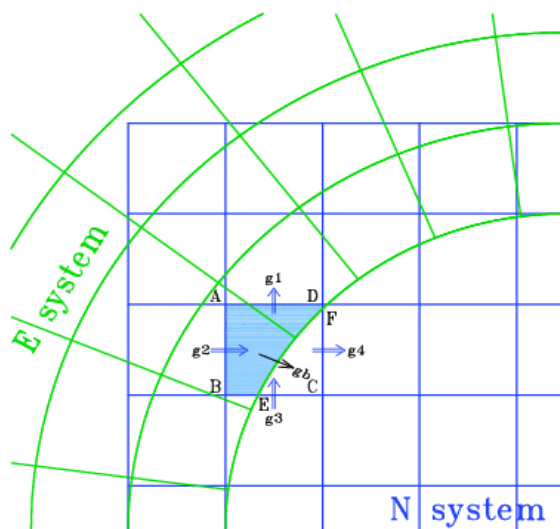


第4図: Arakawa-Cグリッド上における変数の定義についての模式図。

Arakawa-Cグリッドは、グリッド点上でスカラー量を計算し、隣り合うグリッドの中心において、ベクトルを計算する(MSSGでは、上図が示すようにフラックス量として定式化している)。

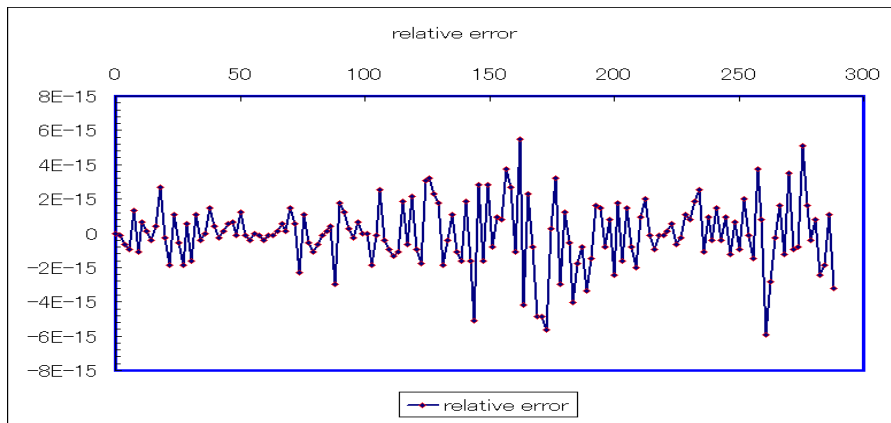
時間方向には、スプリットイクスプリシット法と呼ばれる HE-VI (Horizontally Explicit - Vertically Implicit scheme) 法を採用している。これは、各項を短い時間刻みで評価する項と長い時間刻みで評価する項に分割して扱う方法であり、長い時間刻みのスキームに Runge-Kutta 法 [15]を短い時間刻みのスキームに Forward-Backward 法を使用している。HE-VI 法を用いることによって、MSSG では、水平方向の音波項は陽解法で計算し、鉛直方向の音波項は陰解法で解く。

Yin-Yang グリッドでは、N 系と E 系の重合部分のインターフェイスでは、グリッド点が規則的ではあるが不連続に存在するため、保存量を精度よく計算するための工夫が必要となる。保存量としてまず考えなければならないのは、インターフェイスにおいて補間を行う際に、物理量を保存するように補間をすることである。質量を精度よく保存するために、Yin-Yang グリッドでは、図 5 に示すような保存スキームを開発した[16][17][18]。保存をするための十分条件は、E 系と N 系の重合部のインターフェイスにおいて、E 系から N 系へ E 系上の円弧 EF を介して流入するフラックスは、N 系のグリッド ABCD のフラックスの流入出フラックスの総量と $\{(ABEFD \text{ の面積}) / (\text{矩形 } ABCD \text{ との面積})\}$ の積であらわされるということに基づいて、定式化されている (図 5) [18][19][20]。この質量保存のためのフラックス修正を用いることによって、図 6 に示すように、時間変化とともに質量変化は、数値計算の丸め誤差以下となり、質量の保存が可能であることが示されている[20]。



第 5 図:MSSG 質量の保存のためのフラックスの補間手法。

Yin-Yang グリッドを構成する N 系と E 系の重合部分では、2つのグリッドシステムを用いることになる。MSSG では、E 系、N 系の境界上のみで補間を行い情報を交換している。図 5 に示すような、N 系への流入フラックスの収支は、円弧 EF 上でのフラックス gb と、グリッドボリューム ABCD に流入あるいは流出するフラックスの収支、つまりフラックス g1,g2, g3, g4 の和が等しければ、質量が保存されることが十分条件として知られている[16, 17]ことから、Yin-Yang グリッドにおいても、この十分条件を適用して、質量を保存する[18]。



第6図: フラックスの補間手法を用いたときの全球の質量総和に対する時系列変化。

Yin-Yang グリッドの N 系と E 系の境界上における補間手法として、図5にも示したフラックス収支を一致させたときの質量の時系列変化である。横軸は時間 (day), 縦軸は初期値における質量の変化からの差を表す。計算の丸め誤差の範囲内で、質量が保存されていることがわかる。

本稿では、特に大気大循環モデルとしての MSSG-A を中心にモデルの概要を述べたが、海洋コンポーネントである海洋大循環モデル MSSG-O も基本的には、同様の定式化と離散化を用いている。Yin-Yang 格子上的基本的な検証については、参考文献[21, 22, 23, 24, 25]を参照していただきたい。MSSG-A と MSSG-O の其々のコンポーネントに用いられているスキームを表1にまとめた。2 つ以上の手法やモデルが記述されているところは、シミュレーションの対象をなる現象や比較検証のために、複数の中から1つを選択できるようなプログラムとなっている。雲微物理(microphysics), 乱流モデル(turbulent model), 放射モデル, 陸面モデル, キャノピーモデル (UCSS urban canopy model) については、参考文献[3, 4, 5, 6, 7, 8, 9,12, 26] , sea-ice model などの非断熱過程による時間変化を解くモデルとその解法については、参考文献 [27]を参照していただきたい。

第1表: MSSG-A(大気コンポーネント)と MSSG-O(海洋コンポーネント)において採用されているモデル, 計算手法の一覧表。

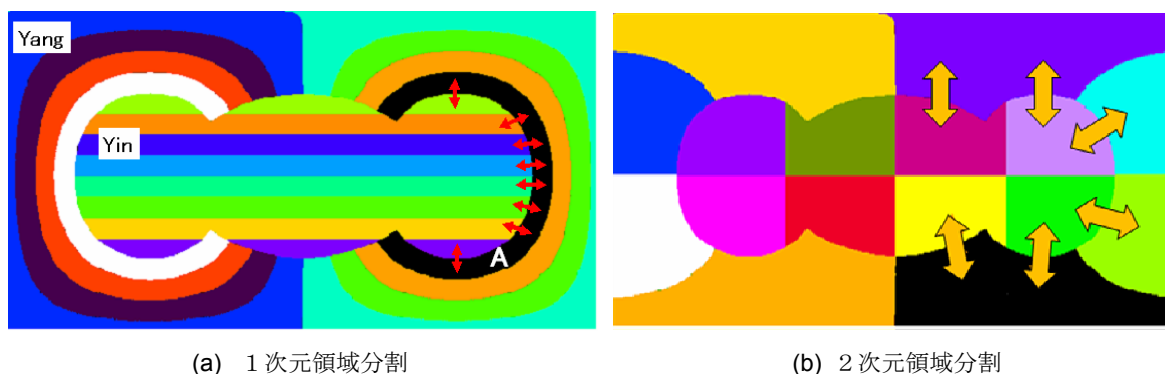
	MSSG-A	MSSG-O
	Non-hydrostatic AGCM	Non-hydrostatic /hydrostatic OGCM
governing eqs.	Fully compressive N-S eqs.	incompressive N-S eqs.
grid system	Yin-Yang grid	Yin-Yang grid
discretization	space	Arakawa-C grid (horizontal), Z* (vertical)
	time	3 rd /4 th Runge-Kutta
adv. schemes	5 th flux form, WAF, CIP-CSLR	5 th flux form
non-adv. schemes	4 th flux form	4 th flux form
sound wave	HEVI, HIVI	Implicit methods (2D, 3D)
microphysics	Bulk method (Qc,Qr,Qi,Qs,Qg)/ hybrid-Bin method	-
turbulence model	static Smagorinsky scheme	static Smagorinsky model
other models	cloud radiation model, bucket land model, UCSS urban canopy model	sea-ice model
parallelization	horizontal 2D decomposition by MPI/ vertical decomposition by micro-task	horizontal 2D decomposition by MPI/ vertical decomposition by micro-task

3. MSSG に用いられている並列化手法と地球シミュレータ上での計算性能

3. 1 計算性能最適化の基本方針

地球シミュレータではノード間、ノード内、プロセッサ内の3段階の並列化を考える必要がある。ノード間は分散メモリ環境であり、MPI か HPF を使って並列化を行うことが可能である。ノード内は共有メモリ環境であるので、並列化はマイクロタスクか OpenMP,あるいはノード間と同様に MPI, HPF によって実装可能である。加えて、プロセッサ内はベクトルプロセッサにより並列実行される。ノード間とノード内の並列化手法においては複数の選択肢が存在する。MSSG では、OpenMP および HPF のライブラリとコンパイラの性能を考慮して、実績のある MPI,とマイクロタスクによる並列化手法を採択した。

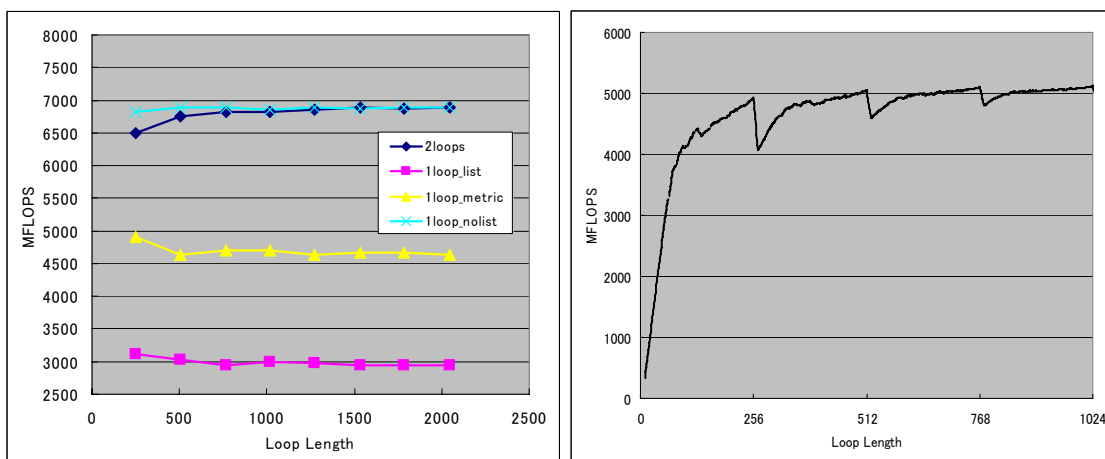
Yin-Yang 格子を形成する N 系と E 系は、基本的には緯度経度格子である。3次元変数は経度 λ , 緯度 φ , 鉛直層数 r として (λ, φ, r) の関数として記述される。この3次元軸において、どの軸を MPI, マイクロタスク, ベクトル化に対応させるかを検討した。全球をシミュレーションの対象とする解像度(水平解像度は 10 km 以下)を想定すると、水平解像度は 10 km としても水平グリッド数は約 400×200 となり、鉛直グリッド数を数 100 層としても、水平グリッド数のほうが鉛直グリッド数に比べてはるかに大きい。計算上の依存関係については、水平方向では、東西南北に相当する4方向の周囲のグリッドにおいて計算された値を用いることになり、情報の通信が必要である。その一方で、鉛直方向のグリッドは、鉛直上下方向のグリッド点の値を用いて計算されることから、鉛直方向のグリッド点は、その水平隣の鉛直グリッド点からの情報を使用して計算されることはない。これらの特徴から、経度緯度に相当する水平方向のループ長が最も長くとることが可能であり、ベクトル化に適していると考えられる。さらに、鉛直上下方向のグリッド点の値を用いて計算される鉛直方向の並列化は、今後の鉛直方向の対流、乱流モデルや雲微物理モデルを拡張した場合においても影響が少ないように、並列ループを柔軟に変えられるマイクロタスクを採用した。以上の検討から、MPI による分割は緯度方向と経度方向の2次元領域分割とし、鉛直方向にはマイクロタスクを用いて並列化することとした。



第7図: Yin-Yang 格子を1次元領域分割した場合と2次元領域分割した時の比較。

16ノード数を例にとり、Yin-Yang 格子を1次元領域分割:(a)、2次元領域分割:(b)したときの、N-E 境界における情報交換量を模式的に示した。分割領域の色の違いは、異なるノードを表す。1次元分割では、ひとつのノードが(黒色のノード)が少量の情報を多くのノード間で交換しなければならないのに対して、2次元分割では、ノード間の情報交換は1次元分割より多く、情報交換が必要となるノードの数も少ないことから、効率的に通信を行うことが可能である。

緯度方向と経度方向の2次元領域分割による並列化を考える際には、まず、ループをどのように実装方法するかを考えなければならない。ソースコードの可読性から考えると、2重ループによる実装が最も適しているといえる。しかしながら、2次元領域分割はベクトルループに対してMPIによる並列化も行うことになるため、ノード数を増やして並列度を向上させると、同時にベクトル長が短くなり、並列化能の低下を招く問題が生じる可能性がある。そこで、2重ループを1重化して実装し、ループ長を確保するする手法も考えられる。基本的には、Yin-Yang グリッドは緯度経度座標系であるので、座標の位置的な情報が容易に得られ、かつ前述のコードの可読性をも考慮すると、2重ループでの実装が望ましい。そこで、実装に関する簡単なテストを行い、2重ループでの実装の際の条件を検討した。図8は、力学過程の計算部分に用いられているループ構造を、2重ループ、リスト構造を持つ1重ループ、リスト構造を持たない1重ループ、メトリック項を排除したリスト構造を持つ1重ループ、の4つの場合を実装し、計算性能を比較した結果、2重ループ構造の実装を行っても、ループ長が十分に確保できる場合は、リスト構造を持たない1重化ループ構造の場合と同等の計算性能が得られることがわかった。このことから、MSSGの実装では、座標の位置的な情報が容易に得られ、可読性も得られる2重ループでの実装を採用した。さらに、図8の結果において、ほぼ計算性能が変わらなくなるループ長が約1000になるまでをより詳しくみると(図8(b))、地球シミュレータの最大ベクトル長である256の倍数でループ長を確保することが、計算性能を低下させることなく2重ループ化できる条件であることが明らかである。



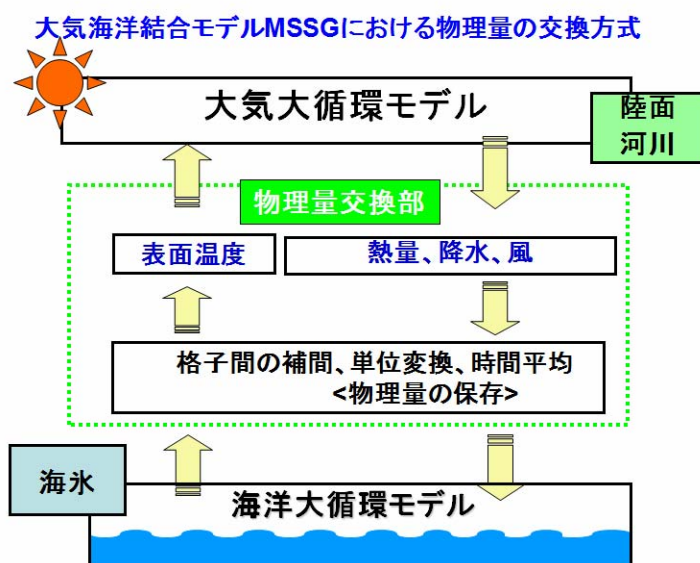
第8図: Yin-Yang 格子2次元領域分割におけるループ構造と計算性能の比較。

力学過程の計算部分に用いられているループ構造を全て、2重ループ(2loops)、リスト構造を持つ1重ループ(1loop_list)、リスト構造を持たない1重ループ(1loop_nolist)、メトリック項を排除したリスト構造を持つ1重ループ(1loop_metric)、のいずれか1つに統一して実装した4つの場合において、計算性能を比較した(左図)。x軸はループ長、y軸はMegaフロップス値を表す。左図は、右図の値のループ長が0から1024までの拡大図である。

3. 2 大気と海洋コンポーネントを結合した MSSG の2次元領域分割

大気コンポーネント MSSG-A と海洋コンポーネント MSSG-O は、並列化についての基本方針も、同様である。MSSG-A と MSSG-O を結合する際には、各ノードが MSSG-A と MSSG-O の同じ領域を分担し並列計算を行う。MSSG-A と MSSG-O の結合際には、通常の結合モデル

で用いられている結合方式を採用している（図9）。大気から海洋へ、大気の最下層での物理量：熱的なフラックス、風による2次元運動量フラックス、降水や河川からの淡水フラックスをMSSG-Aから計算された値から診断してMSSG-Oの境界条件として与える。また、MSSG-Oからは、あたら得られた大気からの境界条件をもとに計算した海洋表面温度を、大気側の境界条件として与え、これをシミュレーション時間の1単位として計算をすすめる。2次元領域分割を採用したことにより、MSSG-AとMSSG-Oは、大気から海洋へは5種類の2次元データを、また逆に海洋から大気へは1種類の2次元データを通信することとなる。



第9図：大気海洋結合モデルMSSGにおける大気、および海洋モデルコンポーネントとそれらのコンポーネント間における物理量交換の模式図。

大気大循環モデル（MSSG-A）から海洋大循環モデル（MSSG-O）へは、短波および長波による熱フラックス、降水や河川からの淡水フラックス、風による運動量フラックスがMSSG-Oへ情報を伝達する。それらの情報を受け取ったMSSG-Oは数値計算を行い、海洋の循環場を決定する。その計算結果のうち海洋表面温度（Sea Surface Temperature: SST）をMSSG-Aに伝達し、シミュレーションの1ステップが進行する。MSSGでは、大気と海洋の解像度は同じであることとしており、物理量交換部では、観測値と整合性をとるような補正を行っていない。

3. 1で述べた並列化の基本方針に加え、計算性能を向上するために、MSSG-A、MSSG-Oそれぞれに対し、以下の手法を徹底的に用いた計算性能最適化を行った。

- 最内側ループのベクトル化によるベクトル効率の向上
- マイクロタスクのオーバーヘッドを削減するためのループ融合の実施
- インライン展開によるベクトル化の促進
- 依存計算のループ外への移動によるループ長の確保
- サブルーチンの統合による小規模処理の減少
- ライブラリ関数の計算アルゴリズム最適化
- 通信部の統合による通信効率の向上

以上の最適化後、大気：MSSG-A、海洋：MSSG-O、結合：MSSGそれぞれに対して、地球シ

ミュレータ上での計算性能を測定した結果を表2に示す。まず、ループ長を確保する戦略が活かされ、ベクトル長が保持されるとともに、徹底的なベクトル化を行った結果が、高いベクトル化率に反映されている。また、ピーク性能比は、大気コンポーネント **MSSG-A** では、約 **60%** を示し、実アプリケーションとして、地球シミュレータの性能を最大限に活用可能な性能を達成した。結合モデル **MSSG** においては、海洋コンポーネント **MSSG-O** における性能の影響を受け、ピーク性能比は **MSSG-A** に比較して低いものの、結合モデルとしては非常に高い計算性能である。**MSSG-O** におけるピーク性能比は、**MSSG-O** のみに用いているポアソンソルバーの性能を反映している。しかしながら、並列化にあまり適さないと一般的に称されるポアソンソルバーに対して、**MSSG-O** では非常に高い計算性能を達成した。本稿では、本ポアソンソルバーについての計算性能最適化についての詳細は割愛するが、**AMGCG** 法の計算効率の向上が本性能の達成に非常に大きく貢献していることを付記しておく。**MSSG, MSSG-A, MSSG-O** における並列化率のいずれもが、**99.99%**以上を実現しており、この性能は、理論的には、地球シミュレータの約 **60** 倍の大きさの計算機まで、計算効率が線形に向上することに相当する。

表 2: **MSSG, MSSG-A** および **MSSG-O** の地球シミュレータ上での計算性能。

表における結合、大気、海洋は、それぞれ、**MSSG, MSSG-A** および **MSSG-O** を表す。注目すべきは、ピーク性能比、加速率、並列化効率であり、実アプリケーションとして地球シミュレータの性能を限界まで使用していると判断できる。

ケース	ノード数	CPU数	格子点数	Mflops/CPU	ベクトル長	ベクトル化率	Tflops	ピーク性能比	加速率	並列化率
結合	512	4096	3662807040	4166.7	229	99.3%	17.07	52.1%	461.0	99.9973
	384	3072		4273.8	229	99.3%	13.13	53.4%	354.6	99.9968
	256	2048		4401.9	229	99.3%	9.02	55.0%	242.6	—
大気	512	4096	2713190400	4575.2	228	99.5%	18.74	57.2%	479.1	99.9983
	384	3072		4606.1	228	99.5%	14.15	57.6%	365.2	99.9969
	256	2048		4692.4	228	99.5%	9.61	58.7%	247.5	—
海洋	498	3984	4718592000	3629.3	240	99.3%	14.46	45.4%	401.3	99.994
	398	3184		3568.5	240	99.3%	11.36	44.6%	333.7	99.989
	207	1656		4234.3	240	99.3%	7.01	52.9%	188.2	—

4. **MSSG** を用いたマルチスケール・マルチフィジックスシミュレーション

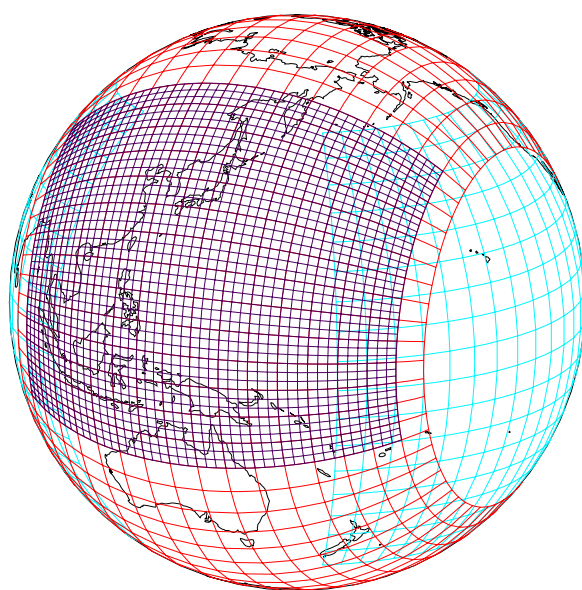
気象現象がどのように気候変動に影響をあたえるのか、また、その逆に、気候変動がどのように個々の気象現象に影響を与えるのか、についての答えは、大きな関心事であることは先に述べたとおりである。そのためには、比較的小さな空間スケールで比較的短い時間スケールの気象現象と、空間的には全球規模で、時間的にも長いスケールの気候変動を同時に扱えるモデルである必要がある。そのためには、大気、海洋の各コンポーネントが非常に大規模な問題を、高速に計算できることが必要不可欠であるので、**MSSG** は、そのために、地球シミュレータ上での計算性能を最大限に向上させた。

気象現象のなかでも台風に関連する集中豪雨などの現象に深いかかわりがある積雲の成り立ちや発達、消滅までの一連の過程の予測には、雲の単位である約 **100m**程度の解像度が求められることが多い。また、近年では、赤道上的様なスケールの雲が、全球への熱的循環に深く関与していることが示唆され、やはり積雲の単位となる **100m**程度の解像度でのシミュレーションと気候変動予測が密接に関連している。

気象と気候変動の異なる時空間スケールの現象の関連性を扱うためには、全球をできる限り

詳細で、一様な解像度でシミュレーションすることがひとつの可能性として存在する。後述するように、MSSG-Aを用いた場合、地球シミュレータ上で全球に対して最も高解像度は、水平1.9 kmである。従って、水平解像度を100m程度にするには、計算時間や大気海洋結合モデルでなければならないことも考慮すると、少なくとも地球シミュレータの約1000倍のスーパーコンピュータが必要となり、2012年に神戸にて稼働予定であるペタスケールスーパーコンピュータの次の世代を待つことになる。

一方、対象とする領域詳細に気象現象と気候変動を結びつけることが可能かどうか、についての県空も進められている。その手法のひとつは、ネスティング手法と呼ばれる手法であり、対象とする領域を詳細なグリッドでシミュレーションし、その詳細領域の外側のグリッドは、比較的粗いグリッドでシミュレーションを行う。図10に、MSSGでネスティング手法を取り入れた場合の例を示す。



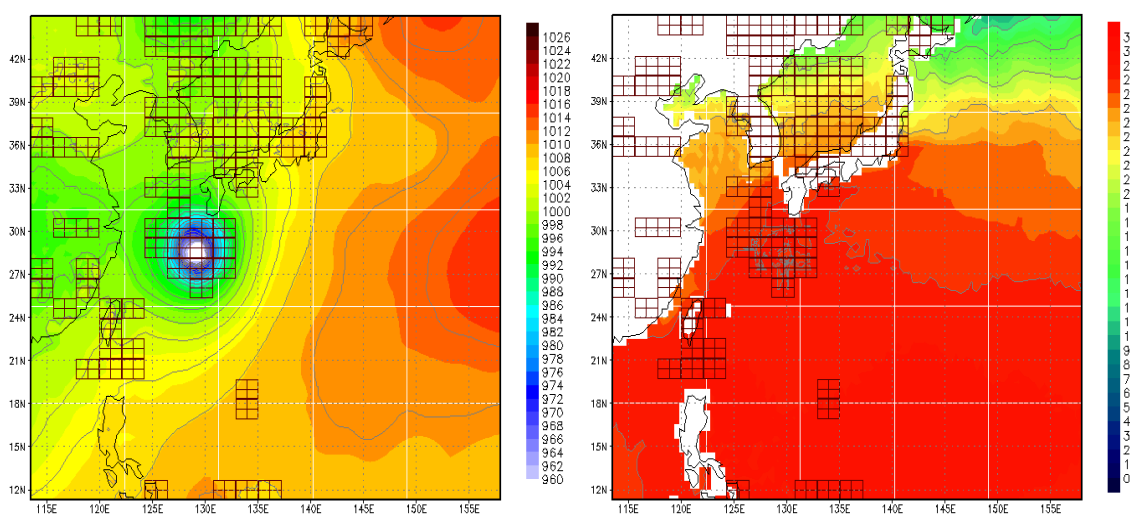
第10図: Yin-Yang グリッド上における全球と詳細解像度領域のネスティング手法による結合の模式図。

群青色で示した詳細グリッド領域は、水平境界値として粗いグリッドの全球シミュレーション結果を補間して与える。全球から領域への境界条件としての情報提供のみを行う結合手法をワンウェイ・ネスティング(1-way nesting)といい、全球から領域へ、同時に領域から全球へ情報を両報告へ交換する場合は、その結合手法をツーウェイ・ネスティング(2-way nesting)という。1-way, 2-wayともに多くの情報の交換手法が提案されている[28][29][30]。

群青色で示した詳細グリッド領域のシミュレーションは、外側の粗い解像度のシミュレーションの結果を境界条件として使用する。境界条件をどのように与えるかについては、様々な手法があるが、スポンジ領域という計算が行われる領域の周辺の補間計算のための領域を設け計算される領域と境界の外側の領域の値をスムーズに保間するための緩衝領域を設けることが多い。この緩衝領域の定義の仕方は、計算精度がよい補間手法を用いる必要があること、流出条件においては、特に注意が必要であることなどが挙げられる。

詳細グリッド領域を用いたシミュレーション手法の長所は、数日単位の比較的短時間のシミ

シミュレーションであれば精度よく、詳細領域の現象を再現、あるいは予測できる点である。しかしながら、数か月以上の予測においては、詳細領域の外側の気象現象に相当するスケールの影響を反映したシミュレーションが困難であるので、気象現象と気候変動との関連メカニズムをシミュレーションすることは難しい。粗いグリッド領域の気象現象も扱うことが可能であり、気候変動シミュレーションも可能とする手法として、適応型動的メッシュ法がその候補として考えられる。適応型動的メッシュ法は、気象現象の単位となる積雲などが発生した場合に、その領域の実を詳細にシミュレーションできるようなグリッドを配置し、気象現象としては変化が速くない領域を粗いグリッドとして計算する手法である。適応型動的メッシュ法により、グリッドを配置した例を図 1 1 に示す。



第 1 1 図: 大気海洋結合モデル MSSG に適応型動的メッシュ法を用いたときの大気、海洋それぞれにおいて設定される格子の模式図。

気圧がある一定以下になることを格子を定義するクライテリオンとしたときの、MSSG の大気コンポーネントモデル MSSG-A (左図) と海洋コンポーネントモデル MSSG-O (右図) において設定された動的適応型格子の例。白線であらわした格子が最も解像度が低く、実線で定義された解像度が最も高い解像度を表す。この例では、クライテリオンに従って台風の周囲では高い解像度で格子が配置されている。MSSG では、大気と海洋の解像度は、物理量を精度よく保存するために同じ解像度を設定しているため、左図と右図は同様の格子配置となっている。台風が移動するとともに、格子も移動して定義され、解像度が高い領域において、降雨量、風の分布などが観測とよく一致して再現されることを確認している。

MSSG には、上記のネスティング手法と適応型動的メッシュ法の両方の手法が実装を完了している。すでに、ネスティング手法と適応型動的メッシュ法の両方について、それぞれの手法を用いた場合の計算性能は、地球シミュレータ上で最適化されている(表 3)。両手法の計算性能最適化についての詳細は本稿では詳述しないが、基本的な方針は、Yin-Ynag グリッド上での計算性能最適化の手法とほとんど変わらない。

本稿の以下の章では、MSSG の様々なスケールの現象の再現性を示し、特に、適応型動的メッシュ法の基盤となるネスティング手法により、比較的短時間の現象を再現可能かどうかについて、検証した結果を紹介する。

表 3：動的適応型格子を用いた時の MSSG の地球シミュレータ上での計算性能。

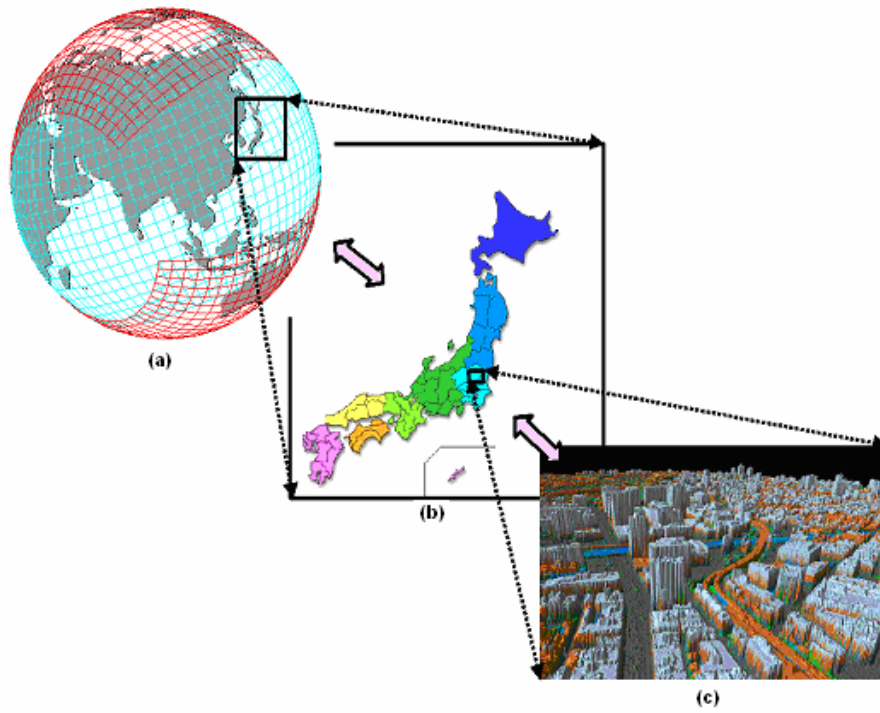
下表において、大気は、動的適応型格子を大気海洋結合モデル MSSG に用いた際の大気コンポーネント MSSG-A、および海洋コンポーネント MSSG-O、MSSG-A と MSSG-O の情報交換、AMR に関連する格子配列に要する計算、それぞれについての計算時間を示したものである。この設定では全球、日本領域、日本領域における動的適応型格子を設定した。全球と日本領域は 1-way nest 手法を用いて結合し、日本領域上で動的適応型格子を配置する。nest 欄の 0.0, 1.1, 2.1 は、それぞれ、0.0 は、全球 11 km, 1.1 は日本領域 5.5 km, 2.1 は 2.75 km の動的適応型格子上で計算時間と演算割合を示す。Mgrid は各解像度のグリッド数、step は計算したステップ数、Mgrid * step は、格子数とステップ数の積和、elaps はそれぞれの設定での計算時間を表す。nest 関連の欄は、全球と日本領域の nest 処理に必要な計算時間の割合、reinit は、動的適応型格子を設定する際に要した計算時間の割合である。計測から、大気と海洋が同じ解像度のときには、大気モデル MSSG-A の計算時間が支配的であること、nest や動的適応型格子の設定に必要な計算量は、ごく少量であることがわかる。

	nest	Mgrids	steps	Mgrids*steps	%	elapse	%	nest関連	%	reinit	%
大気	0.0	181	17738	3210578	69%	15781.9					
	1.1	34	17738	603092	13%	2964.6					
	2.1	46	17670	812820	18%	3995.5					
				4626490		22742.0	90.9%	1028.7	5%	1.1	0.0%
海洋	0.0	77	22140	1704780	43%	964.3					
	1.1	45	22140	996300	25%	563.5					
	2.1	61	21280	1298080	32%	734.2					
				3999160		2262.0	9.0%	363.9	16%	1.4	0.1%
結合								0.7		0.0	
全体						25015.2		1393.3	6%	2.5	0.0%

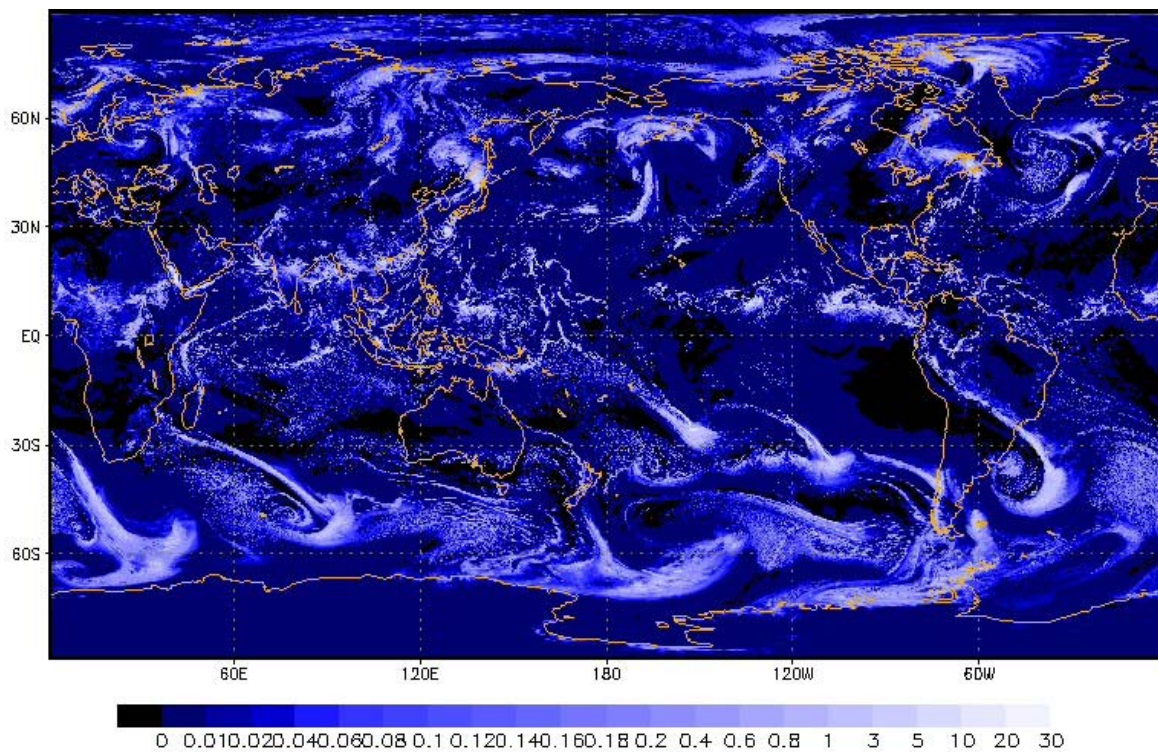
5. MSSG を用いた事例シミュレーション結果

MSSG の開発の本来の目的は、気候変動などの全球スケールの現象や、その影響が領域のスケールにどのような影響をあたえるか、また、さらに詳細な都市スケールの気象や気候変動にどのような影響があるのかを統一的に扱えるような大気海洋モデルを構築することである。図 1 2 に示すように、全球、領域、都市スケールのそれぞれのスケールの現象に焦点をあてて、詳細化することによりその目的を達成した。本章では、それぞれのスケールに焦点をあて、MSSG を用いて実施したいくつかの事例シミュレーションの結果を紹介する。

まず、大気コンポーネント MSSG-A を使用した全球シミュレーションの結果を紹介する。図 1 3 に示すシミュレーション結果は、水平解像度 1.9km、鉛直層としては 30 km までの高度を不等間隔で 32 層に設定した、総計約 50 億グリッドポイント上のシミュレーションである。この規模の 3 次元全球大気のシミュレーションは世界的にも前例がない超大規模なシミュレーションであり、現実的な地形を設定して、全球大気全体の特長と地球シミュレータ上での計算の限界を同定することを目的に行ったものである。このシミュレーション結果から、台風や前線の降雨の様子や、前線構造に沿った降雨分布が日変化としてとらえられ、地形の影響を大きく受ける熱帯域の降水分布などがとらえられていることを確認している (図 1 4)。

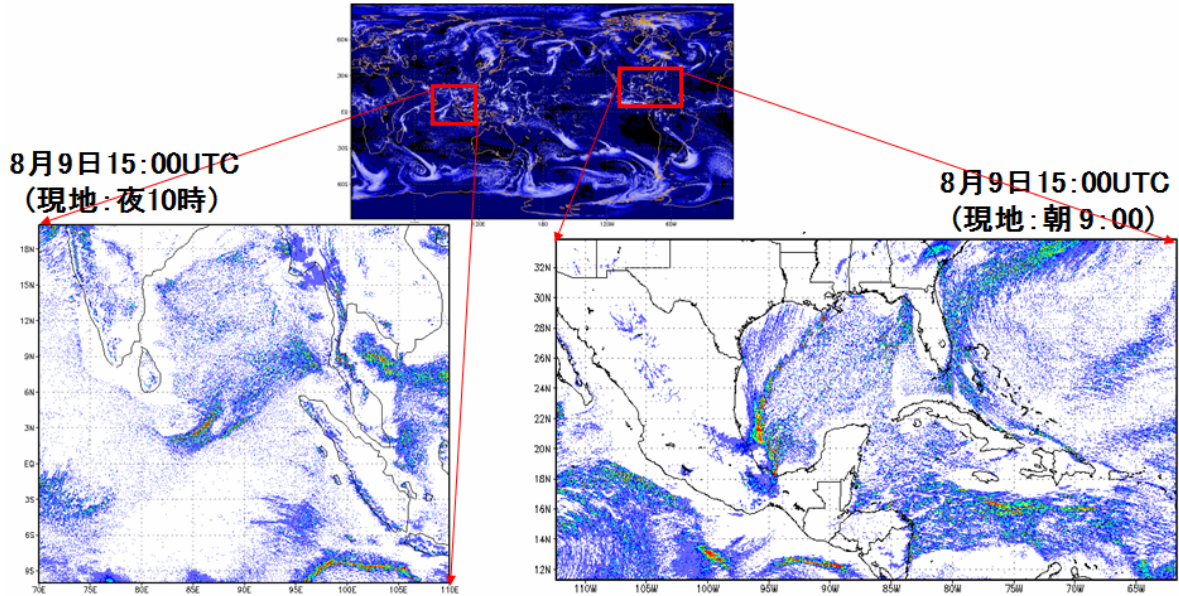


第12図: MSSG が対象とするスケールの模式図。



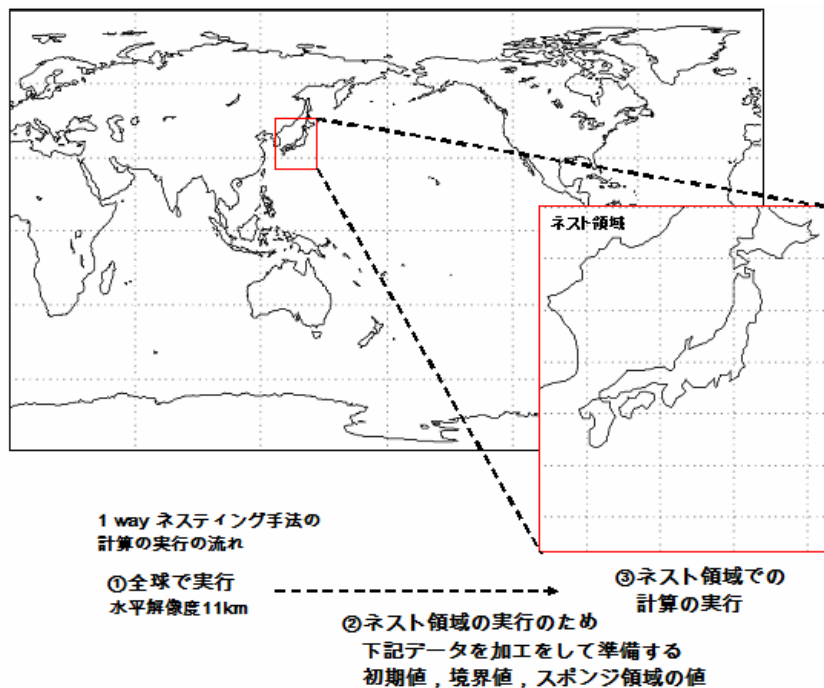
第13図: 水平解像度 1.9 km 鉛直 32 層の全球 MSSG-A を用いたシミュレーションにおける降水分布。

水平1.9km, 鉛直32層の全球シミュレーション



第14図: 図13をインドネシア領域, 北アメリカ南部を拡大した降水量分。

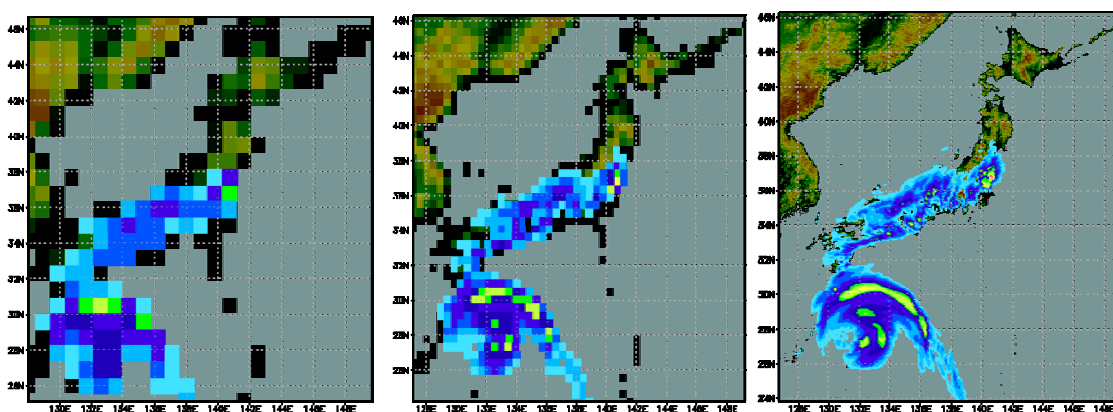
左図は, インドネシア領域, 右図はメキシコ湾を中心とした領域の降水分布のスナップショットを示す。左図でスマトラ島の沿岸に沿って分布する降水, 右図ではメキシコ湾内に発生した前線による詳細な降水分布がとらえられている。



第15図:MSSG-A を使用して全球と日本領域を 1-way nest 手法 (全球から領域への1方向情報伝達手法) を用いたときの計算の概要。

全球で水平解像度 11 km, 鉛直 32 層の計算を行い, nest する領域 (ここでは, 赤線で囲む日本領域が nest 領域に相当する) の初期値, 境界値を求める。nest 領域の境界では, スポンジ領域とよばれる緩和領域を設け nest 領域で計算される値と全球でもとめた境界値との計算値との不連続性を各ステップごとに緩和する。日本領域では, 緩和された境界値を使用して, 計算を進める。

水平解像度の違いが、気象現象の再現や予測精度に大きな影響があることは、従来より指摘されてきた。それは MSSG-A においても同様である。MSSG-A では、図 1 5 に示すような手法で全球を計算し、その結果を対象領域の境界条件として与えることで、領域内のグリッドを詳細にしてシミュレーションを行うことが可能であるので、地球シミュレータ上で 2003 年台風 10 号を、日本領域を対象領域として水平解像度が異なる 100km, 50km, 5km の設定で予測シミュレーションを行った。図 1 6 のシミュレーション結果は、水平解像度を詳細にすることで、台風の特徴的な構造の再現が大きく異なることを示している。水平解像度が 5km であるシミュレーション結果は、台風の目や豪雨をもたらす筋状の雲の構造（レインバンド）を、観測から得られるような現実に近い構造として再現している。ここでは、紙面の関係上、結果を示していないが、水平解像度が 5km 以下であるほうが、最低気圧も現実の観測値に近い値を予測できることもわかってきた。



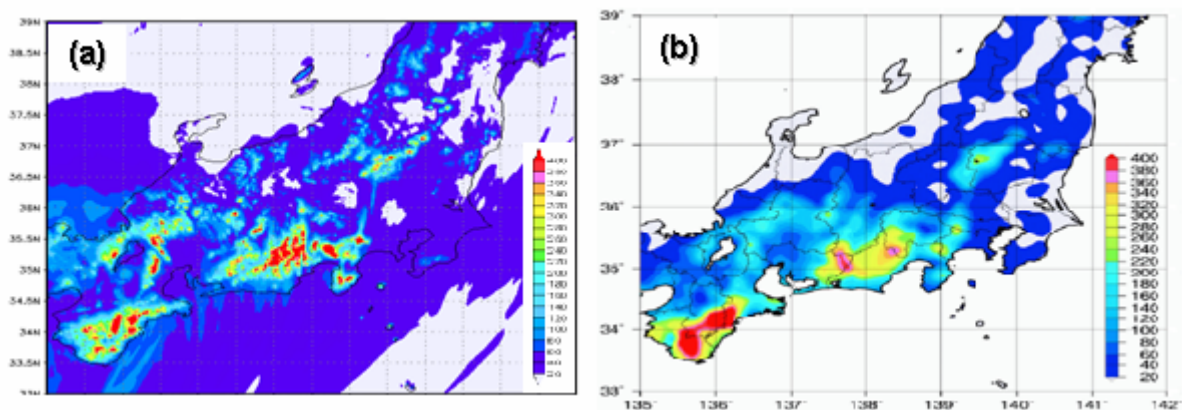
第 1 6 図: 解像度の違いによる降雨量の違いと台風構造の再現性の違い。

2003 年台風 10 号を、日本領域において異なる水平解像度で、鉛直層はいずれも 32 層として設定したときの予測シミュレーション結果を示す。左図は、水平解像度 100km、中央の図と右図は、水平解像度がそれぞれ 50km, 5km にした時の予測シミュレーション結果であり、降雨量 (mm/hour) 分布のスナップショットである。

図 1 7 は、同じ 2003 年台風 10 号に対して、日本領域をさらに詳細に設定し、水平解像度 1.13km、鉛直層数 32 層で再現シミュレーション結果から算出した積算降雨量である。MSSG-A によるシミュレーション結果は、観測データのアメダスの設置間隔である約 17km より詳細な情報を提供していることになる。今後、左図の結果を各観測地点ごとの積算降雨状況と詳細に比較検証する必要があるが、水平解像度を詳細にすることにより、地形も詳細に表現されることから、地形性降雨もより詳細に再現できることになる。

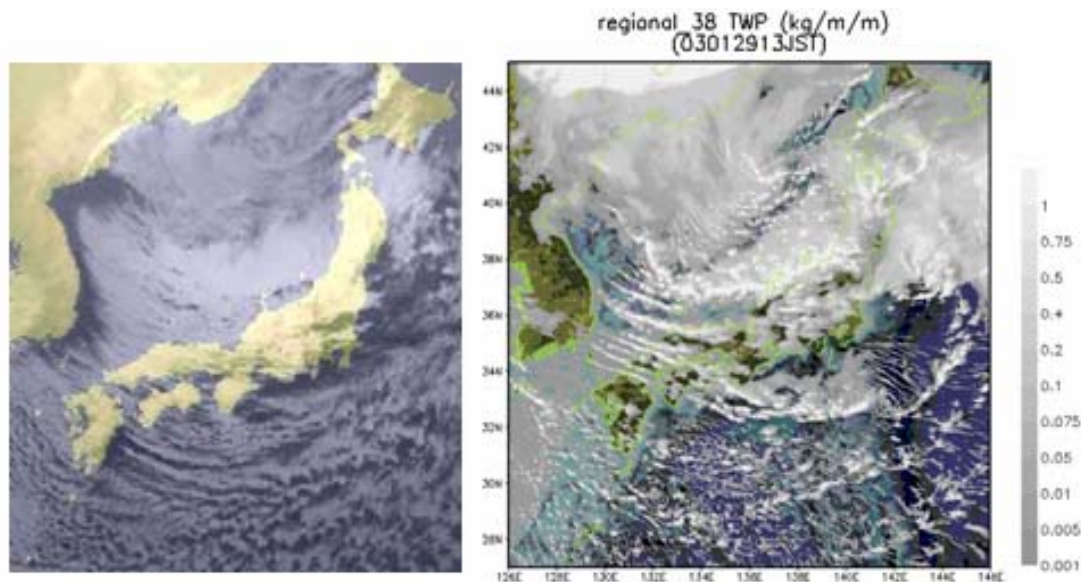
図 1 8 は、図 1 7 とは異なる季節の冬季において、日本領域で観測される顕著な現象のひとつである日本海の寒気吹き出しの再現シミュレーション結果である。日本領域の水平解像度は 1.3km、鉛直は 30 km 高度までを不等間隔で 32 層とした設定は、図 1 7 と同様である。図 1 8 は、衛星赤外面像とシミュレーションによって得られた光学的な雲の厚さをそれぞれ示す比較した結果である。日本海側における発達した帯状雲と太平洋側の筋状雲がよく再現できている。

図17, 図18に示した, MSSG-A を用いて日本領域を詳細に計算したシミュレーションには, 初期値は, (財) 気象業務支援センターから配布されている全球を約 100 kmの水平格子間隔のグリッドデータ (GPV:Grid Point Value; 格子点値) と, 日本領域には, やはり (財) 気象業務支援センターから配布されているメソ数値予報モデル GPV (MSM) の格子間隔 20kmメッシュに細分化されたデータを, それぞれ線形補間を行い, より詳細な解像度の MSSG-A の初期値として使用した。全球 11km の結果を日本領域に相当するネスト領域の境界値として使用し, 18 時間のシミュレーションを行った。



第17図: 2003年台風10号に対して, 日本領域を水平解像度 1.13km, 鉛直層数 32 層で再現シミュレーションした結果から算出した積算降雨量と観測との比較。

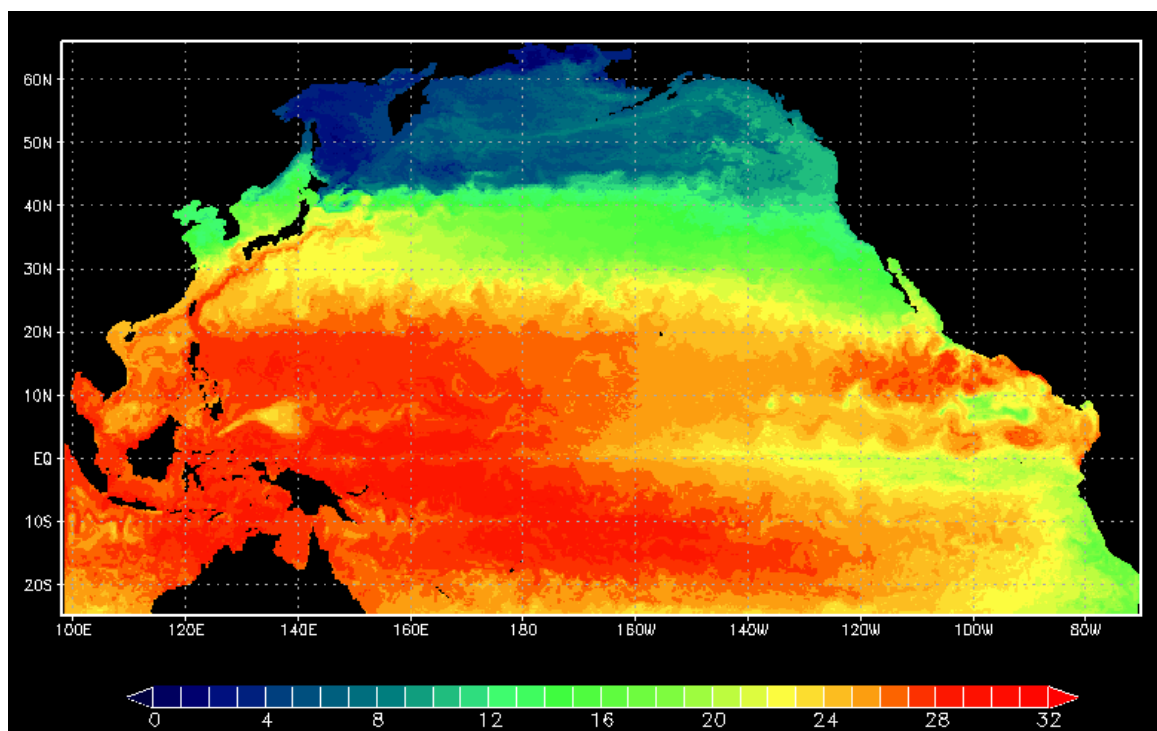
左図(a)は, 対象領域の水平解像度を 1.13km に設定したときのシミュレーション結果より算出した積算降水量, 右図(b)は東京管区気象台から発表されたアメダスのデータによる観測から得られた積算降水量である。

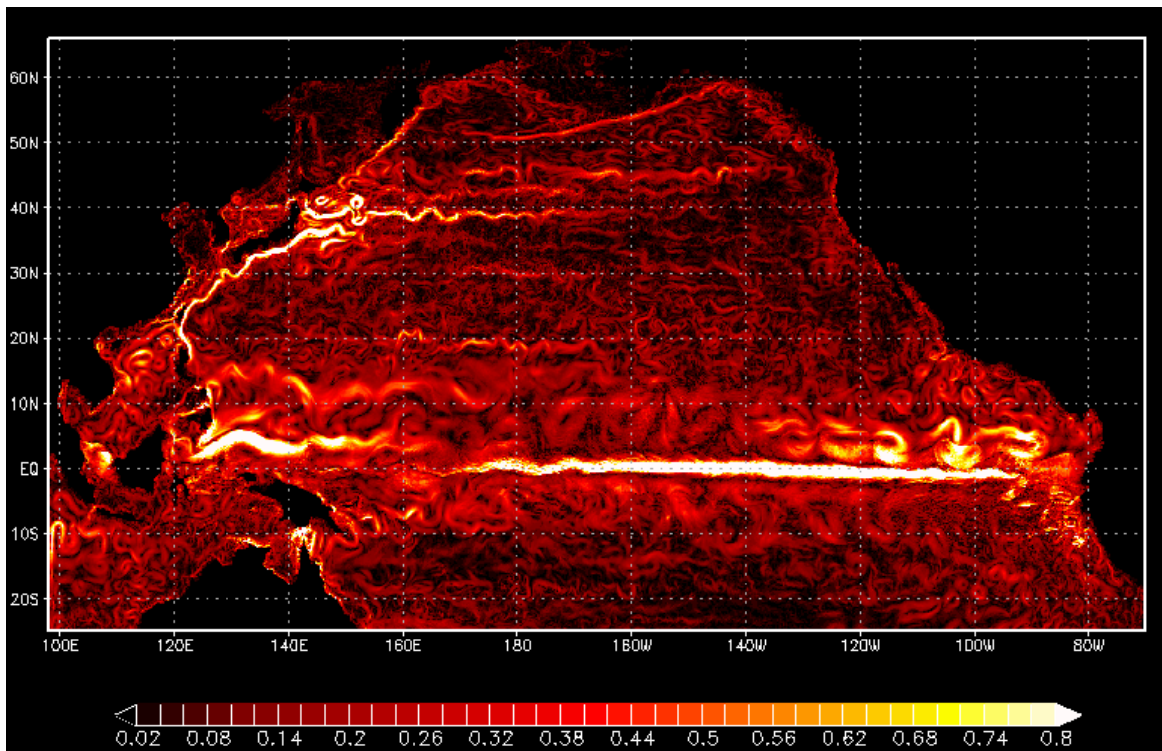


第18図: 冬季日本海における寒気吹き出しの再現シミュレーション結果。

2003年1月29日13時における衛星赤外面像(左図)とシミュレーションによって得られた光学的な雲の厚さ(右図)を比較した結果である。冬季に特徴的な日本海側における発達した帯状雲と太平洋側の筋状雲が, シミュレーション結果においてよく再現されている。

MSSGの大気コンポーネントMSSG-Aと同様に、海洋コンポーネントMSSG-Oにおいても、大気の状態を境界値データとして与えることで、全球の海洋や領域に焦点をあてたシミュレーションが可能である。ここでは、北太平洋の領域に対するシミュレーション結果を紹介する。格子間隔は水平方向に約11km、鉛直層数は水深5kmの深さまでを40層に設定してシミュレーションを実行した。水平方向の境界条件、つまり陸と海洋との境界は、速度にディリクレ境界条件、温度、塩分にノイマン境界条件を設定している。海面における境界条件は、大気の状態から算出した風応力による運動量と流速、および温度、降雨や河川からの流入による淡水フラックス、塩分のフラックスの流入と流出を設定した。風応力として、National Centers for Environmental Prediction (NCEP)が配布している観測データから換算した再解析データのなかから、1年の12ヶ月分に相当する各月の平均データを時間方向に線形補間して、シミュレーションの各時刻における境界データとして用いた。温度・塩分には、WOA98と称される1989年から1999年の11年間を平均して作成した月平均気候値データを用いた。初期条件は、温度・塩分には、WOA98の1月気候値をMSSG-Oに適合する解像度になるように線形補間して与え、初期の速度場は静止状態とした。16年シミュレーションした後の、水深5mの温度(図19上図)と100m深の流速の水平分布(図19下図)を示す。初期の流速は0であったが、風応力で海表面が駆動され、赤道反流、黒潮、親潮などいくつかの特徴的な海流が再現されている。北緯30以上は、16年の積分期間では、十分に表現されないことが知られていることから、日本に沿って流れる黒潮流域は、銚子より北で西向きに流れの方向を変えておりオーバーシュートが見られる。温度場についても、水平解像度11kmで得られることが知られている渦構造が得られており、定性的に観測値と一致した結果が得られた。

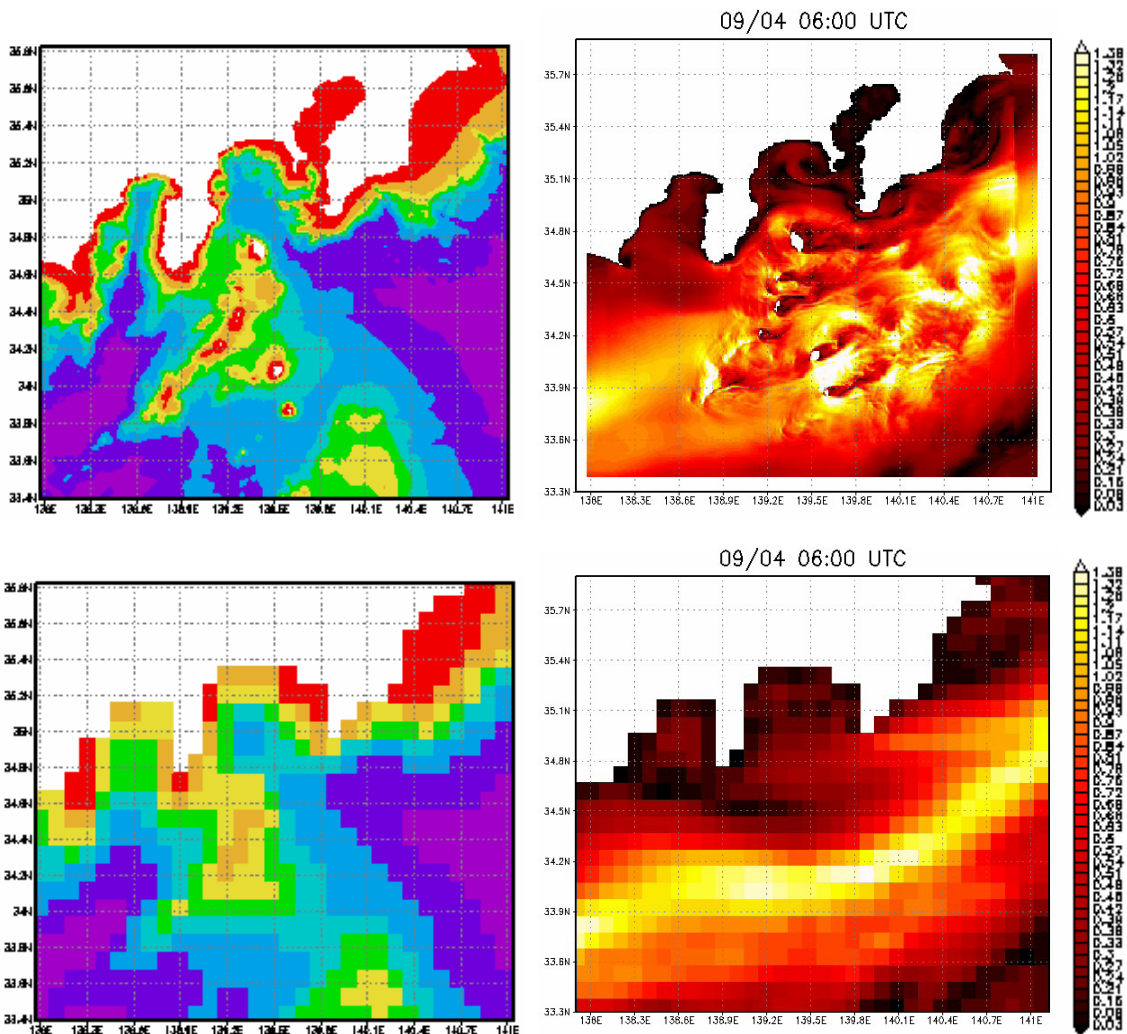




第19図: MSSG-Oにより再現された北太平洋における水温と流の場。

16年間積分した最後の10年について平均をとった夏の北太平洋の5m深の海面水温分布(上図)と100m深の流の速さを示している(下図)。上図では、日本海域の黒潮の様子、インドシナ域の多島海の東側の warm pool (ワーム・プール) とよばれる30度にも達する温かい海域の様子、東北太平洋赤道域にある cold tongue (コールド・タン) と称される温度の低い深い海水の湧昇を期限とする冷たい海域の分布が、観測データと比較してもリアリスティックに表現されている。下図からは、海洋の内部が様々なスケールの渦が構成されていることがわかる。これらの渦は、熱を伝達するために大きな役割を果たしている。赤道域に東から西に流れる焼く1 m/secの速さで流れる赤道潜流もよく再現されている。

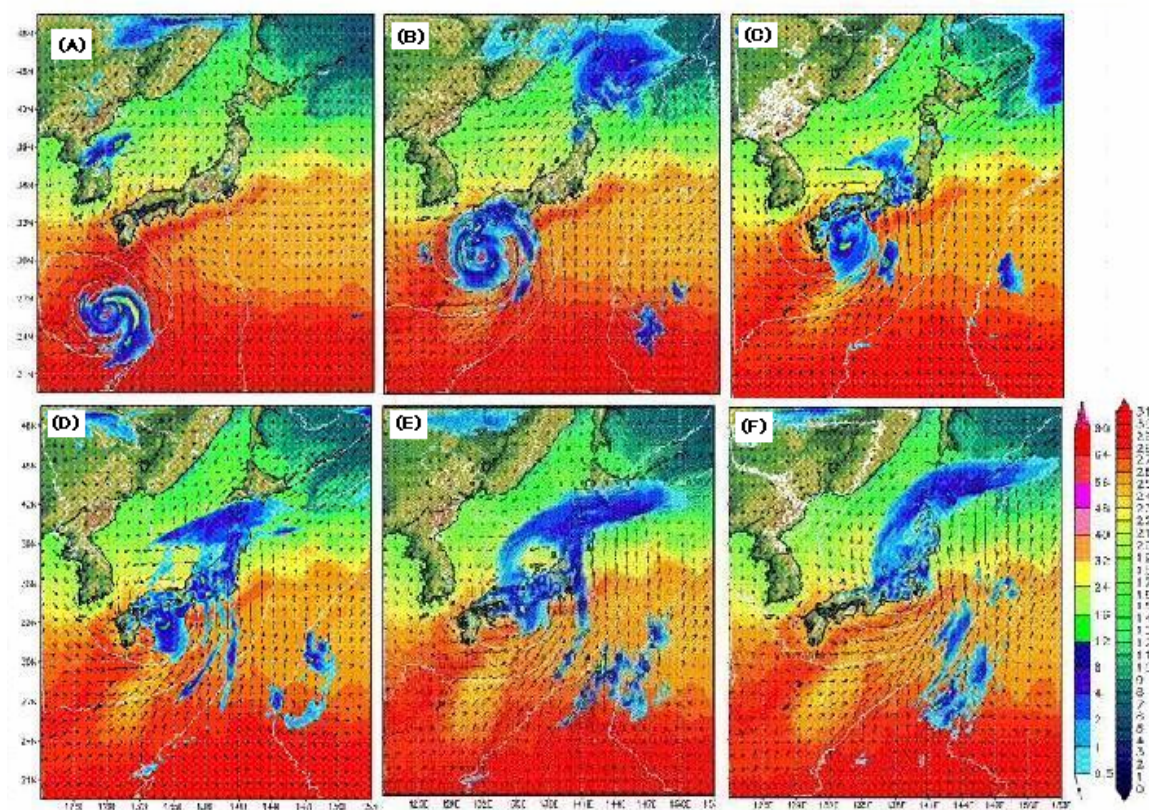
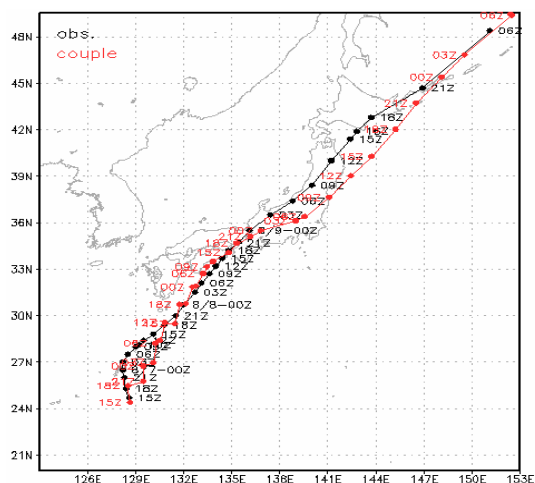
MSSG-Oを用いて、さらに詳細な海洋の流れをシミュレーションすることが可能である。日本沿岸における海洋の流れの状況についてのシミュレーションや予測は、さまざまな産業を支える大きな社会的な意義がある。例えば、東京湾、相模湾、駿河湾と黒潮流域との相互影響を調べるために、より詳細な解像度でシミュレーションが必要であり、MSSG-Oはそれを可能にする。図20(a)は、水平750m解像度でのMSSG-Oを用いたシミュレーション結果であり、図20(b)は水平11km解像度でのシミュレーション結果である。これらのシミュレーションを比較すると、湾内や湾岸域のシミュレーションには、11kmの水平解像度では不十分であることがわかる。水平750mの水平解像度を設定したシミュレーションでは、海岸線や海底地形が詳細に表現でき、島や湾も解像可能である。海岸線と海底地形が詳細になったことに起因して、流れ場が劇的に変わることがシミュレーション結果から明らかである(図20)。島々の存在は、流れの下流方向に剥離渦を形成し、岬の先端からも剥離渦が形成されることが確認できる。これらの渦の存在は、海表面温度の分布や流れ場の違いを生起する原因のひとつである。数100mのオーダーの水平解像度と数kmの水平解像度のシミュレーション結果とのドラステイックな違いは、雲の生成過程や降雨分布に影響を与える可能性がある。



第 20 図: MSSG-O を用いた解像度水平 750m と 11km でのシミュレーション結果。

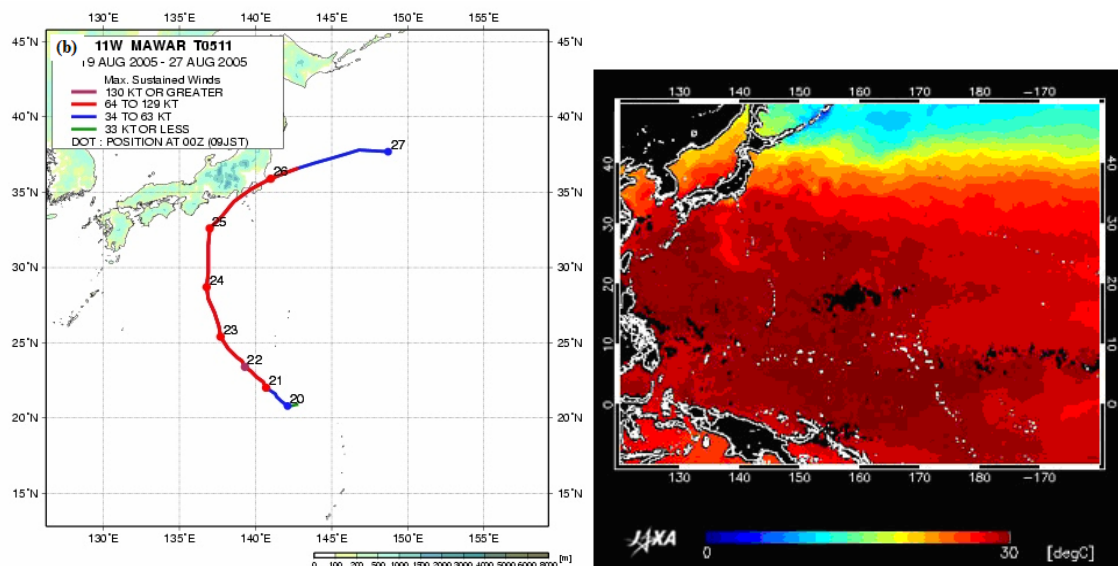
上段の 2 つの図は水平 750m 解像度での海岸線と海底地形の深さ分布 (左図), と MSSG-O によるシミュレーション結果 (右図) である。下段上段の 2 つの図は水平 11km 解像度での海岸線と海底地形の深さ分布 (左図), と MSSG-O によるシミュレーション結果 (右図) である。シミュレーション結果は、いずれも 2003 年 9 月 4 日日本時間 15:00 の海表面の速さを示しており、約 1 m/sec の速さの黒潮の分布がとらえられている。左列の海岸線と海底地形の深さ分布を比較すると、海流の再現におおきな影響を与える日本域の沿岸や湾内、海底地形の様子は、11km の解像度では不足しており、少なくとも 1km より高い解像度のシミュレーションが必要であることがわかる。11km の解像度では、東京湾、相模湾、駿河湾を解像することができない。また、上段右図の海流の速さの分布から、島々の影響で、島後部の黒潮の流れは、渦の剥離による影響を受けていることがわかる。沿岸域の海流の様子は、これらの島々の影響を考慮したシミュレーションが必要であることが示唆される。

日本領域において、大気と海洋をともに水平 2.78km, 鉛直 64 層で設定した大気海洋結合モデル MSSG によって予測シミュレーションした結果を紹介する。大気海洋結合モデルにより、2003 年台風 10 号の進路と強度を 120 時間 (5 日間) 予測シミュレーションした結果を図 21 に示す。大気と海洋の相互作用は、台風の強度予測に不可欠である。海洋は、台風に熱的エネルギーを供給するだけでなく、熱的供給によって発達する台風からも強い風から水平方向、鉛直方向にも運動エネルギーを受けて、流れ自体が変化する。これらのことから、台風の強度、



第 2 1 図: 大気海洋結合モデル MSSG による 2003 年台風 10 号の 8 月 6 日 9 時から 120 時間予測シミュレーション結果。

上図は、2003 年台風 10 号の進路を示す。黒実線は気象庁から発表されたベストトラックを示し、赤実線は予測シミュレーションの結果である。下図は、台風による強い風に応答する海洋表面温度の変化と降雨分布を 3 時間ごとの海水表面温度 (、降雨分布 (青色の分布)、風速ベクトル分布の変化を示す。台風の眼を含む特徴的な構造が、時間経過とともに衰退し変化する様子が、シミュレーションの降雨分布からとらえられていることがわかる。また、海洋表面温度の変化の時系列応答から、台風の進路の東側後方の温度が、台風の移動に伴ってより低い温度へ変化していることがわかる。これは、台風の風からの運動量の伝搬によって、海洋表面の流れ場が変化し発散するながれの傾向となり、より深い海洋の温度の低い海水が表面に現れたことによるものである。海表面の温度は、台風への熱的供給源であることから、台風と海洋との相互作用は、台風の強度 (風と降雨) に影響を与えることから、予測精度へ影響を与えることを示唆している。



第 2 2 図：衛星観測からとらえられた 2005 年台風 11 号の大気海洋相互作用（JAXA 提供）。

左図は、観測から得られた台風 11 号の進路、右図は、衛星がとらえた台風 11 号が通過したのちの海表面温度の分布を示す。左図の台風が通過した経路に沿って、右図における北緯 20–35°，経度 135–145° で囲まれる領域に、海表面の温度が周辺域の温度に比べて低い領域がある。これは、台風の通過にともなう強風によって、海表面の温度変化がもたらされたものである。

台風の風の強さや降雨量の予測には、大気と海洋の相互作用を考慮することが重要であると広く認識されており、大気海洋結合モデルによる予測は、さらに精度のよい予測シミュレーションを可能にするものと期待されている。MSSG によって予測シミュレーションした結果である図 2 1 下図は、台風の強風によって海洋表面上の温度が変化し、台風の進路や強度に影響を与えていることを示唆している。これらの温度変化は衛星観測からもとらえられており（図 2 2），海洋表面の温度が変化することで、大気への熱的エネルギーの供給が変化することから、台風の強度に変化を与えることを示唆している。今後は、さらに、高解像度の大気海洋結合モデルの予測シミュレーションには、これまで観測データからでは困難であった台風の発生や発達、維持、消滅に関する科学的な解明が進むことが期待できる。また、あらたに、高解像度における大気海洋相互作用をどのようにモデル化すべきであるか、あるいは、予測精度の向上にどれほどのインパクトがあるのか、などについての新たなテーマに取り組む必要性があることも明記しなければならない。

さらに詳細なスケール、都市スケールでのシミュレーションを MSSG-A を用いてシミュレーションした結果を示す。東京の都心のある街区（有楽町周辺域：図 2 3）を代表例として、土地利用データ（図 2 3）、道路、交通量、人口排熱等のデータを整備し、気象データをシミュレーション対象領域の境界値として与えた水平、鉛直ともに 5m の解像度でのシミュレーションを行った。このシミュレーションは、ヒートアイランドなどに代表される都市域における熱的循環を再現し、都市域の蓄熱効果や今後のクーリング環境を効果的に選定するための基礎情報都市手利用することが可能である。水平には、気象業務支援センターから配布されている MSM データ（前述）を線形補間して 5m 解像度に適合するようにして初期値を設定した。MSSG-A は、この初期状態から、時間変化を伴う非定常な気温変化、風速の変化などを再現できる。図 2 4 は、2005 年 8 月 5 日 15 時から 10 分間の変化をシミュレーションした結果のうち、異な

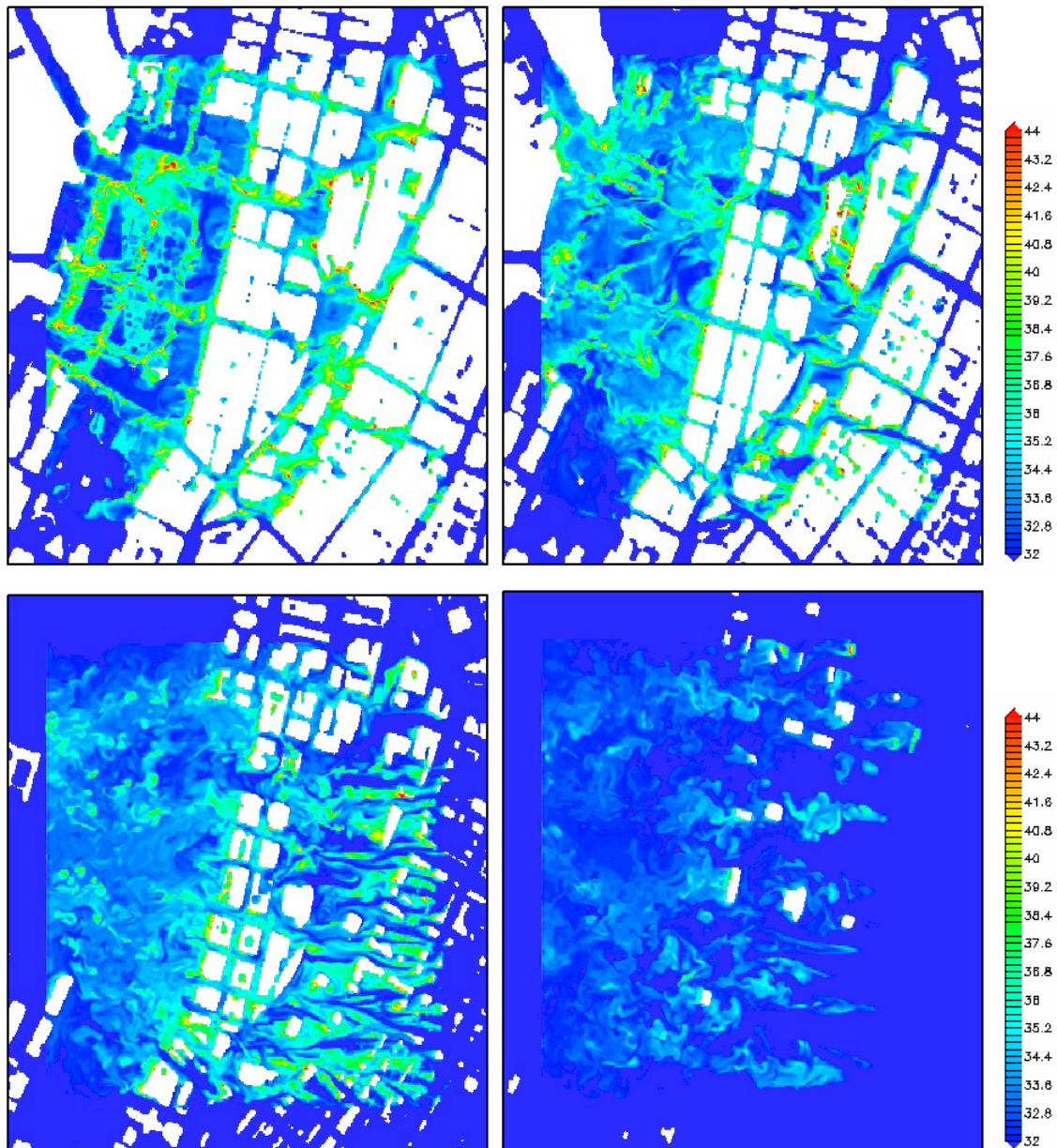
る高さでの温度分布のスナップショットである。地表面から暖められた大気がプルームとなって上昇する様子や、建物形状により風の非定常な流れのようすなどがダイナミックに再現されている。

図25は、ある鉛直断面の流れの様子と温度分布のスナップショットを示している。ビルの風上、風下には渦が生成され、そこで比較的温度的の高い大気が渦を形成している様子がとらえられている。また、ビルより高い上空での大気は、ビルで囲まれる比較的低い大気の流れとは一線を画して、滑らかに流れており、ちょうど低層の温かい大気をパッキングするような状況がとらえられている。このような、ビルとビルに囲まれた大気の流れと、ビルより上空の大気の流れの相互作用は、都市域において、どのように熱エネルギーが蓄積されるかを理解する上で重要であり、そのメカニズムをとらえることが、今後、効果的に都市をクーリングさせるかの手法とその選択を裏付けるための、重要な情報として活用されることが期待できる。

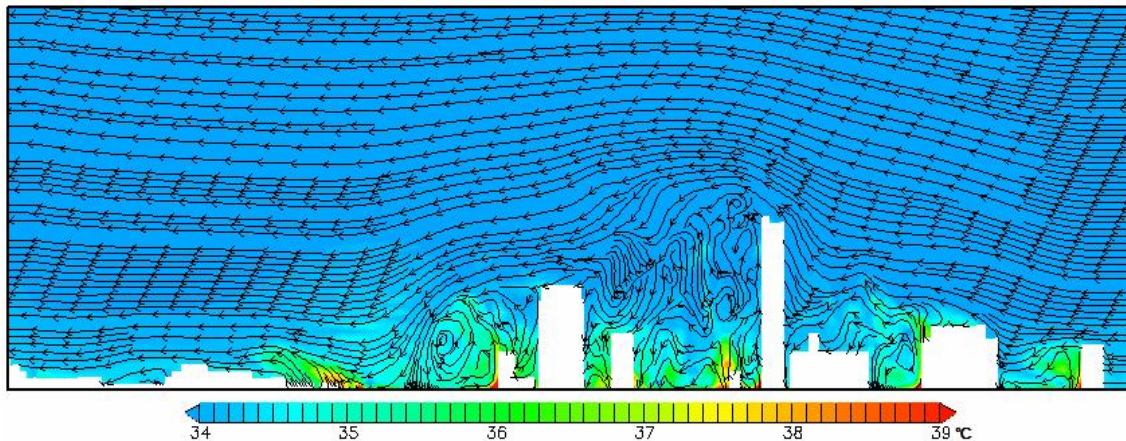


第23図: シミュレーションの対象領域である東京の都心街区の航空写真と土地利用データによる分布。

左図は、シミュレーションの対象領域の航空写真、右図はその領域の土地利用データである。シミュレーションでは、土地利用データに従って、熱の反射や水分含量を一定の境界値として与え、その影響を反映した。道路上の熱的データ、交通量や人口排熱等のデータも整備し、これらも固定した一定の境界値として与えている。対象領域の外側の境界値としては、気象業務支援センターから配布されているMSMデータ(前述)を補間して気象データとして与え、シミュレーションを行った。



第24図: 2005年8月5日15時から10分間のシミュレーションを行った10分後の結果。高さ7.5m(上段左図), 12.5m(上段右図), 32.5m(下段左図), 102.5m(下段右図)における温度分布の15時10分におけるスナップショットを示す。道路上で熱せられた大気塊が、各高さにおいて、シミュレーション時の東風に移流される様子がとらえられている。高さ7.5m, 12.5mに分布する暖められた40度を超える大気塊は、一定の場所に留まることなく、揺らぎながら剥離を伴って移流する。32.5m, 102.5mの温度分布において、建物に付着するような高温ポイントは、人工排熱による境界条件による大気の高温化をとらえている。高温の大気塊の分布やその移流による変化が、都市のヒートアイランドにどのような影響を及ぼすかについては、さらに広域のシミュレーションの統計的な解析が必要である。



第25図: 鉛直断面の流れの様子と温度分布のスナップショット。

2005年8月5日15時は、東風が卓越していた。建物の風上、風下には渦が生成され、高温の道路上やビル壁面からの影響を受けた、比較的湿度の高い大気が渦を形成している様子がとらえられている。また、ビルより高い上空での大気は、ビルで囲まれる大気の流れとは一線を画して、滑らかに流れており、ちょうど低層の温かい大気をパッキングするような状況がとらえられている。

6. MSSGによるシームレスシミュレーション

この10年間に、台風の進路予測の精度は飛躍的に進歩したという評価は、多くの気象学者に受け入れられている。しかし、台風の強度予測についてはまだまだこれからの課題である、という評価も、気象学者のほとんどが認めるところである。台風の強度は最低気圧、最大風速、降雨量によって計測され、災害にもっとも直接的に関係するこの台風の強度の予測には、何が足りないのか。大きく分けて3つの原因が考えられている。ひとつは、予測を行う際のシミュレーションにおける解像度が十分にとれなかったことにより、台風の特徴的な構造を再現できなかったこと、2つ目は、大気と陸地、あるいは大気と海洋の相互作用についての知見不足により、そのメカニズムが予測モデルに十分に反映できていないこと、3つ目は、大気と海洋と陸地の相互作用を取り入れた結合モデルによる予測が、これまで比較的少なかったこと、などが挙げられている。これらは、まさしくマルチスケール・マルチフィジックスな気象シミュレーションの重要性を言い換えるものであり、異なるスケールのシームレスなシミュレーションの進歩が、予測精度の向上に結び付くことを示唆している。

加えて、地球温暖化に伴い、エルニーニョ、インド洋ダイポールなどの気候変動や気候変化に伴って、台風などの局所的な災害をもたらす気象現象がどのような変化を受けるのか、については、社会的な関心が高く、これまでも日本を含む世界の各機関が取り組んできた課題である。近年、IPCC レポートによる人為的な温暖化の進行が確実であることの報告を受け、温暖化時における台風の強度についての関心はさらに高まっており、これまでも増して台風の強度予測に関する積極的な研究開発が展開されている。

上記のような世界的な潮流の中、地球シミュレータセンターでは、地球シミュレータを最大限に活用してこそ可能な、大気海洋結合モデルによる予測シミュレーションの問題規模はどれほどか、また、その予測精度はどうか、についての研究開発をMSSGの開発を介して推進してきた。本稿では、MSSGのモデルとしての概要、超並列かつ高速計算を実現した基本的方針と地球シミュレータ上での計算性能、また、台風の強度予測などの地球シミュレータ上でのみ可

能な超大規模な再現シミュレーションや予測シミュレーションの事例を示し、現時点における予測研究の一端を紹介した。台風や豪雨、ヒートアイランド現象の解明など、私たちの生活に直結するような現象の解明は、雲の生成の単位である数 10m から数 100m のスケールと、それを取り巻く領域の気象や気候変動の変化を同時にとらえてゆく必要があり、また、温暖化の影響をも考えると、全球と領域、さらに都市スケールの気象や気候変動までを考えてゆく必要がある。これは、まさしく、マルチスケールの気象、気候変動予測をシームレスにつなげることが可能なモデルによる予測が必要であることを意味している。

このシームレスなシミュレーションは、非常に大規模なシミュレーションが必要であることと同時に、高い解像度の必要性のみならず、高解像度に対応した物理モデル、高精度の計算手法、高速計算を併せて開発、実現してゆく必要性があり、MSSG はそのための有望なひとつのツールとして位置づけることができる。地球シミュレータ上の MSSG は、地球シミュレータ上において最大限の超高解像度でのシミュレーションを可能とし、10km 水平解像度でのシミュレーションという当初目標に比較すると、さらにその約 100 倍もの規模のシミュレーションを可能とした。しかしながら、地球シミュレータをもってしても、雲の単位のスケールの現象を解像し、さらに複数のスケールの相互影響を扱うには非力であることも明らかとなった。今後は、神戸に完成され 2011 年度から稼働するスーパーコンピュータが、シームレスシミュレーションの実現を担うことになるだろう。そのための基盤的なツールは、揃ってきたということもできる。今後のマルチスケール・マルチフィジックスを扱うことが可能なシームレスシミュレーションによる気象、気候変動現象の予測シミュレーションの実現への挑戦と進展から、目を離せない状況が続くことは間違いない。

参 考 文 献

- [1] Palmer, T. N., Doblas-Reyes, F. J., Weisheimer, A. and Rodwell, M. J., 2008, Toward seamless prediction: Calibration of Climate Change Projections Using Seasonal Forecasts, *BAMS*, April 2008, 459-470.
- [2] Ohfuchi, W., H. Sasaki, Y. Masumoto, and H. Nakamura, 2007: "Virtual" Atmospheric and Oceanic Circulation in the Earth Simulator. *Bull. Amer. Meteor. Soc.*, **88**, 861-866.
- [3] Satomura, T. and Akiba, S. 2003. Development of high-precision nonhydrostatic atmospheric model (1): Governing equations, *Annuals of Disas. Prev. Res. Inst.*, Kyoto Univ., 46B, 331-336.
- [4] Lilly, D. K., 1962. On the numerical simulation of buoyant convection. *Tellus*, **14**, 148-172.
- [5] Smagorinsky, J., Manabe, S. and Holloway, J. L. Jr. 1965. Numerical results from a nine level general circulation model of the atmosphere. *Monthly Weather Review*, **93**, 727-768.
- [6] Zhang, D. and Anthes, R.A. 1982. A High-Resolution Model of the Planetary Boundary Layer - Sensitivity Tests and Comparisons with SESAME-79 Data. *Journal of Applied Meteorology*, **21**, 1594-1609.
- [7] Blackadar, A. K. 1979. High resolution models of the planetary boundary layer. *Advances in Environmental Science and Engineering*, **1**, Pfafflin and Ziegler, Eds., Gordon and Breach

- Publ. Group, Newark, 50-85.
- [8] Reisner, J., Ramussen R. J., and Brientjes, R. T. 1998. Explicit forecasting of supercooled liquid water in winter storms using the MM5 mesoscale model. *Quart. J. Roy. Meteor. Soc.*,
- [9] Davies, H. C. 1976. A lateral boundary formulation for multi-level prediction models, *Quart. J. R. Met. Soc.*, 102, 405-418.
- [10] Marshall, J., Hill, C., Perelman, L. and Adcroft, A. 1997. Hydrostatic, quasi-hydrostatic, and nonhydrostatic ocean modeling. *Journal of Geophysical Research*, 102, 5733-5752.
- [11] Marshall, J., Adcroft, A., Hill, C., Perelman, L. and Heisey, C. 1997. A finite-volume, incompressible Navier Stokes model for studies of the ocean on parallel computers. *Journal of Geophysical Research*, 102, 5753-5766.
- [12] A. Gill, 1982, *Atmosphere-Ocean dynamics*, Academic Press Inc. .
- [13] Mellor, G. L. and Yamada, T. 1974. A hierarchy of turbulence closure models for planetary boundary layers. *Journal of Atmospheric Sciences*, 31, 1791-1806.
- [14] Kageyama, A. and Sato, T. 2004. The "Yin-Yang Grid" : An Overset Grid in Spherical Geometry. *Geochem.Geophys.Geosyst.*, 5, Q09005, doi:10.1029/2004GC000734.
- [15] Wicker, L. J., Skamarock, W.C, 2002. Time-splitting methods for elastic models using forward time schemes, *Monthly Weather Review*, 130, 2088-2097.
- [16] Z. Wang, A fully conservative interface algorithm for overlapped grid, *J. Comput. Phys.*, 122, 96-106, 1995.
- [17] Berger, M. J., On Conservation at Grid Interfaces, *SIAM J. NUMER. ANAL.* 24, 967-984, 1987.
- [18] Peng, X., Takahashi, K., et al. Global mass conservation scheme on Yin-Yang Grid system. Proc. of 2004 Spring Meeting of the MSJ, May 2004 (in Japanese).
- [19] Gal-Chen, T. and Somerville, R. C. J. 1975. On the use of a coordinate transformation for the solution of the Navier-Stokes equations. *Journal of Computational Physics*, 17, 209-228.
- [20] Peng, X., Xiao, F., Takahashi K. and Yabe., T. 2004. CIP transport in meteorological models. *JSME international Journal (Series B)*, 47(4), 725-734.
- [21] Komine, K. 2005. Validation Results from Non-hydrostatic Atmospheric Simulations. 2005 *SIAM Conference on Computational Science and Engineering*, Florida, USA, Feb 2005.
- [22] Ohdaira, M., Takahashi, K. and Watanabe, K. 2004. Validation for the Solution of Shallow Water Equations in Spherical Geometry with Overset Grid System" in Spherical Geometry. *The 2004 Workshop on the Solution of Partial Differential Equations on the Sphere*, Yokohama, July 2004, 71.
- [23] Takahashi, K. et al., Proc. 7th International Conference on High Performance Computing and Grid in Asia Pacific Region, 487 (2004).
- [24] Takahashi, K. et al., 2004. "Non-Hydrostatic Atmospheric GCM Development and its computational performance",
http://www.ecmwf.int/newsevents/meetings/workshops/2004/high_performance_computing-11th/presentations.html

- [25] 足永靖信, ヴ, “空間平均処理を施した3次元都市キャノピーモデルの開発 都市建築計画における都市気候予測システムの開発 その2,”日本建築学会環境系論文集, 17, pp.45-51, 2004.
- [26] 佐川玄輝, 氷盤衝突を考慮した海氷力学モデルの開発とそれを用いたオホーツク海の海氷変動数値計算, 博士(工学)論文(東京大学), 2006.
- [27] Burk, S. D., and W. T. Thompson : A vertically nested regional numerical prediction model with second-order closure physics. *Mon. Wea. Rev.*, 117, 2305-2324.境界条件 p 13, 1989.
- [28] Clark, T.L., and W.D. Hall,: Multi-domain simulations of the time dependent Navier-Stokes equations: Benchmark error analysis of some nesting procedures. *J. Comput. Phys.*, 92, 456-481, 1991.
- [29] Clark, T.L., and R.D. Farley: Severe downslope windstorm calculations in two and three spatial dimensions using anelastic interactive grid nesting: A possible mechanism for gustiness. *J. Atmos. Sci.*, 41, 329-350, 1984.

地震波伝播と強震動の大規模並列 FDM シミュレーション

古村 孝志

東京大学 大学院情報学環総合防災情報研究センター/地震研究所地震火山災害部門

1. はじめに

地球の表面を覆う硬い岩盤（プレート）が複雑に衝突し合う日本列島は、世界の大地震の1割以上が発生する大地震多発地帯である。これまで日本ではマグニチュード（M）7級の内陸地震が年間1～2回、そしてM8級の海溝型巨大地震が10年に1回程度の割合で起きてきている。国の中央防災会議によると、今後30年以内にM7級の首都直下地震が発生する確率は70%、そしてM8級の東海地震の発生確率は87%にもなるという。現在の地震学は、このように地震発生の「長期評価」に関しては高い知見を与えるが、地震がいつどこで起きるかを数日間の短い時間スケールで予測する「短期予知」は依然として困難な課題である。地震被害の軽減には、まず過去の被害地震や将来の大地震の強い揺れ（強震動）の特徴をコンピュータシミュレーションにより再現・予測し、その結果を構造物の耐震設計に生かすことが大切である。

地震は、地下のプレート境界や内部に蓄積した歪みエネルギーが解放される現象である。そして、プレートがずれ動き、プレート内部が破壊した結果、強い地震波が放射される。地震波は、弾性波として地下を伝わり、不均質な地殻・マントルを伝わるにつれ複雑で長い揺れに成長する。地震波が地表に到達すると、柔らかい表層地盤で何倍も強く増幅される、そして堆積平野に地震波が閉じ込められるため長く揺れが続く。このような、複雑な地震動の挙動の評価には、不均質な断層運動と地下構造を適切にモデル化した、地震波伝播の数値シミュレーションが必要である。

地震波伝播の数値計算の歴史は古く、差分法（Finite-Difference Method: FDM）や有限要素法（Finite-Element Method; FEM）を用いた2次元モデルでの計算は、1960年代後半に既に開始されている。しかしFDM計算では複雑な地下構造のモデル化に大きな計算機資源を必要とするため、現実的な3次元地下構造モデルを用いた計算が実用化したのは、1990年代後半以降のことである。

本稿では、まず始めに、3次元不均質場での地震波伝播を記述する基本方程式と、FDM計算による地震波伝播シミュレーションの具体的な数値計算手順を説明する。次に、地球シミュレータに代表されるベクトル型計算機を用いた大規模3次元計算と、近年のT2Kオープンスパコン(東大)*に代表される新しいスカラー型並列計算機の可能性を検討する。そして、地震動シミュレーションの有効性を示すために、近年の日本の大地震のシミュレーションと波動場の可視化事例をいくつか紹介する。同時に、シミュレーションに用いる支配方程式の地震現象への整合性と、計算結果の信頼性の確認を、近年日本列島に高密度に展開された強震観測データと比較から検討する。最後に、現在のシミュレーション技術の成果と限界、そして今後5～10年を考えた次世代地震動シミュレーションの到達点を考える。

*<http://www.cc.u-tokyo.ac.jp/ha8000>

2. 地震波伝播の FDM 計算

2-1. 運動方程式の FDM 計算

地震による断層破壊時の歪みのレベルは高々 $10^{-6} \sim 10^{-5}$ 程度に小さく、これに伴って弾性体中に現れる地震波動が作り出す歪みレベルも同程度である。したがって、地中を伝わる地震波動は弾性体中の無限小歪みの現象と考えて良い。表層地盤では剛性率が小さく、強い地震動により大変形が起きるため、媒質の非線形応答を考慮する必要がある。しかし、地盤の非線形性応答はごく表層部分（地下数メートル以浅）に限られるため、地殻・マントルを伝播する地震波を線形計算で求めてから、さらに必要に応じて地盤の非線形応答の評価を行うのが一般的である。

2-1-1. 運動方程式

3次元直交座標における線形の地震波動場は、以下の釣り合いの式（運動方程式）により記述できる：

$$\rho \ddot{u}_p = \frac{\partial \sigma_{xp}}{\partial x} + \frac{\partial \sigma_{yp}}{\partial y} + \frac{\partial \sigma_{zp}}{\partial z} + f_p, \quad (p = x, y, z) \quad (1)$$

ここで、 \ddot{u}_p は加速度、 σ_{pq} は応力、 ρ は密度、そして f_x は外力を現す。すなわち、微小の媒質に働く加速度は、媒質周辺での応力差と外力（震源により起こされる力）により発生する。

等方完全弾性体では、応力と歪み e_{pq} は、Lame の定数 λ, μ を用いて次の構成方程式で結びつけられる：

$$\sigma_{pq} = \lambda(e_{xx} + e_{yy} + e_{zz})\delta_{pq} + 2\mu e_{pq}, \quad (p, q = x, y, z) \quad (2)$$

ここで、 δ_{pq} はクロネッカのデルタを現す。いっぽう、歪みは変位の空間微分から次式で求められる：

$$e_{pq} = \frac{1}{2} \left(\frac{\partial u_p}{\partial q} + \frac{\partial u_q}{\partial p} \right), \quad (p, q = x, y, z) \quad (3)$$

FDM 計算では、3次元領域全体を細かな格子間隔($\Delta x, \Delta y, \Delta z$)で分割し、各格子点上に地下構造に対応する物性定数(λ, μ, ρ)を独立に置くことにより、任意の不均質構造をモデル化することができる。また、これらの物性値は、人工地震探査や深層ボーリング坑を利用した音波探査などから求められる。

地震波計算でFDMが多用される最大の理由は、地下構造の境界面の形状や物性値が厳密に定義されていないため、たとえばFEMやBEMのように境界面に沿った格子生成により境界条件を高精度に導入する利点が小さいためである。地震動シミュレーションの重要な目的の一つとして、計算された地震波形と地震観測データとの比較を繰り返して、モデルの物性値の修正を行うことがある。このような繰り返し計算において格子形成を毎回やり直すことは困難であり、規則格子を用いるFDMはかえって使い勝手が良い。

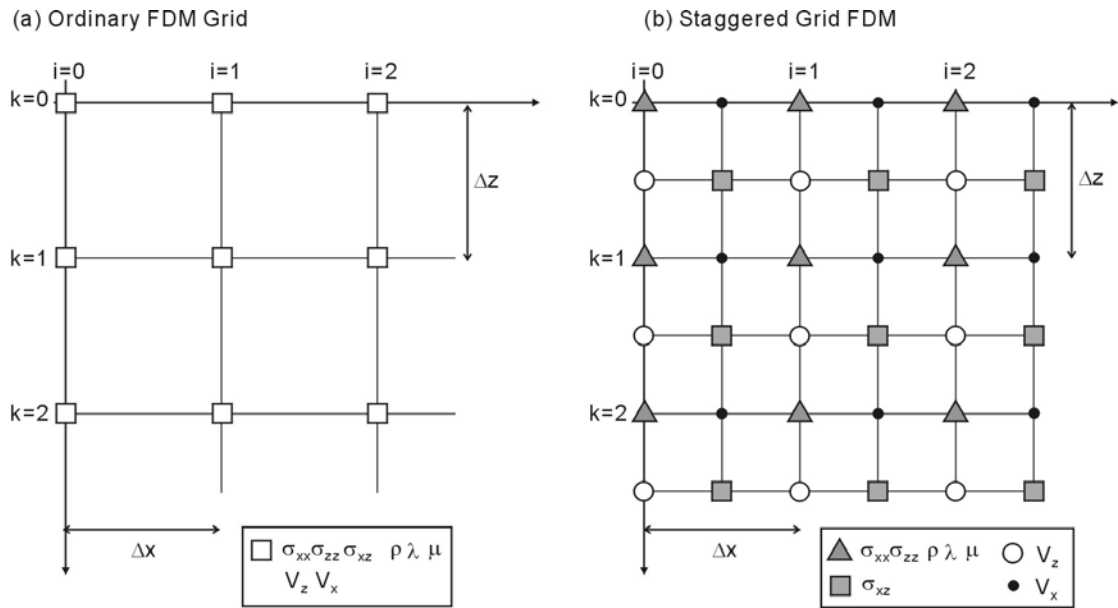


図1 地震波伝播のFDM計算に用いられる格子モデル。(a)一般格子，(b)食い違い格子モデル。食い違い格子（スタガード格子）では，応力と速度，弾性定数の変数が互いに半格子ずれて定義される。

地震断層運動に伴って放射される地震波は，断層運動と等価な力（**物体力**）を，式(1)の右辺に加えることにより実現する。なお，震源からは，縦波（**P波**）と横波（**S波**）が放射されるが，ダイナマイトや核爆発のような人工地震ではP波のみが，また回転（トルク）成分を持つような人工震源からはS波のみが放射される。数値シミュレーションでは，このようにして**自然地震**や**人工地震**に対応した任意の震源をモデルに組み込むことができる。

2-1-2. 差分近似計算

地震波伝播のFDM計算では，かつて（1970～80年代）変位や応力，物性値など全ての変数を同一格子点上に配置する格子モデルが多用されたが（図1a），現在は図1bに示されるように変数どうしを半格子ずれた位置に定義する「**食い違い格子（スタガード格子）**」が一般に用いられるようになった。スタガード格子では，変数の微分が定義される位置に関連の変数が位置するため計算の都合が良く，また計算精度が良いためである。

スタガード格子では，たとえば式(1)に現れる $\sigma_{pq}(x, y, z)$ の x 微分を，隣接の変数値

$\sigma_{pq}(x \pm \Delta x/2, y, z), \sigma_{pq}(x \pm 3\Delta x/2, y, z), \sigma_{pq}(x \pm 5\Delta x/2, y, z), \dots$ を用いて次より評価され

る：

$$\begin{aligned} & \frac{d}{dx} \sigma_{pq}(x, y, z) \\ &= \frac{1}{\Delta x} \sum_{m=1}^{M/2} c_m [\sigma_{pq} \{x + (m + \frac{1}{2})\Delta x, y, z\} - \sigma_{pq} \{x - (m - \frac{1}{2})\Delta x, y, z\}], \end{aligned} \quad (4)$$

ここで， c_m は中央差分の係数を表し，たとえば2次精度のスタガード格子中央差分では $c_1 = 1$ ，4次精度では $(c_1, c_2) = (9/8, -1/24)$ ，そして8次精度では $(c_1, c_2, c_3, c_4) = (9/8, -1/24, 49/5120, -5/7168)$ の値が用いられる。

これらの FDM 演算公式は、微分演算を **Taylor 展開**した後に高次項の打ち切りにより作成される。打ち切り誤差の影響は、短波長の地震波ほど大きく、地震波の伝播速度が周波数毎に異なり波形が崩れる現象 (**数値分散**) が発生する。その影響は、低次の FDM 演算子ほど大きい。数値分散の影響を十分に小さく抑え、高い FDM 演算精度を得るためには、計算で扱う必要のある地震波の波長に比べて、 Δx を十分に小さく選ぶ必要がある。なお、スタガード格子 FDM では、全ての変数を同一格子点上に定義する場合と比べて、格子間隔を半分 ($\Delta x/2$) に選んだ場合と同等の精度が得られる [1]。

3次元 FDM 計算では、式(1)と(3)に合計 18 回の微分演算が現れる。この計算を、3次元領域全体にわたって x, y , あるいは z 方向に沿って連続的に実行される。実は、この演算が FDM 計算全体の 8～9 割以上の時間を占めている。大規模 3次元 FDM 計算では、FDM 演算自体の高速化が最優先課題である。

2-1-3. FDM 演算の高精度化・高速化

計算メモリと計算時間の節約のために、格子間隔が粗くとも高い精度が得られる最適化 FDM 演算子の研究がこれまで進められた [1-3]。一般的な FDM 演算子の誤差は高周波地震動ほど影響が大きくなるが、これを低周波～高周波全域に一樣に広げることにより、数値分散の発生を弱めようというものである。また、2組の FDM 演算子を用いて FDM 計算を 2 回実行し、計算結果を足し合わせることで、数値分散をちょうど打ち消す画期的なアイデア [4-6] も提案された。

高次 FDM 計算を、**高速フーリエ変換** (FFT) を用いて効率良く実行する「**擬似スペクトル法**」も、3次元 FDM 計算でよく用いられた [7-8]。擬似スペクトル法では FFT の周期境界条件が暗黙に科せられる。擬似スペクトル法の一つであり、離散化チェビシェフ変換を用いる **チェビシェフスペクトル法** [9] では、開放端 (自由境界条件) と固定端 (剛体) の 2 つの境界条件を選ぶことができるため、応用範囲が広い。

2-1-4. 地震波伝播の陽的計算

時間発展により地震波伝播を進めるために、式(1)を速度変数を中間変数とする以下の方程式に置き換え、また式(2)と(3)を結合した以下の式を用いて、中央差分に基づく時間積分が行われる：

$$\dot{u}_p^{n+1/2} = \dot{u}_p^{n-1/2} + \frac{1}{\rho} \left(\frac{\partial \sigma_{xp}^n}{\partial x} + \frac{\partial \sigma_{yp}^n}{\partial y} + \frac{\partial \sigma_{zp}^n}{\partial z} + f_p^n \right) \Delta t, (p = x, y, z) \quad (5)$$

$$\sigma_{pq}^{n+1} = \sigma_{pq}^n + \left[\lambda \left(\frac{\partial \dot{u}_x^{n+1/2}}{\partial x} + \frac{\partial \dot{u}_y^{n+1/2}}{\partial y} + \frac{\partial \dot{u}_z^{n+1/2}}{\partial z} \right) \delta_{pq} + \mu \left(\frac{\partial \dot{u}_p^{n+1/2}}{\partial q} + \frac{\partial \dot{u}_q^{n+1/2}}{\partial p} \right) \right] \Delta t, (p, q) = (x, y, z)$$

(6)

すなわち、現時刻 ($t = n\Delta t$) の応力値 σ_{pq}^n を用いて、時刻 ($t = (n+1/2)\Delta t$) の速度値 $\dot{u}_p^{n+1/2}$ を求め、次にこの値を用いて次の時刻 $t = (n+1)\Delta t$ の σ_{pq}^{n+1} を求める。ここでは、速度変数 $\dot{u}_p^{n+1/2}$ と応力変数 σ_{pq}^n は $\Delta t/2$ だけずれた時刻に定義されていることを意味している。

こうして、式(5)と(6)を繰り返し計算することにより、地震波伝播計算が進行し、地震波が震源から計算領域全域に広がる。ここで、時間積分の精度の要請により、1タイムステップ(Δt)に地震波が進む距離は、格子サイズよりも十分に小さいことが要請される。そこで、 Δt はモデル中の地震波の最大伝播速度 V_{max} と最小格子間隔を用いて、以下より定める：

$$\Delta t < a \frac{\min[\Delta x, \Delta y, \Delta z]}{V_{max}}, \quad (7)$$

ここで、係数 a (< 1) が小さいほど式(5)(6)の時間積分の精度を高くすることができるが、計算全体に必要なタイムステップ数が増え、計算時間が増大する。そこで、FDM 計算の効率化の観点から、空間微分の演算精度との兼ね合いで a の最適値が決められる。たとえば4次精度のFDM 計算では $a=0.45$ 程度の値が、またより高精度の16次以上の高次FDM 計算では $a=0.24$ 程度の値が用いられる [10]。

2-2. 物性境界条件

地震波が物質境界に入射すると、反射波と屈折波が発生するほか、P波からS波へ、あるいはS波からP波への変換波も次々発生して次第に複雑な波形が形成される。地震波が気体／固体境界である、地表面に入射すると、地表面に沿って伝播する「**表面波**」と呼ばれる波も発生する。これらの地震波動現象をFD計算から正確に再現するためには、物質境界における変位と応力に関する境界条件の適切なモデル化が必要である。

まず、地表面 ($z=0$) では応力値がゼロとなるような**自由表面境界条件**が課せられる：

$$\sigma_{pz} \Big|_{z=0} = 0, \quad (p = x, y, z) \quad (8)$$

スタガード格子では、応力変数 σ_{pz} の一部が地表 ($z=0$) から $z=\pm\Delta z/2$ だけ離れた位置に定義されるため、(8)式をFDM計算に直接組み込むことはできないので、応力変数の空間補間等の操作を行なう必要がある [11]。

自由表面をFDM格子に組み込む簡便な方法として、空気中の格子に空気(真空)に相当する物性値 ($\lambda = \mu = 0$) を置く簡単な手法もある。この手法は、不規則な地表地形のモデル化に有効であるが、式(8)の**ゼロ応力境界条件**が厳密に満たされないため、格子間隔 (Δz) が大きい場合には、表面波の波形が崩れるなど影響が大きい。表面波は地表面に沿って伝わるため、自由表面境界条件の精度が特に重要になる。座標変換を用いて傾斜面での自由表面境界条件を高精度に評価する工夫 [12] も提案されている。海水などの流体は、剛性率をゼロ ($\mu = 0$) とする物性値を置くことにより、**固体／液体境界**に特別な境界条件を置かなくとも、そのままFDM計算が可能である [13]。ただし、境界条件の精度を高めるためには、剪断応力の変数 [$\sigma_{pq}, (p \neq q)$] がちょうど固体／液体境界に置かれることが望ましい [13] が、現実の複雑な海底では適用がなかなか難しい。

2-3. 無反射境界条件

FDM計算などの領域型解法では、計算領域の周囲の物理境界に**無反射境界条件**を置く必要がある。無反射境界には、1) 運動方程式を一方向伝播の波動方程式に置き換え、地震波を外部

に透過させるもの(透過境界；たとえば[14])， 2) 減衰領域を設け，ここを通過する地震波を徐々に消失させるもの(吸収境界；たとえば，[15])の 2 種類が良く用いられてきた．これらの無反射境界条件は，地震波の種類（P 波，S 波，表面波など）やモデルの不均質性の強さ，そして境界への地震波の入射角により効力が大きく異なるため，万能な手法はない．

近年，物理境界で音響インピーダンスを整合させることにより，地震波を外部に透過させる新しい無反射境界条件（PML；Perfectly Matched Layer）が提案された．PML は，電磁波シミュレーションにおいて最初に提案[16]されて以来，地震波計算への有効性が多くの研究により検証された(たとえば，[17-19])．

2-4. 非弾性減衰 (Q)

地下を伝わる地震波は，震源距離に比例した減衰（幾何減衰）に加えて，地下物性の非弾性的性質による熱エネルギーへの変換（内部減衰；intrinsic Q ）も発生する．加えて，周期1秒以下の短周期地震動では，地殻・マントルの短波長不均質構造における散乱減衰(scattering Q)の影響も大きい(たとえば，[20]を参照のこと)．このうち，散乱減衰は不均質な地下構造モデルの利用により FDM 計算で自動的に評価されるが，内部減衰の評価には特別な工夫が必要である．

このような，媒質の内部減衰 (Q) の効果を FDM 計算に組み込むために，簡便かつこれまで広く用いられてきた手順に，以下の減衰（ダンピング）係数 a ：

$$a = \exp\left(\frac{-\pi f_0 t}{Q}\right) \quad (9)$$

を，式(5)(6)の右辺第2項に乗じる手法がある．このとき，以下に示される周波数依存型の Q が FDM 計算に組み込まれる：

$$Q(f) = Q_0 \frac{f}{f_0} \quad (10)$$

ここで， f_0 と Q_0 は，評価しようとする地震波の中心周波数とその Q 値を表す．評価する地震波の周波数が f_0 から離れると Q 値が大きく変化するため，広い帯域の地震動を計算する場合には， f_0 の設定に注意が必要である．

一般の岩石では，P 波の非弾性減衰定数 (Q_p) は，S 波のもの (Q_s) よりずっと大きい ($Q_p > Q_s$) が，式(10)方法では Q_p と Q_s を区別して評価することができない．たとえば強震動計算のように大きな振幅を持つ S 波の評価を優先する場合には， $Q=Q_s$ と置いて計算を行う必要がある．

非弾性減衰の導入に関するもう一つの方法に，応力一歪みの構成関係に時間遅れ（緩和時間）を持たせる手法「メモリ変数」がある(たとえば，[21-23])．先に述べた構成方程式 [式(2)] 中に，以前のタイムステップの応力変数（メモリ変数）を組み込み，減衰の効果を FDM 計算により評価するものである．この方法では，主応力と剪断応力にそれぞれ別のメモリ変数を割り当て，緩和時間を適切に設定することにより， Q_p と Q_s を独立に評価することができる．さらに，数個のメモリ変数を用いることにより，広い帯域にわたって減衰の効果を適切に設定することも可能である．

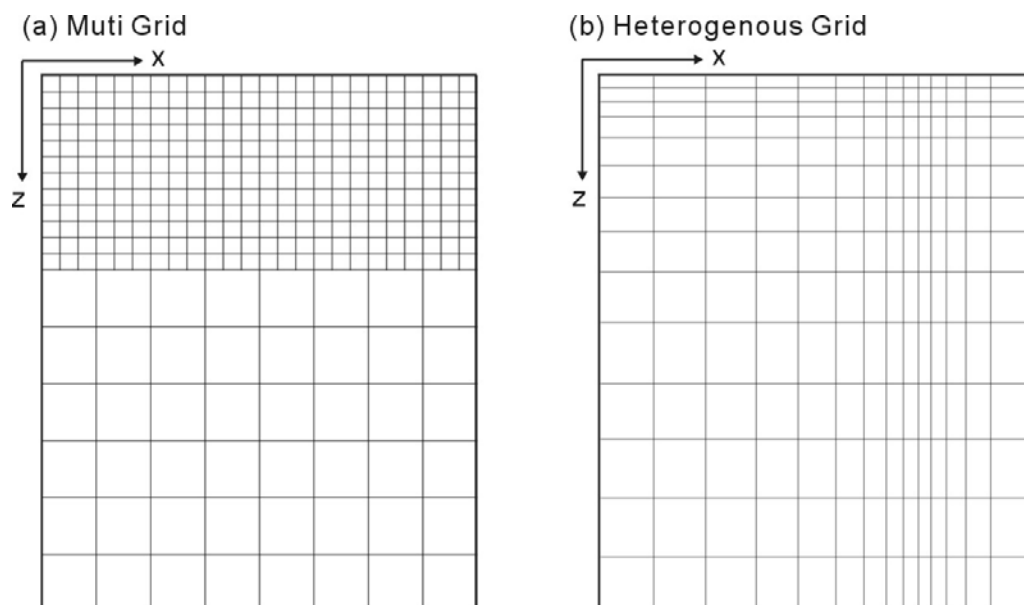


図2 地震動シミュレーションで用いられる格子. (a)不均一格子モデル, (b)不規則格子モデル.

2-5. 不規則・不均一格子の利用

地震波は、伝播速度が遅い地表付近では波長が短く、地中深部では長くなる。このため、FDM 計算において計算領域全域を一様に細かく離散化するのは不経済である。加えて、地表付近は特に構造不均質性が強く、かつ不均質分布の空間スケールが小さいため、より細かな格子間隔が求められる。

このような考えから、地表付近を細かな FDM 格子でモデル化し、地中の粗い FDM 格子モデルと結合した「不均一格子 FDM」の利用 [24-25] や、格子サイズを不均質性に対応して連続的に変化させる「不規則格子 FDM」 [26-29] が提案されている (図 2)。これらの格子モデルを用いた FDM 計算は、領域全体を一様に細かく離散化した場合に比べて、計算メモリと計算時間をそれぞれ 1/10 程度以下に抑えることができるため、3 次元計算の開始時に多く用いられた。しかし、異なる格子間における波動場の補間・間引きに伴う誤差や、格子形状の空間変化に伴う FDM 計算の誤差が最終的に地震波動場に与える評価は難しく、精度のチェックが困難である。少なくとも、格子形状を変えた計算を複数実行し、計算結果が等しいことは確認する必要がある。

3. 並列計算による大規模地震波動計算

大規模な地震波動伝播計算の実用化には、多数の計算機 (CPU) を用いた並列計算が欠かせない。並列 FDM 計算には、1) 複数の CPU と一つの共有メモリを持つ専用並列計算機において、自動並列化コンパイラや Open MP などの並列化技術を用いて FDM プログラムの「並列化」を行うもの (共有メモリモデル)、そして、2) 計算領域をプログラマが「領域分割」し、多数のコンピュータ (CPU) 間でデータ通信を行いながら、計算を並列的に実行するもの (メッセージ通信モデル)、の 2 つがある。後者の「領域分割並列 FDM 計算」は、安価な PC を高速 LAN で接続した PC クラスターの普及とともに、1990 年後半から急速に広まった。

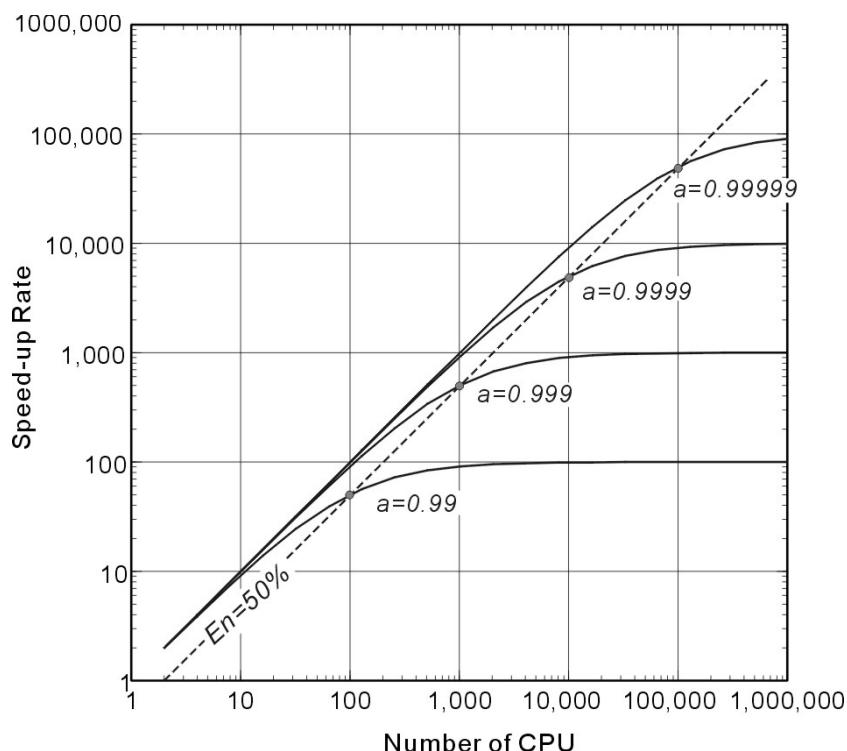


図3 並列計算による計算速度の上昇.

CPU数に対する計算速度の変化を、並列化率 ($a=0.99, 0.999, \dots, 0.99999$) 毎に示す. 点線は、並列化効率 (全CPU性能の総和に対する演算性能) が50%となるCPU数の限界.

3-1. 並列 FDM 計算の効率

領域分割並列 FDM 計算では、3次元計算領域を1方向ないし、2、3次元方向に分割し、これを複数の計算機 (CPU) に割り当てる. 各CPUでは割り当てられた小領域の FDM 計算を独立して進め、タイムステップ毎に計算領域の両側に設けられた「袖領域」の波動場を隣接CPUとのデータ通信を用いて交換し、隣接領域へ地震波伝播させる. データ通信の手順は、MPI (Message Passing Interface) プログラム等を用いて FDM 計算プログラム中に明示的に記述する. MPI の仕様は標準化されており、MPI で書かれた並列 FDM プログラムは PC クラスタから専用並列計算機まで移植性が高い.

プログラム中には、並列化により計算時間が短縮する「並列化部分」と、たとえば入出力処理やデータ通信のように、並列化による時間短縮が期待できない「非並列化部分」がいくらか含まれる. 並列 FDM 計算では、プログラム全体における並列化部分の割合は「並列化率; α 」は $\alpha=0.99\sim 0.999$ 以上の高い値を持ち、多数のCPUを用いた並列計算が比較的容易である.

3-2. 地球シミュレータによる並列 FDM 計算

先に述べたように、FDM 計算では3次元配列全体にわたって x, y , および z 方向に沿って空間微分演算が繰り返し進められる. 地球シミュレータのように「ベクトル計算機」と呼ばれるスーパーコンピュータは、CPU の処理速度と同等の高速メモリ読み書き性能を持ち、そしてベクトル演算器による一括処理により3次元 FDM 計算を非常に高速に実行することができる.

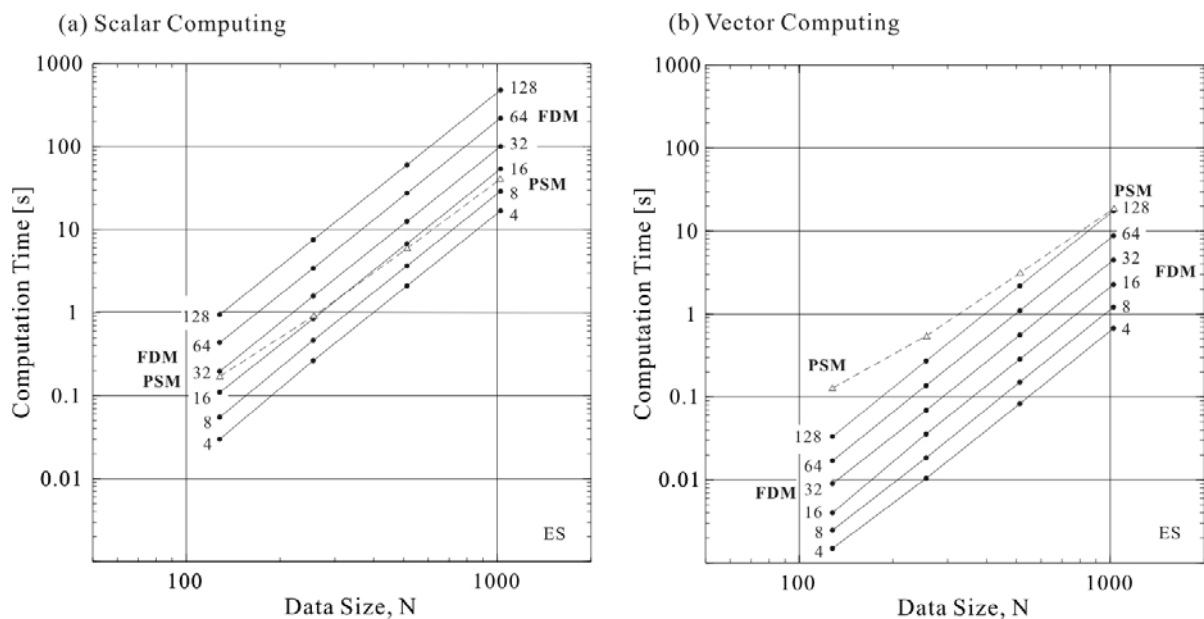


図4 3次元FDM計算のモデルサイズと計算時間。(a)スカラー計算，(b)ベクトル計算の場合。
4, 8, 16, …, 128 次のFDM計算による計算時間とPSM計算時間を，地球シミュレータのベクトル計算ユニットを使用する場合（右）と使用しない場合（左）で比較。

図4aは，地球シミュレータにおける3次元FDMの規模と計算時間を調べたものである。まず，コンパイラオプションによりベクトル演算器の利用を抑制した場合の計算時間を調べた。FDMの次数が2, 4, …, 32次精度の場合について，モデルの一片の大きさ(N)と計算時間の(1タイムステップの経過時間)を測定した。モデルサイズの増加とともに，FDM計算の時間は傾き3の直線で増加し，またFDMの次数を高めるにつれて，計算時間が2倍ずつ増えていくことがわかる。地球シミュレータでは，問題の規模によらず計算量に比例した計算時間が保たれる。FFTを用いるフーリエスペクトル法(Pseudo Spectral Method; PSM)は，低次FDMよりは低速であるが，32次～16次精度以上のFDM計算では逆転する。

次に，FDM計算におけるベクトル演算器の効果を確認する(図4b)。ベクトル演算器を使用しなかった場合(図4a)と比べて，FDM計算の速度が25～40倍以上加速することがわかる。この一方で，PSMの加速率はせいぜい1.5～2倍程度にすぎず，ベクトル計算の効果は見られない。ベクトル演算器は，FDM計算のような単純な繰り返し計算には大きな力を発揮するが，FFTのような複雑なデータ並べ替え操作(バタフライ演算)では性能が発揮できないためである。このように，ベクトル計算機ではPSMよりも高次FDMが圧倒的に有利である。

図5に地球シミュレータを用いて測定した3次元FDMの並列計算性能を示す。ここでは，3つのFDMモデル；小モデル(512x256x256)，中モデル(1024x512x512格子)，および大モデル(2048x1024x1024)について，地球シミュレータの128CPUと1024CPUを用いた2回の並列計算の時間測定から並列化率を求め，1万CPUまでの演算速度(FLOPS値)を推定した。得られた並列化率は，大，中，小モデルについてそれぞれ， $\alpha=0.99716, 0.999527, 0.999887$ であり，モデルサイズが大きいほど並列計算の効率は高くなる。また，並列化効率>50%(CPU数の50%の加速率)を満足する最大CPU数は，3つの計算モデルにおいて353, 2,116，および8,820CPUと見積もられる。

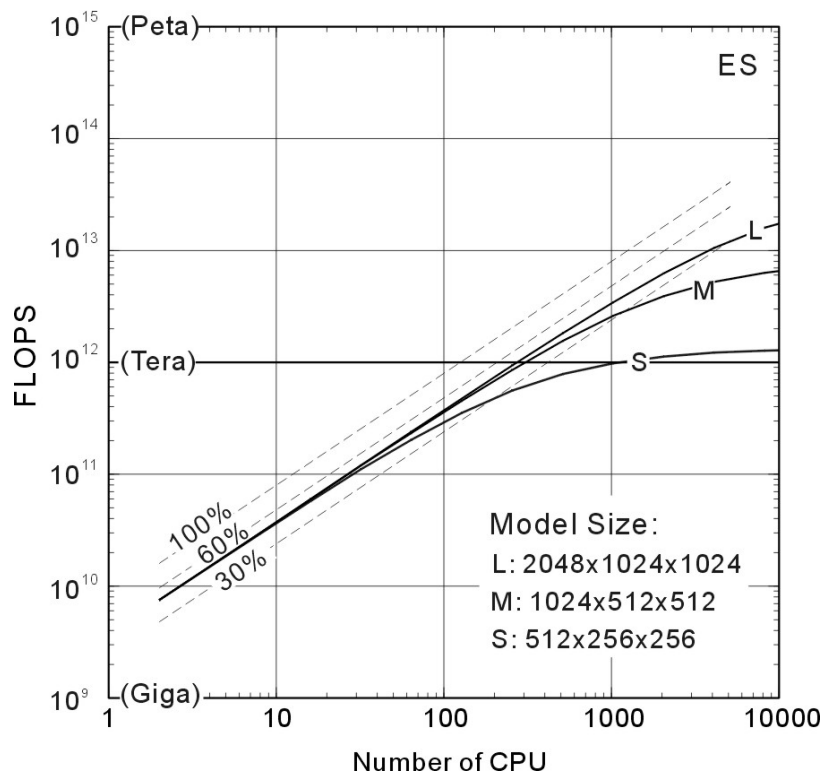


図5 地球シミュレータにおける，並列 FDM 計算の演算性能.

3つのモデル(S, M, L)サイズでの並列 FDM 計算の速度向上率を地球シミュレータで測定. 点線は地球シミュレータの理論性能に対する実効性能比 (%) を表す.

なお，本テストで得られた並列 FDM 計算の実効性能 (FLOPS 値) は，計算に使った CPU 総数 (8GFLOPS/CPU) から期待される地球シミュレータの理論性能の 40~50%になる. このように，FDM 計算はベクトル計算機に良く適合し，演算器の高い実効性能を引き出すことができる.

3-3. T2K オープンスパコン (東大) による並列 FDM 計算

一般的な PC や近年の超並列計算機で一般的に用いられる，**スカラー型 CPU** は数 GHz という高い周波数で内部演算が行われるのに対して，CPU とメモリのデータ読み書き速度 (**メモリバンド**) はその数十分の 1 と遅ことが FDM 計算の高速化に大きな障害となる. スカラー計算機における FDM の計算性能は，CPU のクロック数ではなく，むしろメモリバンド性能に依存する.

そこで，スカラー型 CPU では高速の**キャッシュメモリ**を CPU とメモリの間に置き，メモリから読み取った一塊のデータを一時的に保存することにより，連続するデータ読み出しの高速化の工夫がはかられている. しかし，キャッシュメモリの容量はせいぜい数百 KB~数 MB にすぎず，メモリ全体 (数百 MB~数 GB) のわずか 1/1000 以下にすぎない. このため，メモリ全体にランダムにアクセスする大規模並列 FDM 計算では，キャッシュメモリによる高速化の効果はほとんど期待できない. 3次元 FDM 計算では，たとえばデータ配列の x 方向の並びがメモリ上に連続して配置されたとしても，y 方向や z 方向のデータ並びはメモリ上に飛び飛びに配置されるため，キャッシュメモリを用いた連続読み込みの効果が出ないためである.

このようなスカラー型計算機の特性を考えると，FDM 計算の高速化のためには，メモリの間接参照などメモリアクセス回数を減らす工夫が必要である. たとえば，先に述べた不均一格子 FDM では，場所毎に異なる FDM 演算子が用いられるため，演算子の係数を納めたテーブルを用

意し、各格子点において適切な係数を用いて FDM 計算を行うのが一般である。このようなテーブルの間接参照では、メモリアクセスが 2 回必要になるため、演算効率が大きく低下する。また、先に述べたベクトル計算機では、演算ループ (D0 ループ) 内になるべく複数の処理を並置し、計算機が持つ全てのベクトル演算器を効率よく稼働させるようなコードチューニングが定石であった。しかし、スカラー計算機では、D0 ループ内をむしろ単純化したほうが、キャッシュメモリ内に連続する変数データを多量に取り入れることができる得策である。このように、ベクトル計算機とスカラー計算機では高速化のためのコードチューニングがいくらか異なることにも注意が必要である。

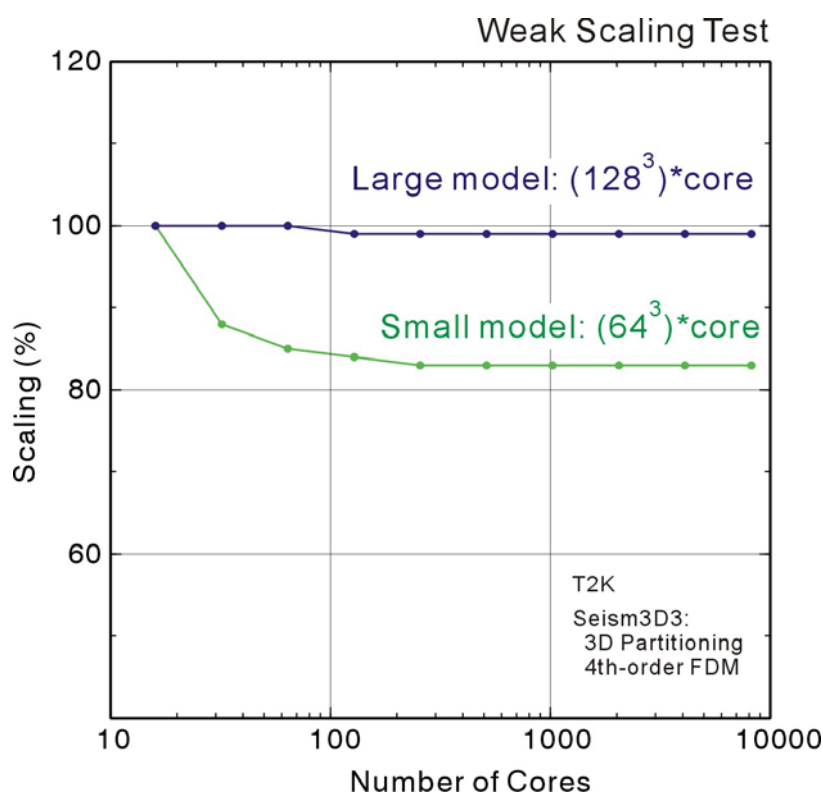


図6 T2K オープンスパコン(東大)における Weak Scaling テスト(2つのモデル;Small/Large).
1 コアあたりの問題規模を固定し、コア数の増加に伴う計算時間の変化を測定(16 コアを基準とする)。

T2K オープンスパコンに代表される、近年の高性能スーパーコンピュータは、スカラー型の CPU を多数用いてベクトル計算機と同等以上の高い演算性能を達成しようという設計思想に基づいている。これには、プログラムの並列化率を高めることが求められる。

図6は、T2K オープンスパコン(東大)での並列 FDM 計算の効率を確認するために、1 CPU (コア) あたりの問題サイズを固定して、CPU 数の増加による計算時間の変化を調べたもの (Weak Scaling Test) である。コア数を 16~8192 まで変化させ、3 次元 FDM 計算の時間を測定した。なお、本テストでは、コア数によらず一定の計算時間となることが期待されるが、モデルが十分に大きな場合 (Large Model;128³ 自由度/コア) には、8192 コアを用いた並列計算まで演算性能が低下しないことが確認できる。問題サイズが小さい (Small Model;64³ 自由度/コア) 場合には、CPU 数とともに計算速度は十数%低下する。これは、データ通信の競合による通信速度の若干の低下により、FDM 計算自体に対する通信時間が相対的に大きくなり、並列化率が悪化するためである。

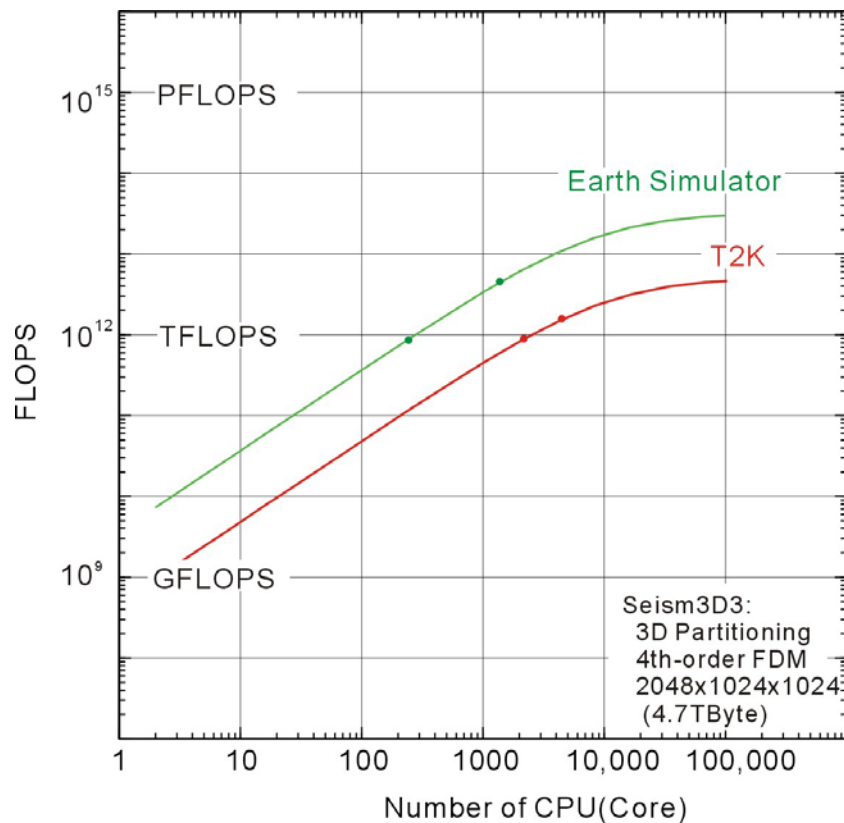


図7 地球シミュレータと T2K オープンスパコン（東大）の並列計算の効率.

並列 FDM 計算の演算速度を，地球シミュレータの 128CPU と 1024CPU の計算時間（T2K では，2048 コアと 4096 コア計算の時間）から推定.

図7は，モデルサイズを固定（2048*1024*1024 自由度モデル）し，CPU（コア）数を変化させた場合の並列 FDM 計算の演算速度の上昇率を調べたものである。ここでは，2048 コアと 4096 コアを用いた並列計算の時間から演算速度（FLOPS）と並列化率を計算し，最大 10,000 コアを用いた並列計算の時間を予測した。同様のテストを地球シミュレータの 128CPU と 1024CPU を用いた並列計算に対しても実施し，2つの演算速度を比較した。3次元並列 FDM 計算では，オープンスパコン（東大）の1コアあたりの実効速度は，スカラーCPUの理論性能（9.2GFLOPS/core）の約8%（0.74GFLOPS）であった。さきに述べたように，地球シミュレータでは，CPUの理論性能（8GFLOPS）の60%（4.8GFLOPS）の実効性能が得られたことから，同等の計算速度を得るためには，地球シミュレータよりもコア（CPU）数を6倍程度増やす必要があることがわかる。今後，スカラー型CPU向けに計算コードの最適化を進め，そしてより多数のCPUを用いた並列計算の実現のために並列化コードの一層のチューニングを進める必要がある。

4. 地震波動伝播の大規模計算と揺れの可視化・検証

近年の大地震による地震波伝播シミュレーションと，高密度地震観測との直接比較による支配方程式の検証の例として，近年日本で発生した3つの大地震の大規模並列シミュレーションを紹介する。

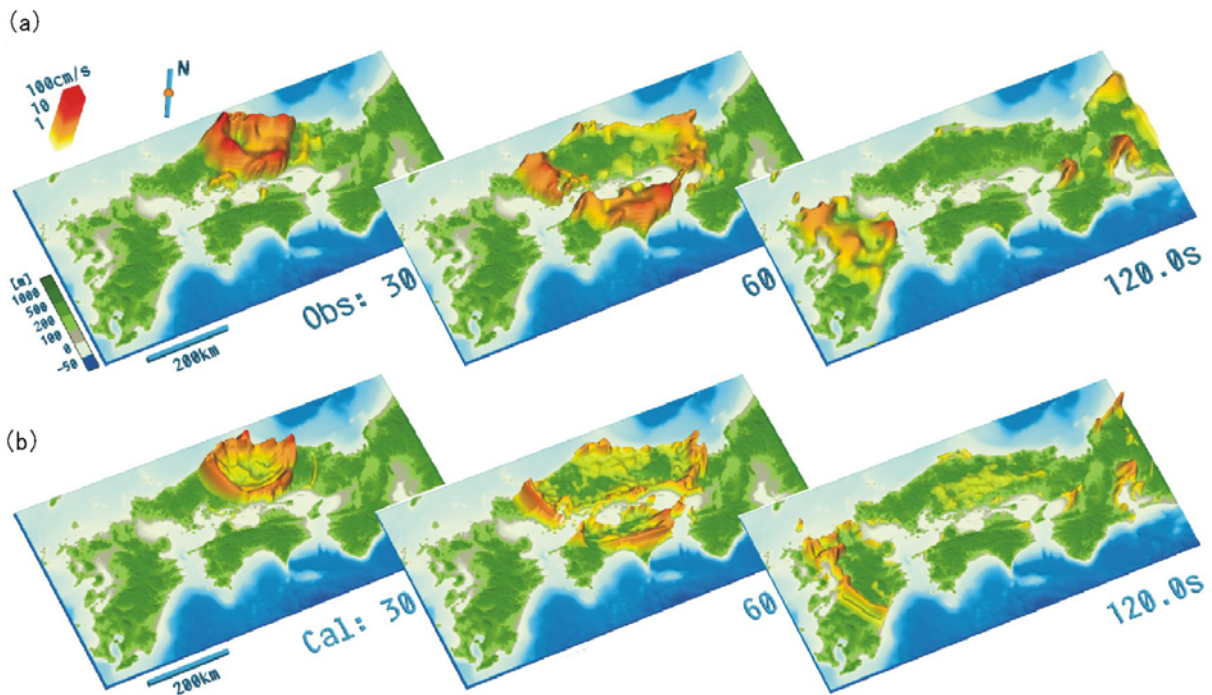


図8 鳥取県西部地震の揺れの伝わり方(地震後30, 60, 120秒後)。(a) 高密度強震観測と, (b) コンピュータシミュレーション結果を比較。

4-1. 2000年鳥取県西部地震のシミュレーションと高密度観測による検証

1995年兵庫県南部地震(阪神淡路大震災)を契機として, 防災科学技術研究所により, 全国1800カ所に20~25kmの高い密度で強震観測網(K-NET, KiK-net)が整備された。2000年10月6日に発生した鳥取県西部地震(M7.3)は, 観測網が設置されて初めての内陸活断層の大地震であり, その強い揺れは西南日本全域にわたって521観測点で記録された。

4-1-1. 高密度地震観測網で記録された地震の揺れ

図8は, 観測点で記録された地震動を時間-空間的に補間して, 震源から揺れが日本列島に広がる様子を表したものである[30]。地震発生から30秒, 60秒, 120秒後の地面の揺れの強さ(横揺れの速度)を表示している。色が濃いほど, また波面の高さが高いほど揺れが強いことを表している。

この地震は, 浅い($h=11\text{km}$)地殻内(ユーラシアプレート)で起きた内陸活断層の地震であり, 横ずれ型断層からはSH波(水平方向に振動するS波)が強く放射された。波動伝播のスナップショットを見ると, SH波から変換して生まれた周期5~8秒の長周期の表面波(Love波)が, 秒速2.8km程度の速度で西南日本全域に広がる様子がわかる。長周期の表面波は, 地震発生から1分後には大阪を, そして2分後には名古屋を通過している。これらの堆積平野に入射した地震動は, 平野の柔らかい堆積層で揺れが数倍以上に増幅され, 強い揺れが2~3分以上にわたって平野を長く揺すったことわかる。

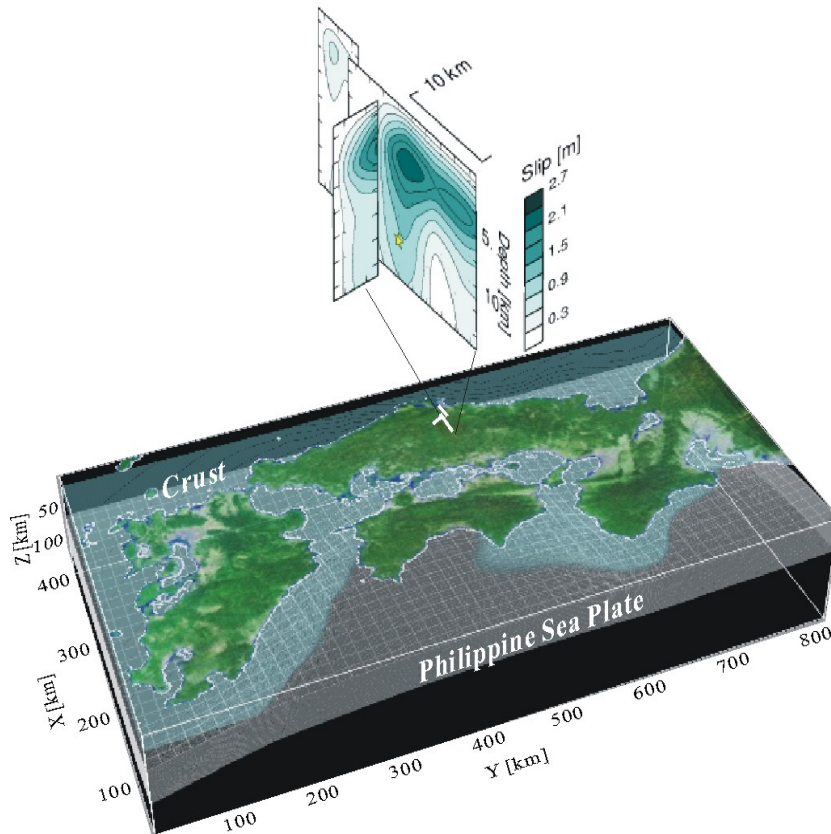


図9 西南日本の地下構造モデル（下）と、2000年鳥取県西部地震の震源断層モデル（上）。

4-1-2. 鳥取県西部地震の揺れのシミュレーション

高密度地震観測で捕らえられた鳥取県西部地震の揺れと表面波の伝播を、地球シミュレータにより再現した[30]。西南日本の820km*410km*128kmの領域を0.4kmの分解能で離散化し、地殻・上部マントル構造、そして低角（約10度）で沈み込むフィリピン海プレートモデルを組み込んだ。最表層のP波およびS波速度は、それぞれ $V_p=3.2\text{km/s}$ と $V_s=1.7\text{km/s}$ である。地殻の厚さは平均33kmであり、 $V_p=5.5\text{km/s}$ 、 $V_s=3.2\text{km/s}$ 、そして上部マントルでは $V_p=7.8\text{km/s}$ 、 $V_s=4.5\text{km/s}$ の地震波速度に対応する物性値を各格子に与えた。沈む込むプレートは、地殻の厚さを5km、そしてマントルの厚さを25kmとして、陸上のプレートよりも地震波速度を5%高速に設定した。また、大阪や名古屋などの平野には、地震波速度の遅い（ $V_s=1.7\text{km/s}$ 、 $V_p=3.4\text{km/s}$ ）堆積層を最大4kmの厚さで与えた。震源モデルには、震源近傍の地震波形記録の逆解析（インバージョン）から求められた震源断層モデル（図9）を用いて断層面から最大周波数1Hzまでの地震波を放射させた。

本計算では、鉛直（z）方向の格子数が少ないため、並列計算の領域分割はz方向に対してのみ行った。並列FDM計算は水平方向には16次精度の高次差分スキームを、また領域分割を行った鉛直方向には4次精度のスキームを用いた。鉛直方向に低次のFDMスキームを用いたのは、隣接領域とのデータ通信量を抑えるためである。FDM計算の精度を水平／鉛直方向で揃えるため、鉛直方向の格子間隔（ D_z ）は、 D_x の半分の大きさのもの（ $D_z=0.1/0.2\text{km}$ ）を採用した。地震発生から3分間の波動伝播の計算は、地球シミュレータの32ノード（256CPU）を用いた計算において40分を要した。

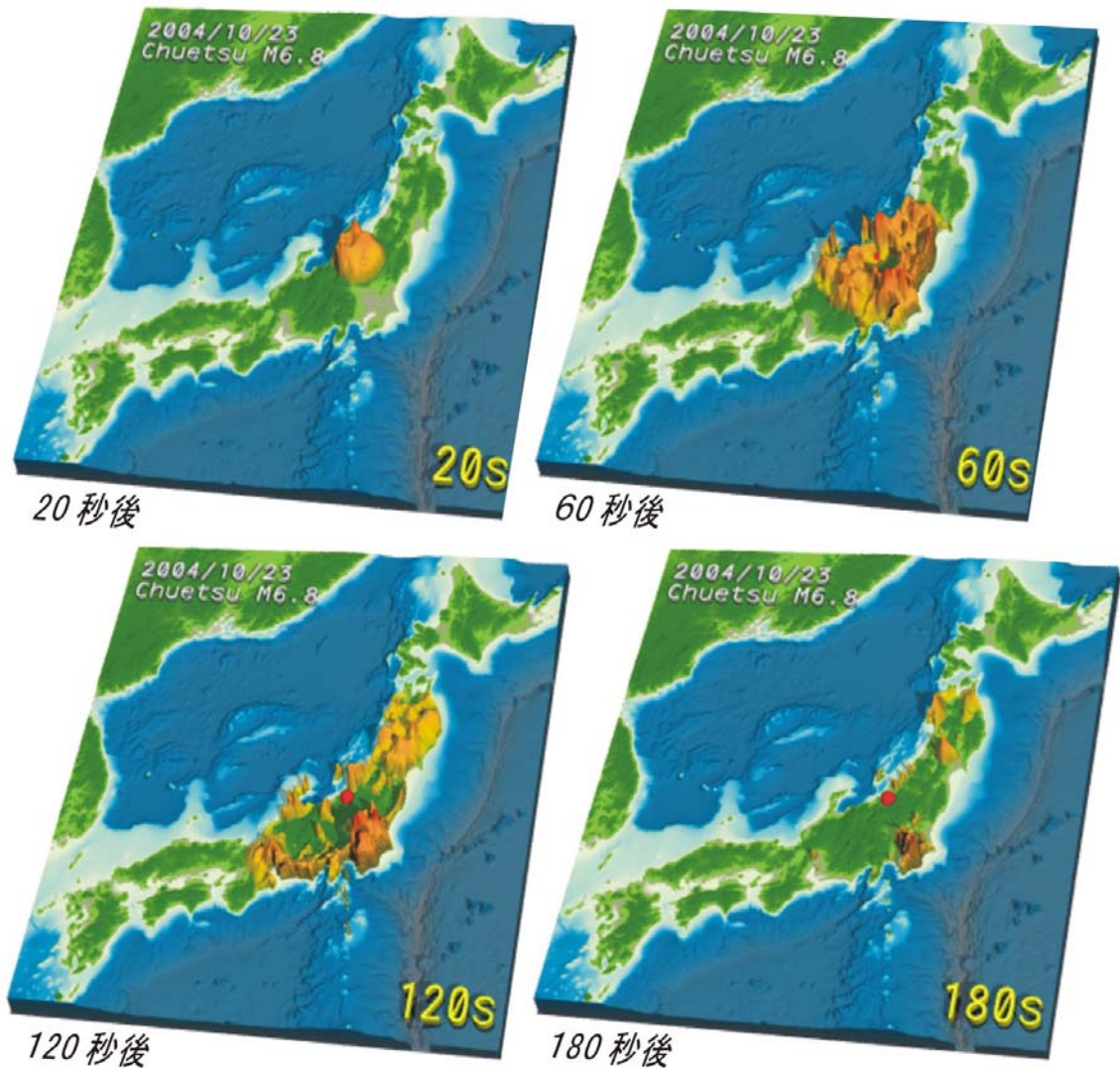


図 10 2004 年新潟県中越地震の揺れの広がり方（地震後，20，60，120，180 秒後）。

全国に高密度に展開された強震観測網の地震計記録を用いて，揺れの広がる様子を可視化．色が濃く表示が高いほど，揺れが強いことを表す．

計算より再現された鳥取県西部地震の揺れの様子を図 8 に示す．波動伝播のスナップショットを見ると，横ずれ断層の断層面の走行とこれに直交する方向で SH 波の放射が強くなる 4 象限型の放射特性が確認できる．大振幅の表面波（Love 波）は，九州，山陰，そして四国地方に向けておよそ 2.8km/s のゆっくりとした速度で伝播していく様子が再現されている．大阪，名古屋，福岡，大分などの堆積平野に入射すると，地震波速度の遅い（剛性率の小さい）表層で揺れが強く増幅され，そして水平不均質構造による分散や散乱現象により波群が長く伸びる様子もわかる．こうして高密度地震観測で見ることができたように，堆積平野で大きな揺れが 2 分以上にわたって長く続く様子が，シミュレーションにより良く再現できた．

4-2. 新潟県中越地震の長周期地震動シミュレーション

2004 年 10 月 23 日に発生した新潟県中越地震（M6.8）では，震源に近い小千谷市や川口町で震度 7 の強い揺れに見舞われたほか，中部・東北日本の広い範囲が有感となった．図 10 は高密度強震計記録を用いて作成した，中越地震の揺れの広がる様子である[31]．2000 年鳥取県西部

地震（図8）で示されたように、人口の集まる堆積平野には地震波の強い増幅が起き、大きな揺れが何分間にもわたって長く続いたことがわかる。

4-2-1. 関東平野で発生した長周期地震動

関東平野は、震源から200km以上離れており震度は3～4程度であったが、都心では周期7秒の長周期地震動が強く発生し、地面は最大5cmの大きさにわたって長く揺れ続けた。関東平野の地下を覆う堆積層の厚さは3000～5000m以上にもなるため（図11）、堅い基盤岩と地表面に挟まれた堆積層で、6～12秒の長い周期を持つ**長周期地震動**が共鳴を起こして強く生成したためである。

周期7秒の長周期地震動は、およそ70階建て以上の超高層ビルや、数十万キロリットルの容量を持つ石油備蓄タンクの**固有周期**に一致する。これら長周期地震動に**共振**した構造物の揺れは、地面の揺れよりも大きく、かつ長時間続くことにより、被害が起きることに注意が必要である。鉄骨造の超高層ビルや石油タンクの浮屋根は減衰が小さいため、いったん揺れ始めるとなかなか止まらない。新潟県中越地震では、新宿の70階建ての超高層ビルが共振して大きく揺れ、エレベーターケーブルの一部が切断する事故が起きた。2003年十勝沖地震（M8.0）において、震源から200km離れた苫小牧の石油タンクが**スロッシング振動**を起こして破損・出火したのも、勇払平野で強く発生した周期7秒の長周期地震動によるものである。

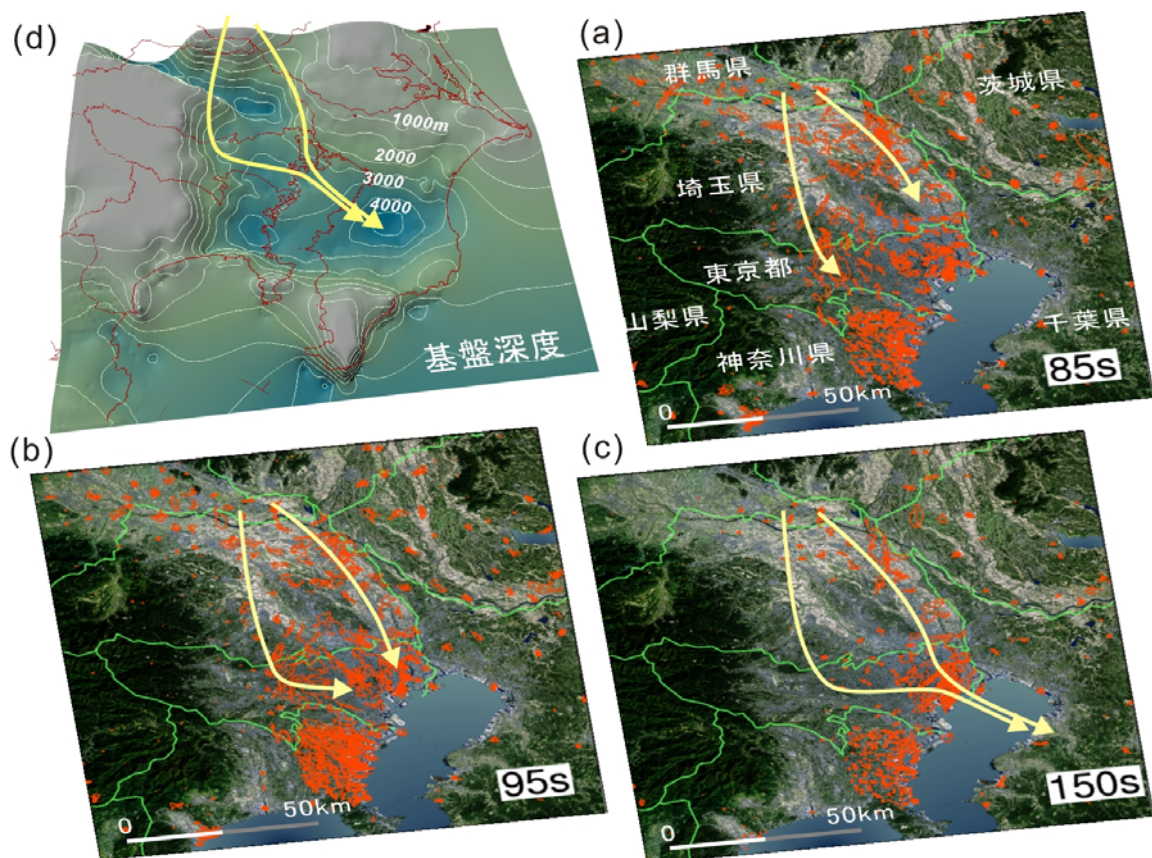


図11 自治体震度計で記録した、新潟県中越地震の揺れの様子（a-c）と関東の基盤構造（d）。地震後85, 95, 150秒後の地震計の動き（振動軌跡）をオレンジ色で表示。黄色の矢印は、観測から推定される地震波の伝播の方向。

4-2-2. 超高密度強震観測で見た都心の揺れ

都心の長い揺れの原因はそれだけではなかった。中越地震の揺れを、関東一甲信越に設置された自治体の震度計データ（SK-net；東大地震研究所）を使って詳しく調べたところ、関東西縁山地に沿って神奈川方面に向かっていた長周期地震動が突然方向転換し、都心に集まってくる様子が見えてきた（図 11）。基盤深度が浅く、表面波の伝播速度が速い周囲の山地から、堆積層が厚く、表面波の伝播速度が遅い都心部へと向けて地震波が一種の**屈折現象**が起こしたらしい。こうして、都心に向かっていろいろな経路から長周期地震動が集まったことにより、何分も続く長い揺れが生まれたことが考えられる。

4-2-3. 地球シミュレータで再現した長周期地震動

このような、関東平野の特異な堆積層構造が作り出す、長周期地震動を関東平野の地下構造モデルを用いた地震波伝播の数値シミュレーションにより確認した[31]。計算対象領域を、甲信越から関東にかけての 440km*240km*160km の範囲に選び、200メートルの規則格子間隔で細かく離散化した。震源モデルは、近地および遠地で記録された地震波形から逆解析により求められた解を用いて、震源断層の格子点から最大周波数 1 Hz の地震波を放射させた。計算は、地球シミュレータの 160 ノード（1280CPU）を用いた並列計算により行った。地震発生から 2 分間の揺れの計算は 30 分を要した。

図 12 は、地球シミュレータで再現された、新潟県中越地震の揺れの伝わる様子である。深海や山地など人家の無い場所には強震計や震度計が設置されていないため、地震観測から確認できる地震波伝播の挙動には限りがある。数値シミュレーションでは、このような地震観測のような制約がなく、任意の場所に「仮想地震計」を置いて、地表や地中の揺れを詳しく見ることができる。

シミュレーション結果を見ると、地震発生から約 20 秒後に、関東平野の北端で長周期地震動が強く増幅・発生し、その後都心に向かって真っ直ぐ伝わる地震動のほかに、平野の西端部を伝わる別の地震動が、国分寺付近で東転して都心へと向かってくる様子が確認できた。その後、都心下の厚い堆積層に集合した長周期地震動は、東京湾を通過して千葉県北西部へと渡り、同じ経路を伝って再び都心に戻ってきたこともわかった。地震観測からは確認できていないが、シミュレーション結果によると東京湾の海底も大きく長く揺れたことが推察される。こうして、関東平野の基盤の最深点を通って都心と千葉県北西部を長周期地震動が行き来したことにより、大きな揺れが長く停滞したと考えられる。

シミュレーション結果の妥当性は、計算から求められた長周期地震動の波形が、観測記録と良く一致することから確認できる（図 12）。特に、S波の後から到着する大震幅かつ継続時間の長い長周期地震動の波形の一致は驚異的である。ただし、長周期地震動が到着する前の、P波とS波の短周期地震動（周期 1～2 秒以下）の再現性は十分ではない。今後、短周期地震動の放射に関わる震源断層運動の不均質性と、震源から関東平野にかけての地下構造の数百メートル以下の微細構造のさらなる知見が必要である。

もちろん、地下構造探査の高精度・高密度化には技術的、そして経済的にも限度がある。シミュレーションモデルの高度化には、既存の地下構造モデルを用いて多数の地震のシミュレーションを実施し、計算結果と観測データ（地震計記録）の一致が良くなるように、境界面の形状や物性値などモデルの修正を繰り返し行うことが求められる。

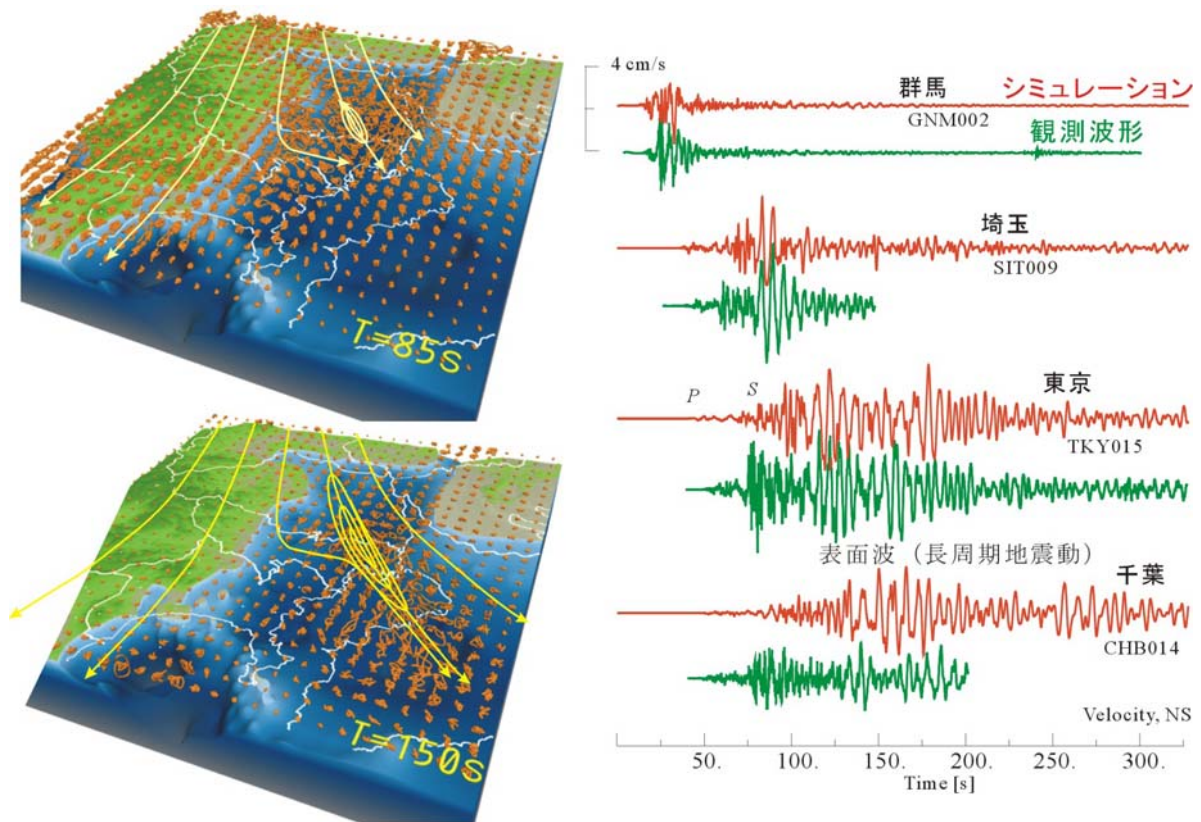


図 12 コンピュータシミュレーションにより再現した，新潟県中越地震の関東平野の揺れ。地震後 85, 150 秒後の地動の振動軌跡と読み取った地震波の伝播方向。右図は群馬，埼玉，東京，千葉地点でのシミュレーション波形（オレンジ）と観測波形（緑）の比較。

4-3. 南海トラフ巨大地震による長周期地震動のシミュレーション

このような長周期地震動の問題は，地震の規模（ M ）の大きい海溝型巨大地震ではより深刻である。御前崎から紀伊半島を経て，四国沖に至る南海トラフでは，フィリピン海プレートの沈み込みに伴い，約 100 年の周期で $M8$ 級の巨大地震が繰り返し発生している。国の中央防災会議によると，今後 30 年以内に東海地震が発生する確率は 87%，東南海地震と南海地震の発生確率は 60～70%にもなるという。

4-3-1. 1944 年東南海地震による長周期地震動

前回の 1944 年東南海地震（ $M8.0$ ）や 1946 年南海地震（ $M8.0$ ）は，第二次世界大戦の戦中・戦後の情勢で被害記録がほとんど残っていない。このため，地震による被害の全貌は良くわかっていない。当時，中央气象台（現気象庁）や東京帝大により横浜，東京大手町，千葉県東金市で強震観測が行われており，その地震計記録紙から地震による地動を求めた[32]。その結果，都心で周期 8 秒前後の長周期地震動が，最大 4cm 以上の大きさで発生し，その後 2cm を越える強い揺れが 15 分間以上にわたって長く続くなど，長周期地震動が強く発生したことがわかった（図 13）。揺れの強さは，2004 年中越地震の揺れとほぼ同等ではあるが，数十分以上にわたって長く揺れ続けたことが特筆される。

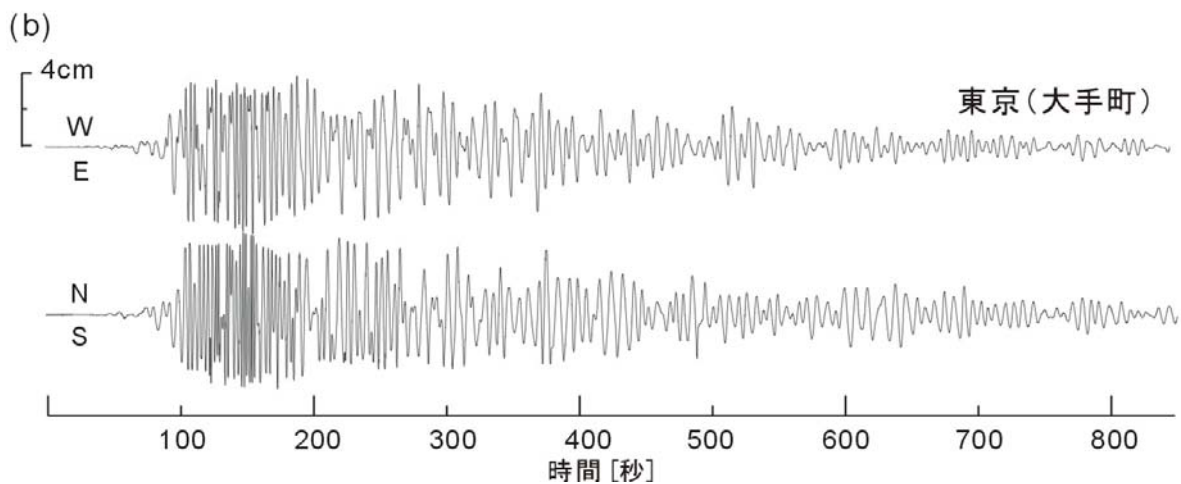
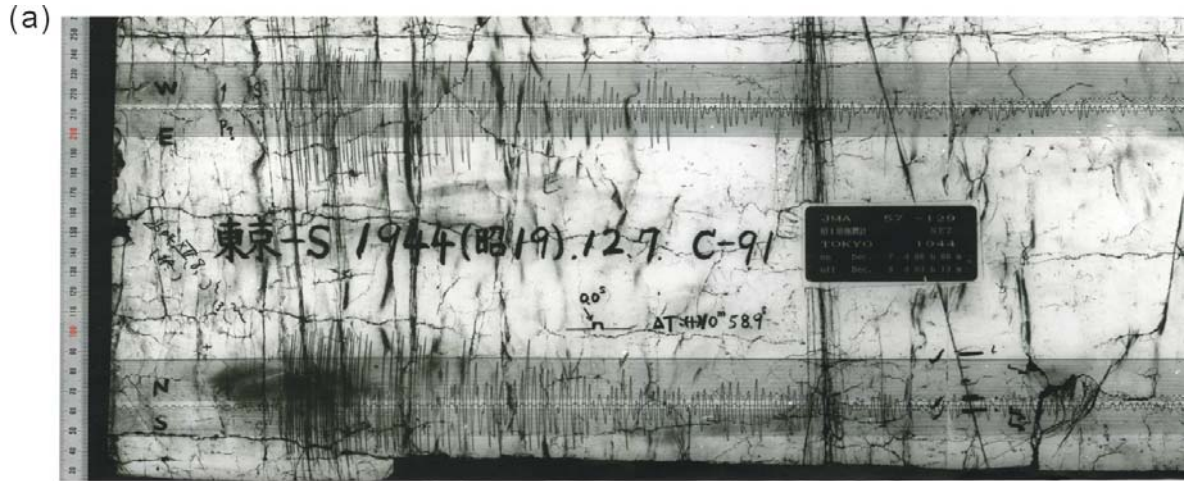


図 13 1944 年東南海地震における東京（大手町）での地震計記録紙と、これから読み取った地動の東西動成分と南北動成分.

南海トラフの地震において、関東に長周期地震動が強く生成する原因として、先に述べた関東平野の堆積層構造に加えて、フィリピン海プレートの上に厚く堆積した、海洋性堆積物（付加体）が重要な役割を担っていることが多くの研究者らにより指摘されている[33]. 低速度の付加体では、平野の堆積層と同様に長周期地震動が共振を起こし強く増幅されるとともに、内部に閉じ込められた長周期地震動が、付加帯の広がりに沿って濃尾平野や関東平野に誘導されるためである. このことは、南海トラフの地震による長周期地震動の予測には、平野の地下構造とあわせて、海域の詳しい地下構造調査が不可欠であり、震源から平野までの数百 km の広い領域を高い分解能でモデル化した大規模 FDM シミュレーションが必要である.

西南日本の地下構造をモデル化したコンピュータシミュレーションから東南海地震の周期 2 秒以上の長周期地震動を再現し、上記観測波形と良い一致を確認することができた[34] (図 14). こうして、観測データとの検証により、将来の南海トラフの巨大地震の長周期地震動の予測シミュレーションへの道が開けた.

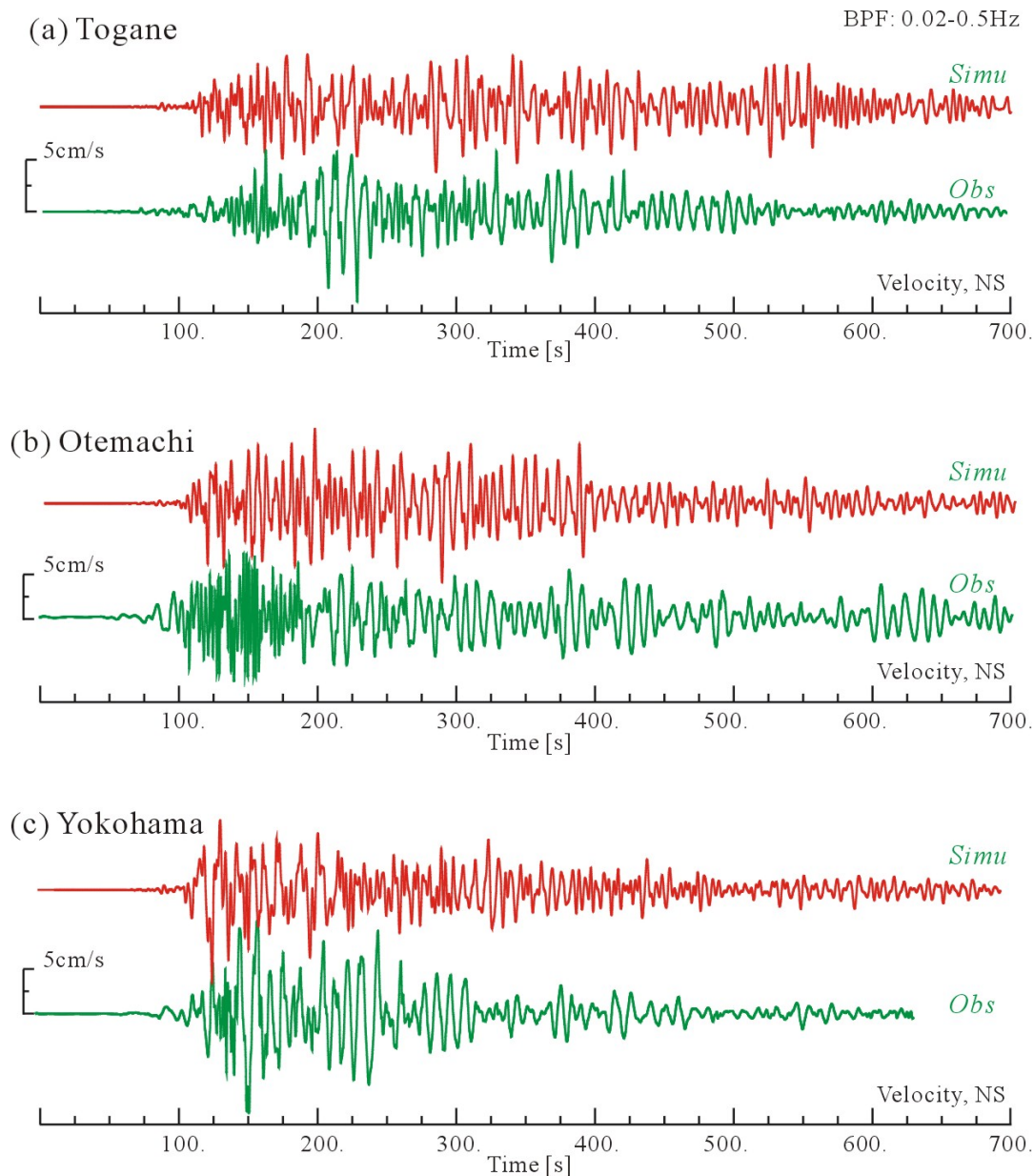


図 14 東南海地震のコンピュータシミュレーションと観測の比較。

(a)千葉県東金市, (b)東京大手町, (c)横浜地点で記録された地震計記録 (緑) とシミュレーション波形 (赤) の比較。

4-3-2. 東南海・東海地震連動発生とその影響

前回の南海トラフ地震 (1944 年東南海, 1946 年南海地震) の規模は, 以前の地震と比べて規模は比較的小さかったほか, 2 つの地震は 2 年間の時間差で単独発生している. また, 東南海地震 (M8.0) は東海地域の手前で停止してしまったため, 残りの東海地震 (M8.0) はいつ起きてもおかしくない状況のまま 60 年余を経過した. 1854 年安政東海地震 (M8.4) では, 東海地震と東南海地震が連続して発生しており, その翌日には南海地震 (M8.4) が起きている. また, 1707 年宝永地震 (M8.4) では, 3 つの地震が同時に発生したと考えられている. このように, 南海トラフ地震の発生様式が毎回大きく異なっている.

今後 30 年以内には東海・東南海, 東海地震の次の地震発生サイクルが到来する. このとき, 3 つの地震が単独で発生するか, あるいは 3 つの地震がどのような時間差で連動して発生する

かは地震防災上重要な課題である。過去の南海トラフ地震の連動発生パターンと、将来の地震の発生パターンの可能性を考え、それぞれの地震波伝播シミュレーションから各地の揺れの違いを明らかにすることが必要である。

たとえば、東南海地震と東海地震が連動発生した場合の揺れのコンピュータシミュレーション結果（図 15）によると、名古屋や東京の長周期地震動は、前回（1944 年）の東南海地震の 2～2.5 倍以上になるおそれがある。このとき、都心の 70 階建ての超高層ビルの揺れは、片震幅が最大 1.5m になり、そして 50cm を越える大きな揺れが 10～30 分以上にわたって長く続く恐れがあるため、耐震性の確認や家具の固定など揺れへの対策が必要である。

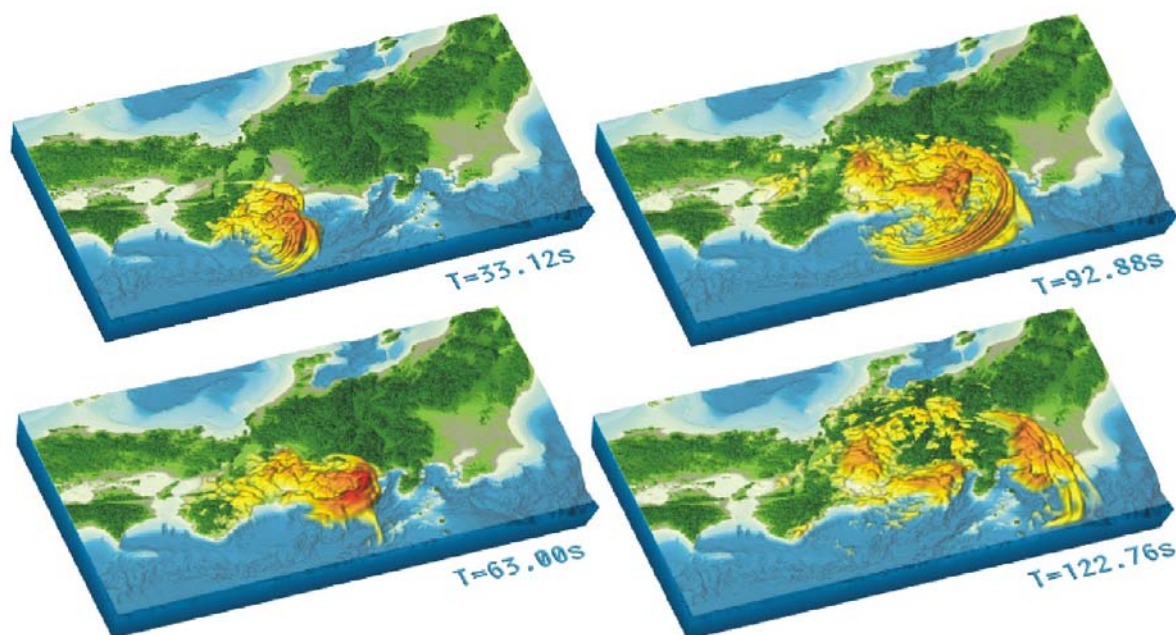


図 15 東海・東南海地震同時発生を想定したシミュレーションの波動伝播の様子（地震後 33, 63, 92, 122 秒後）。

4-4. 深発地震と異常震域の短周期地震動シミュレーション

これまで、陸のプレート内で起きる「内陸活断層地震」（たとえば、2004 年中越地震）や、陸のプレート/海洋プレート境界で起きる「プレート境界地震」（たとえば、南海トラフ地震）とそのシミュレーション例を紹介した。日本では、これらの地震のほかにも日本列島下に沈み込むプレート（スラブ）内で発生する「スラブ内地震」も発生し、被害が起きる。深いスラブ内地震では浅い地震に比べて、震源から短周期地震動が強く放射され、そして短周期地震動が堅いプレート内を伝わって遠地まで大きな揺れ伝える特徴がある。

4-4-1. 深発地震に見られる異常震域

スラブ内地震による特異な震度分布として、京都府沖の深さ 347km で発生した M6.7 の深発地震と、三重県の深さ 395km で発生した M6.5 の深発地震による震度分布を図 16 に例示する。浅い、内陸活断層の地震やプレート境界の地震では、深度分布は震央を中心とする同心円分布を示すことが多いが、これらの地震の震度分布は大きくゆがんでおり、震央より数百 km 以上離れた場所に大きな震度が現れていることがわかる。たとえば、京都府沖の地震（図 16a）では、

震源に近い京都や福井の震度は1にすぎないが、1000km以上離れた帯広が最大震度4になっている。三重県沖の地震（図16b）でも、500km以上離れた福島と栃木が最大震度4であった。こうして関東～東北～北海道の太平洋岸の広い範囲が有感となった。

このように、太平洋プレートの深発地震において太平洋岸の広い範囲の震度が大きくなる現象は「異常震域」として古くから知られている。日本列島下に沈み込む、冷えた硬いプレートは短周期地震動の減衰が小さく、プレートに沿って日本海側を遠くまで揺れが伝わるのに対して、プレートから抜け出して日本海側に向かった地震波は、温度が高いマントルを通過するときに急激に減衰するためである。こうして、日本列島の日本海側と太平洋側の地下の減衰構造の大きな違いが、特異な震度分布を作り出すのである。

4-4-2. 散乱による短周期地震波の導波効果のシミュレーション

硬く、冷えたプレートは地震波の減衰は周囲のマントルよりも10倍以上小さいが、地震波速度がマントルよりも10%程度速いため、このままでは地震波がマントルに向けて屈折を起こして抜け出してしまおうという矛盾が生じる。

異常震域が発生した太平洋側で記録された地震波形を詳しく調べたところ、この地震動には周期0.5秒以下の短周期の揺れが強く含まれ、0.5秒以上の長周期地震動は日本海側と同等か、逆に弱まっていることがわかった。また、短周期地震動揺れが長時間にわたって長く続いている特徴も見えてきた。このことは、プレートは短周期の地震波のみを良く伝え、長周期地震動は外に逃がすこと、短周期地震動の長い波群を作り出す原因がプレート内部に存在することを意味している。

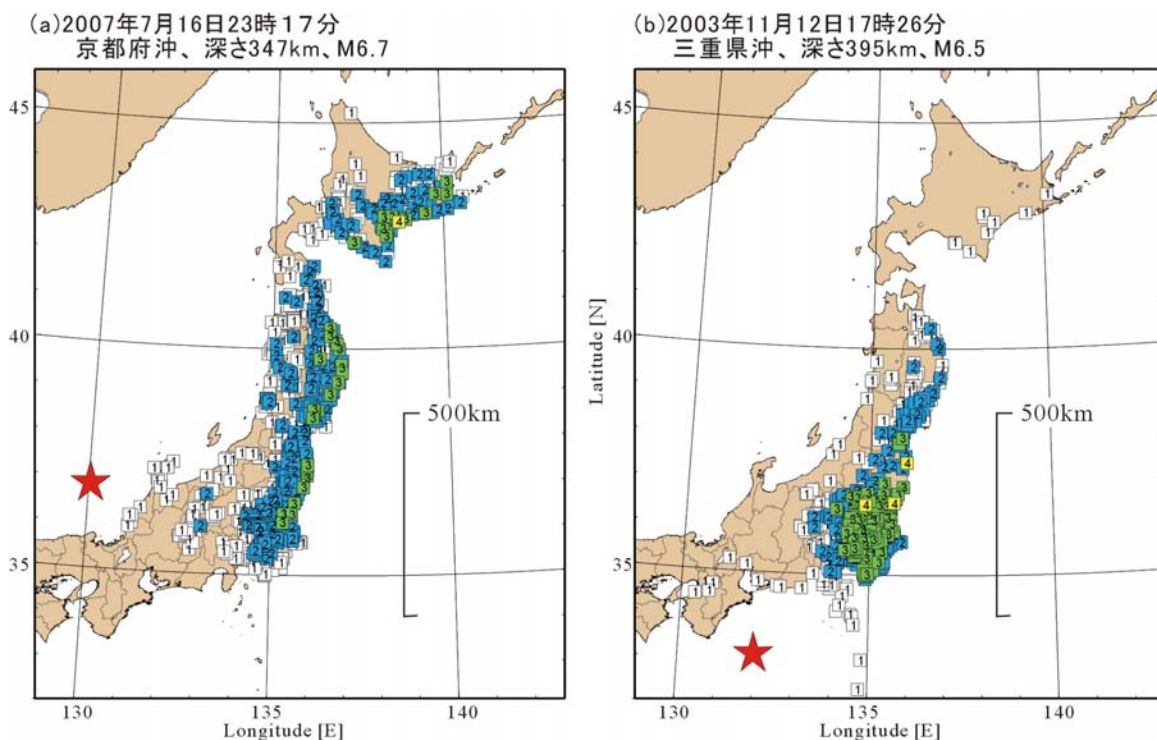


図16 深発地震で見られる震度分布の異常（異常震域）。(a)京都府沖の地震、(b)三重県沖の地震。沈み込むプレートの深部で起きる深発地震の震度分布は同心円状とはならず、地震波が太平洋側にプレートを伝って良く伝わるために、歪な震度分布となる。

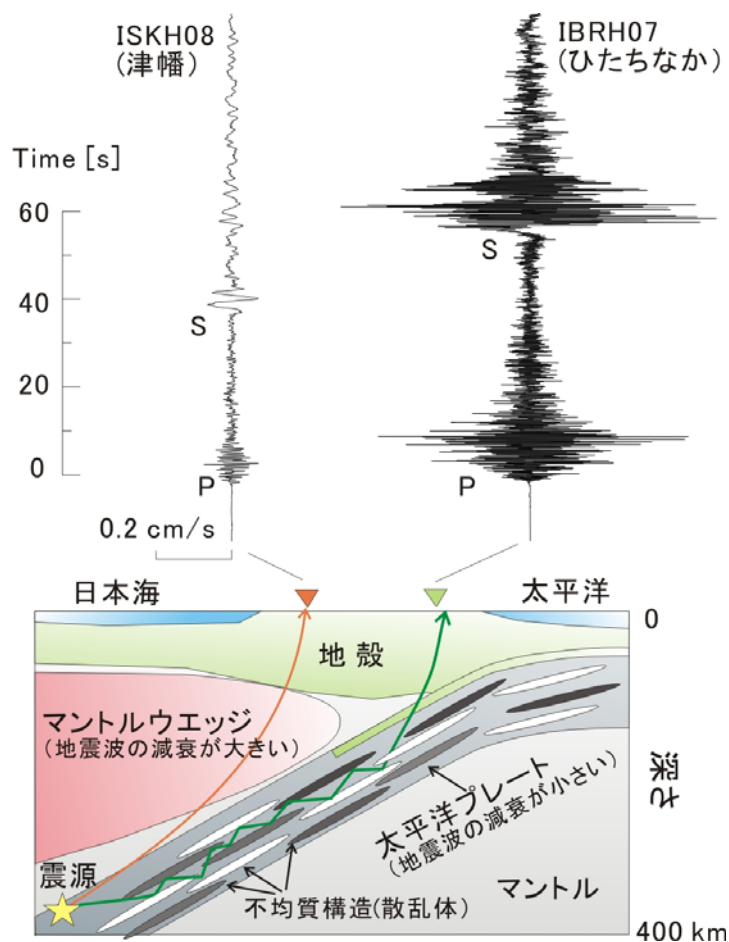


図 17 京都府沖の深発地震（図 16）で観測された地震波形の違い（日本海側と大平側）。

太平洋側（ひたちなか）では、短周期地震動が強く、揺れが長く続く。これは、プレート内の不均質性により強い前方散乱を起こして太平洋側に地震波が良く伝わったためである。いっぽう、マントルを通過して日本海側（津幡）に到達した地震波は短周期地震動の減衰が大きい。

そこで、新しいシミュレーションモデルとして、プレート内が一様な物質ではなく、高速度（硬い）と低速度（柔らかい）の岩石が薄く、交互に積み重なった「ラミナ」構造になっているモデルを考える。このようなラミナ構造を考えると短周期地震波が強い前方散乱を起こすことにより、地震波がプレート内に閉じ込められるように遠地まで良く伝わることを期待できる（図 17）。強い散乱を繰り返し起こすことにより、P波とS波の長い後続層も発生する。いっぽう、波長の長い長周期地震動は不均質構造の影響を受けないので、そのまま屈折を起こしてプレートの外に抜け出すことが考えられる。

4-4-3. 異常震域の生成シミュレーション

不均質なプレート構造により短周期地震波がプレート内部に閉じ込められ遠地まで良く伝播し、そして地表に異常震域が生成する過程を、2003の紀伊半島南東沖の地震のコンピュータシミュレーションにより確認した[34]。

計算領域を 512km*1024km*480km に選び、水平方向に 0.5km、鉛直方向に 0.25km の大きさに 40 億格子に分割した。計算モデルの周囲には PML に基づく吸収境界を設けた。モデル内には、太平洋プレートとフィリピン海プレートを組み込み、太平洋プレートの厚さは 80km、フィリピン

ン海プレートは20kmとし、またプレート上面には厚さ5kmの海洋性地殻を置いた。プレート内部の短波長不均質構造は、統計モデルに基づくランダムな分布を作成し、これを用いてモデル内の物性値に揺らぎを与えた。ランダム分布モデルは、水平方向（プレートの走行と沈み込み方向）に**相関距離**が長く（ $A_x=10\text{km}$ ）、そしてプレートの厚み方向には相関距離が短い（ $A_z=0.5\text{km}$ ）特性を持ち、揺らぎの標準偏差が2%の大きさを持つ、**フォン・カルマン型**分布のものをモデル化した。

地震波伝播計算は、水平方向に16次精度のFDM、鉛直方向には4次精度のFDMを用いて行った。モデル中の最小S波速度は $V_s=3.4\text{km/s}$ とし、最大周波数3Hzまでの波動伝播を評価した。震源モデルはダブルカップル型の点震源で近似し、深さ360kmのプレート内部に置いた。計算は、地球シミュレータの176ノード（1408CPU）を用いた並列計算により行った。地震発生から250秒間の波動伝播を25,000タイムステップで計算した。本計算では、1.3TBのメモリと約1時間のCPU時間を要した。

計算から求められた波動伝播のスナップショットを図19に示す。まず、地震発生から26秒後のフレームを見ると、深さ360kmの深発地震から地震波がマントルとプレート内に向けて放射される様子が確認できる。56秒後にはP波は関東に到達し、S波も紀伊半島を通過して関東に向かってくる。プレート内を地震波が伝わることにより、太平洋岸の揺れが大きくなる（図18b）。揺れは、地震発生から118秒後にはS波が東北にまで到達する。プレートの影響はS波で特に顕著であり、大きく長い揺れがプレートに沿って東北日本の太平洋側をよく伝わっていく様子を見ることができる。また、減衰の大きな**マントルウェッジ**を通過して地震波が到達する日本海側（火山フロントの**背弧側**）では地震波の減衰が大きく、太平洋側（**前弧側**）に比べて揺れが小さいこともわかる。

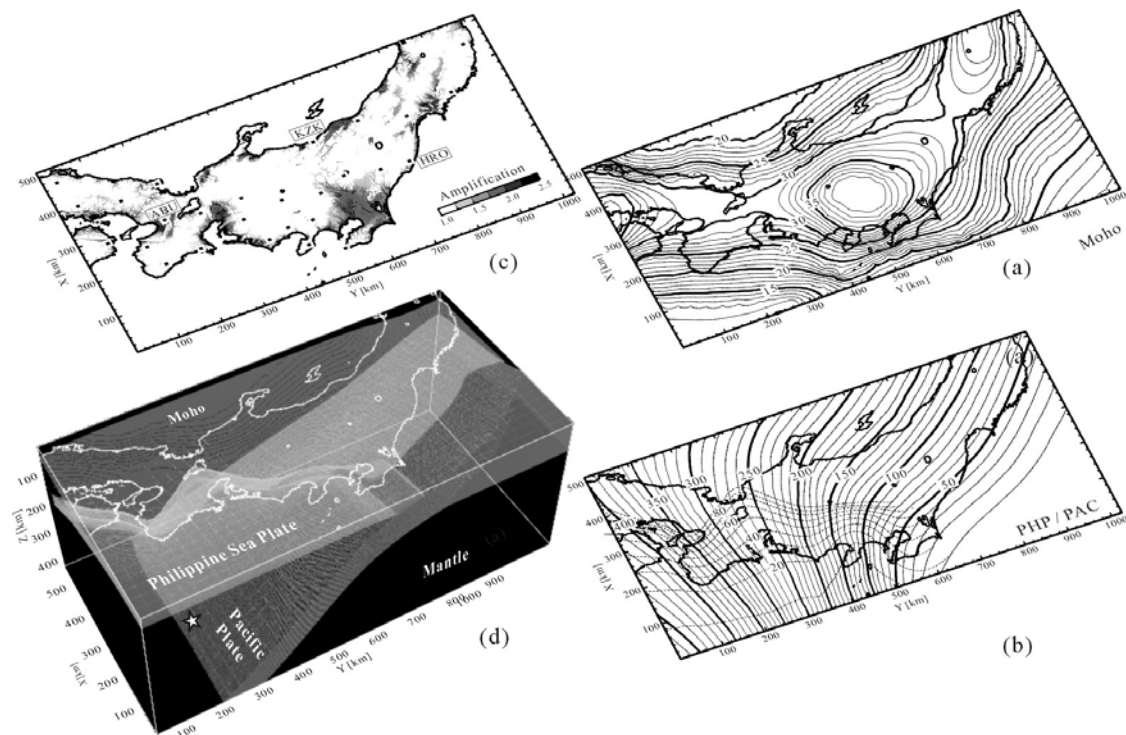


図18 異常震域の3次元数値シミュレーションに用いた中部日本下の地下構造モデル。

(a) 表層の地形分類から求められた表層地盤増幅係数。(b) モホ面の深度分布。(c) 太平洋プレートとフィリピン海プレートの上面の深度。(d) プレート・マントル構造の3次元形状。震源の位置を☆印で表す。

プレートなし，あり（均一，不均質）の3つのモデル計算を行い，計算結果から地表の震度分布を求め比較した（図 20）．まず，プレートが無い（すなわち，高減衰のマントルウェッジもない）モデルでは，震央付を中心に最大震度2の揺れが現れ，ここからほぼ同心円上に震度が弱まる様子が再現された（図 20a）．関東や名古屋では局所的に震度が大きくなっているが，これらの震度の飛び地は，表層地盤による地震動の増幅によるものである．次に，プレートをモデルに組み込むと，地震波がプレートに沿って関東から東北にかけて良く伝わるようになり，南関東では最大震度3が現れた（図 20b）．また日本海側では地震波の減衰が大きくなり，震度が小さくなった．最後に，プレート内の媒質に不均質ゆらぎを与えると，プレート内で地震波の強い散乱が発生し，東北地域に大きな震度の範囲が延びた異常震域が強化された（図 20c）．計算結果は観測された震度分布（図 20d）に良く対応することが確認できた．以上の計算より，周期1秒以下の短周期地震動の伝播計算には，先に述べた長周期地震動の計算のための層構造モデルに加えて，層内の短波長の不均質構造の適切なモデル化が重要であることが示された．

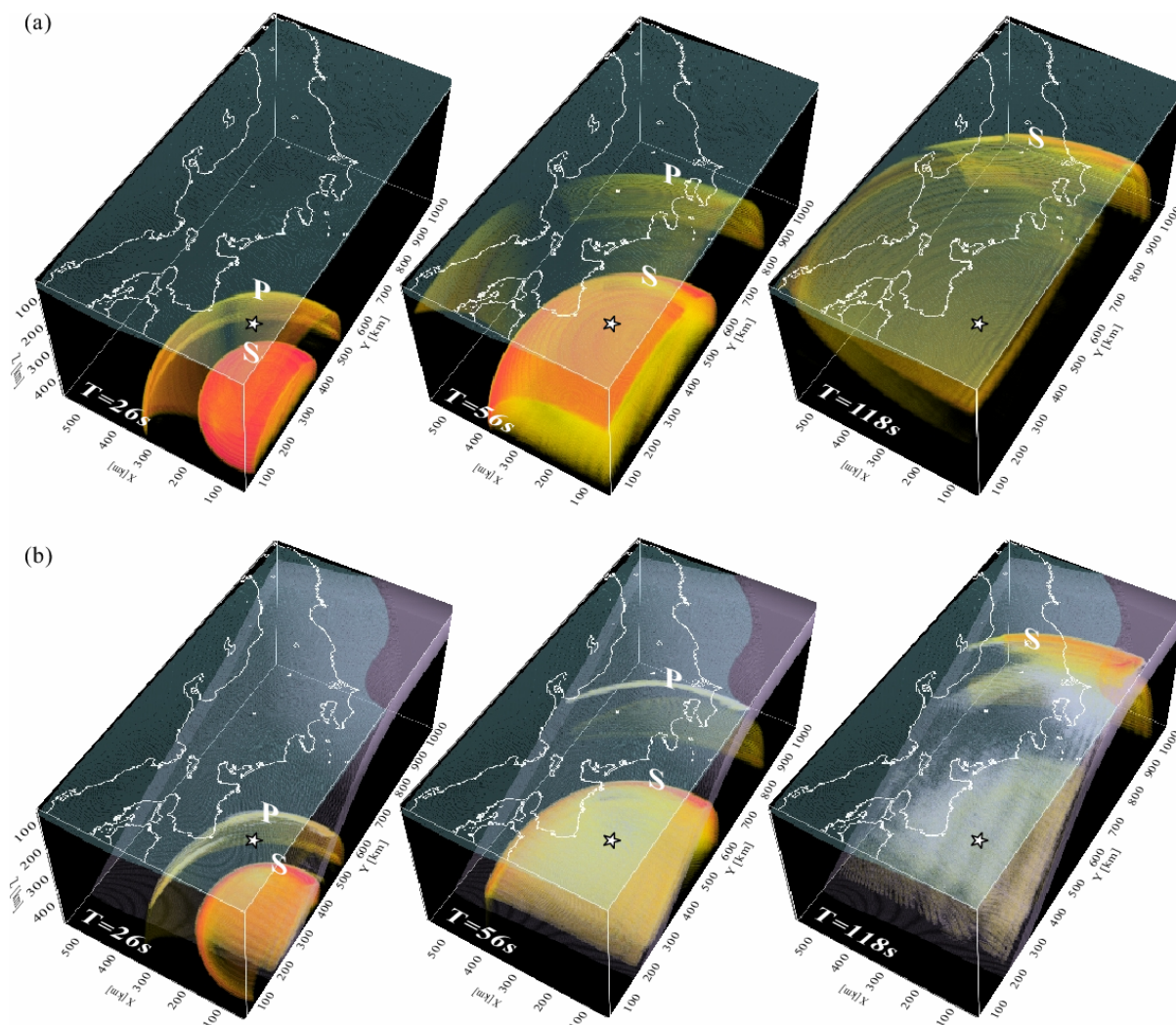


図 19 3次元FDMシミュレーションから求められた地震波動場のスナップショット．

地震発生から26，56，118秒後の揺れの広がる様子をVolume Rendering法により可視化．(a)プレートの無い単純地下構造モデル (b) 不均質なプレートが存在する場合のモデルを用いた計算．

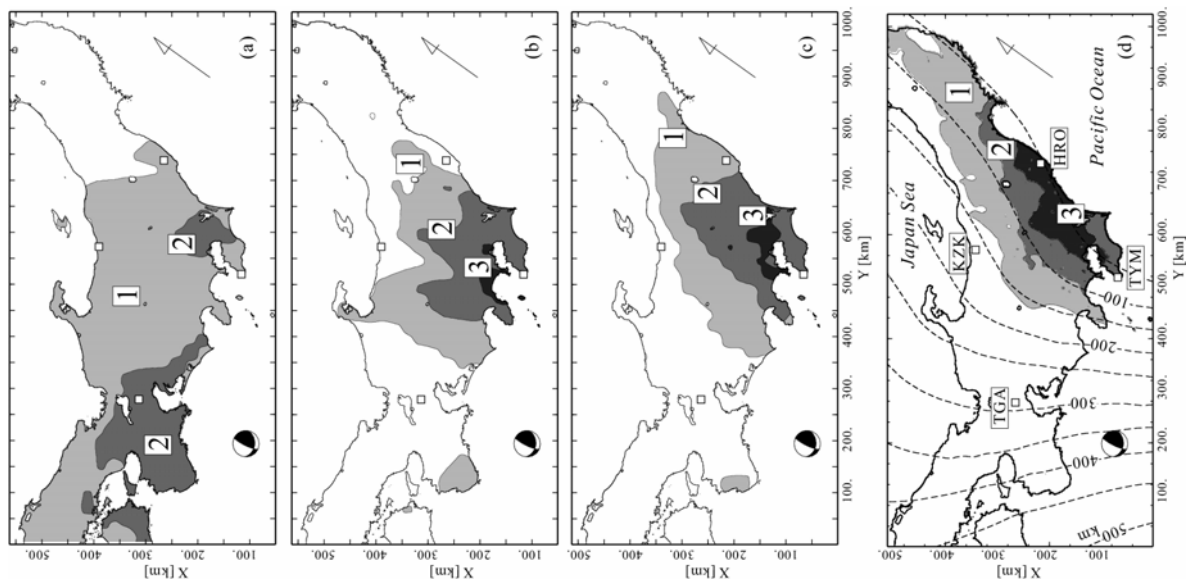


図 20 シミュレーションにより求められた深度分布と観測との比較。

(a)プレートなしモデル，(b)プレートモデル，(c)不均質なプレートモデル，(d)観測された2003年11月12日の紀伊半島沖の地震の震度分布。

5. おわりに

3次元不均質場での地震波伝播の計算は，過去10～20年の計算機の急速な発展に支えられ，これまで急成長を遂げた[35]．地震動シミュレーションの実用化を支える背景には，平野部地下構造探査による地下構造モデルの整備と，高密度の強震観測によるシミュレーション結果の信頼性の検証とモデルの改良の成果が大きい．

ここで紹介した新潟県中越地震や東南海地震のシミュレーションのように，現在の地震波伝播シミュレーションは，周期1～2秒以上のやや長周期地震動をターゲットとするものが一般的である．この地震波周期は10～20階建て以上の超高層ビルなどの長大構造物の被害を考える上では十分であるが，今後，強震動による木造家屋の被害を具体的かつ定量的に評価するためには，計算で扱う地震動の周期帯を，周期0.5～2秒以下まで下げる必要がある．さらに精密機械や工場の生産ラインなど，周期0.5秒以下の固有周期を持つ重要構造物が急激に増えており，現代および未来都市の地震被害予測には，地震動シミュレーションの広帯域化が必須である．

広帯域シミュレーションでは，1)長周期～短周期地震動の放射に関わる震源断層運動の不均質性の理解，2)短周期地震動の伝播と散乱を強く支配する，短波長不均質構造のモデル化，3)大規模並列計算のための計算コードの整備，の3つの課題の解決が急務である．深発地震による異常震域生成のシミュレーション結果から示されたように，周期1秒以下の短周期地震動の評価モデルは，従来の成層モデルに加えて層内の短波長不均質構造を適切にモデル化することが必要である．このような広帯域地震動のシミュレーションモデルの有効性の確認には，観測データとの比較検証のための高密度地震観測網におけるデータの蓄積が望まれる．高性能の並列計算コードの開発への努力もこれまで以上の努力が必要であろう．従来のベクトル型計算機では，ベクトル計算に良く適合するようなプログラミングを心がけることにより，たとえ一般の計算機利用者であっても演算性能を引き出すことが容易であった．しかし，先に述べた

ように今後のスカラ型 CPU の高い演算性能を引き出し、そして大規模な並列計算を実現するためには、コンパイラやハードウェア技術者と計算科学者らの協力が不可欠に思われる。

次に、高精度の地震動シミュレーションの結果を受けて、建物の被害を具体的に評価し、地震災害軽減に役立てるためには、たとえば強震動予測結果を入力データとして地盤と建物の揺れの相互作用を評価し、あるいは超高層ビルの揺れに伴う家具の移動や転倒を評価するなど、地震動と被害予測シミュレーションの強い連携も必要である。また、海の浅い地震による海底地殻変動は津波も発生させるため、海溝型巨大地震による被害予測には、地震と津波シミュレーションの連携も必須である。

これまで、地震や津波、そして構造物の揺れと被害のシミュレーションは、個別に進められてきた。個々のシミュレーション要素技術が十分に実用化した現在では、関連するシミュレーションモデルを結びつけ、地震動から被害までを統合的に評価する**被害予測シミュレーション**へステップアップする段階にきている。災害予測から被害軽減に結びつけるには、**避難シミュレーション**の連成までで視野に入れる必要もあろう。このような、異なる支配方程式と時間スケールで記述される複数のモデルを結合した**マルチフィジックス・マルチスケールシミュレーション**の実現には、計算機の高速度性はもちろんのこと、さらに各計算モデルに適合した異種の計算機システムを高速ネットワークで結合した**グリッド技術**の利用など、利用環境の柔軟性が期待される。

謝 辞

地球シミュレータによる計算は、海洋研究開発機構地球シミュレータセンターと東京大学地震研究所の平成 19 年共同利用プロジェクト研究「3次元不均質場での波動伝播と強震動のシミュレーション」により行われました。また、T2K オープンスパコン（東大）は、東京大学情報基盤センターの HPC 特別プロジェクト「3次元不均質場での地震波伝播の大規模シミュレーション」により行われました。ここに記して関係者の皆様に感謝申し上げます。

参 考 文 献

- [1] 古村孝志・瀨瀬一起・竹中博士, 2000, 大規模 3次元地震波動場（音響場）モデリングのための PSM/FDM ハイブリッド型並列計算, 物理探査, **53**, 294-308.
- [2] Holberg, O., 1987, Computational aspects of the choice of operator and damping interval for numerical differentiation in large-scale simulation of wave phenomena, *Geophys. Prosp.*, **35**, 629-655.
- [3] Kindelan, M., A. Kamel and P. Sguazzero, 1990, On the construction and efficiency of staggered numerical differentiations for the wave equation, *Geophysics*, **55**, 107-110.
- [4] Geller, R.J. and N. Takeuchi, 1998, Optimally accurate second-order time-domain finite difference scheme for the elastic equation of motion; one-dimensional case, *Geophys. J. Int.*, **135**, 48-62.
- [5] Mizutani, H., Geller, R.J. and N. Takeuchi, 2000, Comparison of accuracy and efficiency of time-domain schemes for calculating synthetic seismograms, *Phys. Earth Planet. Int.*, **119**, 75-97.
- [6] Takeuchi, N. and R. J. Geller, 2000, Optimally accurate second order time-domain finite-difference scheme for computing synthetic seismograms in 2-D and 3-D media, *Phys. Earth and Planet. Int.*,

- 119**, 99-131. Virieux, J. 1986, P-SV wave propagation in heterogeneous media: Velocity-stress finite-difference method, *Geophysics*, **51**, 889-901.
- [7] Reshef, M., D. Kosloff, M. Edwards and C. Hsiung, 1988, Three-dimensional elastic modeling by the Fourier method, *Geophysics*, **53**, 1184-1193.
- [8] Furumura, T., B. L. N. Kennett and H. Takenaka, 1998, Parallel 3-D pseudospectral simulation of seismic wave propagation, *Geophysics*, **63**, 279-288.
- [9] Kosloff D., D. Kessler, A.Q. Filho, E. Tessmer, A. Behle and R. Strahilevitz, 1990, Solution of the equations of dynamic elasticity by a Chebyshev spectral method, *Geophysics*, **55**, 734-748.
- [10] Daudt, C. R., Braile, L. W., Nowack, R. L. and C. S. Chiang, 1989, A comparison of finite-difference and Fourier method calculations of synthetic seismograms, *Bull. Seism. Soc. Am.*, **79**, 1210-1230.
- [11] Gottschämmmer, E. and K. B. Olsen, 2001, Accuracy of the explicit planar free-surface boundary condition, implemented in a fourth-order staggered-grid velocity-stress finite-difference scheme. *Bull. Seism. Soc. Am.*, **91**, 617-623.
- [12] Jih, R.S., McLaughlin, K.L. and Z.A.Der, 1988, Free-boundary conditions of arbitrary polygonal topography in a two-dimensional explicit elastic finite-difference scheme, *Geophysics*, **53**, 1045-1055.
- [13] 岡元太郎・竹中博士, 2005, 速度・応力型差分法での固体・流体境界の扱いについて, *地震*, **57**, 355-364.
- [14] Clayton, R. and B. Engquist, 1977, Absorbing boundary conditions for acoustic and elastic wave equations, *Bull. Seism. Soc. Am.*, **67**, 1529-1540.
- [15] Cerjan, C., D. Kosloff, R. Kosloff and M. Reshef, 1985, A nonreflecting boundary condition for discrete acoustic and elastic wave equations, *Geophysics*, **50**, 705-708.
- [16] Berenger, J.P., 1994, A perfectly matched layer for the absorption of electromagnetic waves, *J. Comp. Phys.*, **114**, 185-200.
- [17] Chew, W.C. and Q. H. Liu, 1996, Perfectly matched layers for elastodynamics: a new absorbing boundary condition, *J. Comp. Acoust.*, **4**, 341-359.
- [18] Zeng, Y., J. Q. He and Q. H. Liu, 2001, The application of the perfectly matched layer in numerical modeling of wave propagation in poroelastic media, *Geophysics*, **66**, 1258-1266.
- [19] Marcinkovich, C. and K. Olsen, 2003, On the implementation of perfectly matched layers in a three-dimensional fourth-order velocity stress finite-difference scheme, *J. Geophys. Res.*, **108**, doi:10.1029/2002JBB002235.
- [20] Sato H. and M. Fehler, 1998, *Seismic Wave Propagation and Scattering in the Heterogeneous Earth*, Springer-Verlag, New York.
- [21] Day, S.M. and J. B. Minster, 1984, Numerical simulation of attenuated wavefields using a Pade approximant method, *Geophys. J. Roy. Astr. Soc.*, **78**, 105-118.
- [22] Emmerich, H. and M. Korn, 1987, Incorporation of attenuation into time-domain computations of seismic wave fields, *Geophysics*, **52**, 1252-1264.
- [23] Day, S. M. and C.R. Badley, 2001, Memory-Efficient Simulation of Anelastic Wave Propagation, *Bull. Seism. Soc. Am.*, **91**, 520-531.

- [24] Moczo, P., P. Labak, J. Kristek and F. Hron, 1996, Amplification and differential motion due to an tilted plane 2D resonance in the sediment valleys embedded in a layer over the half-space, *Bull. Seism. Soc. Am.*, **86**, 1434-1446.
- [25] Aoi, S. and H. Fujiwara, 1999, 3-D finite difference method using discontinuous grids, *Bull. Seism. Soc. Am.*, **89**, 918-930.
- [26] Miyatake, T., 1980, Numerical simulations of earthquake source process by a three-dimensional crack model. Part I. Rupture process, *J. Phys. Earth*, **28**, 565-598.
- [28] Moczo, P., 1989, Finite-difference technique for SH-waves in 2-D media using irregular grids-application to the seismic response problem, *Geophys. J. Int.*, **99**, 321-329.
- [29] Pitarka, A., 1999, 3D elastic finite-difference modeling of seismic motion using staggered grid with nonuniform spacing, *Bull. Seism. Soc. Am.*, **89**, 54-68.
- [30] Furumura, T., B.L.N. Kennett and K. Koketsu, 2003, Visualization of 3-D wave propagation from the 2000 Tottori-ken Seibu, Japan earthquake: Observation and numerical simulation, *Bull. Seism. Soc. Am.*, **91**, 4, 667-682.
- [31] Furumura, T. and T. Hayakawa, 2007, Anomalous propagation of long-period ground motions recorded in Tokyo during the 23 October 2004 Niigata-ken Chuetsu (Mw6.6) earthquake, Japan, *Bull. Seism. Soc. Am.*, **97**, 3, 863-880.
- [32] 古村孝志・中村 操, 2006, 1944年東南海地震記録の復元と関東の長周期地震動, *物理探査*, **59**, 4, 337-351.
- [33] 古村孝志・早川俊彦・瀨瀨一起・馬場俊孝・中村 操, 2006, 1944年東南海地震の強震動シミュレーションと関東平野の長周期地震動, *月刊地球*, **55**, 130-137.
- [34] Furumura, T. Hayakawa, M. Nakamura, K. Koketsu, and T. Baba, 2008, Development of long-period ground motions from the Nankai Trough, Japan, earthquakes: Observations and computer simulation of the 1944 Tonankai (Mw8.1) and the 2004 SE Off-Kii Peninsula (Mw7) Earthquakes, *Pure Appl. Geophys.*, **165**, 3, 585-607.
- [35] 古村孝志, 2008, 差分法による3次元不均質場での地震波伝播の大規模計算, *地震2*, 印刷中.

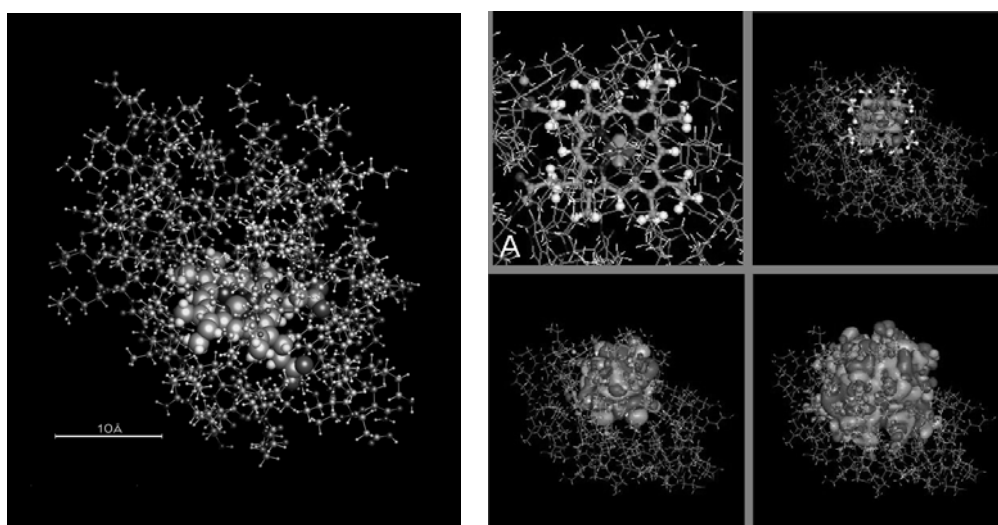
密度汎関数法によるタンパク質全電子シミュレーション

佐藤 文俊

東京大学生産技術研究所

1. はじめに

タンパク質はミクロな世界で生命活動を支える重要な生体分子である。栄養となるもの、構造を形成するもの、運動に関与するもの、免疫や触媒として働くものなど、実に多様な機能を持つ。図1はシトクロム *c* という約 2000 原子からなるタンパク質の構造である。1つの電子を分子間でやり取りするのが役目であるが、ご覧の通りその構造はたいへん複雑である。



第1図: シトクロム *c* の構造(左)と最高被占分子軌道(右)。

(左)ミトコンドリアの中で電子を運ぶタンパク質。アミノ酸残基104個にヘム分子が結合し、原子数は1738、鉄2価のときの電子数は6586である。下部のスペースフィルモデルで示す分子がヘム。その中央の大きな球が鉄イオンを表し、ここを中心に電子のやり取りを行う。(右) $3d_x$ 軌道を主成分とする分子軌道の等値面表示¹。等値面の関数値は左上, 右上, 左下, 右下の順に ± 0.05 , ± 0.005 , ± 0.0005 , ± 0.00005 。濃淡の差はプラス, マイナスを表す。左上Aはスケールを2倍にして描いている。

電子をやり取りすることができる化学分子ならば様々あるのに、なぜ生体はこのような巨大な分子を使うのだろうか、物理化学的な視点から簡単に考察してみよう。

シトクロム *c* は活性中心にヘム(鉄-ポルフィリン)という分子を持っている。仮にもし1つの電子をやり取りする機能を鉄イオンで行うならば 1000 kcal/mol も必要となってしまう。続いて、鉄-ポルフィリンならば 100 kcal/mol で行えるが、まだまだ要するエネルギーは大きい。しかし、これにペプチド鎖が結合したシトクロム *c* ならば 10 kcal/mol 以下になる。電子の授受の前後で、広範囲にわたり少しずつ電子分布を変えることによって、エネルギー変化を小さく

¹ F. Sato, T. Yoshihiro, M. Era, H. Kashiwagi, Chem. Phys. Lett., 341 (2001) p. 645; *ibid.*, 392 (2004) p. 565.

くしていると考えられる。

反応の方向も重要である。シトクロム *c* が活躍するミトコンドリア内膜では、望ましい順番に反応が効率よく進むよう、また逆反応を抑えるように酸化還元電位が順次異なる物質が配置されている。電子のやり取りを行う物質間で形状や静電場などの相補性を利用している場合もある。酵素タンパク質が行っている触媒機能の戦略も基本的にこれと同じである。

以上のように、なぜ生体はタンパク質を働き手に使うのかは、生理温度という穏やかな環境下のもと、わずかなエネルギーで効率よく、望ましい方向に反応を進めるよう最適な場を与えるためと考えると良さそうだ。したがって、複雑な構造を持っていることこそがタンパク質の多様で最適な機能発現を支えているわけで、これらは密接不可分の関係にある。

自然は実に精巧な仕組みを私たちに授けた。しかし、逆にいえば、タンパク質の反応を解析するためには、タンパク質の全構造を考慮に入れ、kcal/mol の精度を保証する手段が必要ということになる。そのうえ、タンパク質の約 3 分の 1 はアミノ酸が連なってできたペプチド鎖以外に、シトクロム *c* のように活性中心や構造安定化モチーフとして金属イオンも利用している。金属イオンとアミノ酸とを同程度の精度で扱うことができ、電子の移動や、触媒反応といった電子が直接かかわる現象を解析する必要がある。

分子が持つ電子の働きを解析する手段は、量子力学の原理を原子分子に適用する量子化学である。その計算方法は分子軌道法と呼ばれ、分子中の各電子の空間分布(分子軌道)を求めるものである。その形と軌道のエネルギー、各種演算子の期待値の計算から分子の全エネルギーや様々な物理量を見積もり、分子物性を解析することができる。

分子軌道計算には Gaussian² や GAMESS³ といったパッケージが整備されており、多くの化学分子で精度の検証が行われていて、kcal/mol 単位でエネルギーが、数%の誤差で物理量の議論ができるようになった。生体分子もこれらの方法を使用して解析できれば、生体反応を解明するための極めて強力な武器となるであろう。

問題は、分子軌道法の計算量は膨大であるため、タンパク質のように数千~数十万原子からなる分子の計算が大変困難となることである。これに加えて、分子軌道法には収束という計算機の能力とはまた別の問題も解決しなければならない。

これらの困難解決に成功したのが、以下の 3 つの方法である。1 つは、一般化学分子の計算で使用される標準的なカノニカル波動関数計算法を高効率に並列化することで、タンパク質分子を丸ごと計算させる、いわゆる力技である^{4,5}。2 つ目は、北浦が開発したエネルギー展開式を利用したフラグメント分子軌道法⁶で、生体分子に適した強力かつ有効な方法である。3 つ目は、量子力学(QM)計算と古典分子力学(MM)計算を組み合わせる QM/MM 法で、古くから生体分子に適用され⁷かつ最も研究が進んでいる。ちなみに、諸熊らの ONIOM 法⁸も QM/MM 法を応用した

² <http://www.gaussian.com/>

³ <http://www.msg.chem.iastate.edu/gamess/>

⁴ F. Sato Y. Shigemitsu, I. Okazaki, S. Yahiro, M. Fukue, S. Kozuru, H. Kashiwagi, Int. J. Quant. Chem., 63 (1997) p. 245.

⁵ K. Tsuda, H. Kaneko, J. Shimada, T. Takada, Comp. Phys. Comm., 142 (2001) p. 140.

⁶ K. Kitaura, T. Sawai, T. Asada, T. Nakano, M. Uebayasi, Chem. Phys. Lett., 312 (1999) p. 319.

⁷ A. Warshel, M. Levitt, J. Mol. Biol., 103 (1976) p. 227.

⁸ M. Svensson, S. Humbel, R. D. J. Froese, T. Mastubara, S. Sieber, K. Morokuma, J. Phys. Chem., 100 (1996) p. 19357.

外挿法である。本拙文は、1 番目の方法の立場から特集号の趣旨に応えようというものである。

生体分子の量子化学計算は、基礎研究のみならず様々な応用も期待されている。ただし、効果的な高速化技法が開発されてはいるものの依然計算量は莫大である。本格的な実用にはペタスケールあるいはその先のエクサスケール規模の計算機が必要となるだろう。

2. 密度汎関数法

分子軌道法は、Hartree-Fock 方程式を基礎方程式とした (*ab initio*) Hartree-Fock 法およびその拡張法や Kohn-Sham 方程式を基礎方程式とした密度汎関数法によるカノニカル(正準)分子軌道法が標準法である。本章では密度汎関数法の基礎について解説する。多原子系に広く適用され、種々の分子で計算精度の検証が行われている堅固な理論で、現在、一般的な化学分子向けの主要な分子軌道計算パッケージで採用されている。加えて、タンパク質のような巨大分子の量子化学計算と高い計算精度という、相反する要求を両立させる方法として注目されている。実際の計算アルゴリズムとタンパク質への適用技法についてはそれぞれ3章と4章で説明する。

2. 1. 密度汎関数法の概要

対象が水素原子であってもタンパク質であっても、原子分子の持つ電子の研究には量子力学が必要であり、その基礎方程式は Schrödinger 方程式である。本稿では、1900 年に始まった量子力学から密度汎関数法にいたる理論を解説するスペースはないため、簡単な説明を述べたのち、密度汎関数法の基礎方程式、計算に用いる方程式を登場させる。これらの詳細に興味がある方は、文末の参考文献に挙げた名著を参照願いたい(筆者らの拙書はご愛敬)。

Schrödinger 方程式は物質などの状態の全波動関数を求める方程式である。原子分子の全電子波動関数を求める式では、電磁波などが外から加えられていない場合、座標と時間の変数分離が達成され、全波動関数が時間変化する式と時間に依存しない方程式が得られる。本稿では主に時間に依存しない方程式を対象に説明する。

続いて、原子分子の物理化学では、電子の質量は1番軽い水素原子核の質量に比べても1800倍軽いという事実を利用して、電子状態は原子核の動きに直ちに追従するという近似が用いられる。これは Born-Oppenheimer 近似と呼ばれる大変重要な仮定である。オングストローム ($\text{Å} = 0.1 \text{ nm}$) 領域の分子を扱う理論研究において、この近似の恩恵は絶大である。分子の電子状態を研究する際は、核の運動を無視し、核座標を固定して電子の振る舞いを計算する。核の運動を議論する際は、核の運動に比べて電子は速やかに平衡分布に達するので、核の運動のみを取り扱う。Born-Oppenheimer 近似はこのような研究方法の基盤となっている。

しかし、Schrödinger 方程式は1電子や非常に特殊な2電子系以外は解析的に解くことができない。そのため、Hartree-Fock 法をはじめとして様々な方法が考案・研究されてきた。密度汎関数法はそれらの方法の1つで、①基底状態のエネルギーが電子密度の汎関数として記述できること、②電子密度を試行関数とする変分原理を満足すること、から導かれた Schrödinger 方程式と同等の解法である。なお、これらの原理は、時間に依存しない Schrödinger 方程式に適用させたのが Hohenberg-Kohn の定理⁹、のちに時間に依存する方程式に適用させたのが Runge-Gross の定理¹⁰と呼ばれている。

⁹ P. Hohenberg, W. Kohn, Phys. Rev., B, 136 (1964) p. 864.

¹⁰ E. Runge, E. K. U. Gross, Phys. Rev. Lett., 52 (1984) p. 997.

密度汎関数法はもう 1 つの標準法である Hartree-Fock 法に対する近似や拡張ではなく、これとは独立した体系である。Hartree-Fock 法の基礎方程式である Hartree-Fock 方程式は、波動関数の汎関数で表された全エネルギーを軌道の規格直交条件のもと、変分することで導かれる。一方、密度汎関数法の骨子は、全エネルギー E が電子密度 $\rho(\mathbf{r})$ の変分可能な汎関数 $E[\rho]$ として与えられるとき、次の変分

$$\delta\left\{E[\rho] - \mu\left(\int \rho(\mathbf{r})d\mathbf{r} - N_{elec}\right)\right\} = 0 \quad (2.1)$$

を求めることである。ここで、 \mathbf{r} は電子の座標である。 μ は電子密度の全空間による積分が全電子数 N_{elec} となる拘束条件を付加するためのラグランジェの未定乗数であり、(2.1)式から

$$\mu = \frac{\delta E[\rho]}{\delta \rho} \quad (2.2)$$

と表される化学ポテンシャルとなる。そのため、まず全エネルギーを電子密度 ρ の汎関数で表す必要があるが、これを保障するのが Hohenberg-Kohn の定理である。

ところで、波動関数 Ψ を汎関数とした変分原理で導かれた Hartree-Fock 系の方法では、分子が作る外場ポテンシャル $v(\mathbf{r})$ が与えられると、その解として Ψ が一意的に決まる。それから電子密度 $\rho(\mathbf{r})$ が一意的に決定される。すなわち、 $v \rightarrow \Psi \rightarrow \rho$ の関係にある。ここで、縮退がなければ、 $\Psi \rightarrow v$ も成り立つ。一方、密度汎関数の定理は $v(\mathbf{r})$ が $\rho(\mathbf{r})$ の 1 価関数であることを保障する。これら全ての結果から、電子密度 $\rho(\mathbf{r})$ と波動関数 Ψ は $v(\mathbf{r})$ を通して 1 対 1 に対応することになる。

$$\rho \leftrightarrow v \leftrightarrow \Psi \quad (2.3)$$

これは真の電子密度で成り立つ。電子密度は波動関数同様、一義的汎関数である。ということは、例えば Hartree-Fock 法では全ての軌道を求める必要があるため、軌道数本分の方程式を連立して解かなければならないが、密度汎関数原理によれば、電子密度による正確な全エネルギーの表現が得られれば、電子密度に関する (2.2) 式のオイラー方程式 1 本で済むことになる。

電子密度もエネルギーも観測量である。そのため、電子密度とエネルギーの対応表が十分にそろえば、十分精密なエネルギー汎関数を構築できる可能性がある。初期の密度汎関数法は、 $E[\rho]$ の記述に明示的な近似汎関数を導入し、著しく簡略化された方程式の導出に関心が持たれた。しかしながら、分子の化学的性質の記述に耐えられるような普遍的なエネルギー汎関数は、今のところ発見されていない。代わりに、その精神は次に述べる Kohn-Sham 法の枠組み中で、実用的な交換相関汎関数を得る問題へと引き継がれ、現在も精力的に研究が行われている。

Kohn-Sham 方程式

Kohn と Sham は汎関数の記述が特にネックとなった運動エネルギー汎関数の表現を放棄し、Hartree-Fock 法と同様の手法を採用した。具体的には、単一のスレーター行列式による波動関数を導入して、多電子問題を有効 1 電子問題に帰着させた¹¹。この試みは見事に成功し、Hartree-Fock 法と同様に軌道数本分の方程式を連立して解かなければならない代わりに、密度汎関数法の精度は驚異的に向上し、多原子分子を解くための現実的な方法となった。これを Kohn-Sham 法と呼ぶ。現在では、密度汎関数法といえば Kohn-Sham 法のことを指し示す言葉と

¹¹ W. Kohn, L. J. Sham, Phys. Rev., A140 (1965) p.1133.

なっている。

それでは、早速 Kohn-Sham 法の支配方程式を見てみよう。本稿では、 \hbar, e^2, m の物理定数値をそれぞれ 1 とする Hartree 原子単位系 au を用いる。この単位系を用いると式がすっきりするため、分子軌道計算ではこの単位を使うことが多い。ただし、実際の計算においては以下のことを念頭に置く必要がある。1 au の距離はボーア半径となり約 0.529 Å である。また、1 au のエネルギーは 27.21 eV, 627.5 kcal/mol であり、タンパク質の反応エネルギーに比べて非常に大きい。分子軌道計算の結果解析の際には十分に注意すべきである。

エネルギー汎関数 $E[\rho]$ は運動エネルギー $T[\rho]$ と核電子間のクーロン引力 $V_{ne}[\rho]$ 、電子間クーロン相互作用 $V_{ee}[\rho]$ の和である。ただし、

$$V_{ne}[\rho] = - \sum_A^{N_{atom}} \int \frac{\rho(\mathbf{r}) Z_A}{|\mathbf{r} - \mathbf{R}_A|} d\mathbf{r} \quad (2.4)$$

である。 Z_A 、 \mathbf{R}_A はそれぞれ原子核 A の電荷、座標である。ここで、単一のスレーター行列式が与える運動エネルギー $T_s[\rho]$ 、古典的な電子間クーロン相互作用を $J[\rho]$ とし、

$$E_{XC}[\rho] \equiv (T[\rho] - T_s[\rho]) + (V_{ee}[\rho] - J[\rho]) \quad (2.5)$$

真の運動エネルギーと近似運動エネルギーとの差、および真の電子間クーロン反発エネルギーと古典の電子間クーロンエネルギーとの差の和を定義する。すなわち、それぞれの項の不明分を押し込めたものである。特に、電子間クーロン反発エネルギーの差が与える物理的意味から、

$E_{XC}[\rho]$ は便宜上、交換相関エネルギーと呼ばれている。

(2.5) 式を用いると、(2.2) 式のオイラー方程式は以下と書ける。

$$\mu = v_{eff}(\mathbf{r}) + \frac{\delta T_s[\rho]}{\delta \rho} \quad (2.6a)$$

$$v_{eff}(\mathbf{r}) = - \sum_A^{N_{atom}} \frac{Z_A}{|\mathbf{r} - \mathbf{R}_A|} + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + \mu_{XC}(\mathbf{r}) \quad (2.6b)$$

ここで、 $\mu_{XC}(\mathbf{r})$ は交換相関ポテンシャルと呼ばれ、 $E_{XC}[\rho]$ が決まれば

$$\mu_{XC}(\mathbf{r}) = \frac{\delta E_{XC}[\rho]}{\delta \rho} \quad (2.7)$$

より直ちに求まる。

単一のスレーター行列式は相互作用のない系における全波動関数を与える。そのため、単一のスレーター行列式の軌道 $\phi_i(\mathbf{r})$ で定義した電子密度

$$\rho(\mathbf{r}) = \sum_{i=1}^{N_{elec}} |\phi_i(\mathbf{r})|^2 \quad (2.8)$$

が厳密な電子密度と等しいとすれば、(2.6a) 式は、(2.6b) 式で表されたポテンシャル $v_{eff}(\mathbf{r})$ のもとで、電子どうしが互いに相互作用無く運動する ($\delta T_s[\rho]/\delta \rho$)、仮想的な電子系の電子密度

を求める方程式に相当する。このように Kohn と Sham は、真のポテンシャルのもとで相互作用する電子系の問題を、(2.6b)式の $v_{eff}(\mathbf{r})$ のもとで電子間の相互作用のない系を解く形式にすり替えたのである。

Kohn-Sham 方程式は、スレーター行列式の軌道の規格直交条件のもとでエルミート性を利用して、正準 Hartree-Fock 方程式と似た 1 電子方程式として与えられる。

$$\left[-\frac{1}{2}\Delta - \sum_A \frac{Z_A}{|\mathbf{r} - \mathbf{R}_A|} + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + \mu_{XC} \right] \phi_i(\mathbf{r}) = \varepsilon_i \phi_i(\mathbf{r}) \quad (2.9a)$$

これを制限付き閉殻 Kohn-Sham 方程式と呼ぶ。 α, β スピンを分離する場合は以下となる。

$$\left[-\frac{1}{2}\Delta - \sum_A \frac{Z_A}{|\mathbf{r} - \mathbf{R}_A|} + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + \mu_{XC}^\sigma \right] \phi_i^\sigma(\mathbf{r}) = \varepsilon_i^\sigma \phi_i^\sigma(\mathbf{r}) \quad (2.9b)$$

これを非制限 Kohn-Sham 方程式と呼ぶ。ここで、 Δ はラプラス演算子、 $\sigma = \alpha, \beta$ である。(2.9)式の左辺の括弧内は Kohn-Sham 演算子と呼ばれ、右辺の ε_i は Kohn-Sham 軌道エネルギーである。Kohn-Sham 法に基づく分子軌道法は、これら正準 Kohn-Sham 方程式を解く方法である。

交換相関エネルギー

Kohn-Sham 方程式と Hartree-Fock 方程式は大変よく似ており、実際(2.9)式において交換相関ポテンシャル演算子 μ_{XC} を Fock 交換演算子に置き換えたものが Hartree-Fock 方程式である。

(2.7)式で μ_{XC} を決める $E_{XC}[\rho]$ は(2.5)式で定義されるように、真の運動エネルギーとの差も繰り込まれているものの、本質的には電子間クーロン相互作用における量子効果が生むエネルギーの総和である。その1つは Pauli の排他原理から要請される電子の交換エネルギーである。これとは別のエネルギーの利得がある。電子はマイナスの電荷を持っているため、電子が互いに接近するとエネルギーが高くなる。そこで平均分布は同じでも互いに避けあう方が安定になる。これを電子相関効果と呼ぶ。Hartree-Fock 法では正確な交換エネルギーを与えるが、異なるスピンを持った電子間のクーロン反発が考慮されていない。したがって、この電子相関効果と運動エネルギー補正を正確に見積もることができれば、Kohn-Sham 方程式は基底状態のシュレディンガー方程式と等価になる。もっとも、先に述べたとおり、 $E_{XC}[\rho]$ の普遍的な汎関数はわかっていない。これが Kohn-Sham 法の最大の特徴でもあり同時に欠点でもある。原理的には、 $E_{XC}[\rho]$ の記述が適切であるほど Kohn-Sham 方程式の精度が向上する。

Hartree-Fock 法の枠組みで不足している電子相関を取り込む場合には、複数のスレーター行列式を用いた系で計算するか、摂動展開でその効果を取り込む。これらの方法では計算量が飛躍的に増えてしまううえに、相関効果の取り込まれ方も非常にゆっくりとしている。Kohn-Sham 法の枠組みでは、 $E_{XC}[\rho]$ の記述の中に相関エネルギーを取り込むことができるため、本質的な

計算量を格段に増やさずに済む可能性が高い。そのため、現在も精力的に新しい交換相関汎関数が提案されている。実際に、Hartree-Fock 法の拡張法よりも遥かに少ない計算時間で、それらと対等またはそれ以上の計算精度を得るまでになっている。今や Kohn-Sham 法は一般化学分子の標準的な分子軌道計算法として認められているが、タンパク質のような巨大で豊富な電子を持つ系においてこそ、まさにその威力が発揮される計算方法である。

2. 2. Kohn-Sham-Roothaan 方程式

Kohn-Sham 方程式の多原子分子の解法には、Hartree-Fock 方程式の解法で開発された Roothaan 法¹²が適用される。分子軌道を基底関数の線形結合で表現するというもので、 i 番目の分子軌道 $\phi_i(\mathbf{r})$ を次のように既知の関数系 ($\mathbf{g}_p(\mathbf{r}): p = 1, 2, \dots, N_{bf}$) によって展開する。

$$\phi_i(\mathbf{r}) = \sum_p^{N_{bf}} C_{pi} \mathbf{g}_p(\mathbf{r}) \quad (2.10)$$

ここで、 $\mathbf{g}_p(\mathbf{r})$ を基底関数と呼ぶ。基底関数はスレーター型やガウス型の関数、さらには平面波関数を用いるものもある。いずれにせよ、実際の計算では、(2.10) 式の展開は有限である。

(2.10) 式の分子軌道 $\phi_i(\mathbf{r})$ の展開式では基底関数は明確に定義されているため、微積分方程式 (2.9) 式を解く問題は未知である展開係数 C_{pi} と ε_i を決める問題になる。これが Roothaan 法最大の特徴である。ちなみに展開係数は当初の発想から LCAO 係数 (Linear Combination of Atomic Orbitals; 原子軌道の線形結合係数) と呼ばれている。

Kohn-Sham-Roothaan 方程式の導出は、軌道を (2.10) 式で展開し、(2.9) 式の左から $\langle \mathbf{g}_p(\mathbf{r}) |$ を演算する。閉殻系の (2.9a) 式の行列表現は

$$\mathbf{FC} = \mathbf{SC}\boldsymbol{\varepsilon} \quad (2.11)$$

である。 $\mathbf{C}, \boldsymbol{\varepsilon}$ は \mathbf{F} を対角化したときに求まる固有ベクトルの行列と固有値であり、 $\boldsymbol{\varepsilon}$ は対角行列である。 \mathbf{S} は重なり積分の行列である。上の式を行列要素で表すと次のようになる。

$$\sum_q^{N_{bf}} F_{pq} C_{qi} = \varepsilon_i \sum_q^{N_{bf}} S_{pq} C_{qi} \quad (2.12a)$$

$$F_{pq} = h_{pq} + \sum_{rs} P_{rs} \langle pq | rs \rangle + \mu_{pq}^{XC} \quad (2.12b)$$

ただし、

$$h_{pq} = \int \mathbf{g}_p^*(\mathbf{r}) \left[-\frac{1}{2} \Delta - \sum_A^{N_{atom}} \frac{Z_A}{|\mathbf{r} - \mathbf{R}_A|} \right] \mathbf{g}_q(\mathbf{r}) d\mathbf{r} \quad (2.13)$$

¹² C. C. J. Roothaan, Rev. Mod. Phys., 23 (1951) p. 69.

$$\langle pq|rs\rangle \equiv \iint g_p^*(\mathbf{r})g_q(\mathbf{r})\frac{1}{|\mathbf{r}-\mathbf{r}'|}g_r^*(\mathbf{r}')g_s(\mathbf{r}')d\mathbf{r}d\mathbf{r}' \quad (2.14)$$

$$P_{rs} = 2\sum_j C_{rj}^* C_{sj} \quad (2.15)$$

$$S_{pq} = \int g_p^*(\mathbf{r})g_q(\mathbf{r})d\mathbf{r} \quad (2.16)$$

$$\mu_{pq}^{XC} = \int g_p^*(\mathbf{r})\mu_{XC}(\mathbf{r})g_q(\mathbf{r})d\mathbf{r} \quad (2.17)$$

である。 \mathbf{F} はKohn-Sham行列と呼ばれる。 \mathbf{h} , $\langle pq|rs\rangle$, \mathbf{S} および \mathbf{P} は,それぞれ1電子積分, 2電子積分, 重なり積分および密度行列と呼ばれ, Hartree-Fock-Roothaan方程式と共通である。Kohn-Sham-Roothaan方程式は, (2.17)式の交換相関積分演算以外は, Hartree-Fock-Roothaan方程式と同じ分子積分演算と一般行列演算からなる。解法そのものもHartree-Fock-Roothaan方程式と共通するものが多い(3.1節参照)。逆に言えば, 化学の世界で培われてきたHartree-Fock-Roothaan方程式を解くための様々な技術が応用できる。

Kohn-Sham-Roothaan方程式の最大の特徴は(2.17)式の μ^{XC} である。Hartree-Fock-Roothaan方程式では, これに対応する項は(2.14)式と同じ形の分子積分を用いて顕に書き下すことができる。現在提案されている $\mu_{XC}(\mathbf{r})$ は一般に有理関数であるため, μ^{XC} の行列要素の計算には分子積分演算とは異なる数値積分計算を用いることになる(3.2節参照)。

このように, 交換相関項の計算に特殊な技術や時間が必要となるものの, Kohn-Sham-Roothaan法はHartree-Fock-Roothaan法と本質的に同程度の計算量で飛躍的に計算精度が増すことができる。これこそが, 小分子のみならず巨大分子であるタンパク質の量子化学計算にとって大変魅力的である所以である。

3. Kohn-Sham-Roothaan法の解法概要

この章では, 2章で示したKohn-Sham-Roothaan方程式を解く基本アルゴリズムと, 巨大なタンパク質の計算に欠かせない高速化と大量データの扱いに関する基礎技術を概説する。なお, これらの方法は1章で紹介したフラグメント分子軌道法やQM/MM法にも有効で, 大きなフラグメントやQM領域を計算する方法に適用できる。

3. 1. Kohn-Sham-Roothaan法のアルゴリズム

まず本節では, タンパク質の座標が固定され, スピンの異なる電子が2個ずつ分子軌道に入る閉殻系の基底状態の計算を例に, もっとも基本的なKohn-Sham-Roothaan法の自己無撞着(SCF)計算アルゴリズムを説明する。一部の計算を除けばHartree-Fock-Roothaan法と共通である。また, 閉殻系の計算は α スピンと β スピンでそれぞれ行えばよい。

解くべき方程式は(2.11)式である。基底関数 $g_p(\mathbf{r})$ は普通規格化されているが, 直交はしていないため, (2.16)式の \mathbf{S} は単位行列ではない。そこで, \mathbf{S} を対角化するユニタリ行列 \mathbf{U} を

使って、直交変換行列 \mathbf{X} を作る。

$$\mathbf{X} = \mathbf{U}\mathbf{s}^{-\frac{1}{2}} \quad (3.1a)$$

$$\mathbf{U}^\dagger \mathbf{S}\mathbf{U} = \mathbf{s} \quad (3.1b)$$

ここで、 \mathbf{s} は \mathbf{S} の固有値を対角要素に持つ対角行列である。 $-1/2$ 乗は対角要素の平方根の逆数をとることを意味する。この \mathbf{X} を用いて (2.11) 式を変形する。

$$\mathbf{F}'\mathbf{C}' = \mathbf{C}'\boldsymbol{\varepsilon} \quad (3.2a)$$

$$\mathbf{F}' = \mathbf{X}^\dagger \mathbf{F}\mathbf{X} \quad (3.2b)$$

$$\mathbf{C}' = \mathbf{X}^{-1}\mathbf{C} \quad (3.2c)$$

これで (2.11) 式は通常の固有値問題 (3.2a) 式になった。得られた \mathbf{C}' から \mathbf{C} を逆変換 $\mathbf{X}\mathbf{C}'$ で計算し、(2.15) 式で \mathbf{P} を計算しなおして全エネルギーが求まる。

$$E = \sum_{pq} P_{pq} h_{pq} + \frac{1}{2} \sum_{pqrs} P_{pq} P_{rs} \langle pq | rs \rangle + E_{XC} + \sum_{A=1}^{N_{atom}} \sum_{B>A}^{N_{atom}} \frac{Z_A Z_B}{|\mathbf{R}_A - \mathbf{R}_B|} \quad (3.3)$$

SCF 計算では、適当な波動関数や電子密度の初期値から計算を始め、(2.12) から (3.3) 式を繰り返し計算し、全エネルギーや各行列要素などの差の絶対値が望みの閾値以下になったときに計算を終了する。これを収束と呼ぶ。

化学分子の Roothaan 法で使用する基底関数は、いくつかのガウス型関数の和で表される実数型の縮約ガウス型関数が一般的である。この場合、これまでの式で複素共役を意味する* (アスタリスク) は取ってかまわない。

Kohn-Sham-Roothaan 方程式の計算手順を、図 2 にしたがって説明する。なお、Hartree-Fock-Roothaan 法と異なるところに図では網掛けを、本文中では○印をつけている。

1) 分子構造、基底関数、計算条件を読み込む。タンパク質の分子軌道の計算を行うにあたり、タンパク質分子の座標や基底関数の組、電子の配置や計算の手法、用いる交換相関汎関数といった情報が必要であり、これを初めに読み込む。
続いて、SCF 計算のための準備計算を行う。SCF の繰り返し計算を通して変化しない物理量をあらかじめ計算しておく。

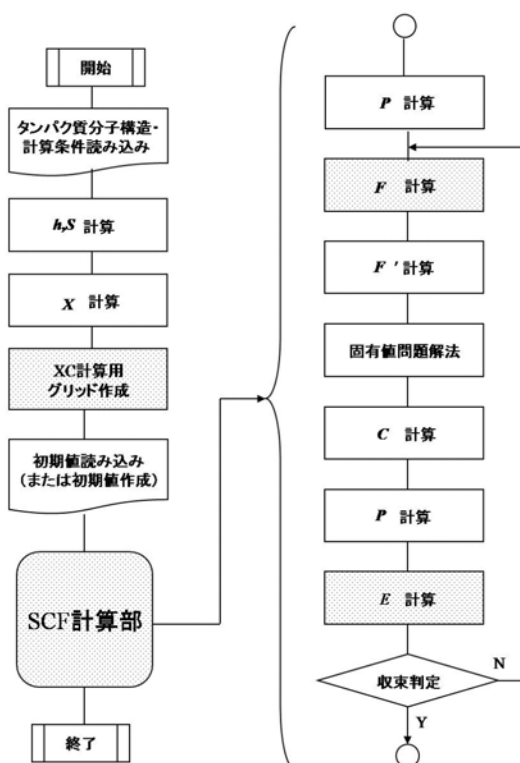
- 2) (2.13) 式の 1 電子積分 \mathbf{h} と (2.16) 式の重なり積分 \mathbf{S} を計算する。
- 3) (3.1a) 式で直交規格化条件を満たすよう基底関数を変換する行列 \mathbf{X} を計算する。この際、基底の線形従属性を取り除くという作業も行っておく。
- 4) (2.17) 式の交換相関項の積分に用いるグリッドとその重みを生成する(○)。
- 5) SCF 計算の初期値を作成する。別に作成したものを読み込んでもよい。LCAO 係数 \mathbf{C} 、密度行列 \mathbf{P} などを与えることができる。初期値作成法に関しては 4 章を参照。

以降が、SCF 計算における繰り返し計算部分である。様々な計算ルーチンが逐次呼び出され計算が実行される。収束条件が満たされるまで計算が繰り返される。

- 6) (2.15) 式から電子密度行列 \mathbf{P} を計算する。電子密度行列が初期値として与えられている場合はスキップする。
- 7) 2 電子積分を計算しながら、(2.12) 式の Kohn-Sham 行列 \mathbf{F} を計算する(○)。
- 8) (3.2b) 式により Kohn-Sham 行列を規格直交基底 \mathbf{F}' に変換する。
- 9) (3.2a) 式の固有値問題を解く。固有ベクトルとして直交基底の Kohn-Sham LCAO 係数 \mathbf{C}' ,

固有値として Kohn-Sham 軌道エネルギー ϵ を得る。

- 10) $C = XC'$ により, LCAO 行列を本来の非直交基底に変換する。
- 11) (2.15) 式で新しい電子密度行列を計算する。
- 12) (3.3) 式の全エネルギー (ϵ) や各原子の電荷などの物理量を計算する。
- 13) SCF 計算の収束条件を満たしたかどうかを判定する。1 回前の SCF で計算された密度行列の全要素と全エネルギーを今回のものと比較して, その差の絶対値が閾値以下になった場合を収束とみなす。収束したら繰り返し計算を終了する。収束条件を満たさない場合は, 4 節の収束法を施して, 7) に戻る。



第 2 図: Kohn-Sham-Roothaan 方程式の計算手順[1]。

Hartree-Fock-Roothaan 法と異なるところに網掛けしている。

以上が Kohn-Sham-Roothaan 方程式の SCF 計算手順である。○印をつけた Hartree-Fock-Roothaan 法との差異はわずかで, 化学の世界で蓄積してきた分子軌道法の技術が利用しやすい形になっている。

なお, SCF 計算は必ず収束するとは限らない。発散したり, 振動したり, また収束しない場合もあるので, 実際の計算には注意が必要である。これはタンパク質の SCF 計算では特に深刻な問題である。この問題への対策は 4 章で説明する。

3. 2. 律速計算のアルゴリズム概要

続いて, 上記の各ルーチンのうち, 律速となる計算のアルゴリズムを概説しておこう。これらは多岐にわたるうえ, 計算機の進歩とともに次々と新しい方法が開発されている。現在もホットなテーマである。本稿ではそれぞれの標準法を簡単に紹介する。

分子積分

分子積分は Roothaan 分子軌道法において最も計算律速となる計算ステップである。分子積分計算の問題点は大きく分けて 3 つある。

第 1 に計算自体が難しいことである。原子軌道型の基底関数には軌道のすその振る舞いからスレーター型の関数 ($\exp(-\alpha|\mathbf{r}-\mathbf{R}|)$) が適しているが、この関数を用いると分子積分の計算に多くの行程が必要になる。そこで、Boys は基底関数にガウス型関数 ($\exp(-\alpha|\mathbf{r}-\mathbf{R}|^2)$) を使用することを提案した。軌道の良好な表現を得るためには、スレーター型の関数に比べて多くの数の関数を使用しなければならないが、分子積分をスレーター型に比べ少ない行程で解析的に計算することができる。現在では、ほとんどの化学分子用分子軌道法プログラムはガウス型基底関数を採用している。

第 2 にデータ量が膨大になることである。中でも、4 つのインデックスを持つ 2 電子積分 $\langle pq|rs\rangle$ のサイズは倍精度浮動小数点型で $8(\text{Byte}) \times N_{bf}^4 / 8 = N_{bf}^4 (\text{Byte})$ になる。ここで、1/8 は分子積分の対称性を考慮した比率である。基底関数が 1 万で 10^{16} Byte, 10 万では 10^{20} Byte になる。保持するには大変な量である。

第 3 に同様に計算量も膨大になることである。2 電子積分の計算量は N_{bf}^4 に比例する。個々の分子積分を高速に計算することができても、分子サイズが大きくなるとすぐに計算できる限界に達してしまう。分子サイズ依存性を低下させるような計算量の削減方法が必要になる。

このような問題点の解決に用いられる現在の標準法を以下に簡単に説明しよう。まず第 1 の問題には、ガウス型関数の分子積分では、漸化式による高速計算法が用いられている。ガウス型関数のもつ基本的な性質として、ガウス型関数どうしの積はやはりガウス型関数になるので非積分関数が簡単なものになる。また、ガウス型関数自体の積分も容易である。さらに、(2. 14) 式で現れるクーロン演算子 $1/|\mathbf{r}-\mathbf{r}'|$ はラプラス変換により、やはりガウス型関数の積分として取り扱うことができる。Boys の提案が受け入れられた背景にこれらの都合な性質がある。実際に使用されている縮約ガウス型関数はガウス型関数の線形結合であるから、分子積分の表現も単に展開項数が増えただけのものにすぎない。

それに加えて、少ない行程で分子積分を計算する実に巧妙な方法を編み出したのが小原と雑賀¹³である。ガウス型基底関数の中心座標に関して微分を計算すると、例えば

$$\frac{\partial}{\partial R_x} \exp(-\alpha(\mathbf{r}-\mathbf{R})^2) = -2\alpha(x-R_x) \exp(-\alpha(\mathbf{r}-\mathbf{R})^2) \quad (3.4a)$$

すなわち

$$\frac{\partial}{\partial R_x} g^s(\mathbf{r}, \mathbf{R}) = -2\alpha g^{p_x}(\mathbf{r}, \mathbf{R}) \quad (3.4b)$$

となる。ここで、 g^s は s 軌道、 g^{p_x} は p_x 軌道である。このように軌道角運動量の量子数が 1 つ大きい軌道が得られる。この特徴から s 軌道の分子積分計算値から出発して、順次 p , d 軌道

¹³ S. Obara, A. Saika, J. Chem. Phys., 84 (1986) p. 3963.

の分子積分を計算してゆく漸化式が導かれた。最も複雑な 2 電子積分の漸化式でも d 軌道までなら高々 8 項程度の和で表現され、演算量が少ないことが最大の利点である。

続いて、第 2 の問題、すなわち分子積分のデータ量とそれに付随する問題について検討する。従来は 2 電子積分の計算に時間がかかったため、SCF 計算の前に 1 度だけ計算してディスクなどの外部記憶に保存し、必要なときにファイルから読み出していた。これを従来法あるいはディスク法と呼ぶ。この方法では Kohn-Sham 行列要素などの計算はファイルアクセスの時間に支配され、計算できる分子サイズがディスクの容量で決められてしまう。タンパク質の計算では全ての値を保存するには膨大なディスク容量が必要になる。

そのため、ディスクに積分値を保持しない方法が考案された。分子積分を SCF 計算で必要になる度に毎回計算して、分子積分を使い捨てにする方法¹⁴である。これをダイレクト法と呼ぶ。この方法が受け入れられた背景として、小原と雑賀法のような効率のよい分子積分計算法が提案されたこと、次に述べるカットオフ法が開発されたことが挙げられる。様々な計算機メーカーが競って高速な CPU を開発していることも見逃せないだろう。現在ダイレクト法は分子軌道計算の標準になっている。しかし、今後、従来法の考え方に再帰する可能性もある。これについては本小節の最後に述べよう。

ダイレクト法では 2 電子積分が必要なときに計算され使用される。クーロン項の形は

$$U_{pq} = \sum_{rs} P_{rs} \langle pq|rs \rangle \quad (3.5)$$

である。この U であれば、 $\langle pq|rs \rangle$ の 2 電子積分を r,s あるいは p,q のインデックスで呼び出しながら、主記憶に保持した電子密度行列 \mathbf{P} との積和計算で求めることができる。これらのインデックスが走る領域を固定長にして、積和が終了したら次の未計算のインデックス群に移行し、インデックスの最後に到達するまで同様の操作を行う。このようなアルゴリズムを用いれば、分子積分に要するメモリを少なく抑えることができる。この方法により、計算できる分子のサイズがメモリやディスクの容量によって大幅に限定されることがなくなった。

ダイレクト法には、SCF 計算の収束過程を利用した高速化法もある。(3.5) 式の行列要素の計算は以下のように書き換えることができる。

$$U_{pq}^{(\lambda)} = U_{pq}^{(\lambda-1)} + \sum_{rs} (P_{rs}^{(\lambda)} - P_{rs}^{(\lambda-1)}) \langle pq|rs \rangle \quad (3.6)$$

ここでインデックス (λ) は SCF 計算の繰り返し回数である。これらの式では、前回の SCF 計算で求められた行列要素に電子密度の差分による補正を加えている。SCF 計算が収束に近づくにつれて $(\mathbf{P}^{(\lambda)} - \mathbf{P}^{(\lambda-1)})$ の要素の絶対値は小さくなるので、 $\langle pq|rs \rangle$ の値が小さいものは計算を省略することができる。これをアップデート法と呼ぶ。この方法は次のカットオフ法と密接な関係があるため、詳細はあとで述べるが、SCF 計算が収束に近づくにつれて計算時間が短縮する。

最後に、分子積分計算量の分子サイズ依存性を劇的に減らすカットオフ法について説明する。この方法とダイレクト法およびアップデート法を組み合わせることにより、分子積分が関与する計算を大幅に高速化することができる。

分子が大きくなればなるほど、2 電子積分の大半はゼロに近い値になる。基底関数 p と q が

¹⁴ J. Almlöf, K. Faegri, Jr., K. Korsell, J. Comp. Chem., 3 (1982) p.385.

離れていれば p と q の重なりは無視できるほど小さくなるからである。カットオフ法とは、分子積分の上限値を事前に簡単な式で見積もり、これが SCF 計算の収束閾値よりも十分に小さい場合は、その分子積分の計算を省略する方法である。2 電子積分は Schwarz の不等式から

$$\langle pq|rs\rangle \leq \sqrt{\langle pq|pq\rangle} \sqrt{\langle rs|rs\rangle} \quad (3.7)$$

である。したがって、 $\langle pq|pq\rangle$ あるいは $\langle rs|rs\rangle$ の値を計算すれば、2 電子積分の上限値を求めることができる。これらは 2 つのインデックスによる計算である。この篩が理想的に働けば計算すべき分子積分の項数を $O(N_{bf}^2)$ にまで減らすことができる。

カットオフ法をダイレクト法と組み合わせるとさらに強力になる。むしろカットオフ法のおかげでダイレクト法が有用となったといってもよい。カットオフ法の上限値の見積もりは (3.5), (3.6) 式の積和の前に行うことができる。これにより、分子積分が関与する U の計算自体が $O(N_{bf}^2)$ に抑えられる。この操作をスクリーニングと呼ぶこともある。

このとき、 $\max|P^{(\lambda)} - P^{(\lambda-1)}| \langle rs|rs\rangle$ をカットオフの判断に使用しても、精度の問題が生じないことがわかる。 $\max|\cdot|$ は行列要素の絶対値の最大値を意味する。すなわち、アップデート法も併用することで、カットオフの効果を高めることができる。SCF 計算の収束値近傍では $\max|P^{(\lambda)} - P^{(\lambda-1)}|$ の値は高々 SCF 収束判定の閾値になるため、 $\langle rs|rs\rangle$ のみをカットオフの判断に使うよりもさらに 3~5 桁大きい分子積分の上限値でカットすることができるからである。

ここで、従来法の考え方に回帰する可能性について述べよう。本稿で紹介した分子積分計算の標準法は目覚しい発展を遂げている CPU の能力に頼った方法である。これらの方法においても分子積分の計算は SCF 計算の大半を占めるので、仮に収束までに SCF 繰り返し計算が 20 回必要になった場合、従来法に比べて 20 倍近く計算時間の損をしていることになる。そこで、全ての分子積分を使い捨てるのではなく、一部をディスクやメモリに保存しておく方法¹⁵が考えられている。セミダイレクト法、部分ダイレクト法と呼ばれるものである。

小原と雑賀の全ての s 軌道タイプの分子積分を保持すると、これまで説明してきたダイレクト法に比べて格段に速くなることが分かっている。ただし、保持しておく分子積分の種類を増やそうとすると複雑な制御が発生し、様々な並列計算機に適合するアルゴリズムは難しいので、それほど普及しているわけではない。もし全ての分子積分を保持することができるとしたら、計算時間は短縮する。ディスクではなく主記憶に保持することができるならば、計算速度は格段に向上する。今では、 10^{15} Byte の主メモリを登載した超並列計算機が誕生しつつある。これならば、タンパク質計算に必要な 2 電子積分がそっくりメモリに収まってしまいうだろう。

交換相関ポテンシャル

この小節では密度汎関数法特有の交換相関項の計算法について概説する。密度汎関数法にお

¹⁵ R. Lindh, J. W. Krogh, M. Schütz, K. Hirao, Theor. Chem. Acc., 110 (2003) p.156.

ける交換相関項の被積分関数は複雑な有理関数 F で記述されている。そのため、上記の分子積分法とは異なる計算方法が要求される。その標準法はグリッドによる数値計算法である。被積分関数が複数の原子核の近傍で大きな値を持つことから、この種の計算は3次元多中心積分と呼ばれる。このような積分には、里子によるボロノイ多面体法¹⁶を発展させた Becke のファジーセル法¹⁷が使用される。

ここで、(2.17)式の交換相関項を一般的な多中心積分として取り扱うことにする。

$$I = \int F(\mathbf{r}) d\mathbf{r} \quad (3.8)$$

通常用いられている数値積分法では、多中心積分をそれぞれの原子核を中心とする1中心積分に展開することから始める。

$$I = \sum_A \int w_A(\mathbf{r}) F(\mathbf{r}) d\mathbf{r} = \sum_A I_A \quad (3.9a)$$

$$I_A = \int w_A(\mathbf{r}) F(\mathbf{r}) d\mathbf{r} \equiv \int F_A(\mathbf{r}) d\mathbf{r} \quad (3.9b)$$

$$\sum_A w_A(\mathbf{r}) = 1 \quad (3.9c)$$

このとき導入した重み関数 $w_A(\mathbf{r})$ は原子核座標 \mathbf{R}_A から作られるボロノイ多面体 Ω_A で定義するのが自然である。すなわち、 \mathbf{R}_A が含まれるボロノイ多面体 Ω_A のみに値を持ち、それ以外の領域は原子核座標 \mathbf{R}_A の勢力圏外とする。これが里子のボロノイ多面体法である。しかし、重み関数 $w_A(\mathbf{r})$ を階段関数で定義してしまうと Ω_A 以外からの効果を取り込めず計算精度に問題が生じるし、座標を変化させながら計算する際に連続性も失われる。そこで、Becke は $w_A(\mathbf{r})$ に階段関数をなました滑らかな関数で定義した。これをファジーセル法と呼ぶ。

続いて、 I_A は原子核座標 \mathbf{R}_A を中心とした極座標系を用いて、

$$\begin{aligned} I_A &= \int F_A(\mathbf{r}) d\mathbf{r} = \int F_A(\mathbf{r}_A + \mathbf{R}_A) d\mathbf{r}_A \\ &\equiv \int f_A(\mathbf{r}_A) d\mathbf{r}_A \\ &= \iiint f_A(r_A, \theta_A, \varphi_A) r_A^2 \sin\theta_A dr_A d\theta_A d\varphi_A \end{aligned} \quad (3.10)$$

と表し、これらの積分はサンプリング点 $\{\mathbf{r}_{Ai}\}$ と重み係数 $\{C_i\}$ を用いた求積法により計算される。

すなわち、

$$I_A \cong \sum_{i=1}^{N_{Grid}} C_i f_A(\mathbf{r}_{Ai}) \quad (3.11)$$

である。ここで N_{Grid} は1中心あたりのサンプリング点の総数である。一般に、球面積分では球面調和関数展開によるガウスの求積法、動径方向の積分にはガウス・ルジャンドルやオイラー・マクローリン法などが用いられる。球面上にサンプリングした1層の点群を異なる半径 r_A で何層にも配置する。これらサンプリング点をグリッドと通称するため、この方法はグリッド法と呼ばれる。したがって、積分 I は(3.9a)式から、

¹⁶ C. Satoko, Phys. Rev., B30 (1984) p.1754.

¹⁷ A. D. Becke, J. Chem. Phys., 88 (1988) p.2547.

$$I \cong \sum_{A=1}^{N_{atom}} \sum_{i=1}^{N_{Grid}} C_i w_A(\mathbf{r}_{Ai}) f_A(\mathbf{r}_{Ai}) \quad (3.12)$$

と計算する。

この方法で、必要な精度を得るためには莫大な数のサンプリング点が必要である。精度を保ちつつなるべくサンプリング点を少なくしたグリッドが提案されており、標準グリッド¹⁸と呼ばれている。1 中心あたり 5000 点ほどのグリッドを用いると小数点以下 7~8 桁程度の計算精度があるが、これで必ずしも十分とはいえない。今後の研究で、分子積分のような解析計算ではなくとも、それと遜色ないほどに十分な精度が得られるようになれば、密度汎関数法の計算上の弱点の 1 つが克服されるだろう。

その他の律速

残る律速計算は、行列積、対角化(計算方法によっては逆行列計算も)といった行列演算である。これらはこれまで説明してきた分子軌道計算特有の計算ではない。個々の計算機種構成でチューニングされ整備されている数値演算ライブラリを使用することが望ましい。

分子積分にしても交換相関積分にしても、カットオフ法もしくは高速多重極展開法(説明は省略)¹⁹などを使用して $O(N_{bf}^2)$ 以下の計算量となる。となると、実は $O(N_{bf}^3)$ の計算サイズ依存性を持つ行列演算こそが深刻で、実際、超巨大分子の計算においては逆転してしまう(次節参照)。

行列演算も SCF 計算毎に繰り返し行われる。計算機システムの能力を最大限に引き出したライブラリの整備が必須である。用途に合わせて、疎行列版と密行列版が必要となるだろう。計算機言語体系別に余計な行列の転置などの操作を要求しないことも大切である。これはコンパイラにもいえることだが、ライブラリの整備なくして高品位シミュレーションの達成はあり得ない。同様に、I/O 性能のチューニングが後回しとなることが多いことも問題である。たとえ先端的な計算機システムが登場し、そのシステム構成に適したアルゴリズムを採用したシミュレーションソフトウェアが存在していても、これらの整備の遅延のため、計算機性能を生かした画期的なシミュレーションがなかなか達成できないことをこれまでしばしば経験している。

3. 3. 計算量と並列計算の方法

タンパク質 Roothaan 方程式の次元

3.2 節のアルゴリズムをもとに、計算律速となるプログラムの演算量を、それぞれ計算規模の指標である原子数、基底関数の数を用いて見積もる。Roothaan 法では非線形方程式の固有値を求める問題を行列方程式の固有値問題に置き換える。そのため、これまでの説明でも登場したように行列のサイズが計算規模を表す指標になる。

タンパク質ではアミノ酸残基の総数と分子量との間に、簡単ではあるが非常によく合う次の経験則が知られている。

$$(\text{タンパク質の分子量}) = 110 \times (\text{アミノ酸残基数}) \quad (3.13)$$

この関係式は、タンパク質は主にペプチド鎖でできていること、アミノ酸をつくる水素、炭素、窒素、酸素、硫黄といった原子の組成比率が巨大な分子で平均化されるためによく合う。

¹⁸ P. M. Gill, B. G. Johnson, J. A. Pople, Chem. Phys. Lett., 209 (1993) p. 506.

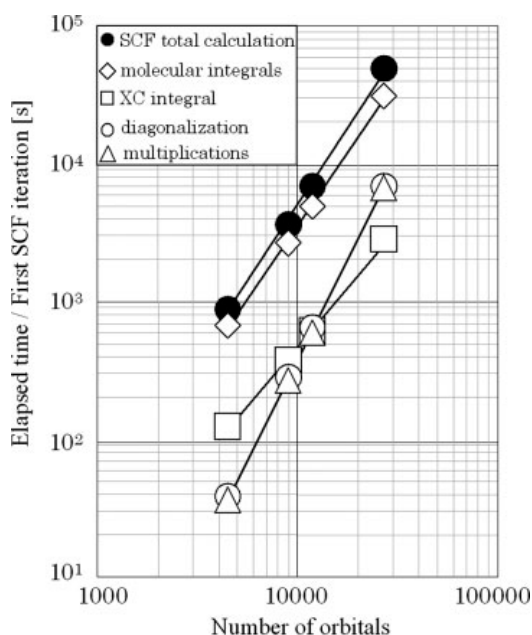
¹⁹ C. A. White, M. Head-Gordon, J. Chem. Phys., 101 (1994) p. 6593.

この経験則を発展させて、タンパク質の総アミノ酸残基数と原子数、電子数、さらには行列の次元との間の関係を導くことができる。この関係は100残基のタンパク質の場合を基準にすると次のようになる。

$$\begin{aligned} (\text{アミノ酸残基数}):(\text{原子数}):(\text{電子数}):(\text{行列次元数}) \\ = 100:1600:6200:10000 \end{aligned} \quad (3.14)$$

100残基中にある1600個の原子のうちおよそ800個が水素原子、残りの800がその他の原子である。行列の次元数は基底関数の総数である。この値は各原子に割り当てる基底関数の種類によって大幅に増減する。上の式の10000という数は、大型分子によく使用されるスプリット・バレンスと呼ばれるタイプの基底関数の場合のおおよその目安の値である。

(3.14)式は、タンパク質の量子化学計算を行う際に用いる大変役に立つ関係式である。100残基のタンパク質であれば1万次元、1000残基のタンパク質ならば10万次元の行列方程式を解くことになる。要素数は次元の2乗となるため、100残基では1億要素、1000残基では100億要素にもなる。このサイズの固有値問題をSCF計算によって何回も解かなければならないので、対象となるタンパク質の計算に必要な演算量や計算時間、I/O量の予測を欠かすことはできない。



第3図: タンパク質密度汎関数法の計算時間。

基底関数の総数とSCF繰り返し計算の1回目の実行時間の両対数プロット²⁰。タンパク質分子は小さいほうから、インスリン単量体、インスリン2量体、インターロイキン4、インスリン6量体。プロットは、SCF全体(●)、分子積分(◇)、交換相関計算(□)、行列対角化(○)および行列積(△)で、それぞれの直線の傾きは2.3、2.1、1.8、2.9および2.9である。計算にはSGI Altix3700の16CPUを使用した。

律速計算の計算量

Kohn-Sham-Roothaan法における律速計算である、分子積分、数値積分、行列対角化、行列積において、高速化を施した実際のタンパク質密度汎関数計算における分子サイズ依存性を見て

²⁰ T. Inaba, F. Sato, J. Comp. Chem., 28 (2007) p.984.

おこう。いろいろなサイズのタンパク質の密度汎関数計算を行い、その基底関数の総数 N_{bf} と SCF 繰り返し計算の 1 回目にかかった実時間を両対数プロットで示したものが図 3 である。サンプルはインスリン単量体、インスリン 2 量体、インターロイキン 4、インスリン 6 量体で、基底関数の総数が約 4000 から 30000 までの本格的な計算である。交換相関ポテンシャルには SVWN を、分子積分の高速化にはカットオフ法と RI 法²¹(説明は省略)を採用している。SCF 計算を、分子積分、交換相関項の数値積分、行列の対角化、行列積の計算に分け、それぞれ◇、□、○、△で示した。●は SCF 計算全体である。

各プロットを結んだ直線の傾きはそれぞれ 2.1, 1.8, 2.9, 2.9 および 2.3 である。この値は、計算にかかる時間が基底関数の総数 N_{bf} の何乗に依存しているかを示している。分子積分および交換相関項の計算はそれぞれ 2.1 乗と 1.8 乗で、カットオフ法による高速化の理想値である 2 乗に近い結果が得られている。一方、行列演算は行列積、行列対角化のいずれもほぼ基底関数の総数の 3 乗に依存している。基底関数の数が 30000 程度までであれば、行列演算はまだ SCF 全体の一部しか占めず、SCF 全体の計算サイズ依存性は $O(N_{bf}^{2.3})$ に抑えられている。

しかし、タンパク質のサイズがさらに大きくなると、行列演算が計算律速となることは明らかである。立体構造が明らかにされているタンパク質の 99%以上が 1000 残基以下である。1000 残基のタンパク質の基底関数の総数は 100000 程度である。このとき、分子積分と行列演算はほぼ同程度の割合となる。

律速計算の並列化

これまで見てきたように、Kohn-Sham 方程式の計算時間の上限は実質 $O(N_{bf}^3)$ である。仮にアミノ酸 1 残基の計算が 1 秒で終わる計算機があった場合、その計算機で 1000 残基のタンパク質正準分子軌道計算にかかる計算時間の上限値は 11574 日 (32 年) になる。個々の計算の最適化や計算の省略で多少高速化できたとしても、今のところ 1 CPU の計算機による計算は実行不可能である。そこで必要となるのが並列計算の技術である。密度汎関数計算の律速計算は、原理的には以下のように全て並列化できる(行列演算の説明は省略)。

計算律速となる 2 電子分子積分は密度汎関数法計算の中で、(2.12b) 式の Kohn-Sham 行列の生成、(3.3) 式の全エネルギー計算で使用される。これらの共通した表現が (3.5) 式であった。これから直ちに、インデックス rs を分割して計算を各計算機に振り分けることで \mathbf{U} の並列計算を行えることがわかる²²。この手法を取れば並列計算時に利用するメモリ量を削減でき、計算開始時のデータ転送量も抑えることができる。またアップデート法を採用することにより、並列計算中のデータ転送量をさらに抑えることができる。

pq のインデックスで計算を各計算機に振り分ける手法もある。インデックス rs で並列化を行うよりも簡単な制御で並列化できるが、(3.5) 式の場合、 \mathbf{P} の行列要素を全てのプロセッサで保持する必要がある。これに対して高島らは Hartree-Fock 計算において、カットオフ法を利

²¹ B. I. Dunlap, J. W. D. Connolly, J. R. Sabin, J. Chem. Phys., 71 (1979) p. 3396.

²² T. Yoshihiro, F. Sato, H. Kashiwagi, Chem. Phys. Lett., 346 (2001) p. 313.

用して \mathbf{P} 行列を圧縮する効果的な並列化を提案し、これを RT 法²³と名付けている。また、稲葉らにより RI 法の分子積分計算をプロセッサに均等分配する手段としてカットオフ法を利用する手法²⁰が提案されている。

数値積分は(2.17)式の交換相関項の積分計算に使用される。これらの被積分関数は原子核上で大きな値を持つため、(3.9)式のように原子核を中心とする部分空間の関数 $F_A(\mathbf{r})$ に分割し、その和として計算する。したがって、これらの数値積分は原子ごとに独立して計算ができるため、並列化は容易である。原子のインデックス A で計算をプロセッサに振り分ければよい。

このような並列計算の手法はあくまでサンプルである。並列計算を行う手法は多種多様なため、用いるアルゴリズムや計算機システムに最適な並列計算手法を選択することが望ましい。

表 1: インスリン 6 量体の SCF 計算 1 回転目での各計算ルーチンにかかった実時間と並列度、並列化効率²⁰

計算ルーチン	8CPUs		16CPUs		32CPUs		64CPUs			
	実時間 (秒)	実時間 (秒)	並列度	並列化 効率(%)	実時間 (秒)	並列度	並列化 効率(%)	実時間 (秒)	並列度	並列化 効率(%)
分子積分	59,353	31,247	15.20	94.97	15,923	29.82	93.19	8,671	54.76	85.56
数値積分	4,225	2,695	12.54	78.39	1,386	24.39	76.21	760	44.47	69.49
行列対角化	11,977	6,901	13.88	86.78	3,558	26.93	84.16	1,899	50.46	78.84
行列積	12,766	6,630	15.40	96.28	3,418	29.88	93.37	1,840	55.50	86.73
その他	3,553	2,350	12.10	75.60	2,207	12.88	40.25	756	37.60	58.75
合計	91,874	49,823	14.75	92.20	26,492	27.74	86.70	13,926	52.78	82.47

並列度: 1CPU の計算時間は 8CPU でかかった計算時間より概算した。

並列化効率 = (並列度 / CPU 数) × 100

表 1 に示すのはタンパク質密度汎関数計算の並列化効率である。計算方法は RI 法で、SVWN 交換相関汎関数を用いた。計算対象はインスリン 6 量体で、アミノ酸 306 残基、原子数 4728、電子数 18552、基底関数の数 26790 である(詳細は 4 章参照)。この計算を 1.3 GHz の Itanium 2 並列計算機のそれぞれ 8, 16, 32, 64 CPU を用いて並列計算を実行した。SCF 繰り返し計算の 1 回目で各種計算にかかる実時間を測定して並列化効率を求めた。

計測した計算はそれぞれ、分子積分、数値積分、行列対角化、行列積で、SCF 繰り返し計算中のディスク入出力時間などはすべてその他にまとめている。並列化アルゴリズムは、分子積分には pq インデックスによる計算分配を、交換相関項の数値計算には原子ごとの計算分配を適用し、行列対角化と行列積は数値演算ライブラリ ScaLAPACK を用いて並列化を行っている。この例では、1 CPU での計算時間が計測できないため、8 CPU での計算時間を 8 倍して 1 CPU の計算時間とし、並列度と並列化効率を求めている。

並列計算では使用する CPU 数が増大すると、Amdahl の法則により、並列度、並列化効率ともに飽和していく傾向が出てくる。この並列計算例でも CPU 数が増えるにつれて並列化効率が若干下がってきているが、SCF 繰り返し計算の合計時間で並列化効率が 80% 以上とおおむね良好な結果が得られている。この計算におけるプログラムの並列化率は 99% 以上と見積もられる。およそ 500 CPU 規模の並列計算機で効率を出すことができるレベルである。10000 CPU 規模の並列計算機に対応するためには、さらにこれらの律速ルーチンでデータ転送やディスク I/O な

²³ H. Takashima, S. Yamada, S. Obara, K. Kitamura, S. Inabata, N. Miyakawa, K. Tanabe, U. Nagashima, J. Comp. Chem., 23 (2002) p.1337.

どに工夫が必要になるだろう。

図2のように多種多数のルーチンを渡り歩く SCF 計算では、やむなく中間データをディスクに書き出すことが多く、これもデータ転送とディスク I/O が問題となる。タンパク質のような大規模計算では顕著である。量子化学計算において並列化効率を上げるためには、各律速ルーチンに対して計算機に適した計算アルゴリズムを採用しチューニングするだけでなく、 h, S, X 行列や SCF 計算ごとに更新される $P^{(\lambda)}, U^{(\lambda)}, F^{(\lambda)}, F'^{(\lambda)}, C^{(\lambda)}, C'^{(\lambda)}$ といった行列は、データ転送を極力抑えながら各律速ルーチンで利用しやすい形式でメモリに保持する必要がある。1000 残基タンパク質ではこれらの行列は1個で 10^{11} Byte にのぼるため、保存するには数 $10^{13} \sim 10^{15}$ Byte のメモリが必要となる。10000 CPU を超える計算機なら十分可能な値である。

4. タンパク質全電子計算に特有の技術

この章ではタンパク質のための Kohn-Sham-Roothaan 方程式の解法について解説する。タンパク質は巨大分子であるため低分子と同じ方法では解くことができない。SCF 計算が収束するための第1条件は良好な初期値を準備することであり、第2は滑らかに収束させることである。タンパク質は細胞の中の複雑な環境にあるので、その効果も取り込まなければならない。

4. 1. 計算シナリオと収束方法

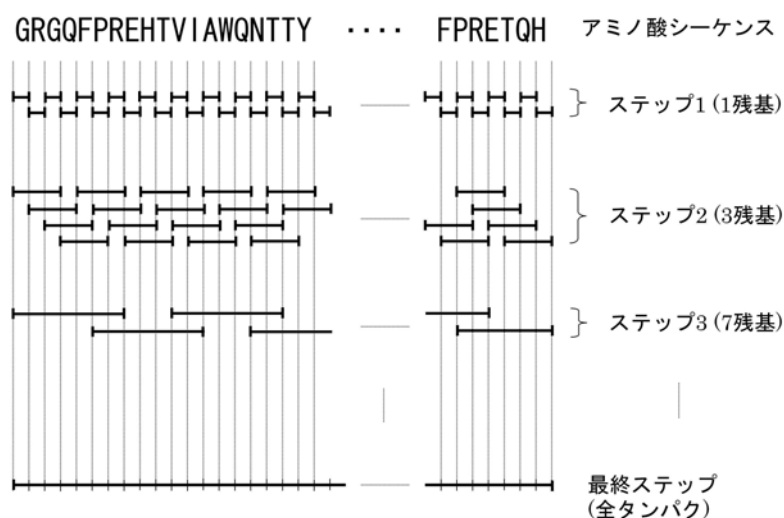
SCF 計算には初期値が必要である。この初期値が悪いと、収束するまでに何百回もの繰り返し計算を行わなければならないかたり、発散して解が得られなかったりする。実は小さな分子の場合でも難しい問題で、タンパク質の場合、大部分は多数のアミノ酸が結合した巨大分子であるためさらに収束が難しい。アミノ酸は多かれ少なかれ互いに似ているので、わずかな環境の変動で軌道成分の混じり方に影響を受ける。タンパク質は半導体的な性質を持っているので、場合によっては電子が遠くまで走ることもある。その上、密度汎関数法では Hartree-Fock 法よりも被占軌道と空軌道のエネルギーが接近しやすいため、これらの成分の入替えが起きる傾向があり、いっそう収束が難しくなる。このように難しい性格の計算を成功させるためには、タンパク質の構造とそれにとまなう性質を充分理解して注意深く計算をすすめるなければならない。

タンパク質の Kohn-Sham-Roothaan 方程式を解く難しさを克服するために、計算シナリオと呼ぶ新しい手続きを導入した¹⁾。初めからタンパク質全体を解こうとはしないで、アミノ酸を1つ1つ解くことから始め、計算する分子を徐々に大きくしていく。図4がその計算シナリオを示したものである。まず、タンパク質のペプチド鎖を形成している個々のアミノ酸残基をそのままの構造で取り出し、窒素には水素を補ってアミノ基としカルボニル炭素には水酸基を付けてカルボン酸にする。原子ごとの電子密度からこのアミノ酸の電子密度を推定して Kohn-Sham-Roothaan 方程式の解を求める。これがステップ1である。

ステップ2では、連続する3個のアミノ酸の解から電子密度を取り出し結合して3残基ペプチドの初期値を作る。この際、付加したアミノ基の水素とカルボン酸の水酸基など不要部分の電子密度は取り除く。この作業はちょうどペプチド結合の生成に当たる。電子密度の推定値が得られたら SCF 計算を行って解を求める。

ステップ3では3残基ペプチドの中央のアミノ酸残基の電子密度のみを取り出し、数残基ペプチドの初期値を作る。このために隣り合う3残基ペプチドは2個のアミノ酸残基が重なるようにしておく。中央のアミノ酸残基の電子密度を両側のアミノ酸残基の影響を取り込んだもの

にするための手続きである。この取り込みは後続のステップでも行う。ペプチドの両端については端のアミノ酸残基の解も取り入れる。この手順を、例えば7残基ペプチド、20残基ペプチドというように次第にペプチドを長くしていき、最後にタンパク質全体のSCF計算を行う。



第4図: タンパク質の計算シナリオ[1]。

タンパク質のSCF計算を滑らかに収束させるための計算手順。図の最上部はアミノ酸の略号である。途中のH型の記号はそれぞれのステップで計算する分子を示す。ステップ1では個々のアミノ酸を、ステップ2では3残基ペプチドをというように徐々に大きくしていき、最後に全タンパク質の計算を行う。

この種の計算にはいくつか難しい問題点がある。前のステップの電子密度はSCF計算の過程でそのステップの分子全体として成り立つように最適化されたものである。このため電子密度の単純な切り貼りは誤差を生み出し、次のステップの良好な初期値を与える保証はない。そこで計算した電子密度の精度を落とさないようにペプチドを長くしていく方法として開発したのがQCLO法²⁴である。

QCLOとは擬カノニカル局在化軌道 Quasi-Canonical Localized Orbital の略称で、1種の局在化分子軌道である。Kohn-Sham 演算子はユニタリ変換によって方程式は不変である。言い換えると、カノニカル分子軌道もそれをユニタリ変換した軌道も表現が違っただけで実質的には同じ解である。適切な基準を選ぶと、カノニカル分子軌道をユニタリ変換して、化学結合や孤立電子対などの局在化軌道を作ることができる。ここで、図4のシナリオの各ステップで計算する分子のことをフレーム分子、分割した軌道空間のことをフラグメントと呼ぶことにする。ステップ2以降はQCLOの計算を行い、これを使って次のステップのフレーム分子の初期値を求める。

フラグメントの分割は結合や孤立電子対などの境で行えば任意であるが、一般にN端-側鎖-ペプチド結合-側鎖-ペプチド結合-側鎖-...-C端のフラグメントに分割し、側鎖のフラグメントに関しては α 炭素を境界に C_α -H結合、 C_α の1sコアおよび側鎖のすべての結合や孤立電子対はこの側鎖のフラグメントに所属させ、N- C_α 結合は左側のペプチドのフラグメントに、 C_α -C結合は右側のペプチド結合にそれぞれ帰属させると一意に分割できる。

²⁴ H. Kashiwagi, H. Iwai, K. Tokieda, M. Era, T. Sumita, T. Yoshihiro, F. Sato, Mol. Phys., 101 (2003) p. 81.

フレーム分子のカノニカル分子軌道が得られたら、まず通常の局在化軌道を求める。1つのフラグメントのQCLOを求めるには、そのフラグメントに所属する局在化軌道を選び出し係数行列 \mathbf{C}_{lo} を作る。この行列の縦の次元はフレーム分子の基底関数の数であり、横の次元はそのフラグメントに所属する局在化軌道の数である。 \mathbf{C}_{lo} を用いて次式のようにフレーム分子全体のKohn-Sham 行列 \mathbf{F} を線形変換すると、そのフラグメントのKohn-Sham 行列 \mathbf{F}' が得られる。

$$\mathbf{C}_{lo}^\dagger \mathbf{F} \mathbf{C}_{lo} = \mathbf{F}' \quad (4.1)$$

この行列の次元はフラグメントに所属する局在化軌道の数になる。これを対角化するとQCLOが得られる。

$$\mathbf{F}' \mathbf{C}' = \mathbf{C}' \mathbf{E} \quad (4.2)$$

1つのフラグメントのQCLOは(4.1)式と(4.2)式を解いて得られる軌道として定義される。フラグメントの局在化軌道のみで展開しているから局在化軌道であり、(4.2)式のカノニカルな解であるから、そのフラグメントのカノニカル軌道としての性格を兼ね備えている。この性質から擬カノニカル局在化軌道と名づけた。

次のステップ3以降では2つの計算方式がある。1つは、フレーム分子の局在化軌道を求めなおして(4.1)式の基底として用いる方法がある。もう1つは基底として前のステップで求めたQCLOを使う方法である。前のステップの小さなフラグメントのQCLOを用いて \mathbf{C}_{lo} を作り、大きなフラグメントのQCLOを計算する。前者の方が精度はよいが計算労力が大きい²⁵。

計算シナリオの途中で得られるQCLOは分割したフレーム分子の解であり、またフラグメントごとに基底関数を限定したものである。したがって、全分子のカノニカル軌道をユニタリ変換して得られる局在化分子軌道とは異なる。しかし、各ステップのQCLOは全分子の局在化軌道のかなりよい近似になっていると考えられる。

QCLOをSCF計算の初期値作成に用いるためには次のような手順をふむ。アミノ酸1個ずつの計算から3残基ペプチドのSCF計算までは前に述べた手順と同じである。ステップ3の数残基ペプチドの電子密度の初期値をステップ2のQCLOから求める。切断と接続はフラグメント単位で行う。以下のステップも同様である。この効果は5.1節で明らかになる。

4. 2. 計算構造

計算に用いる分子構造の決定は大変重要である。化学分子の場合、基準となる分子構造は実験で得られた構造以外にも、典型的な結合長や結合角、二面角を与えて組み上げた模型構造を採用したり、シミュレーションで用いるKohn-Sham方程式のもとでエネルギー最小構造を求めたりする。これらの構造で分子の持つ基本的な性質を理解することに成功してきた。

タンパク質も分子であるから同じことが期待できるが、タンパク質には基準となるあるいはシミュレーションに適切な構造を決定する際に、特に考慮に入れなければならない問題がある。以下に簡単に整理してみよう[2]。

1. 実験構造

タンパク質立体構造は比較的容易にX線構造解析や多次元NMR法などの実験で求めることができるようになった。多くの構造はPDB²⁶に登録されている。これをシミュレーション構造に

²⁵ N. Nishino-Uemura, T. Hirano, F. Sato, J. Chem. Phys. 127 (2007) p.184106.

²⁶ <http://www.rcsb.org/pdb/home/home.do>

採用することが最もよく行われている。しかし、X線構造解析では一般に水素の位置が見えないし、窮屈で位相の揃った結晶場中の構造である。NMR法にしても、原子間距離から構造推定するためモデルがたくさんできてしまう。両者に共通で、原子核位置を決定するリファインメントには経験的な力場を用いたモデル作りの過程が入り込む。

2. 系の取り扱い

化学分子では孤立系のシミュレーションが標準的に行われるが、タンパク質の性質や機能を調べるのに適するかどうかは疑わしい。少なくとも古典法では溶媒を考慮に入れる。ただし、その古典法にしても完璧ではない。例えば、0.05 M 緩衝液の中に 5 μM のタンパク質が入っているとすると。水は 55 M だから、タンパク質とイオンと水の分子数比は 1:20000:10000000 である。このような環境をシミュレーションで作出すことは難しい。溶媒分子の取り扱いもモデル化が必要となる。タンパク質の表面をどう扱うかも悩ましい問題である。

3. 最適化

タンパク質は最安定構造を求めることも困難である。取り扱う原子数を N_{atom} とすると、 $(3N_{atom} - 6)$ 次元空間の最適化問題となるからである。タンパク質は数千~数十万原子からなる分子であるし、もしこれに溶媒分子を含むならば N_{atom} に水分子とイオンを加えなければならない。当面は N_{atom} を小さく抑える局所構造最適化が現実的であるが、その手法にしても注意が必要である。タンパク質に水素を付けてこれを最適化することを考えてみよう。よく活性中心となるセリン、トレオニン、チロシン、システインの OH 基や SH 基の水素の位置は、他の原子が邪魔をして 2 面角が 360° 回れず、最初に付加した位置に依存してしまうことが多々ある。ヒスチジンのイミダゾール環は ϵ 位と δ 位に窒素がある。他に情報が無ければ、どちらに水素が付くか分からないし、両方に結合してカチオンとなることもある。適切なモデルでないとタンパク質の反応機構を間違えて推定してしまうかもしれない。

4. 量子計算固有の問題

PDB に登録されている立体構造は、化学的な見地からは異常な構造歪みや原子間衝突などを残したままリファインメントが行われているケースが多々見られる。水素付加を行うと、しばしば原子間衝突が起こってしまう。量子化学計算結果はこのような構造に敏感で、反応性の高い軌道を与える。タンパク質は N 末端、C 末端以外にも多くの解離基を持っている。孤立系でタンパク質の密度汎関数計算を収束させることはほぼ絶望的である。

結局、タンパク質の計算構造は一意には決まらない。上記の問題点を考慮しつつ、シミュレーションによって明らかにしたい内容に基づき、適切な構造を決定すべきである。これこそがタンパク質シミュレーションの最も重要なプロセスの 1 つなのである。

これまでの経験で著者らが開発した、現在の計算機環境でできるモデリング法を「おまけ」として紹介しておこう [2]。タンパク質密度汎関数計算を比較的安全に達成できる計算構造を構築するテクニカルな話であり、決定版でもない。

本来、分子の立体構造は、シミュレーションを行う支配方程式、すなわち本ケースならば Kohn-Sham 方程式に基づいて最適化されることが望ましい。将来は実現されるであろうが、今のところ構造最適化は古典分子力場を用いて行うのが現実的である。精度は劣るが、立体構造の歪みや原子間距離の異常の修正に大いに役に立つ。化学分子でも、第 1 原理でなんなく構造最適化できるようになったのはごく最近の話で、古典モデリングは標準的に行われていた。

構造最適化には、分子動力学 (MD) 法やモンテカルロ (MC) 法といった系の状態をシミュレーシ

オンする手法と、最急降下法や共役勾配法といったエネルギーの極小値を解く手法がある。前者は大域的な構造探索がなされるため、立体構造の歪みを修正するばかりでなく初期構造を大きく変化させてしまうことがある。後者は初期構造近傍のエネルギー極小点へ収束するため、初期構造はそれほど大きくは変化しない。実験データがあり、これになるべく近い構造で立体構造の歪みや局所構造異常を修正したい場合は、後者を選ぶ方が無難かもしれない。またこのとき、溶媒に連続近似を選択することも一考である。

それでも、タンパク質の構造最適化はなかなか収束しないことが多い。膨大な自由度が原因である。ただし、古典分子力場での構造最適化であるから、不必要なまでに精度を高め、厳しい収束判定を課す必要は無い。適切な結果に至らない場合は、その都度原因を解析し、立体構造を修正する。その妥当性は最終的には計算者の判断に委ねられる。

表面の取り扱いには様々なモデルが考えられるが、その1つに解離基の周りにこれと反対の電荷を持ったイオンを近づける緩衝液モデルがある。実際には、これらのイオンは水分子同様、動き回っていると考えられるが、溶媒が近づける表面にある解離基の近くにカウンターイオンを付加するモデルで比較的安定に密度汎関数計算を達成できる。

タンパク質を取り巻く水は極性を持った液体の分子としては特殊なものである。当然ながら、タンパク質に影響を与えている。溶媒としての水分子のモデルは様々ある。もしあらわに扱うのであれば、その構造最適化にもタンパク質と同等のモデルを採用し、溶媒水分子の構造を再現させることが望ましい。また、タンパク質の特定の部位に取り込まれて動きが鈍い水分子がある。これらは少なくともタンパク質の一部としてあらわに扱うべきである。

これらの方法で決定された構造によるタンパク質の全電子計算では、この系全てを量子系として扱うこともあるし、量子系はタンパク質のみでイオンや水分子は点電荷として取り扱うこともある。このような選択もシミュレーションの目的次第である。とはいえ、溶媒分子や細胞膜などの環境まで量子化学計算を行うのは大きな負担である。複数のサブユニットから構成される大型のタンパク質の場合には、量子化学計算を活性中心を含むサブユニットに限定せざるを得ないこともある。このような場合に用いるのがQM/MM法である。

一般的なQM/MM法はその名称が示すとおり、全系を量子化学的に取り扱う部分と古典的な分子力場による部分に分割して計算する方法で、量子化学計算の計算労力が大きいため考え出されたものである。基本的なアイデアは、計算機の能力が低分子を計算するのにも不十分であった時代に提唱されたものである。1976年にはWarshelとLevittによるタンパク質の先駆的な計算²⁷が行われ、今でもよく利用されている。

QM/MM領域の切り分け方は原理的には任意であるが、量子系領域にタンパク質全体、古典系領域に溶媒を指定すれば、比較的うまく行くだろう。ここでは、タンパク質の全電子計算を行う際の周囲の水の効果を取り込む便宜的な取り扱いとして、タンパク質のまわりの水分子を点電荷として扱う方法を紹介する。水の水素原子と酸素原子はそれぞれいくらかの正、負の電荷を帯びている。そこで、水分子の水素、酸素に点電荷を割り当てるのである。この水のモデルはタンパク質に静電場を与えるのみである。一般に点電荷と点双極子により任意の場のポテンシャルは表現できる²⁷ため、溶媒の水が外場としてタンパク質に与える効果を記述できる。

この静電場は、タンパク質の全電子計算にはKohn-Sham行列の1電子項(2.13)式の第2項である原子核からのポテンシャル・エネルギーの形で、外場として点電荷を追加すればよい。

²⁷ H. Nakamura, J. Mol. Graphics, 11 (1993) p. 30.

$$h_{pq} = \int g_p^*(\mathbf{r}) \left[-\frac{1}{2} \Delta - \sum_A^{Natom} \frac{Z_A}{|\mathbf{r} - \mathbf{R}_A|} + \sum_B^{Next} \frac{q_B}{|\mathbf{r} - \mathbf{R}_B|} \right] g_q(\mathbf{r}) d\mathbf{r} \quad (4.3)$$

第 3 項が外場で、 B は外場のすべての点電荷を走り、 q_B はその電荷である。この式には分子軌道の係数も電子密度も含まれていないので、SCF の繰り返しで再計算の必要がない。追加された 1 回分の積分計算のコストが増えるのみで、経済的に溶媒効果を取り込むことができる。

4. 3. 収束法

タンパク質など、巨大分子の SCF 計算で使用される収束法について簡単にコメントしておこう。紙面の都合から文献を挙げる程度にとどめる。これまで述べてきたように、SCF 計算が成功する鍵は適切な計算シナリオと良い初期値の作成である。それに加えて、収束法は SCF 計算の過程を制御し計算の達成を補助する技術である。収束法には大きく分けて 2 つの役目がある。1 つは SCF 計算を安全に収束へと導くことであり、もう 1 つは収束を加速する役目である。プロジェクション法²⁸、軌道の重なり法、レベルシフト法²⁹やシンプルミキシング法はどちらかといえば前者の方法で、Anderson 法³⁰や DIIS 法³¹は後者の性格が強い。それぞれ特徴があるので適切な場面で用いると効果を示す。

SCF 計算を安全に収束へと導く技術は特にタンパク質のような巨大分子で欠かすことができない。やさしく注意深く収束解に導かないかぎり SCF 計算は振動したり発散したりしてしまう。

分子軌道計算では、あらかじめ計算する分子の電子数と同時に電子の軌道への詰め方を決めておかなければならない。閉殻の基底状態では、Pauli の排他原理に従って、普通は、エネルギーの低い軌道から順に 2 つずつ電子を詰めていく。それから密度行列を計算し、SCF 計算が繰り返される。ところが、タンパク質の軌道エネルギーは固体のバンドのように密に分布し、被占軌道と空軌道とのギャップが狭い。その実測値は実験方法や測定条件によって値が大きく変わるが、おおよそ 3~7 eV と見積もられている^{32,33}。これを Hartree-Fock 法で計算すると、しばしば 10 eV 以上のギャップが開いて³⁴絶縁体となってしまう。Kohn-Sham 法、特に局所密度法ではかえって 1~2 eV と狭くなる^{1,35}。SCF 計算の始めでは、被占軌道と空軌道とのギャップがより狭くなり、密度汎関数法ではしばしば軌道の順番も入れ替わる。このような症状は SCF 計算の初期値が悪いほど顕著となる。開殻系の計算では、閉殻と空軌道の狭いギャップの間に開殻軌道がくるので、ますます収束が難しくなる。このように軌道エネルギーが密に分布している場合に正しく電子を詰める方法が、SCF n 回転目の閉殻・開殻・空軌道がどの程度 $n-1$ 回目の軌道の空間に含まれているかを示す指標を計算するプロジェクション法、今回の分子軌道と前回の軌道の重なり大きさで対応関係を判断する軌道の重なり法、特定の軌道について望みの値だけ軌道エネルギーをずらすレベルシフト法や、電子状態の急激な乱れを防ぐため今回

²⁸ 岡崎功, 九州工業大学情報工学部博士論文(2000).

²⁹ R. Carbo, J. A. Hernandez, F. Sanz, Chem. Phys. Lett., 47 (1980) p. 581.

³⁰ D. G. Anderson, J. Assoc. Comp. Machinery, 12 (1965) p. 547.

³¹ P. Pulay, Chem. Phys. Lett., 73 (1980) p. 393.

³² D. D. Eley, D. I. Spivey, Trans. Faraday Soc., 56 (1960) p. 1432.

³³ J. Ladik, F. Bogar, B. Penke, Int. J. Quant. Chem., 98 (2004) p. 522.

³⁴ S. Suhai, J. Kaspar, J. Ladik, Int. J. Quant. Chem., 17 (1980) p. 995.

³⁵ T. Inaba, N. Tsunekawa, T. Hirano, T. Yoshihiro, H. Kashiwagi, F. Sato, Chem. Phys. Lett., 434 (2007) p. 331.

の SCF 計算で求めた物理量を前回の物理量で薄めるシンプルミキシング法である。

一方、SCF 計算の繰り返し回数が 100 回、200 回かかっても正しく収束するならば、計算結果の議論には問題はない。しかし、巨大分子の量子計算はコストが高く時間がかかる。混ぜ合わせ率を状況に合わせて変化させる動的なミキシング法である Anderson 法や、過去の物理量を使って残差を最小にする外挿を行う DIIS 法などの収束加速はそのまま計算時間の短縮、しいては計算の実現に貢献する。したがって、収束加速法は実際の計算達成には欠かせない技術である。

適用の仕方は様々で、シンプルミキシング法、Anderson 法、DIIS 法は同時に適用できないが、軌道の重なり法やプロジェクション法、レベルシフト法はこれらと組み合わせることができる。

5. 将来展望：タンパク質反応シミュレーション

5. 1. タンパク質全電子計算の現状

1997 年の Challacombe と Schwegler の先駆的な P53 の Hartree-Fock 法による計算 (3836 軌道)³⁶に端を発し、2000 年には著者らのヘムタンパク質シトクロム *c* の密度汎関数計算 (9600 軌道)¹と高田らの PKC の Hartree-Fock 計算 (8672 軌道)⁵が達成された。本格的な生体分子量子シミュレーションの歴史はまだほんの 10 年である。

この節では、これまで解説してきたいろいろな技法を使ったインスリン 6 量体の密度汎関数法計算^{20,35}を紹介する。その計算規模は 3 章で示したが、執筆時点で世界最大のカノニカル分子軌道計算であり、本稿が展望する将来のタンパク質計算への一里塚である。

インスリンは血液中の糖분을制御するタンパク質ホルモンである。インスリン単量体は 21 残基の A 鎖と 30 残基の B 鎖がジスフィルド結合でつながった小型のタンパク質である。単量体の表面には解離基が多く、凝集して 6 量体になる傾向が強い。糖尿病のための製剤は高濃度のため 6 量体が形成された状態で投与されるが、ホルモンとして働くのはインスリン単量体である。体内で単量体に分解する時間が長くなり効き目がすぐに現れないことが問題になっていて、アミノ酸残基の置換などにより分解しやすい製剤も開発されている。

実は、インスリン 6 量体の計算は稲葉らによって 2005 年に成功していた。その後、この巨大系で安定した解法を探るべく、様々な技法が考案されテストされた。その結果、最終的に採用されたのは以下で説明する方法である。まず、分子構造についての予備計算を行った。NMR 法で得られた分子構造を用いているが、データベースの構造には局所的な歪みがあって量子化学計算に耐えられない。また水溶液中にあるタンパク質の計算をするために、水和イオンや溶媒の配置も重要である。この過程で 4.2 節のような手順が開発された。

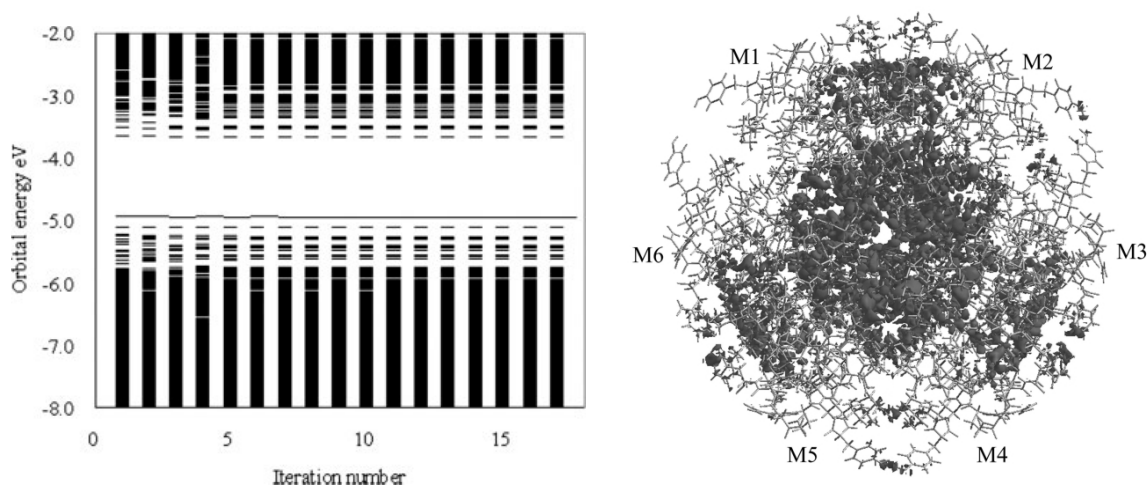
全電子計算は、4.1 節の QCL0 法を採用するとともに、各単量体の全体計算が終わった段階で、他の単量体の原子および周辺の水分子を点電荷として計算に取り込むことが安全な収束へのもう 1 つのポイントとなった。最終ステップの全電子計算の収束加速には Anderson 法が有効であることを発見し、これを用いて SCF 繰り返し計算 17 回で収束した。1.3 GHz の Itanium 2 プロセッサ 64 台 (理論ピーク 333 GFLOPS) で 65 時間であった。

図 5 にインスリン 6 量体の最終ステップの軌道エネルギー変化を示した。SCF 計算の最初の 3 ~ 4 回にわずかな変化が見られるが、それ以降の軌道エネルギーのパターンはほとんど変化していない。良好な初期値と収束制御によりスムーズに収束に向かっている様子が見られる。バ

³⁶ M. Challacombe, E. Schwegler, J. Chem. Phys., 106 (1997) p. 5526.

バンドギャップも約 1.3 eV で安定していて、被占軌道が空軌道で置き換わるような気配は一切ない。全エネルギーの初期値と収束値の差は 0.15 au であった。

単量体で計算を止めても電子密度は十分よいように思えるが、実際は広く電子分布が変化する。図 6 は単量体の計算と 6 量体の計算結果における電子密度分布の差を示したものである。このような結果から、1 章で述べたように、タンパク質は凝集などの摂動に対し、広範囲にわたって少しずつ電子分布を変えることも立証された。



第 5 図: インスリン 6 量体の最終ステップの軌道エネルギー変化²⁰(左)。

縦軸は軌道エネルギー，横軸は SCF 計算の繰り返し回数。下部が占有軌道のエネルギーで，一番上のつながっている横線は最高被占軌道のエネルギー，その上は空軌道のエネルギーである。

第 6 図: インスリンの 6 つの単量体と 6 量体の電子密度分布差(右)。

M1, M2 などは単量体である。6 つの単量体と 6 量体との電子密度分布の差を図示した。単量体ごとの計算では他の単量体の原子にはマリケン電荷を置いている。濃淡の差はプラス，マイナスを表す。

本稿で解説したタンパク質の Kohn-Sham-Roothaan 方程式を解くための技法を用いて、すでに 100 種類のタンパク質の計算が行われている。さらに改良が進めば、タンパク質波動関数計算の完全自動化も夢ではないだろう。

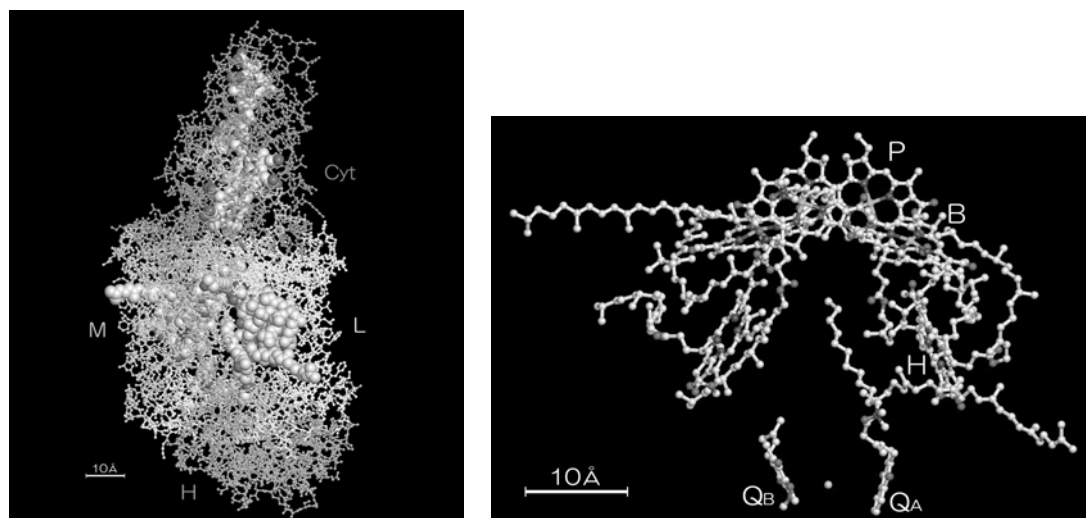
6. 2. タンパク質反応シミュレーション

最後に、今後飛躍的な発展が期待されるタンパク質密度汎関数法による反応シミュレーションの将来を展望する。

タンパク質で特徴的な反応に分子内電子移動や酵素による触媒反応がある。ミトコンドリア内膜の呼吸鎖では、タンパク質複合体が整然と並配置されている。これらのタンパク質はアミノ酸残基以外の補欠分子族を豊富に持っていて、電子はタンパク質内でも 10 Å を超える長距離を移動し、さらに複合体から複合体へと電子を伝達し、最終的には酸素を水に変換してエネルギー物質の生産を行う。高度に組織化された、無駄のない生体のエネルギー生産工場である。

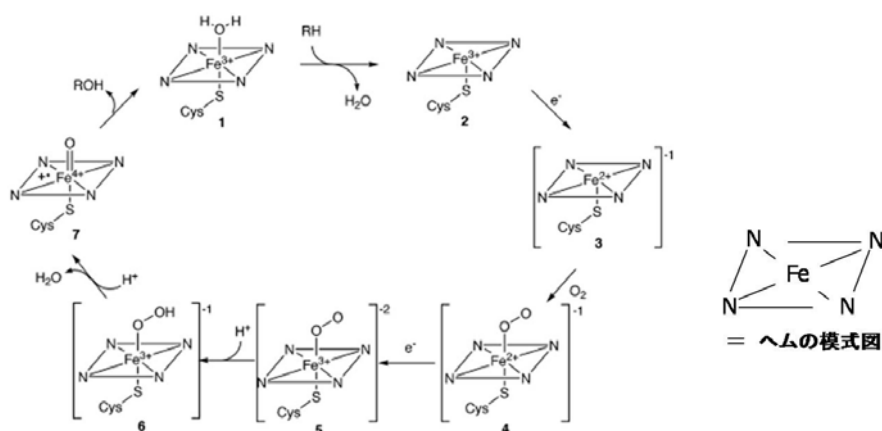
光合成システムは生物界の大部分のエネルギー生産を行っている重要なエネルギー工場である。図 7 は、Deisenhofer らによって初めて明らかにされた紅色光合成細菌の光合成反応中心タンパク質の構造である。様々な色素分子を持ち、高度に組織化されている点で呼吸鎖と同じ

である。2つのクロロフィルがスペシャルペアと呼ばれる2量体を構成している。これが集光を受け持つアンテナタンパク質からの光エネルギーを受けて励起し、電荷分離反応の基点となる。色素群はほぼ2回対称的に配置されているが、電子は図の向かって右側の経路のみを移動する。電子はスペシャルペアからサテライト・クロロフィルを経てフェオフィチンに移り、さらにその下のキノンに移動する。光合成反応中心は光のエネルギーを高い効率で化学エネルギーに変換する初期過程を担う精密分子である。



第7図：紅色細菌の光合成反応中心タンパク質の構造(左)とM, L, H鎖中の色素(右) [1]³⁷

(左)紅色細菌 *Rhodospseudomonas viridis* の光合成反応中心タンパク質。太陽光からのエネルギーによって電荷分離を起こし、太陽電池のような働きをする。シトクロム鎖 Cyt, M鎖, L鎖, H鎖と呼ばれる4つのサブユニットからできていて、スペースフィルモデルで表された多数のヘテロ分子を持っている。(右) M, L, Hサブユニットの色素, P: スペシャルペア (バクテリオクロフィル2量体), B: サテライト・バクテリオクロフィル, H: バクテリオフェオフィチン, QA: メナキノン, QB: ユビキノン。Pが励起状態に上がると電子が飛び出し、0.65 psで右側のBへ移動, そこから3.5 psでHへ, 200 psでQAに, 100 μsでQBに移動する。



第8図：へムタンパク質P450の1原子付加反応サイクル。

³⁷ J. Deisenhofer, O. Epp, K. Miki, R. Huber, H. Michel, J. Mol. Biol., 180 (1984) p. 385.

酵素は化学反応を生体温度で効率よく進行させる触媒である。酵素反応では基質と酵素が鍵と鍵穴の関係にある。形状の整合性は酵素反応の始まりを左右するが、化学反応の第1ステップにすぎない。形が似ていても酵素反応を停めてしまうのが阻害剤である。酵素反応が促進されるか抑制されるのか、またその効率ほどの程度かを解析するためには化学反応シミュレーションが必要である。酵素タンパク質は基質、水分子などを取り込んで特有の反応を進める。図8はヘムタンパク質 P450 が行う 1 原子付加反応のサイクルを示している。全体としてみれば



という反応であるが、実に様々な反応中間体を經由していることが分かる。これはわずかな活性化エネルギーで効率よく反応を進めるための生体の基本戦略である。

タンパク質反応シミュレーションにとっては反応そのものが的確に記述できることが大切である。電子移動や化学反応は電子が関わるため記述には量子化学計算が欠かせない。kcal/mol の活性化エネルギーで進行するタンパク質の反応を定量的にシミュレーションするためには、電子相関効果の取り込みが欠かせない。数千～数万原子からなる巨大分子の量子化学計算と高い計算精度という、相反する要求を両立させる方法として、本稿ではタンパク質の密度汎関数法を解説してきた。その最終的な目的はタンパク質反応の定量的な解析である。

以下に2つのターゲットについて簡単に説明しよう。

電子移動シミュレーション

電子移動とはある2つの系の間で電子が移動する現象である。豊富で詳細な実験をもとに、Marcus らの研究³⁸に端を発し、電子移動独自の理論体系が編み出されている。時間が関与する反応の記述には、時間に依存する Schrödinger 方程式の解法が必要である。その近似解法として摂動法により、電子移動反応速度 k_{ET} は、電子が単位時間あたりに始状態から終状態へと遷移する確率として表現される。これは Fermi の黄金律として知られており、

$$k_{ET} = \frac{2\pi}{\hbar} |V|^2 (FC) \quad (6.2)$$

と書かれる。ここで、 \hbar は Planck の定数を 2π で割った定数、 V は電子カップリング行列要素と呼ばれ、電子の波動関数から計算される項である。 (FC) は Frank-Condon 因子と呼ばれ、原子核の波動関数などから計算される。

反応速度が非常に速い電子移動反応では Born-Oppenheimer 近似よりもさらに近似を進めた原子核の波動関数と電子波動関数を切り離す Condon 近似を採用することによって、(6.2) 式のように電子移動反応速度 k_{ET} は $|V|^2$ と (FC) の積で表されるため、速度因子を別々に議論することができる。タンパク質の量子計算による本格的な (FC) の見積もりはしばらく待たねばならないだろうが、古典的分子動力学計算による知見は集まりつつある。上記の光合成反応中心タンパク質の (P^*-H) から (P^+-H) への電荷分離反応において、右側周りの経路では最高速に移動できる配置を持っており、左側周り経路ではこの条件から外れているらしいこと。さらに、遅く見積もられすぎる $(\text{P}^+-\text{H}) \rightarrow (\text{P}-\text{H})$ の電荷再結合反応速度が、構造変化を加味すれば実験値に近づくといった報告³⁹がある。シミュレーションには仮定が多いが、タンパク質の構造や構造変

³⁸ R. A. Marcus, N. Sutin, Biochim. Biophys. Acta, 811 (1985) p.265.

³⁹ M. Massimo, J. N. Gehlen, D. Chandler, M. D. Newton, J. Am. Chem. Soc., 115 (1993)

化に基づく因子の究明が可能になりつつある。

タンパク質の電子移動反応の電子による因子 V に関しては、これまでのところタンパク質波動関数を直接用いた計算は行われていない。とはいえタンパク質の密度汎関数計算から精密な V の値が見積もられるようになるのはもうすぐである。様々な計算法の諸考察については、これを先導的に理論構築してきた Newton の総説⁴⁰を参考にするとよい。

電子波動関数による直接計算法の考え方は簡単である。光合成反応中心タンパク質の光励起による電子移動過程の電子カップリング要素 V は z 方向の振動電場 $F\sin(\omega t)$ と電子の双極子モーメント $(-e\mathbf{r})$ ($\mathbf{r} = (x, y, z)$) との相互作用から以下と表される。

$$V = \int \Psi_f \frac{zF}{2} \Psi_i d\mathbf{r} \quad (6.3)$$

すなわち原理的に始状態波動関数 $\Psi_i(P^* - H)$ と終状態波動関数 $\Psi_f(P^+ - H)$ の電子波動関数が精度良く見積もられればよいことがわかる。これらの波動関数の計算に配置間相互作用法を用い、電子カップリング要素を計算するのが分子軌道計算の正攻法である。

現在のところ、配置間相互作用計算法の拡張である SAC-CI 法を用いた長谷川・中辻らの計算⁴¹が先駆的な例である。光合成反応中心タンパク質における、クロロフィル 2 量体-フェオフィチン、続くフェオフィチン-キノン間における色素間で電子カップリング行列要素を計算し、光合成色素の励起構造に関する有用な知見を与えた。ところで、タンパク質の軌道は図 1 のシトクロム c のように、主に色素の成分を持つ軌道でもタンパク質全域に広がる。タンパク質電子移動の解明にはこの広がりや考慮に入れることが重要である。タンパク質の励起波動関数計算では、1 電子励起の記述が有効な密度汎関数法-配置間相互作用 (DFT-CIS) 法や Runge-Gross の定理から発展した時間依存密度汎関数 (TDDFT) 法が有力な候補である。タンパク質の軌道を考慮に入れた計算法は、光合成反応中心タンパク質における (P^*-H) から (P^+-H) への電荷分離反応がサテライト・クロロフィルを経由する理由の解明に欠かせないだろう。

しかし、タンパク質のすべての軌道における励起を考えると 1 電子励起ですら、ハミルトニアン行列の次元は天文学的な数になってしまう。タンパク質を、例えばクロロフィル 2 量体を含む空間やフェオフィチンを囲む空間などのいくつかの軌道空間に分割する必要があるだろう。これらの空間の波動関数の記述には 4 章で説明した QClO を用いるとよい。QClO は分子のある部分に局在し、その部分のカノニカル分子軌道の性格を持つ直交局在化軌道であった。このように分割しておくとならば必要な軌道の数を絞ることができ、計算量が大幅に減る。空間内の CIS 計算をすれば励起エネルギーや振動子強度が求まり、QClO の直交性を利用して各空間の間の励起計算をすれば電子移動の確率を求めることができる。このような手法を用いて、近い将来、本格的なタンパク質電子移動シミュレーションが行われるようになるだろう。

化学反応シミュレーション

化学反応は、物質がそれ自身か他の物質と作用し合って、元の物質とは異なる物質を作り出す現象である。分子の構造と作用する過程で理論づけられる。そのため、タンパク質の化学反

p. 4178.

⁴⁰ M. D. Newton, Chem. Rev., 91 (1991) p. 767.

⁴¹ J. Hasegawa, H. Nakatsuji, J. Phys. Chem., B102 (1998) p. 10420.

応シミュレーションには、溶媒、温度の効果などを考慮に入れながら、化学結合の組み換えなどの分子の電子状態変化を精度よく追跡できる計算方法が必要になる。1964年にHohenbergとKohnが密度汎関数原理を発表した⁹が、これは温度0における理論である。その翌年、Merminが有限温度への拡張を発表した⁴²。これにより、密度汎関数法—分子動力学(DFT-MD)法、密度汎関数法—モンテカルロ(DFT-MC)法が有力な計算方法となることが保証された。

その目的は、反応系から遷移状態を通して生成系に至る反応経路(反応座標)の解析である。どのように分子が構造変化し、どのように結合が組み変わるのかを追跡する反応機構モデル解析の出発点は、遷移状態の存在を仮定した遷移状態理論である。化学反応速度の差異は活性化自由エネルギー ΔF を通して現れるため ΔF が最も重要な物理量である。 ΔF の正確な見積もりにには、電子相関を取り込んだ密度汎関数法が役立つ。

タンパク質反応の ΔF および反応経路の求め方の基本的な原理は一般的な化学分子の方法と同じである。電子移動反応に比べて格段に遅い化学反応にはBorn-Oppenheimer近似が適用限界である。この近似では分子構造が変化するとすぐに電子状態が追従する。言い換えれば、電子状態は常に分子構造の明示的な関数として表してよい。そのため、様々な分子構造でエネルギーを計算することによって、最も起こりやすい軌跡と登るべき山の高さを決定すればよい。

化学反応に関わる原子の数は反応系を構成する全原子である。この事情はタンパク質においても変わらない。酵素が基質と結合する際、その前後で結合部位の形状は変化しタンパク質全体に影響をおよぼす。構造の自由度を底面の軸にして、ポテンシャル・エネルギーを縦軸で示した俯瞰図で考えると、地形には複数の谷が存在する。すなわちこのエネルギーの極小点における原子核座標を持つ分子の構造が安定構造である。これらは化学反応の反応系と生成系のエネルギー状態の候補である。反応が起きるためには谷から谷へと分子が渡る必要がある。最も省エネルギーな行き方は尾根の峠を通る道筋である。峠の辺りは馬の鞍の形をしているため鞍部あるいは鞍点と呼ばれる。この鞍点が遷移状態である。したがって、反応系の安定点と鞍点のエネルギー差が活性化自由エネルギーである。これが低いほど反応速度は速い。

反応系、生成系、そして遷移状態はエネルギーの原子核座標による1次微分(グラジエント)が零となる点である。鞍点を極小点、極大点から識別するためには、2次微分の値を計算して判定を行う。エネルギーの2次微分行列はヘッシアンと呼ばれ、これを対角化すれば力の定数を得る。安定構造すなわち極小点は全ての固有値が正のもので、遷移状態すなわち鞍点は負の固有値を1つだけ持つものである。負の固有値の固有ベクトルは反応経路の方向を向いており、この方向のみポテンシャル・エネルギー超曲面は降っていて、他の方向は全て登っている。

以上のように、原理的には反応系および生成系の安定構造と遷移状態構造を求め、活性化自由エネルギー ΔF を見積もることができれば、化学反応シミュレーションの大半の目的が達成される。さらに、反応系から遷移状態、生成系に至る分子構造と作用の変遷、すなわち反応経路を追跡できれば、その化学反応機構が十分に解析できたことになる。その代表がIRC・DRC法^{43,44}で、量子化学計算で化学反応経路を求める標準的な手法である。これは、反応系・生成系の安定構造や遷移状態が求められた場合、それぞれの構造が一連の反応に対応しているかを調べる手段である。タンパク質の化学反応も原理は同じである。

⁴² N. D. Mermin, Phys. Rev., A137 (1965) p.1441.

⁴³ C. Gonzalez, H. B. Schlegel, J. Chem. Phys., 90 (1989) p.2154.

⁴⁴ K. M. Dieter, J. P. Stewart, J. Mol. Struct. (Theochem), 163 (1988) p.143.

IRC 法や DRC 法は、反応系—遷移状態—生成系の素反応において、作用最小の法則にしたがう反応経路を探す試みであり、これらは理想的な反応経路なので重要である。しかし、実際の反応は必ずしもこのような作用最小の経路を進むわけではない。他の峠も迂回して反応が進む確率は決して小さくない。しかも、そもそも安定構造が 1 つに決まる保障すらない。

このような性質を持つタンパク質の化学反応解析を行うためには、分子動力学法で温度や初期構造などの計算条件を変えてタンパク質の運動を追いかけること、モンテカルロ法で様々な条件によるカノニカルアンサンブルを計算することが直接的な方法である。これらの方法をより広い反応空間すなわち位相空間に適用できれば、大域的な反応座標探索や熱力学量の統計平均を得ることが期待できる。すなわち、究極の化学反応シミュレーションは様々な熱統計力学サンプリングによる自由エネルギー計算に他ならない。

自由エネルギーの差は、

$$F(\lambda_j) - F(\lambda_i) = \int_{\lambda_i}^{\lambda_j} d\lambda \left\langle \frac{\partial E^{(\lambda)}}{\partial \lambda} \right\rangle_{NVT} \quad (6.4)$$

と計算される。ここで、 λ は反応経路を示すパラメータで、被積分部分は Kohn-Sham エネルギーの λ に対する偏微分係数のカノニカル平均である。古典系のエネルギーを用いて精力的に研究が進められてきた自由エネルギー計算法の量子版である。系こそまだ小さいが、密度汎関数法においても様々なシミュレーションが行われ、良好な結果が得られつつある⁴⁵。

大きな系が取り扱える古典系では、反応座標の周辺を効果的にサンプリングする様々な方法が次々と開発されている。そのほとんどが再重み付け (reweighting) 技術を応用したもので、バイアスをかけたり、束縛条件を与えたりして、サンプリングする位相空間の範囲を狭めている。量子計算に基づくタンパク質のシミュレーションにも適用可能なものが多い。

また、反応時間の短い化学反応は、効果的に位相空間の探索を狭めることができるためシミュレーション達成は近いだろう。代表的な例はロドプシンの励起状態ダイナミクスである。ロドプシンは色素レチナールがオプシントタンパク質 (約 350 残基) とシッフ塩基結合した光受容膜タンパク質で、光によって 11-*cis*-レチナールから all-*trans*-レチナールに異性化する。この反応は、溶液中よりもオプシン中の方が高速で、わずか数百 fs といった時間スケールで反応が進む。その理由や反応機構は、近い将来、タンパク質量子化学計算で精度良く解析することができるだろう。

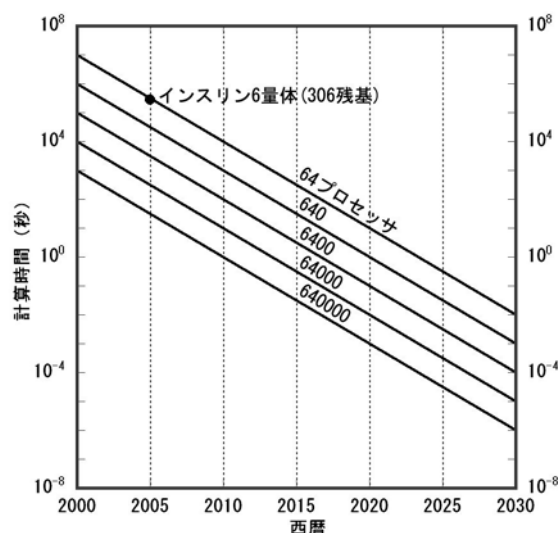
まとめ

図 9 はタンパク質密度汎関数計算の計算時間予測グラフである。図の横軸は西暦年で、2000 年から 2030 年までふつである。縦軸はタンパク質 1 構造の計算時間を秒単位の対数軸で表したものである。Moore の経験則が維持され、アルゴリズムとソフトウェアの改良によって 10 年で 10 倍高速化するという効果を加えると、計算は 10 年で 1000 倍速くなる。さらに、計算機の超並列化によっても計算時間が短縮される。このグラフを見ると、2030 年には 64 万台 (コア数でもよい) の超並列計算機を使用して、300 残基タンパク質の固定構造の計算が 10^6 秒で計算できる。これは 2000 年における古典力場による計算時間とほとんど変わらない。

タンパク質密度汎関数法による本格的な反応シミュレーションはこれからの研究課題である

⁴⁵ B. L. Trout, M. Parrinello, J. Phys. Chem., B, 103 (1999) p. 7340.

が、2030年までにはタンパク質の励起や運動が確実に追跡できるようになっているはずである。タンパク質の本格的な励起状態や電子移動の計算は2010年ごろまでにはスーパーコンピュータで達成され、2020年には個人のコンピュータで難なく実行できるだろう。同様に、タンパク質の化学反応のシミュレーションも2020年までには本格的なものが登場し、2030年にはもはや普通に行われるようになっていくに違いない。



第9図:密度汎関数法によるタンパク質の計算時間予測[1]。

プロセッサの台数および計算時間は2005年のインスリン6量体の計算をほぼ基準とし、それぞれの斜め線は $10^{-3}/10$ 年の傾きで引かれている。 10^{-2} はプロセッサ高速化のMooreの経験則によるもの、 10^{-1} はアルゴリズムやソフトウェア技術の向上によるものである。

タンパク質の励起エネルギー、振動子強度の計算や自由エネルギー計算の結果は、実験結果と直接比較することができる。密度汎関数の原理によりさらに高品位な交換相関汎関数が見出され、それをを用いた計算結果が実験値に接近するという好ましいサイクルが生まれるだろう。様々なタンパク質についてのシミュレーションが蓄積されれば、タンパク質の反応を制御している要因を電子レベルで分析することが可能になるだろう。生命の分子機構を解明し理解することは人類最大の科学的テーマの一つである。

光合成反応中心や光異性化反応の機構が完全に理解されれば、これらを基にした太陽電池、光センサー、非線形光学材料の設計へと応用されることになるだろう。天然や人工光合成システムから効率的に電気エネルギーを獲得し、超伝導エネルギー貯蔵装置に大電力を備蓄できるというような時代が来るかもしれない。

タンパク質の化学反応機構の解析は創薬の分野で多くの注目が集まっている。例えば、ヘムタンパク質のP450 (CYP)で代謝分解できない薬剤は体内に残り続けるため、薬剤の設計にはP450との相互作用解析が欠かせない。代謝産物が、DNAや細胞組織に作用して、癌や壊死を引き起こす問題にも関係する。タンパク質の化学反応シミュレーションにより、基質の結合性や活性化エネルギー、反応速度などが予測できる。系統的にシミュレーションを行うことによって知見を蓄積し、基質特異性や反応特性を制御できるようになるだろう。

生体触媒は工業製品の生産にも利用されつつある。これまでの工業触媒では重金属が多く使

われており、毒性のある副産物が問題になる。これに対して、P450には過酸化水素を酸化剤とし副産物は水だけというような特性がある。このように、生体が行う触媒反応はクリーンな反応を推進する。生体触媒の設計にもタンパク質の化学反応シミュレーションが威力を発揮するだろう。

低分子ではすでに実績のある量子化学シミュレーションの信頼性は高く、反応予測、反応設計ができるレベルに達している。密度汎関数法によるタンパク質反応シミュレーションが実用的になれば、基礎科学のみならず様々な応用分野でも大きな貢献をすることだろう。

本稿で使用した一部のデータは、文科省のいくつかのプロジェクト(革新的シミュレーションソフトウェアの研究開発、次世代生命体統合シミュレーションの研究開発、イノベーション基盤シミュレーションソフトウェアの研究開発)で得られた成果を利用した。また、一部の計算には統計数理研究所の統計科学スーパーコンピュータシステムを利用した。

参 考 文 献

最後にさらに詳しい内容に興味を持たれた方のために、関連する著書を数遍リストアップした。

生体分子量子化学計算関連

[1] 佐藤文俊, 恒川直樹, 吉廣保, 平野敏行, 井原直樹, 柏木浩:”タンパク質密度汎関数法”, 柏木浩監修, 森北出版, (2008). 【注: 本稿は本拙書の抄録的位置づけである。】

[2] 上村(西野)典子, 佐藤文俊, 恒川直樹, 西村康幸, 平野敏行, 吉廣保, 甘利真司, 加藤昭史, 小林将人, 田中成典, 中野達也, 福澤薫, 望月祐志: “プログラムで実践する 生体分子量子化学計算 –ProteinDF/ABINIT-MP の基礎と応用”, 佐藤文俊, 中野達也, 望月祐志編, 森北出版, (2008).

密度汎関数法

[3] R. G. Parr, W. Yang 著, 狩野覚, 吉田元二, 関元訳: “原子・分子の密度汎関数法”, シュプリンガー・フェアラーク (1996).

[4] 里子允敏, 大西楯平: “密度汎関数法とその応用—分子・クラスターの電子状態”, 講談社 (1994).

量子化学計算・分子軌道法全般

[5] 藤永茂: “分子軌道法”, 岩波書店 (1980).

[6] A. Szabo, N. S. Ostlund 著, 大野公男, 阪井健男, 望月祐志訳: “新しい量子化学—電子構造の理論入門<上・下>”, 東京大学出版会 (1987).

[7] T. Helgaker, P. Jorgensen, J. Olsen: “Molecular Electronic-Structure Theory”, John Wiley & Sons, (2000).

有限要素法に基づく第一原理分子動力学法について

土田 英二

産業技術総合研究所 計算科学研究部門

E-mail: Eiji.Tsuchida@aist.go.jp

1 はじめに

人間の眼に見えるようなマクロな世界は通常ニュートン方程式やマクスウェル方程式によって記述されるが、電子の波動性が重要な意味を持つようなミクロ・ナノ系の振る舞いは量子物理学の基礎方程式であるシュレーディンガー方程式によって決定されている [1]。この方程式自体は電子の質量などの自然定数のみを含み、経験的なパラメータを一切必要としない。従って、もしある系のシュレーディンガー方程式を十分な精度で解くことができれば、任意の物理・化学的性質を非常に正確に知ることが出来る。このようなアプローチは入力パラメータとして実験値を一切使用しないことから、第一原理計算と呼ばれる [2]。一昔前までは小さな分子や完全結晶を扱うことができる程度であったが、現在では計算機およびアルゴリズムの発展のお陰で、1000 原子を超える系を十分な精度で調べることができるようになった。当然ながら第一原理計算の応用範囲は広く、孤立分子に始まり、固体・液体やそれらの表面・界面等の複雑な系まで適用され、他の方法では得難い情報を与えてくれる [3]。特に最近では材料科学 [4] や地球科学 [5]、生命科学 [6, 7] など、より応用的な分野へと適用範囲を広げつつある [8]。

実際にシュレーディンガー方程式、もしくはその近似式を解く場合、何らかの方法で離散化して数値的に扱う必要があるが、通常は平面波基底 [2] や原子基底 [9] のような（物理的に意味のある）基底関数を使用するのが標準手法となっている。歴史的に言えば、平面波は周期性が本質的な意味を持つような系、即ち完全結晶の電子状態計算において最も自然な選択であり、計算機の進歩と共に適用対象を広げてきた。また、原子軌道は少数の原子から成るような孤立分子の電子状態を効率よく表現するために広く使われてきた。

一方、有限要素法は主に流体解析や構造解析といった用途に適用されることから「マクロな系の解法」というイメージが定着しており [10, 11]、上述のようなミクロ系における数値解法として採用されることは非常に稀であった。しかしここ 10 年の間に、こういった系に対しても有限要素法や差分法のようなマクロ系における標準手法が有用であるということが認められるようになってきた [12, 13]。これは主に、大規模な並列計算機との相性の良さや、オーダー N 法 [14, 15]・ハイブリッド法 [16] 等の新しいアルゴリズムへの拡張性が期待されているためである。以下では、特に有限要素法に基づく第一原理計算の現状について紹介する。

ここで、まず手短かに量子物理学における有限要素法の歴史を振り返っておきたい。一・二次元問題における有限要素法の登場は少なくとも 1970 年代に遡る [12, 13]。一方、応用上重要な系は対称性を持たないことが多いため、三次元問題として扱わざるを得ない。

当然ながら三次元の場合は一・二次元の場合と比べて遥かに計算量が大きく、より洗練された手法が必要とされる。三次元問題における本格的な応用が現れたのは80年代であり、Whiteらによりハートリー近似 [9] に基づく計算がヘリウム原子や水素分子等の系について行われた [17]。この計算では一様な格子を用いていたが、場所によって基底関数の数を変えることで計算の効率を上げていた。

90年代に入ると、より有限要素法の特長を生かした応用が行われるようになった。まず、四面体要素と局所的分割を組み合わせた手法により、いくつかの一電子系について計算が行われた [18]。また、我々のグループでは、六面体要素と局所的分割を用いて水素分子やシリコン結晶の第一原理計算を行い、期待通りの結果を得ることができた [19]。ただし、計算コストの点から見てこの時点での有限要素法の性能は平面波や原子基底を用いた場合と比べると不十分であった。この点を改善するために、我々は一次微分まで連続な滑らかな基底関数と曲線座標系を組み合わせた方法を提案した [20]。更に、スーパーコンピュータ向けにこの方法を実装したコード FEMTECK の開発を続けている [21]。以下ではこのアプローチについて他の方法との比較も交えて紹介したい。なお、我々以外にも国内外で多くの研究者が現在有限要素法を用いた第一原理計算の研究を活発に進めている [22, 23, 24, 25, 26, 27]。

2 問題の定式化

まず、本稿で扱う問題をここで一通り紹介する。今考えている系の中に含まれる原子の数を N 、それぞれの三次元座標を $\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N$ で表す。そうすると、原子座標 $(\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N)$ の関数として系の全エネルギー E_{total} 及び各原子に働く力

$$\mathbf{F}_\alpha = -\frac{\partial E_{\text{total}}}{\partial \mathbf{R}_\alpha}, \quad \alpha = 1, 2, \dots, N$$

を求めることが最も基本的な作業となる。実際に E_{total} を計算する手続きは使用するアルゴリズムによって大きく異なるが、いずれにしても \mathbf{F}_α を求めるためには \mathbf{R}_α の滑らかな関数でなくてはならない。ひとたび E_{total} と \mathbf{F}_α が得られれば後は同じであり、それらの情報を利用して系の安定構造を決定したり、分子動力学計算を行うことができる。また、安定構造に対して振動数や振動モードを求めることも可能である。

例えばアルゴンのような希ガス系の振る舞いはレナード・ジョーンズ型と呼ばれる古典的な二体ポテンシャルで比較的正確に記述されることが知られており、

$$E_{\text{total}} = \sum_{\alpha < \beta} \phi(|\mathbf{R}_\alpha - \mathbf{R}_\beta|)$$

のように、 E_{total} を $\mathbf{R}_1, \dots, \mathbf{R}_N$ の関数として陽に与えることができる [28]。この場合には \mathbf{F}_α も容易に計算することができることは明らかであろう。共有結合のようなより複雑な相互作用で記述される系では球対称でなく方向性を持つポテンシャルがしばしば使われるが、その場合でもやはり E_{total} は \mathbf{R}_α の陽な関数として与えられることが多い [28]。

一方、第一原理計算の場合はもっと複雑であり、以下に述べるように原子が移動する度に電子状態を計算する必要がある。原理的には、シュレーディンガー方程式を正確に解けば E_{total} や F_{α} 等必要な情報は全て得られるのであるが、複数の原子から成り立つような系、特に数十原子以上を含むような系において、これは事実上不可能である。従って、何らかの近似を導入する必要がある訳であるが、大雑把に言って一電子近似に基づく密度汎関数法 [29, 30, 31] と、ハートリー・フォック近似を基本とする伝統的な量子化学的手法 [9] が二本柱として普及している。それぞれ様々な長所・短所があって使い分けられる傾向があるが、応用する立場から見ると、周期境界条件の下で液体や固体などの系を扱う場合には、密度汎関数法の方がかなりコストパフォーマンスが良い。一方、量子化学的な手法は一電子近似では不十分な場合に多体効果を段階的に取り入れることが可能であり、小分子であれば極めて高精度の計算を行うことができる。我々はできる限り大規模な系を一電子近似の範囲で扱うことを目指しているため、以下では密度汎関数法のみを扱うこととする。

まず密度汎関数法における全エネルギーの表式を原子単位系で書き下すと、

$$E_{\text{total}} = \sum_{i=1}^M \int \psi_i(\mathbf{r}) \left[-\frac{1}{2} \nabla^2 + V(\mathbf{r}) \right] \psi_i(\mathbf{r}) d\mathbf{r} + \frac{1}{2} \int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}d\mathbf{r}' \\ + \int n(\mathbf{r}) \epsilon_{xc}(n(\mathbf{r})) d\mathbf{r} + \sum_{\alpha < \beta} \frac{Z_{\alpha} Z_{\beta}}{|\mathbf{R}_{\alpha} - \mathbf{R}_{\beta}|} \quad (1)$$

のようになる [2]。ただし M は電子数、 $\psi_1, \psi_2, \dots, \psi_M$ は各電子の波動関数であり、正規直交条件

$$\int \psi_i(\mathbf{r}) \cdot \psi_j(\mathbf{r}) d\mathbf{r} = \delta_{ij}, \quad i, j = 1, \dots, M \quad (2)$$

を満たすものとする。また、 $n(\mathbf{r})$ は電子密度であり

$$n(\mathbf{r}) = \sum_{i=1}^M |\psi_i(\mathbf{r})|^2 \quad (3)$$

で与えられる。 Z_{α} は各原子の電荷数を表し、 $V(\mathbf{r})$ は各原子が電子に及ぼすポテンシャルの総和である。なお、 ϵ_{xc} は交換相関エネルギー密度と呼ばれる $n(\mathbf{r})$ の関数であるが、具体的な関数形は分かっていない。従って実際に第一原理計算を行う際には何らかの近似式を使用する必要がある。良く使われるものとしては、局所密度近似あるいは密度勾配近似が挙げられる [2, 3]。

上記の E_{total} は一見 ψ_i と R_{α} の両方に依存しているように見えるが、実際には拘束条件 (2) の下で E_{total} を最小にする ψ_i を求め、その ψ_i に対する E_{total} を系の全エネルギーと見なす。従って、 E_{total} は R_{α} のみの関数として一意に定まる。又、こうして得られた ψ_i を使って F_{α} も容易に計算できる。この手順を以下では電子状態計算と呼び、具体的な計算方法については 3.3 章で説明する。

この時、各波動関数は以下のコーン・シャム方程式を満たすことが知られている [29, 30]

:

$$H_{\text{KS}} \psi_i = \epsilon_i \psi_i, \quad i = 1, \dots, M \quad (4)$$

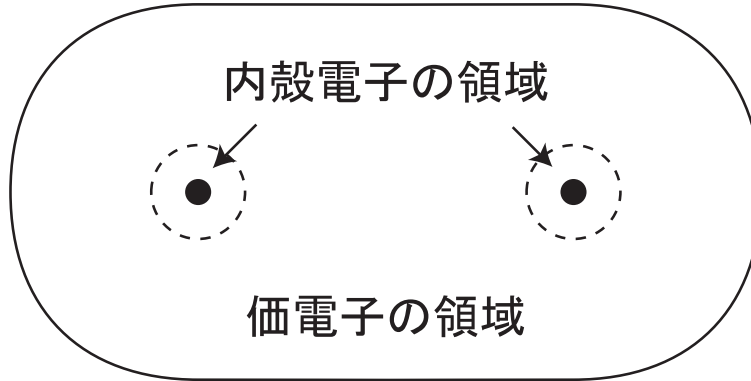


図 1: 内殻電子の波動関数は各原子核 (●) のごく近傍に局在しているが、価電子の波動関数は主にその外側に分布している。

ただし、

$$H_{\text{KS}} = -\frac{1}{2}\nabla^2 + V(\mathbf{r}) + \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + \mu_{xc}(\mathbf{r}) \quad (5)$$

及び

$$\mu_{xc}(\mathbf{r}) = \frac{\partial}{\partial n(\mathbf{r})} (n(\mathbf{r})\epsilon_{xc}(n(\mathbf{r}))) \quad (6)$$

であり、 ϵ_i は各波動関数の固有エネルギーとする [2]。この方程式は一電子に対するシュレーディンガー方程式と同じ形をしているが、他の電子の影響を全てポテンシャルの中に取り込んでいるため、 H_{KS} が ($n(\mathbf{r})$ 経由で) 解である ψ_i にも依存している点が異なる。

境界条件については通常、孤立境界条件 (遠方で $\psi_i(\mathbf{r}) \rightarrow 0$) か、周期境界条件 (L を計算領域のサイズとして $\psi_i(\mathbf{r} + \mathbf{L}) = \psi_i(\mathbf{r})$) が使われる。前者は分子等の孤立した系向けであり、後者は固体等の凝縮系に向いているが、十分計算領域を大きく取れば (場合によっては補正を加えることにより) 周期境界条件で孤立系を扱うことも可能である。数値計算上都合が良いため、以下では常に周期境界条件を満たすものとする。また、周期境界条件の場合には波動関数に位相因子 $\exp(i\mathbf{k}\mathbf{r})$ が付くが、大規模系への応用を重視しているため、ガンマ点 ($\mathbf{k} = 0$) のみを考え、かつ波動関数は実数であるものとして扱う。

更に、ここでもう一つの重要な近似について説明しておく。 E_{total} の表式中に現れる $V(\mathbf{r})$ について考察すると、これは各原子が電子に及ぼす影響を表しており、近似しない場合には $-\sum_{\alpha} Z_{\alpha}/|\mathbf{r} - \mathbf{R}_{\alpha}|$ という形をしている。すなわち、生のクーロン・ポテンシャルの和で与えられる。この形のまま計算を進める場合、全電子計算と呼ばれる。このような相互作用を受けた結果、各電子は原子核の近傍に局在する傾向があるが、その中でも特に強く原子核に束縛された内殻電子と、その外側にいて他の原子との結合に寄与する価電子に分けることができる (図 1)。

通常、内殻電子は非常に安定であり、原子のいる環境にはほとんど影響を受けない [2, 32]。従って、価電子のみを陽に扱い、 $V(\mathbf{r})$ の式を「原子核 + 内殻電子」が価電子に及ぼすポテンシャルに変更することで内殻電子を計算から除外することが良い近似として成り立

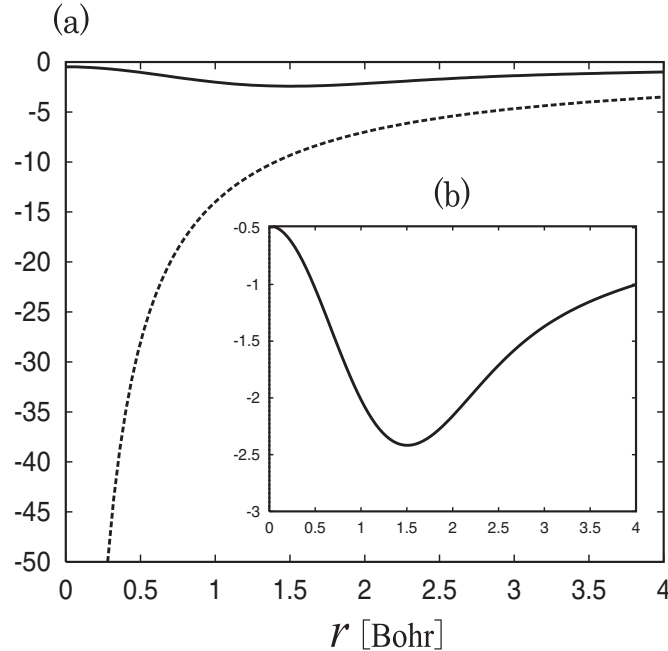


図 2: (a) シリコンに対する擬ポテンシャル [33] (実線) 及び生のクーロン・ポテンシャル (破線)。 (b) 擬ポテンシャルを拡大したもの。

つ。この変更後の $V(r)$ を擬ポテンシャルと呼ぶ。擬ポテンシャルを使うことにより、全電子計算より遥かに少ない数の要素で十分収束した解を得ることができる。また、内殻電子を計算から除くことができるため、陽に扱う電子の数を大幅に減らすことができる。従って、ある程度大きな系の第一原理計算を行う場合 (特に重元素を含む時) には擬ポテンシャルは必須の道具であり、以下でも特に明示しない限り常に擬ポテンシャルを使用するものとする。すなわち、全エネルギー表式中の $V(r)$ は擬ポテンシャルであり、また価電子の波動関数のみを考えるものとする。図 2 に擬ポテンシャルの例を示す。

なお、実際に高精度でかつ扱いやすい擬ポテンシャルを作成するのはかなり職人芸的な知識と経験を必要とする作業である。最近は多くの元素について十分なチェックを受けた擬ポテンシャルのリストが出版されており [34]、我々のコードもこれを利用している。

3 FEMTECK の構成

3.1 基底関数

計算機上で電子状態計算を行うためには何らかの基底関数を用いて波動関数を展開する必要がある。我々はエルミート補間式 [35] と呼ばれる区分多項式を基底関数として使用している。この補間式は任意の奇数次に対して定義できるが、FEMTECK では三次のも

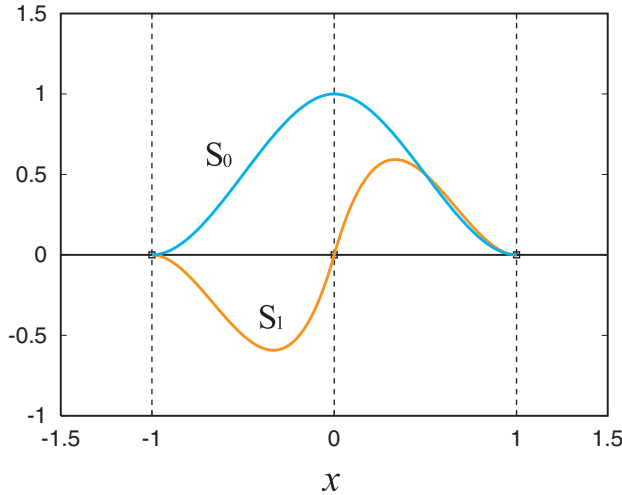


図 3: 原点に属する三次のエルミート補間式 $S_0(x), S_1(x)$ 。ただし $S_1(x)$ は見易いように拡大している。

のに特化している。まず簡単のため、一次元の格子上で 1 の間隔で節点が並んでいるものとする。この時、原点には 2 個の基底関数 $S_0(x), S_1(x)$ が割り当てられ、これらは条件

$$S_0(0) = 1, \quad S_0'(0) = 0, \quad S_0(\pm 1) = S_0'(\pm 1) = 0 \quad (7)$$

および

$$S_1(0) = 0, \quad S_1'(0) = 1, \quad S_1(\pm 1) = S_1'(\pm 1) = 0 \quad (8)$$

を満たすように定義される。これらを満たす三次の区分多項式として、

$$S_0(x) = \begin{cases} 1 - 3x^2 + 2x^3, & 0 \leq x \leq 1 \\ S_0(-x), & -1 \leq x < 0 \\ 0, & \text{その他} \end{cases} \quad (9)$$

および

$$S_1(x) = \begin{cases} x - 2x^2 + x^3, & 0 \leq x \leq 1 \\ -S_1(-x), & -1 \leq x < 0 \\ 0, & \text{その他} \end{cases} \quad (10)$$

が得られる。参考のため $S_0(x)$ 及び $S_1(x)$ を図 3 に示す。また、他の節点における基底関数も単純にこれらを平行移動することで得られる。そして、波動関数などの未知関数は全ての節点におけるこれらの基底関数の線形結合で近似される。式で表現すると、

$$\psi(x) = \sum_n (a_n \cdot S_0(x - n) + b_n \cdot S_1(x - n)) \quad (11)$$

のように展開される。ここで、 a_n および b_n は未知変数であり、それぞれ $x = n$ における波動関数の値と微分に対応している。これらの値はコーン・シャム方程式を満たすように決定される (3.3 章を参照)。

三次元問題の場合には、一次元の基底関数の積として以下のように与えられる：

$$\left\{ S_i \left(\frac{x - n_x}{h} \right) \cdot S_j \left(\frac{y - n_y}{h} \right) \cdot S_k \left(\frac{z - n_z}{h} \right) \right\}, \quad i, j, k = 0, 1 \quad (12)$$

ただし、各節点は間隔 h の一様な三次元格子にあり、それぞれの座標を (n_x, n_y, n_z) で表すものとする。これは一辺が h の立方体要素に分割していることに相当する。この時、各節点には 8 個の基底関数が割り当てられ、各基底関数は隣接する節点に属する基底関数のみと重なりを持つことになる。この結果、基底関数間の行列は行ごとに 216 個の非零要素を持つ疎行列となる。この性質は後述の曲線座標系を使用した場合にも保たれる。計算の精度を上げる場合には格子間隔 h を小さくして解像度を上げてゆけば良い。

また、実際に波動関数から E_{total} を計算する場合にはいくつかの項で数値積分を行う必要があるため [21]、多項式向けによく使われるガウス・ルジャンドル積分法 [36] を用いる。三次の基底関数の場合には通常五次の積分点を用いれば十分な精度が得られる [21]。

3.2 適合曲線座標系

2 章で説明したように、電子は原子からのポテンシャルを受けて運動している訳であるが、ポテンシャルの形から想像できるように電子の波動関数は原子付近では比較的激しく変化し、原子から離れるに従って速やかに減衰する。従って、なるべく少数の自由度で高精度の解を得るためには何らかの方法で原子付近の解像度を上げることが望ましい。このためには色々な方法があるが、良く知られたものとしては以下の三つが挙げられる [11]：

- p 法 (要素高次化法)：要素自体は固定し、その代わり誤差が大きい要素から優先的に基底関数の次数を上げてゆくことで精度を上げる。
- h 法 (要素細分化法)：メッシュ及び基底関数の次数は一定にし、誤差が大きい要素を分割してゆくことで精度を上げる。
- r 法 (節点移動法)：基底関数の次数は要素によらず一定にし、誤差が大きい要素が相対的に小さくなるように要素の形状を連続的に変形することで精度を上げる。

初期の実装においては、p 法に従って原子付近の要素における基底関数の数を増やすか [17]、あるいは h 法のように原子付近の要素を細分化する [18, 19] という方法が提案されてきた。しかし、これらの方法で局所的に自由度を増やした場合、原子がよほど一様に分布していない限り、並列計算機上での負荷分散が悪化し、通信量も増大することが予想される。また、例えば各原子の位置が時間と共に変化する場合を考えると、細分化する領域をどこかで不連続に変更しなくてはならない。そうすると E_{total} が R_α の不連続な関数になってしまい、非常に都合が悪い。一方、r 法の場合には常に連続的に変化するため、

こういった問題に遭遇することなく原子付近の解像度を上げることが出来る。また 3.5 章で説明するように、並列計算機上での性能を損なうこともない。従って、以下では r 法を中心として説明を進める。

ここで r 法を実装する上で必要となる座標変換の方法について紹介する。このような工夫を加えることによって強く局在した軌道を持つ元素（例えば酸素やフッ素）を扱う際に基底関数の数を一桁程度減らす事ができる。我々が使用している適合曲線座標系は平面波基底と共に初めて電子状態計算に適用され [37]、次いで差分法 [38]、有限要素法 [20, 21] へと拡張された。座標変換の式としては原子座標 R_α の関数として陽に与えられる形を使用している。具体的に説明すると、ユークリッド座標を r 、曲線座標を ξ で表した時、座標変換の式は

$$\xi = r + \sum_{\alpha} (r - R_{\alpha}) g_{\alpha}(|r - R_{\alpha}|) \quad (13)$$

で与えられる。ここで、 $g_{\alpha}(r)$ は以下のような性質を持つ関数である：

1. 任意の $r \geq 0$ に対して $g_{\alpha}(r) \geq 0$.
2. 原点付近で最大値を取り、 $r \rightarrow \infty$ で速やかに減衰する
3. 少なくとも三次微分まで計算可能

このような条件を満たす関数形は色々考えられるが、実用上は少数のパラメータで指定できる形が便利である。例えば、二個のパラメータで決まる場合の形としては

$$g_{\alpha}(r) = A_{\alpha} \exp\left(-\left(\frac{r}{a_{\alpha}}\right)^2\right), \quad (14)$$

のような式が考えられる（実際にはもう少し複雑な式を採用している）。この場合には、原子 α の位置における解像度がおよそ $(1 + A_{\alpha})$ 倍に拡大され、 a_{α} が拡大する範囲を決定する。適切なパラメータの値は元素ごとに大体決まっており、例えば酸素原子のように強く局在した電子状態を持つ場合は A_{α} を大きく取る必要があるが（1程度）、アルミニウム原子のように広がった電子状態の場合には小さい値（もしくはゼロ）で十分である。

実際の計算は全て ξ 空間における一様なメッシュ上で行うため、原子が動く度に各節点（及び積分点）を r 空間に変換する必要があるが、ニュートン・ラフソン法 [36] で 2 ~ 3 回反復すれば十分な精度で式 (13) を逆変換し、 r 空間での座標を求めることができるため、特に問題はない。ただしあまり極端に解像度を上げると解が存在しなくなるため、注意が必要である。例として、窒素分子が大きな周期セルの中央で振動・回転している場合に、 ξ 空間の一様なメッシュを r 空間へ射影した結果を図 4 に示す。価電子密度（図 5）と見比べると、波動関数が最も激しく変化する原子核の近傍でほぼ等方的に解像度が拡大されていることが分かる。また、原子が動くにつれてメッシュも変化するため、 F_{α} を計算する際にはこの影響も考慮に入れている。これにより、分子動力学計算を行った場合でも系の全エネルギーが高い精度で保存される [21]。

また、曲線座標系を使用した場合としない場合について窒素分子の全エネルギーを原子間距離の関数として計算したものを図 6 に示す。これから分かるように、適切な座標変換を行うことによって一様なメッシュを使用した場合と比べて精度が大幅に向上する。

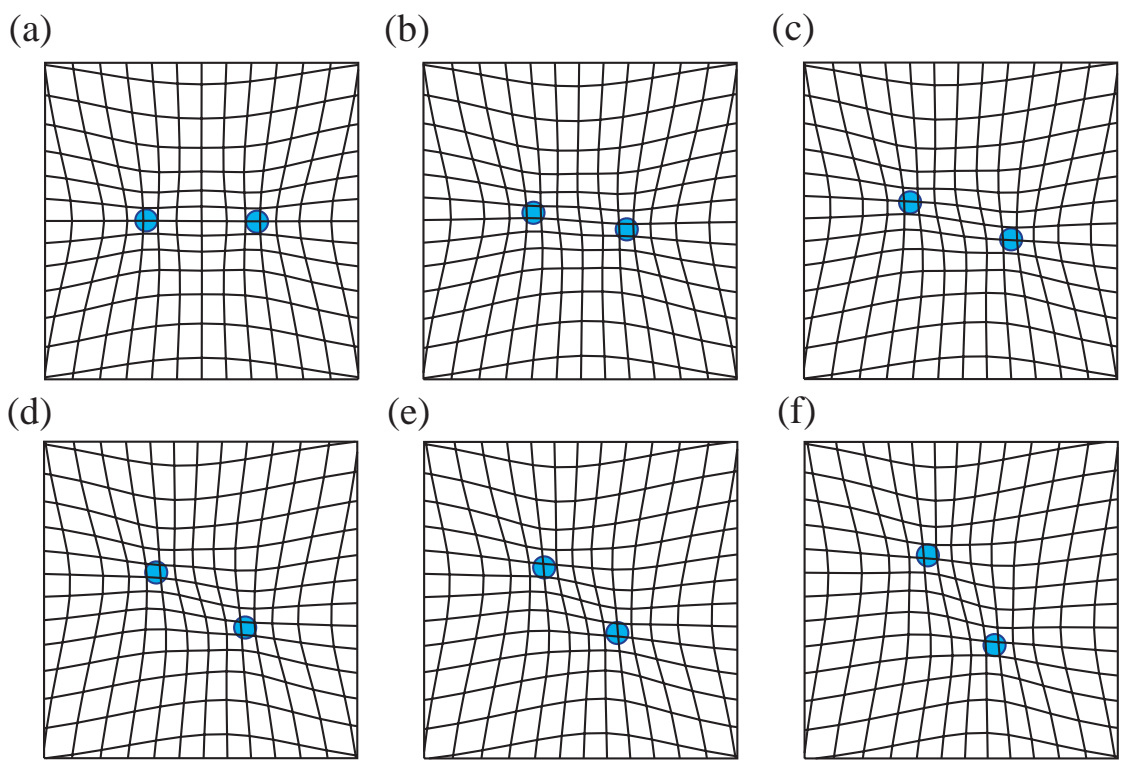


図 4: 窒素分子 (N_2) が振動・回転している場合の曲線座標系の時間変化を示す (ユニットセルの中心付近のみ表示)。

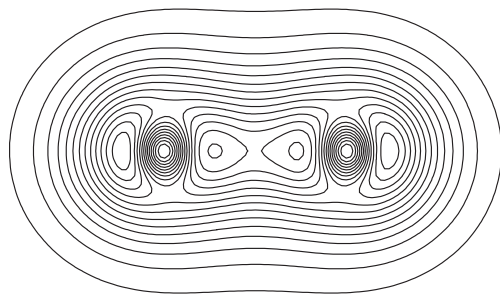


図 5: 窒素分子の価電子密度

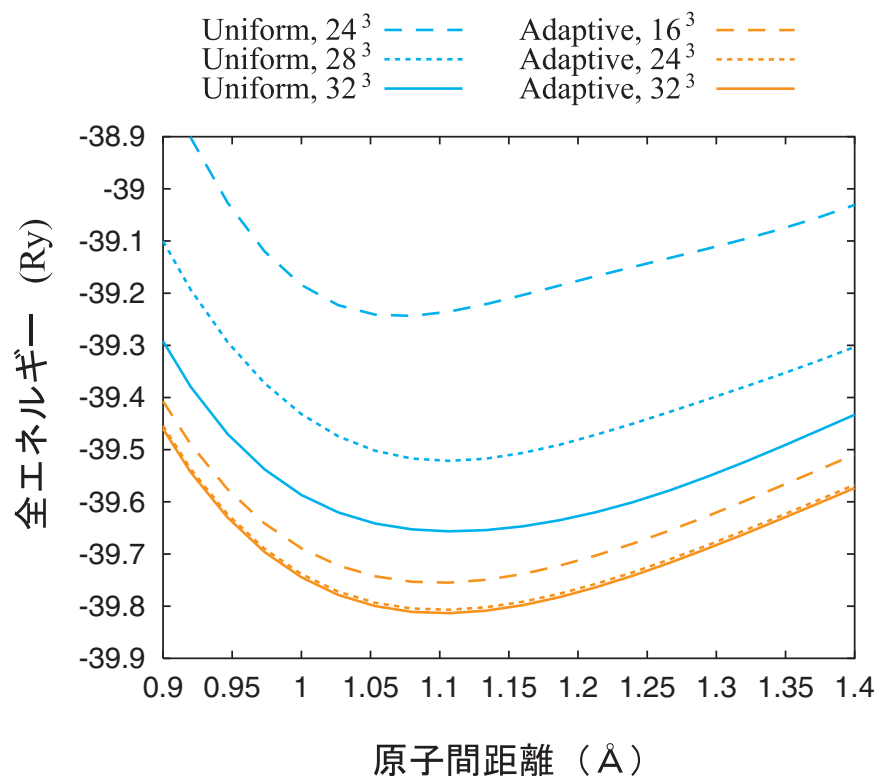


図 6: 一辺が 12 Å のスーパーセル中に置いた窒素分子のエネルギーを原子間距離の関数として計算したもの。Uniform: 一様なメッシュの場合、Adaptive: 適合曲線座標系を使用して各窒素原子の付近で 2 倍強程度解像度を上げた場合。結合長の実験値は 1.10 Å 程度であり、計算値と良く一致している。

3.3 電子状態計算の方法

ここでは、与えられた R_α に対して電子状態計算を行い、基底状態の波動関数及び E_{total} を得る方法について具体的に説明する。まず、波動関数は全て ξ 空間における一様なメッシュ上の各節点に属するエルミート補間式 (式 (12)) で展開される。すなわち、 $\psi_1, \psi_2, \dots, \psi_M$ は式 (11) を三次元に拡張した形で近似される。このとき、コーン・シャム方程式 (式 (4)) は次の式のように離散化される：

$$A \vec{\psi}_i = \epsilon_i S \vec{\psi}_i, \quad i = 1, \dots, M \quad (15)$$

ここで A は H_{KS} を基底関数で挟んで積分した行列であり、 S は基底関数間の重なり行列である。また、 $\vec{\psi}_i$ は波動関数を展開した際の係数をベクトルとして並べたものである。通常、有限要素法の基底関数は非直交であるため、離散化されたコーン・シャム方程式は一般化固有値問題になる。必要な固有状態の数 M は大体 10^2 から 10^3 程度である。また、 A 及び S は非常に規則性の高い対称な疎行列であり、その次元は M の $10^2 \sim 10^3$ 倍程度であるため、系によっては 10^6 を超える巨大な行列になることもある。更に、波動関数が増えると H_{KS} 自体も変更されるため、注意を要する。式 (15) を解く方法は色々提案されており、十分な精度で計算すればどの方法でも同じ結果が得られるが、 A や S の性質を活用する為には何らかの反復法を使うことが望ましい。

我々は固有値問題を直接解く代わりに、正規直交条件

$${}^t \vec{\psi}_i \cdot S \cdot \psi_j = \delta_{ij}$$

の下で

$$E = \sum_{i=1}^M {}^t \vec{\psi}_i \cdot A \cdot \psi_i \quad (16)$$

を最小にするような $\vec{\psi}_i$ を求めるという一種の最適化問題に変換し、これを準ニュートン法 [16, 39] で反復して解いている。FEMTECK では特に limited memory BFGS 法と呼ばれる準ニュートン法の省メモリ版 [40, 41] を実装し、大規模系でも効率良く扱うことができる。

一方、電子状態計算のコストを抑えるためには最適化の方法だけでなく、波動関数の良い初期値を用意することも非常に重要である。分子動力学を行う場合を考えると、通常各原子の位置は時間と共に滑らかに変化する。従って、過去 2 ~ 3 ステップの原子配置における波動関数の値から外挿することにより、現在の原子配置に対する波動関数の近似値を求める方法が非常に効果的である [42]。一番簡単な場合の外挿式を示すと、

$$\psi_n^{\text{ini}} = \psi_{n-1} + (\psi_{n-1} - \psi_{n-2}) = 2\psi_{n-1} - \psi_{n-2} \quad (17)$$

のようになる。ただし、 ψ_{n-1}, ψ_{n-2} はそれぞれ 1, 2 ステップ前の波動関数の (収束した) 値であり、 ψ_n^{ini} が最新の原子配置に対する波動関数の初期値である。このような外挿を行った場合、 R_α が変わるたびに大体 10 ~ 15 回程度反復すれば十分な精度を持った波動関数が得られる。

なお、電子状態計算を行うのに必要な計算量は原子数 N と共に

$$c_0 + c_1N + c_2N^2 + c_3N^3$$

のように増加する。 c_0, \dots, c_3 の値は実装方法に依存するが、全て正である。 N が数百の場合は $O(N)$ 及び $O(N^2)$ の項が最も計算時間を消費するが、1000 原子を超えるような大規模系では $O(N^3)$ の項が計算時間の大半を占める。また、必要なメモリ量は $O(N^2)$ のように増加する。計算時間ほど急激に増加する訳ではないが、メモリ量が足りないとそもそも動作させること自体できないため、こちらも深刻な問題である。

ところで、シミュレーションを行う系がエネルギー・ギャップを持つような場合には、精度を損ねることなく電子状態を強く局在させることが可能であることが知られている。このような性質を生かして計算量・メモリが共に $O(N)$ で済むような（近似的な）電子状態計算法が最近数多く提案されている [14, 15]。これらのオーダー N 法と呼ばれる手法には色々なバリエーションがあるが、大部分は実空間における電子状態の局在性を利用しているため、空間全体に広がった平面波基底よりも有限要素法や差分法のような実空間基底の方が適している。オーダー N 法は今まで安定性に問題があるという理由で広く使われることはなかったが、我々は最近この問題を回避できることを発見し、FEMTECK 中に実装した [43]。更にこの方法の信頼性を実証するため、エタノール₁₂₅ 分子（1125 原子）から成る大規模な溶液系で第一原理分子動力学計算を行った（図 7） [44]。十分な平衡化を行った後、300 K で 10 ピコ秒程度のシミュレーションを行い、動径分布関数や振動スペクトル等に関して十分な精度で実験値を再現することができた。また、水素結合の統計データ等、実験からは求めにくい物理量についても計算した。この系のシミュレーションでは、オーダー N 法を使用することによって従来の方法と比べて 5 倍近い高速化を実現することができた。

3.4 ポアソン方程式の解法

系の全エネルギー（式 (1)）の第二項、すなわち

$$\frac{1}{2} \int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}d\mathbf{r}'$$

は電子間の相互作用を表しているが、これを計算するためには電子密度 $n(\mathbf{r})$ が変化する度にハートリー・ポテンシャル

$$V_H(\mathbf{r}) = \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}'$$

を求める必要がある。周期境界条件の場合には、この式をそのまま計算するよりもポアソン方程式

$$\nabla^2 V_H(\mathbf{r}) = -4\pi(n(\mathbf{r}) - \bar{n})$$

の解として求める方が都合が良い。ただし \bar{n} は電子密度の平均値である。実際には V_H を ξ 空間で基底関数展開することによって上記のポアソン方程式を離散化し、一次方程式の

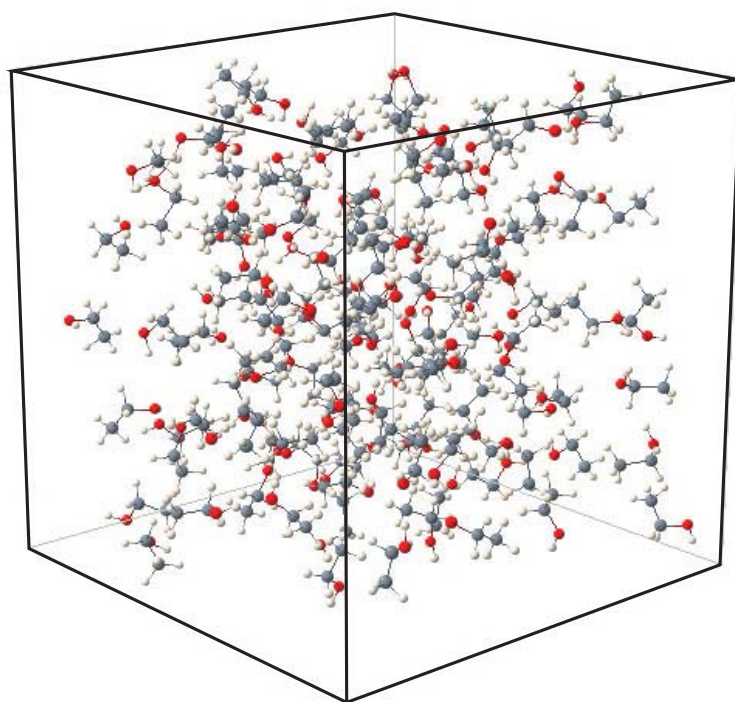


図 7: エタノール 125 分子から成るスーパーセル

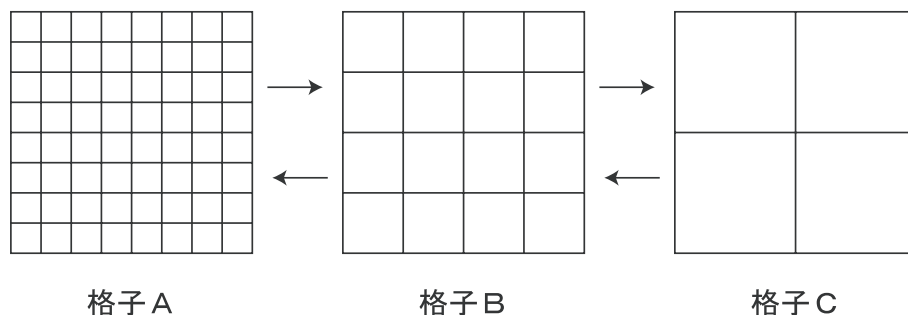


図 8: 多重格子法で格子 A 上での近似解を改善したい場合、残差を A B C B A の順に移動させ、それぞれの格子上で緩和法を適用する。

形にして解く。これを

$$Lv = n \quad (18)$$

と書くと、 L は ∇^2 に対応する行列であり、行ごとに 216 個の非零要素を持つ疎行列になる。系が小さい時はコレスキー分解のような直接法によってこの方程式を解くことが可能であるが、系が大きくなると行列の次元が非常に大きくなるために何らかの反復法に頼らざるをえない。標準的な反復法としては共役勾配法が良く知られているが、反復回数が系のサイズとともに増加するという難点がある。これは、系が大きくなるにつれて残差 ($= n - Lv$) の長波長成分の減衰がどんどん遅くなっていくことに起因している。このような場合、最も有力な解法は多重格子法と呼ばれる方法であり、系の大きさによらず一定の反復回数で同じ精度の解を与えることが知られている [45]。多重格子法のポイントは、解像度の異なる複数の格子を用意し、それぞれの格子毎に残差を減少させるような緩和法を適用することである (図 8)。これによって長波長から短波長の残差まで偏りなく減少させることができるため、系が大きくなっても一定の収束率を保つことが可能である [45]。

今の所、各格子上での緩和法としては、ベクトル化及び並列化の容易な最小二乗多項式の種類 [46] を採用し、異なる格子間の移動については V サイクルを使用している。また、主に並列化との兼ねいで、最も粗い格子上でも厳密に解くことはせずに、緩和法を数回適用している。なお、多重格子法は単独で用いても良いが、共役勾配法などの前処理として使用した方が、より効果を発揮できることが多い [21, 47]。例として図 9 にエタノール 125 分子の系でポアソン方程式の解の収束性を比較した結果を示す。共役勾配法の前処理として多重格子法を使用すると、対角行列の場合と比べて 3 倍以上の高速化が実現できることが分かる。系が大きくなるに連れてこれらの差は開くため、多重格子法は大規模計算には必須の技法である。

3.5 並列化の方法

ある程度大規模な第一原理計算を行う場合には、メモリ使用量及び計算コストの面から並列計算機を活用することが不可欠であるが、ここまで説明してきたアルゴリズムを並列

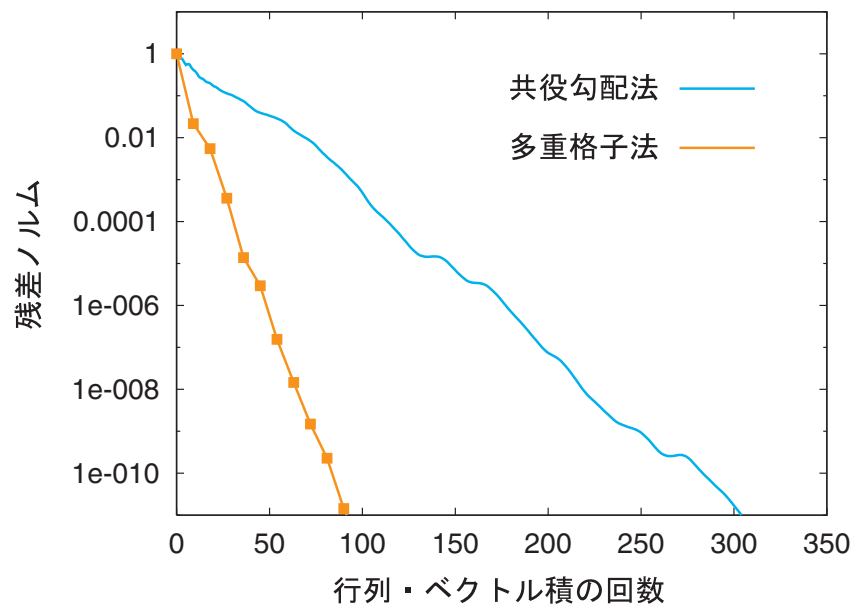


図 9: ポアソン方程式を通常の共役勾配法（前処理 = 対角行列）と多重格子法で解いた場合の収束性を示す。一反復当たりのコストが異なるため、行列・ベクトル積の回数を横軸としている。

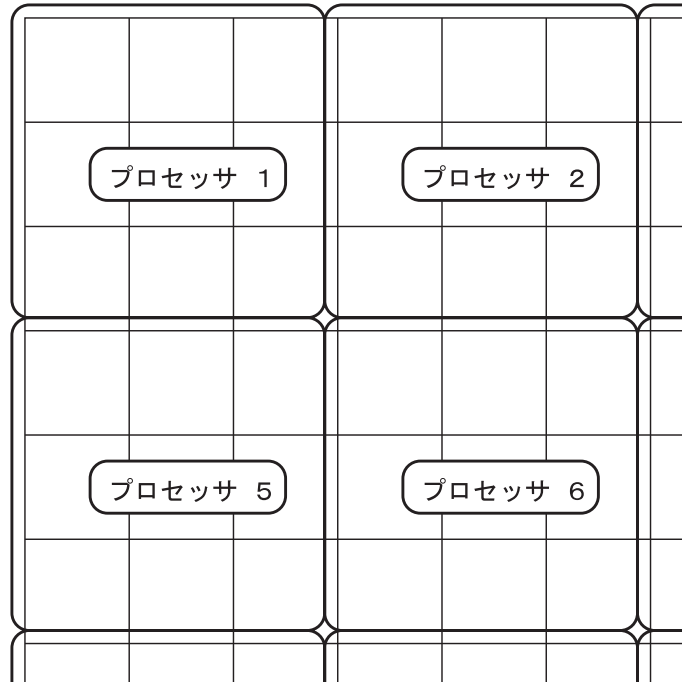


図 10: 並列化のイメージ図

化することは極めて容易である。即ち、我々の実装は r 法に基づいているため、曲線座標系 (ξ 空間) における一様なメッシュを使用している。これを同じサイズの領域に三次元分割して各プロセッサに割り振るだけで十分である (図 10)。この時各プロセッサには同数の要素が割り当てられるため、負荷分散はほぼ完全に等しくなる。また、第一原理計算の場合には古典分子動力学法と異なり、各プロセッサが全ての原子の座標を持つことが普通である。

電子状態計算を行う際に必要な通信のタイプは主として (i) 波動関数等の境界における値を隣接プロセッサ間で共有する為のローカルな通信、及び (ii) 電子数 M 程度の次元を持つ行列のグローバルな通信、の二種類に分類される。言い換えれば、最も情報量の多い波動関数の通信をローカルで済ませ、通信負荷の高いグローバルな通信はほぼ必要最小限に抑えることができる。従って、プロセッサ間の通信によるオーバーヘッドは比較的小さい。こういった事情は有限要素法に限らず差分法でも同様であり [48, 49]、実空間法に共通の利点の一つとすることができる。

FEMTECK は 1997 年頃に、本センターに設置された直後の SR-2201 上で並列化を行った。一時期、SR-8000 向けに最適化していた際には共有メモリを活用するためにマルチスレッド化を施していたが、(特に PC クラスタ上では) 色々な制約が付くため、現在は MPI 通信のみを使用して並列化している。上述の領域分割に基づく並列化の効率を HA8000 ク

ラスタシステム (東京大学情報基盤センター)¹ 上で測定した結果については6章で紹介する。

3.6 分子動力学

2章で触れたように、我々の計算では原子が移動する度に電子状態計算を行い、各原子に働く力を求めている。これは Born-Oppenheimer molecular dynamics (BOMD) 法と呼ばれる手法であり、非常に多くの系 (特にエネルギー・ギャップを持つような系) において正確な記述を与える [3]。ここでは原子核の量子性は考慮せず、全て古典的に振舞うものとする。そうすると、各原子 α の位置 R_α 及び速度 V_α はニュートン方程式

$$\frac{dR_\alpha}{dt} = V_\alpha \quad (19)$$

$$\frac{dV_\alpha}{dt} = \frac{F_\alpha}{m_\alpha} \quad (20)$$

に従って変化する。実際の計算ではこれらの式を離散化した速度ベルレ法 [50, 51]

$$R_\alpha(t + \Delta t) = R_\alpha(t) + \Delta t V_\alpha(t) + \frac{(\Delta t)^2}{2m_\alpha} F_\alpha(t) \quad (21)$$

$$V_\alpha(t + \Delta t) = V_\alpha(t) + \frac{\Delta t}{2m_\alpha} (F_\alpha(t + \Delta t) + F_\alpha(t)) \quad (22)$$

を用いて数値積分することで時間発展を追うことができる。ここで Δt は時間刻みを表す。なお、式 (21), (22) はミクロカノニカル集合に対応したものであるが、原子系の温度をコントロールする必要がある場合には Berendsen 熱浴 [52] を加えている。また、シミュレーションの初期段階などでは分子内の共有結合長を一定に保つような拘束条件の下で時間発展を追うことも多い。このような場合には RATTLE と呼ばれるアルゴリズムを実装することで対応している [53]。これにより初期配置が不自然であったり系が高温である等の理由で共有結合が切れることを防ぐ事ができ、更に時間刻みを2~3倍大きく取ることができる。

4 他の基底関数との比較

ここまで有限要素法を用いて波動関数を展開し、電子状態を計算する方法について説明してきたが、他にも様々な基底関数があり、それぞれ得手不得手がある。本章ではこのような視点から、代表的な基底関数について長所・短所を手短に考察する。

まず、いわゆる原子基底²の長所としては、以下のような点が挙げられる：

- 原子数の2~3倍の基底関数でも半定量的な計算ができる。従って、机上のパソコンでもある程度のサイズの系の第一原理計算を手軽に行うことができる。

¹ <http://www.cc.u-tokyo.ac.jp/ha8000/>

² ガウス型、スレーター型、あるいは数値型を指すことが多い。

- 豊富な機能を持つ有償・無償のプログラムが多数開発されており、専門家でなくとも容易に計算を開始することが出来る³。

一方、原子基底には次のような問題点もある：

- (特に基底関数が不十分な場合、) 基底関数重なり誤差の影響で原子間相互作用が過大評価され易い。
- 多原子系の場合、基底関数の選び方は必ずしも自明ではない。従って、いくら基底関数の数を増やしても波動関数が完全に収束していることを保証することは容易ではない。また、数を増やし過ぎると線形従属に近くなり、数値的な安定性が落ちる。

次に平面波基底について考察すると、まず長所としては

- 正規直交基底であり、また微分・積分が容易なため、 E_{total} のうち ϵ_{xc} を含む項以外は全て解析的に計算することができる。
- 原子基底と同様、完成度の高いプログラムが多い⁴。

といった点があり、短所としては

- フーリエ変換を頻繁に行う必要があり、並列化を行う際に大きな負担となる。
- 本質的に広がった基底関数であり、オーダー N 法やハイブリッド法のように局在性が重要なアルゴリズムへの適用は容易ではない。

といった点が挙げられる。一方、差分法は有限要素法と同様に実空間で定義されているため、高い並列性やオーダー N 法・ハイブリッド法への適合性など、多くの共通する特長を持つが、

- 有限要素法とは異なり、正規直交系を形成している。
- 差分法では波動関数を基底関数で展開する代わりに E_{total} 中の微分演算子を差分近似している。従って変分原理を満たさず、その結果 E_{total} が真の値より低くなり得る。

という点で異なっている。

なお、原子基底以外のもの(平面波、差分法、有限要素法)は機械的に計算精度を上げることが可能であり、系統的な基底関数と呼ばれる。例えば有限要素法であれば、実空間メッシュの間隔を小さくしていけば常に収束した波動関数が得られる。これらの基底関数の場合、原子数が一定であれば、計算量はほぼ基底関数の数に比例する。一方、原子基底の場合には基底関数を増やしてゆくと計算時間は急激に増加するため、ある程度精度の高い計算を行う場合には系統的な基底関数の方が高速であると考えられる。

以上、代表的な基底関数を幾つか紹介して来たが、実際には必要な精度・系のサイズ・使用する計算機などに応じて最適なものを選ぶのが現実的である。例えば膨大な数の原子

³ Gaussian, Q-Chem, GAMESS, SIESTA, OpenMX 等が有名である。

⁴ 代表的なものとしては CPMD, ABINIT, VASP, CASTEP 等がある。

配置（あるいは原子組成）から有望なものを選び出すような場合には、まず最小限の原子基底を用いてある程度絞り込みを行い、残ったものに対して系統的な基底関数を使用して高精度計算を行う、といった使い分けが考えられる。

5 応用例

ここでは、FEMTECK を用いた第一原理計算の応用例をいくつか紹介する。

常温における水素結合性の分子性液体は非常に複雑な構造を持ち、頻繁に水素結合の生成・消滅を繰り返している。このような系を定量的に理解するため、メタノール (CH_3OH) [54], ホルムアミド (HCONH_2) [55], エタノール ($\text{C}_2\text{H}_5\text{OH}$) [44] 等の分子性液体について数ピコ～数十ピコ秒程度の第一原理分子動力学計算を行い、構造（動径分布関数、水素結合数）や動的性質（拡散係数、振動数分布）などの物性を計算した。特に、実験と直接比べることが可能な量については非常に良い一致を示すことを確認している。

また最近、我々のグループでは燃料電池に関連する系のシミュレーションに力を注いでおり、濃リン酸 (H_3PO_4) [56], 硫酸水溶液 ($\text{H}_2\text{SO}_4 / \text{H}_2\text{O}$) [57]、ナフィオン/水界面 [58] 等の系について研究を進めている。これらの系では頻繁にプロトン移動が起きるため、古典ポテンシャルで定量的に記述することは非常に困難であり、第一原理計算で扱う意義が大きい。また、ベリー位相を利用して外部電場を加えることでプロトン伝導を促進することも行っている [59, 60]。

6 ベンチマーク計算

最後に、本センターに設置された HA8000 上における FEMTECK のベンチマーク結果について紹介したい。以下のような3種類の規模の異なる系についてそれぞれ第一原理分子動力学計算を行い、1ステップ進めるのに要する時間を測定した。これらの計算では十分な数の基底関数を使用し、またオーダーN法は使用していない。

名称	系	原子数	波動関数の数
システム A	ダイヤモンド	64	128
システム B	水	375	500
システム C	水	1920	2560

システム A を使って基本的な部分を詰めるため、まず1ノード中の1コアのみを使用して、単体性能を測定した。コンパイラ（言語は Fortran）及び数値ライブラリ（BLAS/LAPACK）としては幾つか選択肢があるが、HA8000 上で問題無く動作した組み合わせは「日立コンパイラ + netlib」及び「インテルコンパイラ + ACML(MKL)」であった。コンパイルの際にはそれぞれ以下のようなオプションを使用した。

```
% f90 -Oss -noparallel *.F -llapack_sc -lblas_sc  
% ifort -O3 -static *.F -lacml
```

また、ACML ライブラリは

<http://developer.amd.com/cpu/Libraries/acml/downloads/Pages/default.aspx>

よりダウンロードした Intel Fortran 用の 64 ビット版 (バージョン 4.1.0) を使用した。比較対象として、普段使用している AIST スーパークラスターの P32 システム (Dual Opteron 2GHz) を 1 コアのみ使用して測定した結果を用いた。この環境では PGI Fortran 90 及び ACML ライブラリを使用した。

これらの条件の下で、システム A で実行時間を測定した結果を以下に示す：

コンパイラ / 数値ライブラリ	システム	実行時間 (秒)
日立+netlib	HA8000	112.7
インテル+ACML	HA8000	87.1
インテル+MKL	HA8000	87.8
PGI+ACML	P32	112.9

まず分かるように、HA8000 上でのパフォーマンスは「インテル+ACML/MKL」の方がかなり速い。ただし、内訳を良く見ると、数値ライブラリを呼び出している部分で差が付いており、それ以外の自作ルーチンに対する性能は大差ない。従って、(リンクが可能であれば)「日立コンパイラ+ ACML/MKL」の組み合わせの方が速くなる可能性はある。また Goto BLAS 等も有力な選択肢であるが、LAPACK や ScaLAPACK を使う際には一手間ありそうだったため、今回は見送った。なお、「インテルコンパイラ + MKL」はやや挙動不審な面があり、また ACML と比べて特に性能上のメリットも見られなかったため、以降のベンチマークは「インテルコンパイラ + ACML」に限定する。

一方、HA8000 上で ACML を使用した場合のパフォーマンスは P32 と比べておよそ 1.3 倍程度高速化されている。HA8000 及び P32 システムの 1 コア当たりの理論ピーク性能はそれぞれ 9.2GFlops と 4GFlops であるため、HA8000 の方が 2.3 倍高速なはずであり、実測値はこれよりもかなり差が小さいことになる。この理由を知るために profile を調べた所、数値ライブラリに依存する部分は大体理論性能に比例したパフォーマンスが出ているが、それ以外の部分、特にイレギュラーな演算を行っているルーチンでほぼ同程度の時間を消費していることが分かった。これは恐らく、P32 の場合には 1 クロック当たり 2 演算を行うことで理論ピーク性能を達成できるが、HA8000 の場合には同時に 4 演算を行う必要があり、規則性が低い演算の場合には性能を出すことがより困難であるためと考えられる。

次に並列性能について考察するため、16 並列のジョブを色々な条件の下で実行し、実行時間を測定した。並列化はフラット MPI で行い、OpenMP は一切使用していない。また数値ライブラリも全てシングル・スレッドモードで使用した。まず 1 ノード 16 コアを使用し、デフォルトの設定で測定した場合と、プロセスが連続したコアに割り振られるように自分で NUMA を指定した場合を比べると、以下のようになった：

計算条件	実行時間 (秒)
NUMA 指定なし	7.85
NUMA 指定あり	7.42

なお、NUMA は以下のようなスクリプトに従って指定した：

```
#!/bin/bash
MYRANK=$MXMPI_ID
TEMP=$(expr $MYRANK % 16)
NODE=$(expr $TEMP / 4)
/usr/bin/numactl --membind $NODE --cpunodebind $NODE $@
```

FEMTECK の場合には、隣接するプロセス間ほど大規模な通信を行う仕様になっているため、このような指定が有効であると考えられる。

ただし、NUMA を指定した場合であっても、1 コアの場合と比べたパフォーマンスは 12 倍弱であり、並列性能としてはやや不満足である。この点を明らかにするため、(かなり贅沢な使い方であるが) 1 ノード当たり 1 コアのみを使用し、16 ノード 16 コアを使用して全く同じジョブを流して実行時間を測定した。この時の実行時間は 5.70 秒であり、15.3 倍の高速化が得られた。この結果から、1 ノード 16 コアを使用した場合には、

- 4 つのコア毎に一つのメモリを共有しているため、バンド幅が不足する
- ノード内の通信が干渉する

といった現象の影響が大きいと推測される。

とはいえ、トータルの性能ではやはり全てのコアを使用した方が速いのは確かであり、以下のベンチマークでは特に断らない限りノード毎に 16 コア全てを使用し、NUMA を指定した上で測定するものとする。このような条件で、システム B を用いて 16 コア (1 ノード) から 512 コア (32 ノード) まで並列性能を測定した：

コア数	実行時間 (秒)	パフォーマンス
16	230.2	1
32	124.0	1.86
64	64.2	3.59
128	34.5	6.66
128 (NUMA なし)	47.1	4.89
256	20.0	11.5
512	12.6	18.3

大体この程度の系が今までの計算資源で十分な統計精度で扱うことができる限界に近いサイズであったが、HA8000 の大規模キューを使用すれば 2 ~ 3 日で 10^4 ステップ程度のデータを集めることができるようになった。また、試しに NUMA を指定せずに測定した所、かなり大きな差が付いた。並列度が高い場合の方が NUMA の影響が大きいようである。

最後に、月に一度の大規模ジョブ実行サービスを利用して、システム C で最大級のベンチマークを取った。なお、メモリの制約上、動作させるためには最低でも 512 コアが必要であった。また、1024 コアの場合は時間の都合でデータを取ることが出来なかった。ベンチマーク結果は以下ようになった：

コア数	実行時間 (秒)	パフォーマンス
512	559.5	1
2048	208.9	2.68
4096	161.7	3.46

基本的には系が大きい方が並列性能が出易いはずであるが、その割にはやや不本意な性能であった。更に profile を見ると、以下のような点が問題であることが分かった：

- 細かい通信を頻繁に行うルーチンのコストが大きい
- 一部、並列化していないルーチンがあり、コア数が多い場合には（相対的に）かなり時間を消費している
- 数値ライブラリに依存する部分は 512 → 4096 コアの際に行列の次元が減る分、実効 Flops が 10% 強低下する

これらの点に配慮し、重複計算している部分を並列化する等の改善を行うことにより、512 → 4096 コアの際に 5 倍程度の高速化は達成できると見積もっている。現状では HA8000 上でのプロダクション・ランで 4096 コアを使うことはあまり現実的ではないが、次世代スパコンでは通常の粒度になる可能性もあり、その場合にはかなり細かい部分まで漏れなく並列化することが重要になりそうである。

参考文献

- [1] 小出 昭一郎 著：「量子力学 I・II」 [改訂版] (裳華房, 1990).
- [2] 金森 順次郎、米沢 富美子、川村 清、寺倉 清之：「固体 - 構造と物性」(岩波書店, 2001).
- [3] R. M. Martin: *Electronic Structure: Basic Theory and Practical Methods* (Cambridge University Press, New York, 2004).
- [4] A. E. Mattsson *et al.*, *Modelling Simul. Mater. Sci. Eng.* **13** (2005) R1.
- [5] M. J. Gillan *et al.*, *Rep. Prog. Phys.* **69**, 2365 (2006).
- [6] S. Rauegi, F. L. Gervasio, and P. Carloni, *Phys. Stat. Sol. B* **243** (2006) 2500.
- [7] *Topics in Current Chemistry* **268** (2007) :「Atomistic Approaches in Modern Biology」特集号.
- [8] *固体物理* **39** (2004) 11月号: 「計算機ナノマテリアルデザイン」特集号.
- [9] A. Szabo, N. S. Ostlund : 「新しい量子化学 (上・下)」(東京大学出版会, 1987).

- [10] 矢川 元基、吉村 忍 著 : 「有限要素法」(培風館, 1991) .
- [11] 手塚 明、土田 英二 著 : 「アダプティブ有限要素法」(丸善, 2003) .
- [12] T. L. Beck: Rev. Mod. Phys. **72** (2000) 1041.
- [13] L. R. Ram-Mohan: *Finite Element and Boundary Element Applications in Quantum Mechanics* (Oxford University Press, New York, 2002).
- [14] S. Goedecker: Rev. Mod. Phys. **71** (1999) 1085.
- [15] G. E. Scuseria: J. Phys. Chem. A **103** (1999) 4782.
- [16] ナノシミュレーション技術ハンドブック委員会 編 : 「ナノシミュレーション技術ハンドブック」(共立出版, 2006).
- [17] S. R. White, J. W. Wilkins, and M. P. Teter, Phys. Rev. B **39** (1989) 5819.
- [18] J. Ackermann, B. Erdmann, and R. Roitzsch, J. Chem. Phys. **101** (1994) 7643.
- [19] E. Tsuchida and M. Tsukada, Phys. Rev. B **52** (1995) 5573.
- [20] E. Tsuchida and M. Tsukada, Phys. Rev. B **54** (1996) 7602.
- [21] E. Tsuchida and M. Tsukada, J. Phys. Soc. Jpn. **67** (1998) 3844.
- [22] S. Yamakawa and S. Hyodo, Phys. Rev. B **71** (2005) 035113.
- [23] Y. Shiihara, O. Kuwazuru, and N. Yoshikawa, J. Solid Mech. Mater. Eng. **2** (2008) 95.
- [24] 櫛田 慶幸、奥田 洋司 : 日本計算工学会論文集 No. 20040023 (2004).
- [25] J. E. Pask and P. A. Sterne: Modelling Simul. Mater. Sci. Eng. **13** (2005) R71.
- [26] T. Torsti *et al.*, Phys. Stat. Sol. B **243** (2006) 1016.
- [27] J. -L. Fattebert, R. D. Hornung, and A. M. Wissink, J. Comp. Phys. **223** (2007) 759.
- [28] 岡田 勲, 大澤 映二編 : 「分子シミュレーション入門」(海文堂, 1989).
- [29] P. Hohenberg and W. Kohn, Phys. Rev. **136** (1964) B864.
- [30] W. Kohn and L. J. Sham, Phys. Rev. **140** (1965) A1133.
- [31] R. Car and M. Parrinello, Phys. Rev. Lett. **55** (1985) 2471.
- [32] W. E. Pickett, Comp. Phys. Rep. **9** (1989) 115.

- [33] J. A. Appelbaum and D. R. Hamann, Phys. Rev. B **8** (1973) 1777.
- [34] S. Goedecker, M. Teter, and J. Hutter: Phys. Rev. B **54** (1996) 1703.
- [35] P. M. Prenter, *Splines and Variational Methods* (Wiley, New York, 1975).
- [36] W. H. Press 他 著、丹慶勝市他 訳、「ニューメリカルレシピ・イン・シー C 言語による数値計算のレシピ」(技術評論社, 1993) .
- [37] F. Gygi, Phys. Rev. B **48** (1993) 11692.
- [38] N. A. Modine, G. Zumbach, and E. Kaxiras, Phys. Rev. B **55** (1997) 10289.
- [39] 茨木俊秀、福島雅夫: 「FORTRAN77 最適化プログラミング」(岩波書店, 1991) .
- [40] D. C. Liu and J. Nocedal, Math. Prog. **45** (1989) 503.
- [41] E. Tsuchida, J. Phys. Soc. Jpn. **71** (2002) 197.
- [42] T. A. Arias, M. C. Payne and J. D. Joannopoulos: Phys. Rev. Lett. **69** (1992) 1077.
- [43] E. Tsuchida, J. Phys. Soc. Jpn. **76** (2007) 034708.
- [44] E. Tsuchida, J. Phys.: Condens. Matter **20** (2008) 294212.
- [45] 荒川 忠一 著 : 数値流体力学 (東京大学出版会, 1994).
- [46] Y. Saad, SIAM J. Sci. Stat. Comput. **6** (1985) 865.
- [47] O. Tatebe and Y. Oyanagi, in: Proceedings of Supercomputing '94 (IEEE Computer Society, 1994).
- [48] J. R. Chelikowsky, N. Troullier, K. Wu and Y. Saad, Phys. Rev. B **50** (1994) 11355.
- [49] E. L. Briggs, D. J. Sullivan, and J. Bernholc, Phys. Rev. B **54** (1996) 14362.
- [50] M. P. Allen and D. J. Tildesley: *Computer Simulation of Liquids* (Oxford Science Publications, Clarendon, Oxford, 1987).
- [51] 上田 顕 著 : 「コンピュータシミュレーション」(朝倉書店, 1990)
- [52] H. J. C. Berendsen *et al.*, J. Chem. Phys. **81** (1984) 3684.
- [53] H. C. Andersen: J. Comp. Phys. **52** (1983) 24.
- [54] E. Tsuchida, Y. Kanada, and M. Tsukada, Chem. Phys. Lett. **311** (1999) 236.
- [55] E. Tsuchida: J. Chem. Phys. **121** (2004) 4740.

- [56] E. Tsuchida: J. Phys. Soc. Jpn. **75** (2006) 054801.
- [57] Y. -K. Choe, E. Tsuchida, and T. Ikeshoji, J. Chem. Phys. **126** (2007) 154510.
- [58] Y. -K. Choe, E. Tsuchida, T. Ikeshoji, S. Yamakawa, and S. Hyodo, J. Phys. Chem. B **112** (2008) 11586.
- [59] P. Umari and A. Pasquarello, Phys. Rev. Lett. **89** (2002) 157602.
- [60] I. Souza, J. Íñiguez, and D. Vanderbilt, Phys. Rev. Lett. **89** (2002) 117602.

粒子法シミュレーションの大規模化と高速化

越塚 誠一

東京大学大学院工学系研究科

1. はじめに

ここでの粒子法は、流体や固体といった連続体に関する力学を粒子の運動に離散化してシミュレーションをおこなう方法を意味しています。具体的には、MPS(Moving Particle Semi-implicit)法やSPH(Smoothed Particle Hydrodynamics)法です。また、粉体を粒子の集まりとして計算するDEM(Discrete Element Method)も加えたいと思います。分子動力学も分子を粒子として計算する方法ですので、プログラミング上は共通するところも多いはずですが、筆者の専門外ですので、ここでは除外することにします。

以下、粒子法シミュレーションのプログラミングに関する特徴、計算例、大規模計算と高速計算、計算結果の信頼性について、最新の研究成果や動向を含めて述べたいと思います。

2. 粒子法と格子法はプログラミングする上で何が違うのか

従来の格子法(有限差分法や有限要素法)でも粒子法でも連続的な空間を離散化することは同じです。図1のように、格子法では空間を格子状に分割して節点に速度などの変数を配置します。一方、粒子法は空間の中に粒子を配置し、その粒子が速度などの変数を持ちます。プログラムではどちらの場合も変数を配列として確保し、その添字が節点番号あるいは粒子番号を意味します。プログラミング言語の構造体の概念を用いれば、各節点あるいは各粒子の持つ様々な変数を1つの構造体の中に記述することもできます。いずれにしても、変数を配列として確保するところは格子法と粒子法で変わりません。

格子法と粒子法の違いは以下になります。(1) 格子法では通常は格子点の座標は固定されていますが、粒子法では粒子の座標は変化します。(2) 格子法では格子点上に配置された変数の関係式を格子線が陽に表していますが、粒子法では近傍の粒子(影響半径 r_e より近い粒子)を時間ステップ毎に探索してこれらとの間に一時的な関係式を作ります。

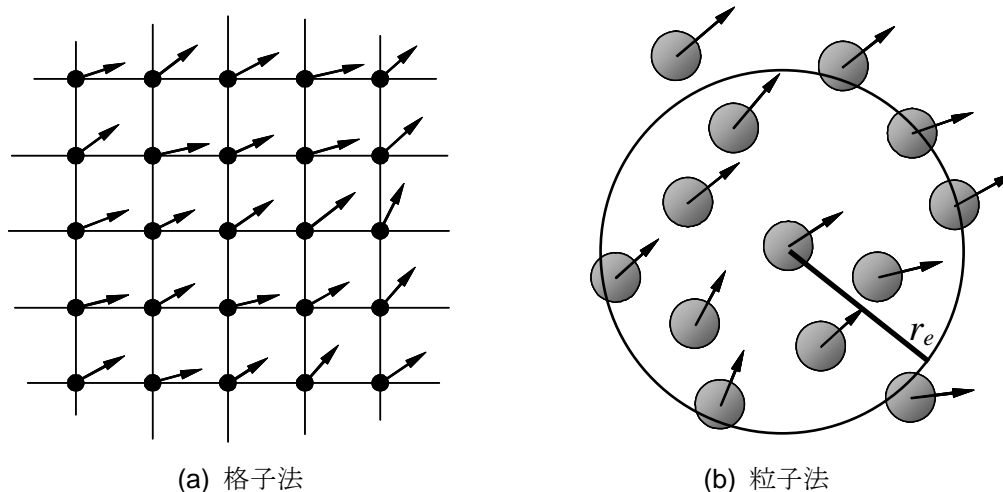


図1 格子法と粒子法

特徴(1)に関しては、粒子の座標を変数として扱う必要があるため計算容量が増加し、その変数の更新のために計算量が増加しますが、これらは相対的に無視できます。特徴(2)に関しては、近傍粒子の探索に粒子間距離を用いますが、この計算には全粒子数を N とすると、 $O(N^2)$ の計算量になります。これは連立一次方程式の解法よりも N の指数が大きいいため、大規模計算では計算時間に対して支配的になります。こうした粒子法の特徴に対する計算アルゴリズムの工夫については後で述べます。

次に、連続体力学の支配方程式は偏微分方程式で記述されますが、これをどのように離散化するかという問題があります。有限差分法では格子幅が差分の基準になりますし、有限要素法では要素の中で積分をおこなって離散式を導出します。こうした格子や要素が無く、単に粒子があるだけで偏微分方程式をどうやって離散化するかが粒子法のアイデアの本質的なところで

す。MPS 法では粒子間の差分に基づいた粒子間相互作用モデルを用います[1]。差分は、2つの粒子があればその間で作ることができます。ある粒子位置の勾配演算子の粒子間相互作用モデルは、まずその粒子と近傍の粒子との間に差分を作り、その差分の重み平均にします(図2)。同様に発散演算子やラプラシアンに対しても粒子間相互作用モデルを用意します。そして、これらを支配方程式の空間微分に適用すれば、粒子が互いに相互作用しながら運動する式が得られます。これが支配方程式の離散式になります。

SPH 法ではカーネル関数に粒子の持つ変数を掛けて重ね合わせます(図3)[2]。こうすると、格子が無くても変数の空間分布を作ることができます。この変数の空間分布を微分すれば空間微分の離散式が得られます。

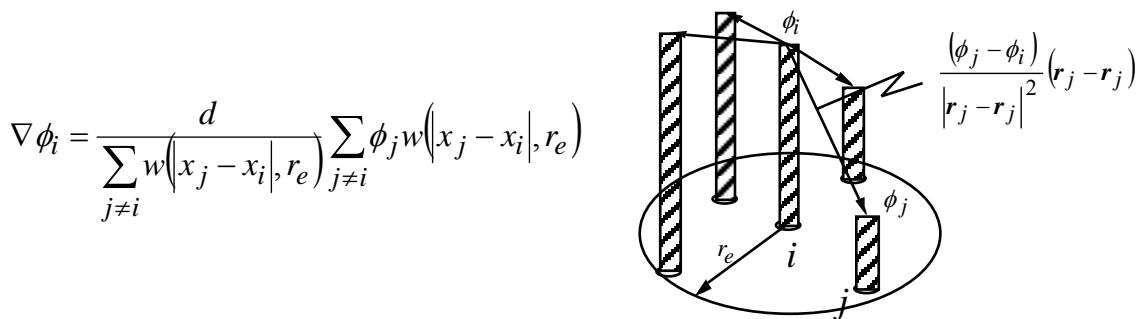


図2 MPS 法における勾配演算子の粒子間相互作用モデル

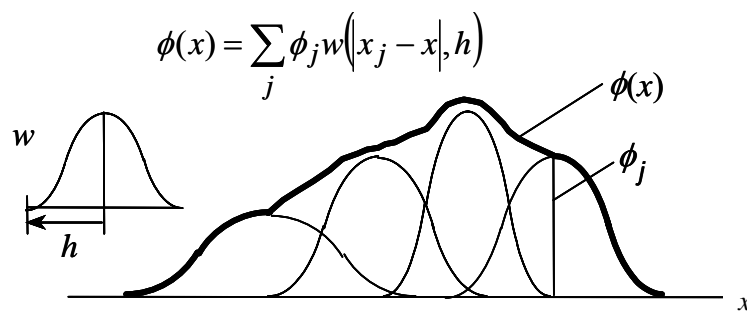


図3 SPH 法における空間分布の構築

MPS法で用いる重み関数 w と、SPH法で用いるカーネル関数 w は、粒子間の離散式を作るうえでほぼ同じ働きがあります。例えばMPS法では多くの場合、以下の重み関数を使います。

$$w(r) = \begin{cases} \frac{r_e}{r} - 1 & 0 \leq r < r_e \\ 0 & r_e \leq r \end{cases} \quad (1)$$

ここで r は粒子間距離、 r_e は影響半径です。 $r < r_e$ であれば w はゼロでない正の値になり、 $r \leq r_e$ ではゼロになります。これによって粒子間相互作用を近傍粒子との間に限定できます。ある粒子に対して、それ以外の全ての粒子との間で相互作用を計算し、重み関数がゼロになることによって遠方の粒子との相互作用を自動的にゼロにできます。ただし、これは $O(N^2)$ の計算量になってしまいます。そこで、近傍粒子を探索し、 $w=0$ とならない近傍粒子とのみ相互作用の計算をすることで、計算量を $O(N^1)$ に落とします。式(1)では $r=0$ で $w \rightarrow \infty$ になりますが、MPS法では同じ粒子との間で粒子間相互作用を計算することは無いので、問題ありません。

MPS法もSPH法も、このように格子を用いることなく微分方程式を離散化する一般的な方法になっています。そのため、流体力学だけでなく固体力学にも適用することができます。粒子法は格子を用いないので、流体力学では自由表面流や気液二相流の液面や界面が大きく変形するような問題、固体力学では大変形や破壊の問題、さらには流体-構造連成問題に適しています。

3. 粒子法の計算例

粒子法がどのような問題に適用されてきたかを紹介します。全体を2種類（エンジニアリングと物理ベースコンピュータグラフィックス）に分けます。それぞれ書籍にまとめられています[3, 4]。

3.1 エンジニアリング

粒子法による非圧縮性流れのシミュレーションでは、自由表面を有する流れを扱うものが最も多いでしょう。その中で基本的なのがダム崩壊問題です。MPS法の最初の論文でもこの問題が解かれています[1]。図4にあるように、初期に水槽の一方の垂直壁に接して水柱を配置し、これが重力で崩れる様子を計算するものです。この問題では、下向きにかかる重力が流体の非圧縮条件を通じて横方向の力に変換され、初期の水柱が広がっていきます。底面上の水の先端の動きは実験と計算で定量的に比較されます(図5)。やがて、もう一方の垂直壁に衝突して水が跳ね上がり、時間の経過とともにその跳ね上がった水が落下します。こうしたふるまいは実験と計算でよく一致しています。

MPS法は自由表面流れに次いで混相流に関する計算例が多くあります。混相流では界面があり、場合によってはそれが大きく変形するので、自由表面流れと同様に粒子法に適した問題です。計算の1例を図6に示します[5]。加熱された底面の上で大気圧の水が沸騰する過程を詳細に計算したものです。加熱壁に配置された初期の気泡が加熱によって成長し、やがて浮力により底面から離脱して上昇します。この問題では壁から離れた位置の水はサブクール（沸点より低い温度）になっているので、加熱壁に接した気泡がある大きさまで成長すると、加熱壁付近での蒸発量と壁から離れた位置での凝縮量がバランスして成長が止まります。また、気泡が加熱壁を離脱して上昇する時には、凝縮によって気泡の体積が減少します。この計算に用いられている粒子法はMPS-MAFL(Meshless Advection Flow-directional Local grid)法で、粒子の

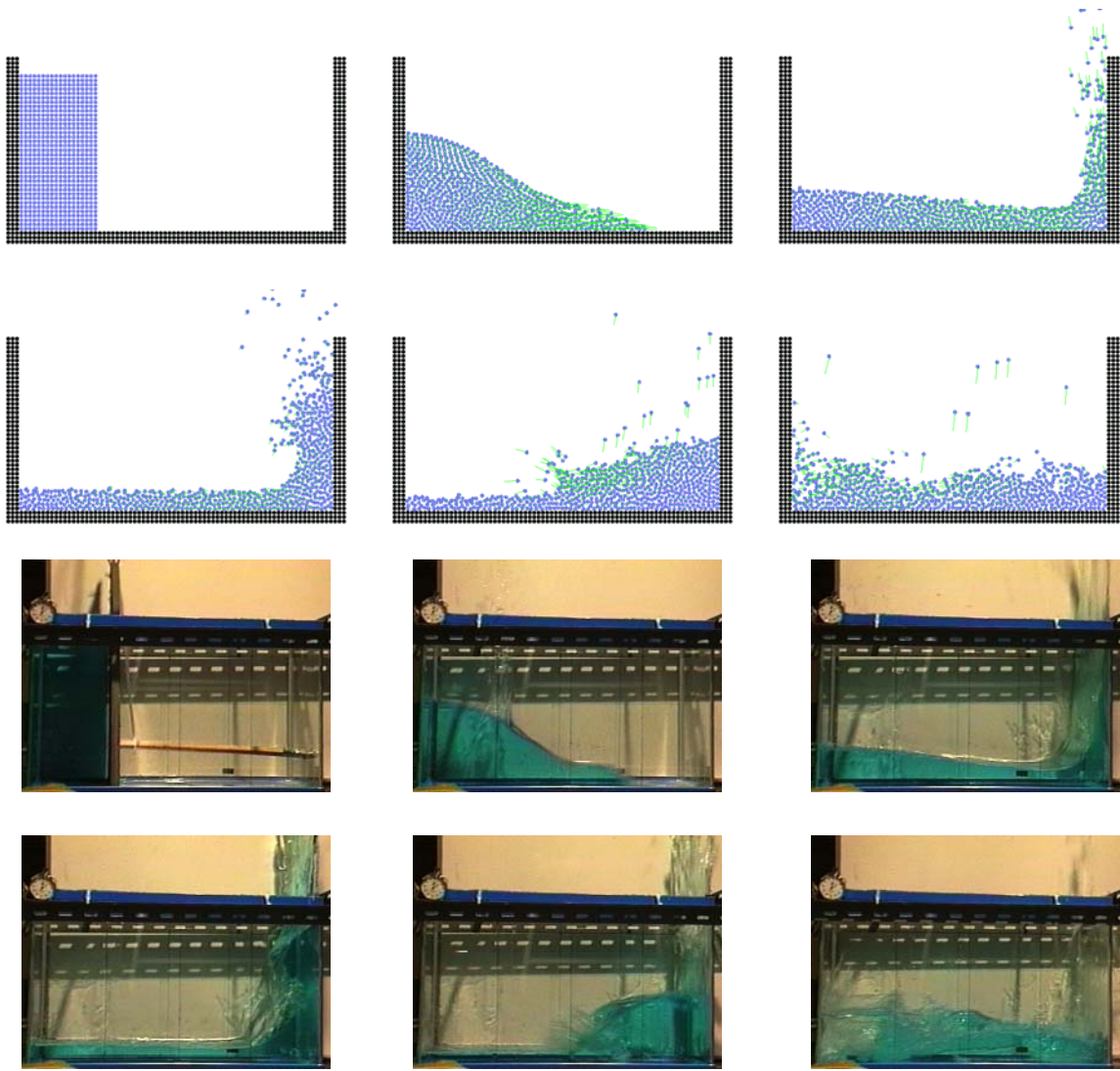


図4 ダム崩壊問題のMPS法による計算と実験

位置を流れとは関係なく任意に決めることができます。これによって、加熱壁と気泡の界面の近傍に小さな粒子を配置し、局所的に空間解像度を上げることができます。もし粒子が流体と一緒には動いてしまうと、最初に小さな粒子を加熱壁近傍に配置したとしても、時間の経過とともに移動して壁から離れてしまいます。粒子位置を流れとは関係なく決められることで、空間解像度を上げたい場所だけに小さな粒子を配置することができます。また、気泡の成長や上昇とともに小さな粒子の数を増やしており、計算中に総粒子数を変えることも自由にできます。定量的な比較では、気泡半径の時間変化や伝熱量が実験によく一致しています。

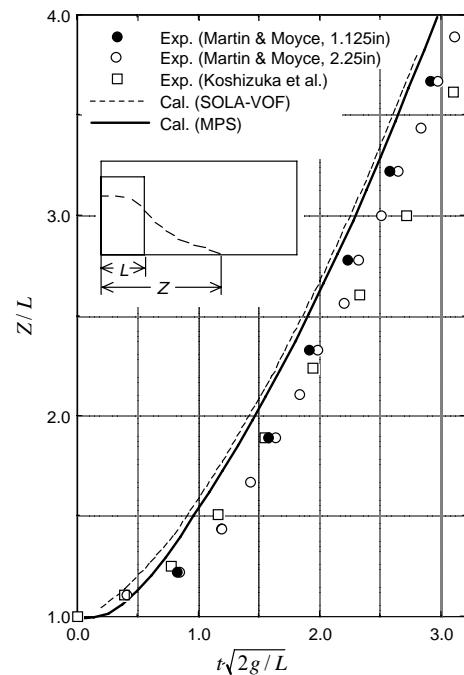


図5 ダム崩壊における底面上の水の先端位置の変化

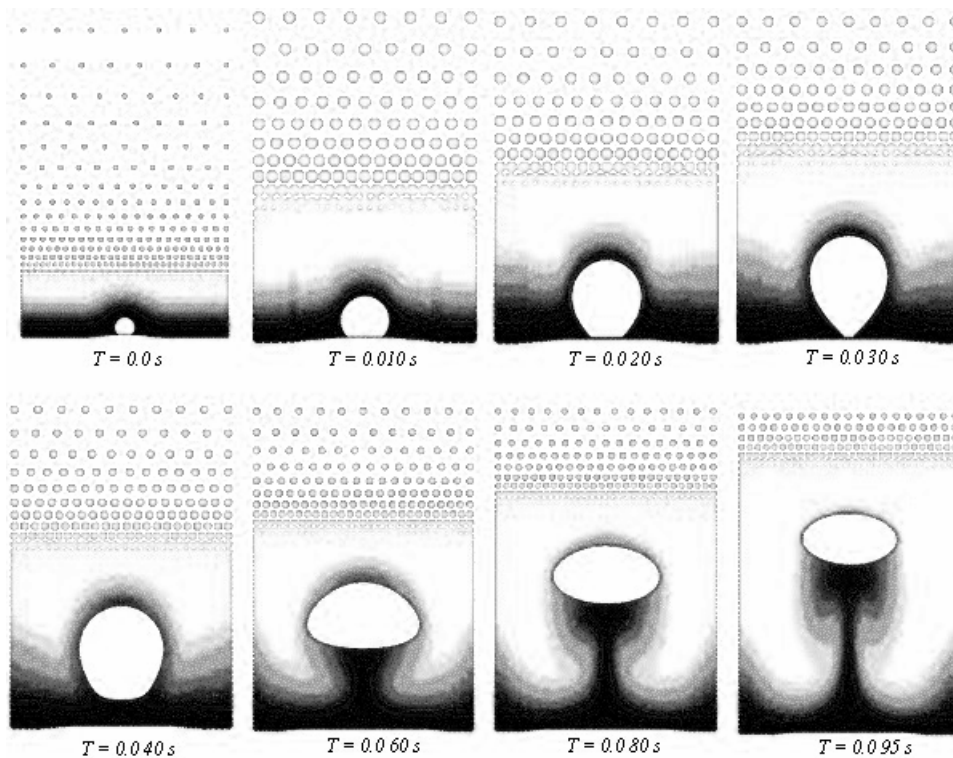
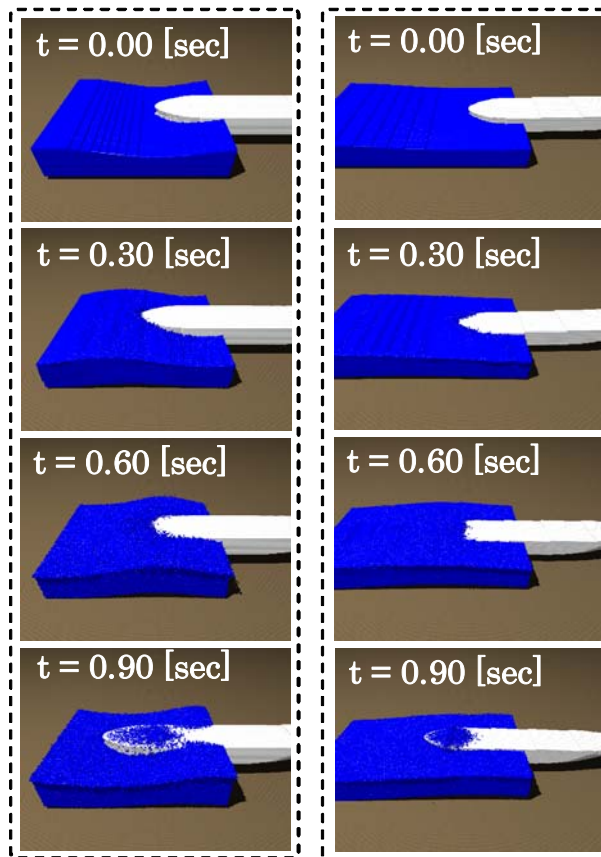


図6 サブクール核沸騰における単一気泡の成長と離脱のシミュレーション（濃淡は温度）

次に、船舶と波浪との相互作用を計算した例を示します。波浪中を航行する船舶では、波が甲板にまで打ち込む現象が起きます。これをMPS法で解析しました[6]。計算結果を図7に示します。この例では模型船を用いた実験が行われていますが、実験で用いた大きな水槽の全体を計算領域に取ることは難しく、ここでは粒子数を節約するために計算領域をできる限り小さくしています。模型船の前方に約1周期分の線形波を作り、側面や底面の壁は線形波の速度分布を用いて動かして、なるべく壁境界の影響が出ないようにしています。模型船の運動は実験では波との相互作用によって生じるのですが、この計算では強制運動として実験で測定されたものを与えています。図7には2例あり、それぞれ波長船長比（=波長／船長）が0.7と1.5の条件になっています。波長船長比が0.7の場合は船体運動が小さく、波頭に



(a) 波長船長比 0.7 (b) 波長船長比 1.5

図7 船舶甲板への打ち込み水挙動の計算結果

船舶の先端が突入するようにして打ち込みが生じます。波長船長比が 1.5 の場合は船体運動が大きくなり、水面をすくい上げるようにして打ち込みが生じます。こうしたふるまいは実験と良く一致しています。

マイクロサイズの流路を用いた装置は μ TAS (micro Total Analysis System) と呼ばれています。例えば、水相と有機相を単純な Y 字型のマイクロ流路によって合流させるだけで、ほぼ均一なマイクロ液滴を生成することができます。しかしながら、この問題は界面の大変形を伴う液-液の二相流で、しかも表面張力が相対的に大きいので計算が難しくなります。そこで MPS 法の適用を試みました。これについても実験がおこなわれた条件で計算しています。計算結果の 1 例が図 8 です [7]。幅 $100\mu\text{m}$ の流路に水とブチルアセテートを流入させています。計算条件は、水相の流量 $Q_{\text{water}}=8.0\mu\text{l}/\text{min}$ 、流入流量比 $R=Q_{\text{org}}/Q_{\text{water}}=1.0$ となっています。合流部で有機相が分裂してマイクロ液滴が生成されています。水相の流量と流入流量比を様々に変えて計算をおこない、液滴サイズに関して計算と実験を比較したところ、定量的に良い一致が得られました。

粒子法を用いたシミュレーションにはこの他にも多くあります。特に、土木工学の分野では後藤仁志先生が精力的に研究を進めています [8]。原子力工学においても高速炉の過酷事故解析のために、MPS 法を用いた多相多成分流れの解析コードの開発が進められています [9]。構造力学の分野では酒井譲先生が SPH 法を使って様々な問題を解いています [10]。MPS 法による例もあります [11]。

3.2 物理ベースコンピュータグラフィックス

物理的に正しいシミュレーションをおこない、その結果にもとづいたコンピュータグラフィックスにより映像を制作することを、物理ベースコンピュータグラフィックス (物理ベース CG) と呼びます。流体解析、固体解析、剛体解析、布の解析、相変化などの様々な物理現象の粒子法シミュレーション手法がこれまでに研究されています。流体解析や固体解析では MPS 法や SPH 法といった方法が使われるのに対し、剛体解析では DEM (Discrete Element Method) が使われます。粒子法では界面の複雑な動きのような激しい現象を扱うことが容易なため、映像制作のためのシミュレーション手法に適しています。また、粒子法では粒子が新たに現れたり消えたりしなければ質量保存が成立します。多少の計算誤差があっても質量保存が崩れないという性質は高速計算の上で利点です。さらに様々な物理現象を粒子間相互作用という統一的な考え方で計算できるため、マルチフィジックスシミュレーションが比較的容易にできます。そのため、物理ベース CG において粒子法は広く用いられるようになってきています。

例えば、粒子法によってコップに水を注ぐシミュレーション

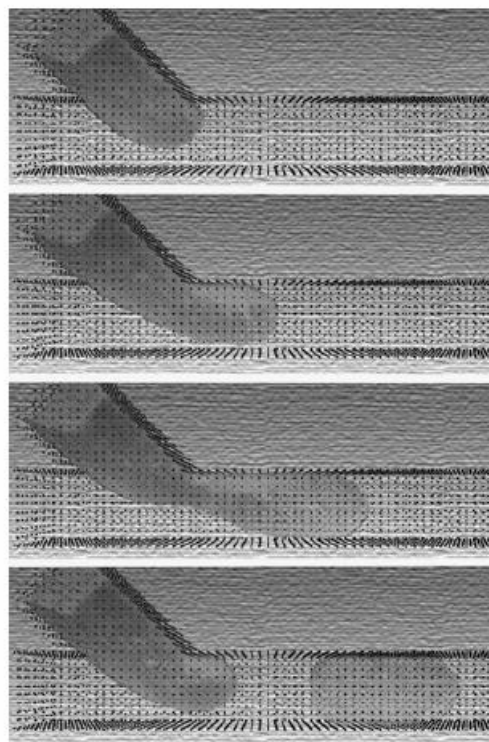


図 8 マイクロ液滴生成の計算結果

をおこない、その結果にもとづいて作成した画像が図9です[12]。この例では気泡の計算モデルが組み込まれていて、これによって画像が本物らしくなっています。図10は剛体の物理ベースCGです[13]。1個の剛体を複数の粒子で形状を模擬し、剛体間の衝突や摩擦は粒子にもとづいて計算します。従来のポリゴンにもとづいた方法と比較して、剛体間の衝突や相互作用をきわめて効率的に計算することができます。描画はポリゴンにもどしておこないます。同様に、弾性体を粒子群に離散化してMPS法を適用した物理ベースCGを図11に示します[14]。この場合、変形に対応して弾性体を構成する粒子間の相対位置が変化します。この変形の情報ポリゴンの頂点の座標に反映させて描画しています。連成解析の例として、流体と剛体を組み合わせた例を図12に示します[15]。積み上げたグラスは剛体



図9 流体の物理ベースCG

として計算され、そこに赤い色の流体を注いでいます。新たなグラスを投げつけると積み上げたグラスが崩れますが、これは剛体としてのシミュレーションをおこなっているため、得られる映像は本物らしくなります。剛体も流体も粒子によって計算されているため連成は容易です。

こうした物理ベースCGの技術は、映画、テレビ番組、コンピュータゲームといった映像産業に既に広く用いられています。また、映像産業だけでなく、災害対策、仮想外科手術、操船シミュレータ、バーチャルファッションショーなど、様々な産業応用も考えられています。

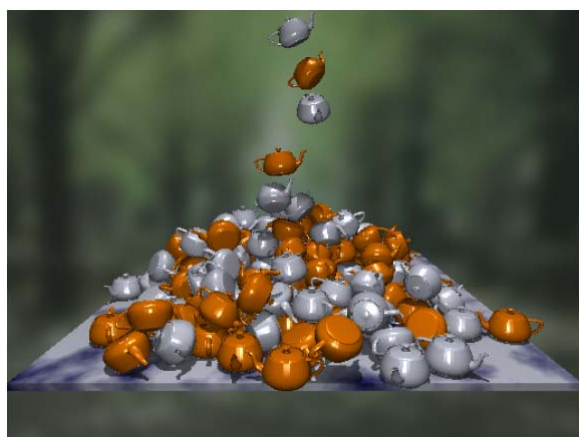


図10 剛体の物理ベースCG

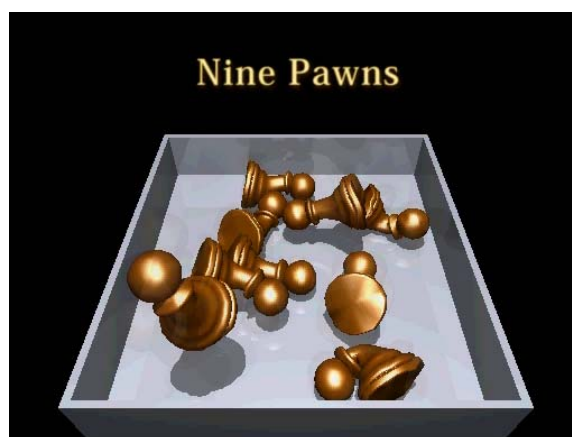


図11 弾性体の物理ベースCG

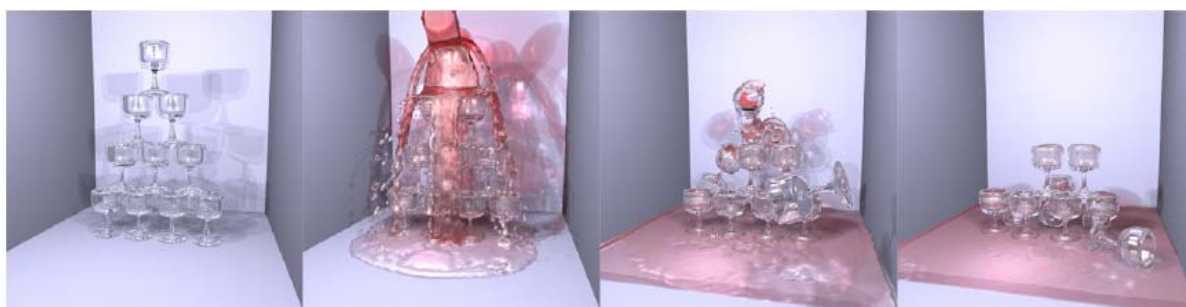


図12 流体と剛体の連成の物理ベースCG

4. 大規模化と高速化

4.1 粒子法特有の問題に対するアルゴリズムの工夫

2章で粒子法のプログラミング上の特徴をまとめました。ここではその特徴に対する並列計算手法などのアルゴリズムの工夫を述べます。図1に戻って、大規模高速計算の観点から粒子法の特徴を整理してみます。

まず第1に粒子の座標が変数で、時々刻々変化するということがありました。この計算は、粒子 i の古い時刻の座標と速度をそれぞれ $r[1][i]$, $u[1][i]$, 新しい時刻の座標と速度をそれぞれ $rn[1][i]$, $un[1][i]$ とすると (1 は空間次元), C 言語では以下のように書けます。

$$rn[1][i]=r[1][i]+dt*un[1][i]; \quad (2)$$

すなわち、古い時刻の粒子座標 $r[1][i]$ に、時間刻み幅 dt と新しい時刻の粒子速度 $un[1][i]$ を掛けたものを足します。粒子座標の更新の計算は、どのような物理現象を解く場合でも基本的には共通です。

新しい時刻の速度 $un[1][i]$ は運動量保存則から計算します。

$$un[1][i]=u[1][i]+dt*f(r, u); \quad (3)$$

運動量保存則における右辺の $f(r, u)$ は流体であれば流体の運動量保存則, 弾性体であれば弾性体の運動量保存則によって計算します。つまり、解きたい物理現象によってここは異なります。式(2)と(3)の順番については、式(3)の計算が先で、その結果である $un[1][i]$ を用いて式(2)を計算します。

さて、粒子の座標が時々刻々変化するのを計算するのは具体的には式(2)であり、相対的には無視できる程度の計算量にしかありません。粒子数を N とすると $O(N)$ の計算量であり、大規模になっても相対的な計算量は増えません。

第2に各時間ステップで近傍粒子を探索する必要があります。図1に示すように粒子法では影響半径 r_e 内の近傍粒子を探索します。この計算のために全ての粒子の間の距離を計算し、これが r_e 以下かどうかを判定すると、 $O(N^2)$ の計算量になります。距離の計算は単純で粒子数が少なければ計算量は少ないのですが、連立一次方程式の解法の計算量が $O(N^{1.5})$ 程度ですので、粒子数が多くなると距離の計算に必要な計算量は相対的に増加し、100万粒子ともなるとほとんど全ての計算時間を距離計算に費やしてしまうこととなります。そこで、これを回避するためにMPS法ではこれまで2種類の方法が使われてきました。

1つは、より大きな影響半径による近傍粒子テーブルをもう1つ作成しておくことです。この広範囲の近傍粒子テーブルから影響半径 r_e の本来の近傍粒子テーブルを作成します。一方、広範囲の近傍粒子テーブルは毎時間ステップに更新する必要はありません。これで計算量は $O(N^2)$ から $O(N^{1.5})$ に落ちます[16]。もう1つの方法は空間を格子状に分割して、粒子がどのバケット(格子線で囲まれた領域, セル)に入っているかを登録しておくものです。ある粒子の近傍粒子は、その粒子の入っているバケットとその周囲のバケットに限定して探索すればよく、計算量は $O(N^{1.0})$ まで落ちます[17]。後者の方法は実装が簡単なので、現在のMPS法の大規模計算では広く使われています。

格子法において離散式を構成する格子点の数は、2次元では図1を見ると線につながっている上下左右および自分を合わせて5点になります。一方、粒子法においては影響半径 r_e 内にある粒子数は10個を数えます。このように、粒子法において離散式を構成する粒子数は、格子法において離散式を構成する格子点よりも多くなります。影響半径 r_e を小さくしていけば近傍

の粒子の数は減るのですが、減りすぎると計算が不安定になるという問題が出てきます。安定に計算ができる程度の大きさの影響半径 r_e を取ると、近傍粒子の数は多くなってしまいます。これは計算時間の増大につながります。

粒子法では離散化式が空間の次元によってほとんど変化しません。粒子間距離の計算に次元の数が影響しますが、粒子法の離散式では粒子間距離を使いますので、粒子間距離を求めた後の計算は空間次元の影響をほとんど受けません。従って、粒子法では3次元計算コードの作成が容易です。

PCクラスタで並列計算をする場合、粒子番号によって分割して異なるCPUに担当させることが考えられます。粒子番号による分割が空間による分割と対応していれば、近傍粒子同士が同じCPUの担当になっている場合が多くなり、CPU間の通信が少なくて済みます。しかしながら、粒子法では粒子が時々刻々移動するため、粒子群がお互いに混ざり合うことになります。そうすると、他のCPUが扱う粒子が近傍粒子となる場合が増え、CPU間の通信が増大します。これは並列計算の効率を悪化させます。そこで、ある程度計算が進んで粒子群が混ざってきたところで、粒子番号を空間分割に合わせて再設定します[18]。こうすると、CPU間の通信量をあまり増大させることなく、粒子法による並列計算を続けることが可能になります。粒子番号の再設定において、近傍粒子探索に用いたバケットを利用することもできます。

また、近傍粒子探索で用いたバケットによって領域分割して、異なるプロセッサに割り当てることも考えられます。この場合、各プロセッサの担当する粒子数が一定ではなくなるという問題があります。複数のGPUを用いた並列計算ではこのバケットを用いた分割の方法が用いられています[19, 20]。

4.2 GPUの利用

GPU(Graphics Processing Unit)は描画を高速におこなうための専用のプロセッサですが、近年プログラムによって動作させることが可能になり、科学技術計算に利用する試みがおこなわれています。プラズマの差分法による計算(ポアソン方程式を解くステップも含む)においてGPUとCPUを比較し20倍の計算速度が得られ、移流項部分だけでは85倍の計算速度が得られたとの報告があります[21]。粒子法のシミュレーションやこれに関連する前処理におけるGPUの利用に関する研究は、原田隆宏氏が多数おこなっています[4, 19, 20, 22]。最近では大規模体系でCPUの約100倍の性能が出たとの報告があります[20]。複数のGPUを用いて粒子法の並列計算をおこない、4GPUで単一GPUの約3倍の性能が得られたと報告されています[19, 20]。この研究ではGPUの並列計算に適した近傍粒子探索アルゴリズムおよび粒子番号の再設定のアルゴリズムが提案されています。GPUは極めて安価ですので、今後ますます科学技術計算に利用されていくのではないかと思います。

4.3 大規模化と高速化のニーズ

シミュレーションの大規模化と高速化には現在においても大きなニーズがあります。シミュレーションの目的は研究、設計、トラブル解明、評価など様々ですが、ここでは実用的なシミュレーション、すなわち産業において使われることを考えてニーズを考察してみます。

まず産業で考えられるのが、製造業における設計や評価への利用です。この場合は計算結果の信頼性が重要です。計算結果の信頼性の中身については5章で議論したいと思います。信頼

性の高い計算をするためには、十分な空間解像度、十分な時間解像度、十分な反復計算の収束が必要であるだけでなく、複雑な形状の製品の模擬性を良くしたり、様々な物理現象の複合的な効果を考慮する必要があります。どの要求に対しても計算負荷の増大につながります。しかも予算と時間の制約の中で実行しなければなりません。現在の計算機的能力では、実際の製品を非常に単純化しなければ設計のための計算ができないという場合がまだ多いでしょう。そもそもシミュレーションでは精度の高い設計ができないという場合もしばしばあると思います。計算の大規模化と高速化が進めば、設計や評価においてシミュレーションが使える対象が増えますし、既にシミュレーションが使われている場合には設計コストの低減、時間短縮、精度向上による設計マージンの低減が進められます。最近ではシミュレーションによる最適化やロバスト設計が試みられるようになっていますが、この場合には入力パラメータを変化させて多数のシミュレーションを実行する必要があり、ここでも計算速度の向上が求められています。

物理ベース CG においても、大規模化と高速化には強いニーズがあります。映画でコンピュータグラフィックスが多用できるのは、計算機のパフォーマンスの向上によってシミュレーションと CG の大規模化と高速化が可能になったためでしょう。映画のための映像制作はオフラインでよいのですが、運転訓練シミュレータやゲームはリアルタイムである必要があります。ここで要求される高速化は非常に厳しいものです。これまでは科学技術計算レベルのシミュレーションはリアルタイムでは不可能でした。しかし最近では、特に GPU を用いたシミュレーションでは、これが現実的になってきています。

4.4 HPC の展望

HPC (High Performance Computing) については、ハードウェア開発および計算機の設置のどちらについても日本の占める位置が相対的に低下しています。例えば、スーパーコンピュータの TOP500 のリストが半年毎に公開されていますが、最新の 2008 年 6 月のリストでは、プロセッサのベンダーは IBM 68, Intel 375, AMD 55, NEC 1, Cray 1, その他 0 となっており、マイクロプロセッサの寡占化が進んでいることがわかります[23]。TOP1 は米国ロスアラモス研究所のマシンで、プロセッサは IBM の PowerXCell 8i, ピーク性能は 1.37PFlops となっています。日本設置のマシンに限定した TOP4 (世界では 16, 20, 24, 34 位) では、いずれも AMD のマイクロプロセッサが使われています。

4.2 で述べたように、最近 GPU を用いた科学技術計算の高速化が注目されています。単精度計算ですが、わずか 2 万円程度のグラフィックカード (例えば、nVIDIA 社の GeForce8800GTS (G92) 512MB, コア数 128 個) によって CPU よりも 2 桁速い計算速度が得られています。GPU の技術の進歩は速く、既に次の世代の GPU (例えば、nVIDIA 社の GeForce GTX280, コア数 240 個) が発売されており、その性能は以前のもの 2 倍近くになっています。こうした GPU はプログラミング可能で、nVIDIA 社は CUDA という言語を提供しています。一方、Intel 社もこれまでのマイクロプロセッサの延長線上に新しい GPU (Larrabee) を独自に開発しています[24]。

こうしたグローバルな状況において、今後の HPC の動向を予想するには、研究のための科学技術計算における HPC だけを見ていたのでは不十分で、産業利用のための安価で使いやすい HPC の市場や、映像制作やゲームにおける研究開発の動向まで広く見る必要があると思います。まず第 1 に、現状では TOP500 のマシンのうち 498 台が 3 社のマイクロプロセッサに占められており、汎用マイクロプロセッサの開発動向に HPC は従属していると言わざるを得ません。次に、

これに対する新しい動きとして GPU が HPC に参入すると予想されます。GPU を用いた高速な科学技術計算は、リアルタイム性を要求されるゲーム分野の物理シミュレーション技術として注目されており、その動向が今後の HPC のありかたに大きな影響を与えることも考えられます。

5. シミュレーションの信頼性

本稿ではシミュレーションの Verification と Validation (V&V) について紹介します。V&V とはシミュレーション結果の信頼性を定量化する方法で、数値流体力学に対しては米国航空宇宙学会 (AIAA) の 1998 年のガイドライン[25]が、計算固体力学に対しては米国機械学会 (ASME) の 2006 年のガイドライン[26]があります。Verification と Validation ですが、どちらも日本語に訳すと「検証」になります。英語の日常会話でも Verification と Validation は類義語で区別がないそうです。しかし、ガイドラインでははっきりと区別します。

工学の目的でシミュレーションをおこなう手順を考えます (図 1 3)。まず、実際に解きたい現実世界の問題に対して、本質的な物理現象だけを抽出した概念モデルを作ります。この概念モデルを元に計算モデル (計算コード) を作成し、計算結果を得ます。AIAA のガイドラインでは、Verification とは概念モデルと計算結果との一致を定量化するものです。具体的作業としては、概念モデルを記述する微分方程式の解析解と、その微分方程式を離散化して得られた計算モデルの数値解との比較です。両者の違いの原因には、空間の離散化誤差、時間の離散化誤差、反復計算の収束誤差、プログラムミスや入力ミスが考えられます。一方、Validation では現実世界の問題と計算結果との一致を定量化するものです。具体的作業としては、現実世界の問題を反映する実験をおこなって、その実験結果と計算結果の違いを比較します。現実世界の問題は複合問題が多いので、それらを階層構造でとらえて、様々なレベルでの Validation の積み重ねが必要になります。ただし、予算などの制約条件によって V&V は影響を受けるというような記述もあって、現実的なガイドラインになっています。

ASME のガイドラインでも Verification と Validation の定義はほぼ同じです。ただし、Verification を Code Verification と Calculation Verification の 2 つに分けています。また、Validation において使用目的に限定されることを強調しており ISO9000 の考え方が入って

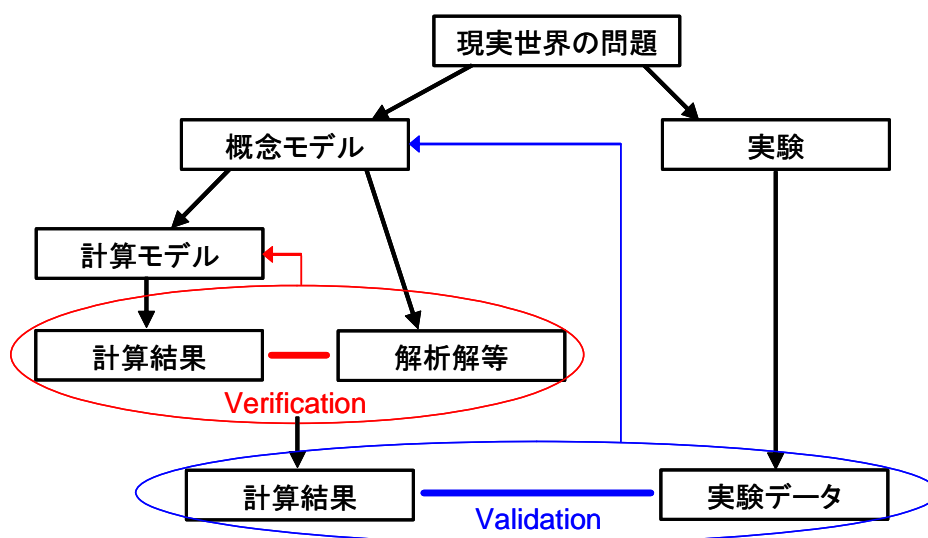


図 1 3 Verification と Validation (V&V) の手順

います。ISO9000 はより一般的なソフトウェア開発における品質保証の技術基準であり、そうした観点から Verification と Validation についての別の定義が示されており、Validation については使用目的に限定されることが書かれています。さらに、ASME のガイドラインでは概念モデルの構築に重要度分類表(PIRT)の利用を推奨しておりますが、PIRT は原子力プラントの統計的な安全解析で用いられている考え方で、それを一般的な V&V に取り入れているというものです。

こうした米国のガイドラインに沿って Verification と Validation をおこなっていくことが、シミュレーション結果の信頼性を定量化するには、現状では最良と思われれます。ただ昨今では、ガイドラインを陽に用いていなくても、これと類似の手順に沿って研究開発がおこなわれている場合が増えていると思います。ガイドラインはそれらを整理して1つの手順にまとめたものと見ることもできます。

ちなみに、これらのガイドラインはシミュレーションや科学技術の本質に関する哲学的な問題にも触れています。例えば AIAA のガイドラインには、prediction は未知の現象の予測であって V&V によって prediction の精度を直接語るができない、と書かれています。この問題は、筆者が 2007 年 7 月に San Francisco で ASME のガイドラインに関する特別セッションに出席したときにも議論になっていました。V&V で語れる精度が postdiction に関するものだけであるとしたら、シミュレーションの有用性はいったい何なのかということになりますね。極論すれば、シミュレーションとは既に分かっていることをコンピュータで再現しているだけだ、ということです。AIAA のガイドラインではこの問題は先送りされていると思います。

一方、ASME のガイドラインには、「科学理論や予測の真実性は演繹的な論理で証明されるものではない。それらは証拠(evidence)の積み重ねによって正しさが試される。証拠は V&V によって系統的に組織化される」という記述があります。これは上記の問題に対する科学技術の態度を適切に言い表していると思います。さらに議論の対象を予測(prediction)だけでなく科学理論(scientific theory)にまで広げていることから、ガイドラインの作成者が科学哲学を意識していると推察できます。

さて、シミュレーションの信頼性は V&V だけで完結するものではありません。プログラムミスや入力ミスは、品質保証の問題と技術者の技量の問題を含んでいます。品質保証の問題には例えば ISO9000 の方法論を適用すればいいですし、技術者の技量に対しては例えば日本機械学会による計算力学技術者認定試験の制度があります。シミュレーションが学術や産業において広く使われるようになってくると、その信頼性が社会にとって極めて重要であることは、少し前の耐震計算に関わる偽装事件を挙げるまでもなく、明らかです。シミュレーションの技術者には結果の信頼性に対して責任があるというような記述が ASME のガイドラインにも書かれています。これは技術者倫理の問題とすることができます。

6. おわりに

本稿では粒子法シミュレーションの特徴や計算例の紹介をおこないました。また、大規模計算と高速計算に関しては、粒子法に関する並列計算手法などのアルゴリズムの工夫と GPU を用いた高速計算を解説した後、大規模化と高速化のニーズと HPC に関する私見を述べました。最後に、計算精度に関しては筆者が最近気になっている米国における V&V のガイドラインを紹介しました。内容が様々なテーマに分散しておりますが、部分的にでも読者の興味に合うところ

があれば幸いです。

参 考 文 献

- [1] Koshizuka, S. and Oka, Y., "Moving-Particle Semi-implicit Method for Fragmentation of Incompressible Fluid," *Nucl. Sci. Eng.*, **123**, 421-434 (1996)
- [2] Lucy, L. B., "A Numerical Approach to the Testing of the Fission Hypothesis", *Astron. J.*, **82**, 1013-1024 (1977)
- [3] 越塚誠一, "粒子法," 丸善 (2005)
- [4] 越塚誠一, 原田隆宏, 田中正幸, 近藤雅裕, "粒子法シミュレーション —物理ベースCG入門—," 培風館 (2008)
- [5] Yoon, H. Y., Koshizuka, S. and Oka, Y., "Direct Calculation of Bubble Growth, Departure and Rise in Nucleate Boiling," *Int. J. Multiphase Flow*, **27**, 277-298 (2001)
- [6] 柴田和也, 越塚誠一, 谷澤克治, "船体の強制運動を伴う海水打ち込みの粒子法による3次元解析," *日本船舶海洋工学講演論文集第1号*, 福岡, November 24-25 (2005) p. 59-60
- [7] 原田隆宏, 鈴木幸人, 越塚誠一, 荒川貴博, 庄子習一, "MPS法を用いたマイクロ混相流3次元解析," *日本機械学会論文集(B)*, **73**, 437-444 (2007)
- [8] 後藤仁志, "数値流砂水理学 —粒子法による混相流と粒状体の計算力学—," 森北出版 (2004)
- [9] Koshizuka, S., Liu, J., Shirakawa, N., Uehara, Y., Naitoh, M. and Yamamoto, Y. "R & D of the Next Generation Safety Analysis Methods for Fast Reactors with New Computational Science and Technology (1) Introduction of the Project and Development of Structural Mechanics Module," *Proc. 16th Int. Conf. on Nuclear Engineering (ICONE16)*, Orlando, May 11-15, 2008, ICONE16-48499
- [10] 酒井讓, "粒子法による構造解析の動向とトピックス," *粒子法研究会 (第十回) 報告書, UTNL-R-0454*, 東京, March 8, 2006, p. 1-14
- [11] 宋武燮, 越塚誠一, 岡芳明, "MPS法による弾性構造体の動的解析," *日本機械学会論文集(A)*, **71**, 16-22 (2005)
- [12] 須藤健太郎, 越塚誠一, "粒子法による泡の運動シミュレーション," *第19回計算力学講演会講演論文集*, 名古屋, November 3-5, 2006, p. 703-704
- [13] Tanaka, M., Sakai, M. and Koshizuka, S., "Rigid Body Simulation Using a Particle Method," *DVD publication at ACM SIGGRAPH posters*, Boston, July 30 - August 3, 2006, 132
- [14] Kondo, M., Tanaka, M., Harada, T. and Koshizuka, S., "Elastic Objects for Computer Graphic Field Using MPS Method," *DVD publication at ACM SIGGRAPH posters*, San Diego, August 5 - 9, 2007, sap-179
- [15] Harada, T., Tanaka, M., Koshizuka S. and Kawaguchi, Y., "Real-time Particle-Based Simulation on GPUs," *DVD publication at ACM SIGGRAPH posters*, San Diego, August 5 - 9, 2007, sap-151
- [16] Koshizuka, S., Nobe, A. and Oka, Y., "Numerical Analysis of Breaking Waves Using the Moving Particle Semi-implicit Method," *Int. J. Numer. Meth. Fluids*, **26**, 751-769

- (1998)
- [17] 後藤仁志, 林稔, 酒井哲郎, “粒子法による直立護岸前面砕波・越波の数値シミュレーション,” *土木学会論文集*, 第 726 号, 87-98 (2003)
 - [18] 入部綱清, 藤澤智光, 越塚誠一, “粒子法による大規模解析におけるノード間通信の低減,” *Trans. Japan Society for Computational Engineering and Science*, Paper No. 20080020 (2008)
 - [19] 原田隆宏, 政家一誠, 越塚誠一, 河口洋一郎, “複数の GPU を用いた粒子法シミュレーションの並列化,” *第 13 回計算工学講演会論文集*, **13**, 仙台, May 19-21, 2008, p. 289-290
 - [20] 原田隆宏, “GPU を使った並列処理で物理シミュレーションを高速化,” *日経エレクトロニクス*, 2008 年 6 月 2 日号, p. 105-115 (2008)
 - [20] 小川慧, 青木尊之, “GPU を用いた CIP 法によるプラズマ 2 流体不安定性の高速シミュレーション,” *第 13 回計算工学講演会論文集*, **13**, 仙台, May 19-21, 2008, p. 837-840
 - [22] Harada, T., “Real-time Rigid Body Simulation on GPUs,” *GPU Gem 3*, Addison-Wesley (2007)
 - [23] <http://www.top500.org/>
 - [24] Seiler, L., Carmean, D., Sprangle, E., Forsyth T., Abrash, M., Dubey, P., Junkins, S., Lake, A., Sugerman, J., Cavin, R., Espasa, R., Grochowski, E., Juan, T. and Hanarahan, P., “Larrabee: A Many Core x86 Architecture for Visual Computing,” *ACM Trans. Graphics*, **27**, Article 18, 2008
 - [25] American Institute of Aeronautics and Astronautics, “Guide for the Verification and Validation of Computational Fluid Dynamics Simulations,” *G-077-1998* (1998)
 - [26] American Society of Mechanical Engineers, “Guide for Verification and Validation in Computational Solid Mechanics,” *ASME V&V10-2006* (2006)

東京大学情報基盤センター・スーパーコンピューティングニュース

Vol.11 No. Special Issue 1 (2009.2)

スーパーコンピューティングニュース編集スタッフ

編集長 中島研吾

編集幹事 西澤明生

編集委員 石川裕, 佐藤周行, 田浦健次朗, 松葉浩也, 堀敦史,
片桐孝洋, 吉廣保, 渡辺宙志, 鴨志田良和, 藤田肇,
平野光敏, 丹下藤夫, 有賀浩, 宮寄洋

編集・発行 東京大学情報基盤センター
スーパーコンピューティング部門

〒113-8658 東京都文京区弥生 2-11-16

(電話) 03-5841-2717 (ダイヤルイン)

(FAX) 03-5841-2708