

生体高分子系の緩和モード解析

高野 宏[†]

(受付 2013 年 12 月 31 日; 改訂 2014 年 7 月 15 日; 採択 7 月 16 日)

要 旨

詳細釣り合いの成り立つシミュレーション法に対して、時間発展演算子の固有値、固有関数から、扱っている系の緩和率、緩和モードをそれぞれ定義することができる。計算機シミュレーションの時系列データから緩和率分布と緩和モードを変分問題に基づいて評価する方法である緩和モード解析の方法について解説する。さらに、ペプチド系のモンテカルロシミュレーション、タンパク質系の分子動力学シミュレーションへ応用した例について紹介する。

キーワード：緩和モード、緩和率、緩和モード解析、シミュレーション、時間相関関数。

1. はじめに

緩和モード解析の方法(Takano and Miyashita, 1995; Koseki et al., 1997; Hirao et al., 1997)は、計算機シミュレーションで得られた時系列データから、扱っている系の緩和モードと緩和率分布を評価する方法である。従来、ランダムスピン系、高分子系に用いられてきた。本稿では、この方法および生体高分子系への応用例(Mitsutake et al., 2011; Nagai et al., 2013)について説明する。

生体高分子であるタンパク質の機能は、その構造の揺らぎと関連があることが知られている。タンパク質のシミュレーションで得られたデータから、構造の揺らぎを解析する方法として、従来、主成分分析の方法が広く用いられている。この方法は、構造揺らぎの分散共分散行列の対角化により、分散の大きい揺らぎのモードを順次求めて行く方法で、構造のデータが時系列としてどのように並んでいるかという情報を用いない静的解析である。これに対し、緩和モード解析の方法は、二つの異なる時間間隔に対する構造揺らぎの時間相関行列に関する一般化固有値問題を解く方法で、時系列としての情報を用いた動的解析である。生体高分子のシミュレーションの同じデータに対して、両方法を適用した結果がどのように異なるかは興味深い問題である。

以下では、第2節で系の時間発展演算子の固有値、固有関数からどのように緩和率、緩和モードが定義されるかについて説明する。第3節では、変分問題に基づいてシミュレーションのデータから緩和モード、緩和率を評価する方法である緩和モード解析の方法について説明する。第4節では、ペプチド系のモンテカルロ・シミュレーションに緩和モード解析を適用した結果について報告し、主成分分析の結果との比較を行う。第5節では、タンパク質系の分子動力学シミュレーションに緩和モード解析を適用した例について報告する。この場合、主成分分析で得られた、揺らぎの大きいいくつかの主成分に基づいて緩和モード解析を行う。その際、速い緩和モードの影響を取り除くために行った工夫についても説明する。第6節でそれまでの内容を

[†]慶應義塾大学 理工学部：〒223-8522 神奈川県横浜市港北区日吉 3-14-1

まとめる。緩和モード解析の方法では重要な物理量の線形結合を変分問題の試行関数としている。固有値方程式と変分問題の関係について付録 A で、試行関数の選び方について付録 B で説明する。

2. 緩和モードと緩和率

詳細釣り合いの条件を満たす三つのシミュレーション方法に対して、時間発展演算子の固有値、固有関数から緩和率、緩和モードがどのように定義されるかについて説明する。

最初のシミュレーション方法は、詳細釣り合いの条件を満たした連続時間のモンテカルロ法である。この方法では、時刻 t に系が状態 Q をとる確率 $P(Q; t)$ の時間発展がマスター方程式

$$(2.1) \quad \frac{\partial}{\partial t} P(Q; t) = - \sum_{Q'} \Gamma(Q|Q') P(Q'; t)$$

で与えられる。ここで、 $\Gamma(Q|Q')$ は時間発展演算子(遷移行列)であり、平衡分布関数 $P_{\text{eq}}(Q)$ に対し詳細釣り合いの条件

$$(2.2) \quad \Gamma(Q|Q') P_{\text{eq}}(Q') = \Gamma(Q'|Q) P_{\text{eq}}(Q)$$

を満たすように選ばれる。

次のシミュレーション方法は、Brown 動力学法で、 N 個の粒子からなる系において、粒子の位置を \mathbf{r}_i ($i = 1, \dots, N$) で表し、粒子間の相互作用のポテンシャルを $U(\{\mathbf{r}_j\})$ とするとき、粒子の位置の時間発展は Langevin 方程式

$$(2.3) \quad \frac{d\mathbf{r}_i}{dt} = \frac{1}{\zeta} \left[- \frac{\partial U(\{\mathbf{r}_j\})}{\partial \mathbf{r}_i} + \mathbf{w}_i \right]$$

で与えられる。ここで \mathbf{w}_i は i 番目の粒子に働くランダム力で、時刻 t におけるその α 成分を $w_{i\alpha}$ と表すと、それらは

$$(2.4) \quad \langle w_{i\alpha}(t) \rangle = 0$$

および

$$(2.5) \quad \langle w_{i\alpha}(t) w_{j\beta}(t') \rangle = 2\zeta k_B T \delta_{i,j} \delta_{\alpha,\beta} \delta(t-t')$$

を満たす白色 Gauss 過程である。ここで、 ζ は抵抗係数で、速度に比例した抵抗力の比例定数である。 k_B は Boltzmann 定数、 T は系の絶対温度を表す。このとき、時刻 t に系が状態 $Q = {}^t(\mathbf{r}_1, \dots, \mathbf{r}_N)$ をとる確率 $P(Q; t)$ の時間発展は Smoluchowski 方程式

$$(2.6) \quad \begin{aligned} \frac{\partial P}{\partial t} &= -\Gamma P \\ &= \sum_i \frac{\partial}{\partial \mathbf{r}_i} \cdot \frac{1}{\zeta} \left\{ k_B T \frac{\partial}{\partial \mathbf{r}_i} + \frac{\partial U}{\partial \mathbf{r}_i} \right\} P \end{aligned}$$

で与えられる。時間発展演算子 Γ は平衡分布 $P_{\text{eq}}(Q) \propto \exp[-U/(k_B T)]$ に対し詳細釣り合いの条件

$$(2.7) \quad P_{\text{eq}}(Q') \Gamma(Q) \delta(Q - Q') = P_{\text{eq}}(Q) \Gamma^\dagger(Q) \delta(Q - Q')$$

を満たす。ここで $\Gamma(Q) \delta(Q - Q')$ と $\Gamma^\dagger(Q) \delta(Q - Q')$ は Γ とその共役 Γ^\dagger が $\delta(Q - Q')$ の Q にもみ作用することを表す。 $\Gamma(Q) \delta(Q - Q') = \Gamma(Q|Q')$ 、 $\Gamma^\dagger(Q) \delta(Q - Q') = \Gamma(Q'|Q)$ と表す演算子の行列表現を用いれば、この詳細釣り合いの条件は式(2.2)と同じである。

三つ目のシミュレーション方法は、Langevin 熱浴を温度制御に用いた分子動力学法で、簡単のため、同じ質量 m 、同じ抵抗係数 ζ をもった N 粒子系を用いて説明する。 i 番目の粒子の位置 \mathbf{r}_i と速度 \mathbf{v}_i の従う Langevin 方程式は

$$(2.8) \quad m \frac{d\mathbf{v}_i}{dt} = -\zeta \mathbf{v}_i - \frac{\partial U(\{\mathbf{r}_j\})}{\partial \mathbf{r}_i} + \mathbf{w}_i, \quad \frac{d\mathbf{r}_i}{dt} = \mathbf{v}_i$$

で与えられる。ここで、 $U(\{\mathbf{r}_j\})$ は粒子間の相互作用のポテンシャルである。 \mathbf{w}_i は i 番目の粒子に働くランダム力で、Brown 動力学法で用いたものと同じ白色 Gauss 過程である。このとき、時刻 t に系が状態 $Q = {}^t(\mathbf{r}_1, \dots, \mathbf{r}_N, \mathbf{v}_1, \dots, \mathbf{v}_N)$ をとる確率 $P(Q;t)$ の時間発展は Kramers 方程式

$$(2.9) \quad \frac{\partial P}{\partial t} = -\Gamma P \\ = \sum_i \left[-\frac{\partial}{\partial \mathbf{r}_i} \cdot \mathbf{v}_i + \frac{\partial}{\partial \mathbf{v}_i} \cdot \frac{1}{m} \frac{\partial V}{\partial \mathbf{r}_i} + \frac{\partial}{\partial \mathbf{v}_i} \cdot \frac{\zeta}{m} \left(\mathbf{v}_i + \frac{k_B T}{m} \frac{\partial}{\partial \mathbf{v}_i} \right) \right] P$$

で与えられる。時間発展演算子 Γ は平衡分布 $P_{\text{eq}}(Q) \propto \exp[-\frac{1}{k_B T}(\frac{m}{2} \sum_i \mathbf{v}_i^2 + U(\{\mathbf{r}_j\}))]$ に対し詳細釣り合いの条件

$$(2.10) \quad P_{\text{eq}}(Q') \Gamma(Q) \delta(Q - Q') = P_{\text{eq}}(\epsilon Q) \Gamma^\dagger(\epsilon Q) \delta(\epsilon Q - \epsilon Q')$$

を満たす (Risken, 1996)。ここで $\epsilon Q = {}^t(\mathbf{r}_1, \dots, \mathbf{r}_N, -\mathbf{v}_1, \dots, -\mathbf{v}_N)$ は状態 Q の速度を反転した状態である。行列表現を用いれば、この詳細釣り合いの条件は

$$(2.11) \quad \Gamma(Q|Q') P_{\text{eq}}(Q') = \Gamma(\epsilon Q'|\epsilon Q) P_{\text{eq}}(\epsilon Q)$$

と書ける。ここで、平衡分布は

$$(2.12) \quad P_{\text{eq}}(\epsilon Q) = P_{\text{eq}}(Q)$$

を満たす。

以上の時間発展演算子の固有関数としての緩和モードについて説明する。時刻 t に系が状態 Q をとる確率 $P(Q;t)$ の時間発展の方程式を行列表現で表すと (2.1) 式となる。モンテカルロ法、Brown 動力学法では、速度の自由度がないため、 $\epsilon Q = Q$ となるので、分子動力学法の場合と合わせて、詳細釣り合いの条件は (2.11) で与えられる。時間発展演算子 $\Gamma(Q|Q')$ の固有値問題を考える。固有値 λ_n に対する左固有関数を $\phi_n(Q)$ 、右固有関数を $\psi_n(Q)$ とする。即ち、

$$(2.13) \quad \sum_Q \phi_n(Q) \Gamma(Q|Q') = \lambda_n \phi_n(Q'),$$

$$(2.14) \quad \sum_{Q'} \Gamma(Q|Q') \psi_n(Q') = \lambda_n \psi_n(Q).$$

右固有関数を $\psi_n(Q) = \hat{\phi}_n(Q) P_{\text{eq}}(Q)$ と書くとき、 $\hat{\phi}_n(Q)$ は左固有関数 $\phi_n(Q)$ を使って $\hat{\phi}_n(Q) = \phi_n(\epsilon Q)$ と書ける。このことは次のようにしてわかる。左固有関数の固有値方程式 (2.13) 式の両辺に $P_{\text{eq}}(Q')$ をかけて

$$(2.15) \quad \sum_Q \phi_n(Q) \Gamma(Q|Q') P_{\text{eq}}(Q') = \lambda_n \phi_n(Q') P_{\text{eq}}(Q')$$

を得る。この左辺は詳細釣り合いの条件 (2.11) 式を使って、

$$(2.16) \quad \sum_Q \phi_n(Q) \Gamma(Q|Q') P_{\text{eq}}(Q') = \sum_Q \phi_n(Q) \Gamma(\epsilon Q'|\epsilon Q) P_{\text{eq}}(\epsilon Q)$$

$$= \sum_{\epsilon Q} \Gamma(\epsilon Q' | \epsilon Q) \phi_n(\epsilon(\epsilon Q)) P_{\text{eq}}(\epsilon Q) = \sum_{Q''} \Gamma(\epsilon Q' | Q'') \phi_n(\epsilon Q'') P_{\text{eq}}(Q'')$$

と変形できる. 一方(2.15)式の右辺は $\epsilon(\epsilon Q) = Q$ と(2.12)式を用いて

$$(2.17) \quad \lambda_n \phi_n(Q') P_{\text{eq}}(Q') = \lambda_n \phi_n(\epsilon(Q')) P_{\text{eq}}(\epsilon(Q'))$$

と変形できる. この式の右辺と(2.16)式の最後の式を等号で結び, $\epsilon Q'$ を Q と, Q'' を Q' と書き換えて

$$(2.18) \quad \sum_{Q'} \Gamma(Q|Q') \phi_n(\epsilon Q') P_{\text{eq}}(Q') = \lambda_n \phi_n(\epsilon Q) P_{\text{eq}}(Q)$$

を得る. これは右固有関数の固有値方程式(2.14)式で, $\psi_n(Q) = \phi_n(\epsilon Q) P_{\text{eq}}(Q)$ であることを示す. 即ち,

$$(2.19) \quad \psi_n(Q) = \hat{\phi}_n(Q) P_{\text{eq}}(Q) = \phi_n(\epsilon Q) P_{\text{eq}}(Q).$$

固有関数は規格直交関係

$$(2.20) \quad \sum_Q \phi_m(Q) \psi_n(Q) = \sum_Q \phi_m(Q) \hat{\phi}_n(Q) P_{\text{eq}}(Q) = \langle \phi_m \hat{\phi}_n \rangle = \delta_{m,n}$$

を満たすようにとる. ここで, $\langle A \rangle$ は物理量 A の平衡状態での平均を表す. このとき, 時間間隔 t だけ離れた $\phi_m(Q)$ と $\hat{\phi}_n(Q)$ の時間相関関数は

$$(2.21) \quad \begin{aligned} \langle \phi_m(t) \hat{\phi}_n(0) \rangle &= \sum_Q \sum_{Q'} \phi_m(Q) T_t(Q|Q') \hat{\phi}_n(Q') P_{\text{eq}}(Q') \\ &= \sum_Q \sum_{Q'} \phi_m(Q) e^{-\Gamma t}(Q|Q') \hat{\phi}_n(Q') P_{\text{eq}}(Q') = \sum_Q \sum_{Q'} \phi_m(Q) e^{-\Gamma t}(Q|Q') \psi_n(Q') \\ &= \sum_Q \phi_m(Q) e^{-\lambda_n t} \psi_n(Q) = \delta_{m,n} e^{-\lambda_n t} \end{aligned}$$

と与えられる. ここで,

$$(2.22) \quad T_\tau(Q|Q') = e^{-\Gamma \tau}(Q|Q')$$

は時刻 t で系が状態 Q' にあるとき時刻 $t + \tau$ で系が状態 Q にある条件付き確率である.

2つの物理量 $A(Q)$ と $B(Q)$ が固有関数 $\phi_n(Q)$ と $\hat{\phi}_n(Q)$ を使って

$$(2.23) \quad A(Q) = \sum_n a_n \phi_n(Q), \quad B(Q) = \sum_n \hat{b}_n \hat{\phi}_n(Q)$$

のように固有関数展開できるとき, $A(Q)$ と $B(Q)$ の時間相関関数は

$$(2.24) \quad \langle A(t) B(0) \rangle = \sum_n a_n \hat{b}_n \exp(-\lambda_n t)$$

と与えられる. この性質から, 固有関数 $\phi_n(Q)$ と $\hat{\phi}_n(Q)$ を緩和モード, 固有値 λ_n をその緩和率と呼ぶことができる.

3. 緩和モード解析の方法

緩和モード解析の方法は, 時間発展演算子に対する固有値問題と等価な変分問題を考え, 適切な試行関数を選ぶことで, 緩和モード, 緩和率を評価する方法である. まず, 時間発展演算

子そのものではなく、条件付き確率(2.22)に対する固有値方程式

$$(3.1) \quad \sum_Q \phi_n(Q) T_\tau(Q|Q') = e^{-\lambda_n \tau} \phi_n(Q'),$$

$$(3.2) \quad \sum_{Q'} T_\tau(Q|Q') \psi_n(Q') = e^{-\lambda_n \tau} \psi_n(Q)$$

を考える。この固有値方程式と等価な変分問題は

$$(3.3) \quad \mathcal{R} = \frac{\sum_Q \sum_{Q'} \phi_n(Q) T_\tau(Q|Q') \psi_n(Q')}{\sum_Q \phi_n(Q) \psi_n(Q)} = \frac{\langle \phi_n(t) \hat{\phi}_n(0) \rangle}{\langle \phi_n \hat{\phi}_n \rangle}$$

に対し、 $\delta\mathcal{R} = 0$ で与えられる (Talkner, 1994; Hirao et al., 1997; Drozdov, 2000)。付録 A で固有値方程式と変分問題について説明した。 $\delta\mathcal{R} = 0$ となる $\phi_n, \hat{\phi}_n$ が固有関数で、そのときの \mathcal{R} の値が $\exp(-\lambda_n \tau)$ となる。この変分問題に対し、重要と思われる物理量の線形結合を試行関数として用いることで、計算機シミュレーションのデータから緩和モードと緩和率を評価することができる。

緩和モードの計算法の例として、 N 個のセグメントからなる孤立線状高分子のシミュレーションを考える。 i 番目のセグメントの位置を \mathbf{r}_i 、速度を \mathbf{v}_i とし、その重心からの相対位置を \mathbf{R}_i 、相対速度を \mathbf{R}_{N+i} とする。即ち、

$$(3.4) \quad \begin{aligned} \mathbf{R} &= {}^t(\mathbf{R}_1, \dots, \mathbf{R}_N, \mathbf{R}_{N+1}, \dots, \mathbf{R}_{2N}) \\ &= {}^t(\mathbf{r}_1 - \mathbf{r}_c, \dots, \mathbf{r}_N - \mathbf{r}_c, \mathbf{v}_1 - \mathbf{v}_c, \dots, \mathbf{v}_N - \mathbf{v}_c). \end{aligned}$$

ここで $\mathbf{r}_c = \frac{1}{N} \sum_i \mathbf{r}_i$ と $\mathbf{v}_c = \frac{1}{N} \sum_i \mathbf{v}_i$ は、それぞれ重心の位置と速度である。 $\phi_n(Q), \hat{\phi}_n(Q)$ に対する試行関数(近似的固有関数)として、それぞれ、

$$(3.5) \quad \mathbf{X}_p(Q) = \sum_i f_{p,i} \mathbf{R}_i(t_0/2; Q), \quad \hat{\mathbf{X}}_p(Q) = \sum_i f_{p,i} \mathbf{R}_i(t_0/2; \epsilon Q)$$

を考える。ここで、

$$(3.6) \quad \mathbf{R}_i(t; Q) = \sum_{Q'} \mathbf{R}_i(Q') T_t(Q'|Q)$$

は時刻 0 で系が状態 Q にあるとき時刻 t での \mathbf{R}_i の期待値である。パラメーター t_0 は、 \mathbf{R}_i の中に含まれる緩和の速いモードの寄与を相対的に小さくするためのものである。付録 B で t_0 の役割について説明した。 $f_{p,i}$ は変分パラメーターである。この試行関数は、左右固有関数の間の関係(2.19)式 $\hat{\mathbf{X}}_p(Q) = \mathbf{X}_p(\epsilon Q)$ を満たしている。

これらの試行関数を(3.3)式に代入すると、 $f_{p,i}$ に対する変分問題 $\delta\mathcal{R} = 0$ は一般化固有値問題

$$(3.7) \quad \sum_i f_{p,i} C_{i,j}(t_0 + \tau) = e^{-\lambda_p \tau} \sum_i f_{p,i} C_{i,j}(t_0),$$

$$(3.8) \quad \sum_j C_{i,j}(t_0 + \tau) \epsilon_j f_{p,j} = e^{-\lambda_p \tau} \sum_j C_{i,j}(t_0) \epsilon_j f_{p,j}$$

と等価になる。ここで、

$$(3.9) \quad \epsilon_i = \begin{cases} 1, & i = 1, \dots, N \\ -1, & i = N + 1, \dots, 2N \end{cases}$$

である。 $C_{i,j}(t) = \frac{1}{3} \langle \mathbf{R}_i(t) \cdot \mathbf{R}_j(0) \rangle$ は時間相関行列であり、詳細釣り合いの条件(2.11)式より、(3.9)式で定義された ϵ_i を用いて $C_{i,j}(t) = \epsilon_i \epsilon_j C_{j,i}(t)$ という対称性を持っている。この対称性

より, $f_{p,i}$ に関する一般化固有値問題(3.7)式と $\epsilon_i f_{p,i}$ に関する一般化固有値問題(3.8)式が等価であることがわかる.

固有関数間の規格直交関係(2.20)式は $\mathbf{X}_p(Q)$, $\hat{\mathbf{X}}_p(Q)$ の α 成分を, それぞれ $X_{p,\alpha}(Q)$, $\hat{X}_{p,\alpha}(Q)$ と書くと

$$(3.10) \quad \langle X_{p,\alpha} \hat{X}_{q,\beta} \rangle = \delta_{p,q} \delta_{\alpha,\beta}$$

で与えられるが, これは $f_{p,i}$ に対する条件

$$(3.11) \quad \sum_i \sum_j f_{p,i} C_{i,j}(t_0) \epsilon_j f_{q,j} = \delta_{p,q}$$

に相当する. このとき, 近似的緩和モードの時間相関関数として

$$(3.12) \quad \langle X_{p,\alpha}(t) \hat{X}_{q,\beta}(0) \rangle \simeq \delta_{p,q} \delta_{\alpha,\beta} \exp(-\lambda_p t)$$

が成立すると考えられる.

規格直交関係(3.11)式より, $\mathbf{R}_i(t_0/2; Q)$ の線形結合で試行関数 $\mathbf{X}_p(Q)$ と $\hat{\mathbf{X}}_p(Q)$ を与える(3.5)式の逆変換は

$$(3.13) \quad \mathbf{R}_i(t_0/2; Q) = \sum_p g_{i,p} \mathbf{X}_p(Q), \quad \epsilon_i \mathbf{R}_i(t_0/2; \epsilon Q) = \sum_p \hat{g}_{i,p} \hat{\mathbf{X}}_p(Q)$$

となる. ここで, $g_{i,p}$, $\hat{g}_{i,p}$ は, それぞれ

$$(3.14) \quad g_{i,p} = \sum_j C_{i,j}(t_0) \epsilon_j f_{p,j}, \quad \hat{g}_{i,p} = \sum_j f_{p,j} C_{j,i}(t_0) = \epsilon_i g_{i,p}$$

で与えられる. これより \mathbf{R}_i の相関行列は, $t \geq t_0$ のとき,

$$\begin{aligned} C_{i,j}(t) &= \frac{1}{3} \langle \mathbf{R}_i(t) \cdot \mathbf{R}_j(0) \rangle = \frac{1}{3} \sum_Q \sum_{Q'} \mathbf{R}_i(Q) \cdot T_t(Q|Q') \mathbf{R}_j(Q') P_{\text{eq}}(Q') \\ &= \frac{1}{3} \sum_Q \sum_{Q'} \sum_{Q''} \sum_{Q'''} \mathbf{R}_i(Q) T_{t_0/2}(Q|Q') \cdot T_{t-t_0}(Q'|Q'') T_{t_0/2}(Q''|Q''') \mathbf{R}_j(Q''') P_{\text{eq}}(Q''') \\ &= \frac{1}{3} \sum_{Q'} \sum_{Q''} \sum_{Q'''} \mathbf{R}_i(t_0/2; Q') \cdot T_{t-t_0}(Q'|Q'') T_{t_0/2}(\epsilon Q''''|\epsilon Q''') \mathbf{R}_j(Q''') P_{\text{eq}}(\epsilon Q''') \\ &= \frac{1}{3} \sum_{Q'} \sum_{Q''} \sum_{Q'''} \mathbf{R}_i(t_0/2; Q') \cdot T_{t-t_0}(Q'|Q'') \{ \epsilon_j \mathbf{R}_j(\epsilon Q''') T_{t_0/2}(\epsilon Q''''|\epsilon Q''') \} P_{\text{eq}}(Q'') \\ &= \frac{1}{3} \sum_{Q'} \sum_{Q''} \sum_{Q'''} \mathbf{R}_i(t_0/2; Q') \cdot T_{t-t_0}(Q'|Q'') \epsilon_j \mathbf{R}_j(t_0/2; \epsilon Q'') P_{\text{eq}}(Q'') \\ &= \frac{1}{3} \sum_{Q'} \sum_{Q''} \sum_{Q'''} \sum_p g_{i,p} \mathbf{X}_p(Q') \cdot T_{t-t_0}(Q'|Q'') \sum_q \hat{g}_{j,q} \hat{\mathbf{X}}_q(Q'') P_{\text{eq}}(Q'') \\ &= \sum_p \sum_q g_{i,p} \hat{g}_{j,q} \frac{1}{3} \langle \mathbf{X}_p(t-t_0) \cdot \hat{\mathbf{X}}_q(0) \rangle \\ &\simeq \sum_p g_{i,p} \hat{g}_{j,p} \exp[-\lambda_p(t-t_0)] = \sum_p \tilde{g}_{i,p} \hat{\tilde{g}}_{j,p} \exp(-\lambda_p t) \end{aligned}$$

のように与えられる. ここで,

$$(3.15) \quad \tilde{g}_{i,p} = e^{\lambda_p t_0/2} \sum_j C_{i,j}(t_0) \hat{f}_{p,j}, \quad \hat{\tilde{g}}_{i,p} = e^{\lambda_p t_0/2} \sum_j f_{p,j} C_{j,i}(t_0) = \epsilon_i \tilde{g}_{i,p}$$

である。途中の変形で詳細釣り合いの条件(2.11)式から得られる結果

$$(3.16) \quad T(Q|Q')P_{\text{eq}}(Q') = T(\epsilon Q'|\epsilon Q)P_{\text{eq}}(\epsilon Q)$$

および(2.12)式を用いている。(3.15)式の最後の行は、 $\mathbf{R}_i(Q)$ の(2.23)式のような展開が

$$(3.17) \quad \mathbf{R}_i(Q) \simeq \sum_p \tilde{g}_{i,p} \mathbf{X}_p(Q) \simeq \sum_p \hat{g}_{i,p} \hat{\mathbf{X}}_p(Q)$$

のように与えられると見なすことができる。

以上、相関行列 $C_{i,j}(t_0)$ と $C_{i,j}(t_0 + \tau)$ をシミュレーションで計算し一般化固有値問題(3.7)式、(3.8)式を解くことで、緩和モード(3.5)式と緩和率 λ_p を評価することができる。その結果に基づいて、(3.15)式を用いて $t \geq t_0$ に対する時間相関行列を計算し、実際に計算機シミュレーションで求めた結果とくらべることで、得られた緩和モード、緩和率が妥当なものであるかどうかを検証することができる。

緩和モード解析の結果の例として、Langevin 熱浴を温度制御に用いた孤立線状高分子の分子動力学シミュレーションへ適用した Hirao et al. (1997)の結果を紹介する。

孤立線状高分子の N 個のセグメント間の相互作用ポテンシャルは Kremer and Grest (1990)と同じもので、全セグメント間に働く斥力相互作用(Lennard-Jones ポテンシャルの斥力部分)と高分子鎖上で隣接するセグメント間に働く引力相互作用から成る。温度や抵抗係数等のパラメータも Kremer and Grest (1990)と同じである。このポテンシャルとパラメータでは、セグメント間の排除体積相互作用が取り入れられることが知られている。流体力学的相互作用を無視して、(2.8)式に基づいて、セグメント数 $N = 10, 20, 40$ の場合にシミュレーションを行った。平衡化後の時間 4×10^4 (単位は斥力ポテンシャルの Lennard-Jones 単位)のデータを用いて時間相関行列 $C_{i,j}(t)$ を計算し、 $t_0 = 1.6$, $\tau = 0.4$ として、一般化固有値問題(3.7)式、(3.8)式を解き、緩和モード、緩和率を評価した。得られた結果から(3.15)式を用いて重心からの相対位置の自己相関関数 $C_{i,i}(t)$ ($i = 1, \dots, N$) と重心からの相対速度の自己相関関数 $C_{i,i}(t)$ ($i = N + 1, \dots, 2N$) とを計算し、実際にシミュレーションから求めた結果と比較したのが図1である。減衰振動的な挙動も良く再現されており、得られた緩和モード、緩和率が妥当であることを示している。分子動力学法の場合、振動的な振舞いも出てくるので、緩和率は複素数となる。その分布を図2に示す。緩和率の実部が小さいときは、緩和率が実数になり、純粹緩和となっていることがわかる。緩和率が実数となっている緩和率について、緩和率が小さい順に緩和率 λ_p の番号 p を1からつける。このとき、 $\lambda_p \propto (p/N)^x$ というべき乗則的振舞いが見られた。フィッティングで得られた指数は $x \simeq 2.2$ となっており、理論の予想値 $x \simeq 2.18$ (Doi and Edwards, 1986)に近い値が得られている。また、(3.17)式の展開係数 $\tilde{g}_{i,p}$ の i 依存性は、理論の予想 $\tilde{g}_{i,p} \propto \cos[(i - \frac{1}{2})p\pi/N]$ に良く合っていた。

4. ペプチド系への応用

生体高分子系への応用例の1つとして、真空中の Met-Enkephalin のモンテカルロ・シミュレーションへ緩和モード解析を適用した Mitsutake et al. (2011)の結果を紹介する。Met-Enkephalin は Tyr-Gly-Gly-Phe-Met の5残基からなるペプチドで、真空中では複数の準安定状態をもつことが知られている(Mitsutake et al., 1998)。そのため、安定状態、準安定状態間の遷移を見るのに適した系である。シミュレーションの詳細は原論文に譲り、ここでは、主成分分析と緩和モード解析という2つの方法の結果を示す。

解析に用いる変数は

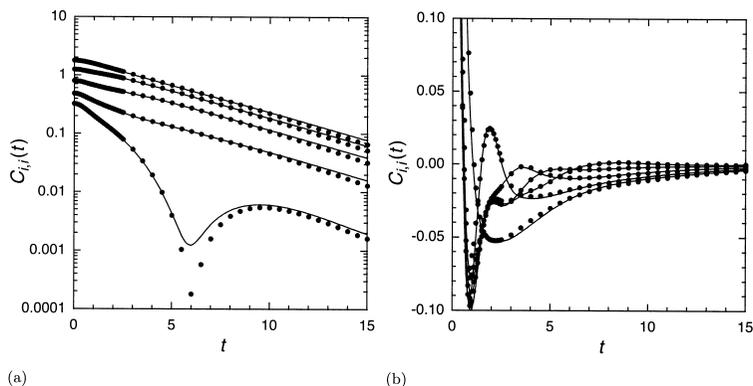


図1. $N = 10$ の場合の $C_{i,i}(t)$ の t 依存性. 実線は緩和モード解析の結果に基づいて(3.15)式を用いて計算した結果を, 黒丸はシミュレーションから直接評価した結果を表す. (a)には重心からの相対位置の自己相関関数 $C_{i,i}(t) = C_{N-i+1, N-i+1}(t)$ (上から $i = 1, 2, 3, 4, 5$) が, (b)には重心からの相対速度の自己相関関数 $C_{i,i}(t) = C_{3N-i+1, 3N-i+1}(t)$ ($t \simeq 8$ で下から $i = 11, 12, 13, 14, 15$) が示されている. Reprinted with permission from H. Hirao, S. Koseki, and H. Takano, J. Phys. Soc. Jpn. **66**, 3399 (1997). Copyright 1997, The Physical Society of Japan.

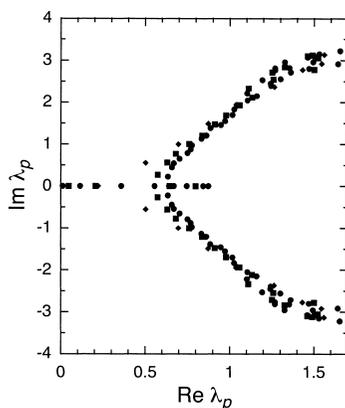


図2. λ_p の分布. 菱形, 正方形, 丸がそれぞれ $N = 10, 20, 40$ の結果を表す. Reprinted with permission from H. Hirao, S. Koseki, and H. Takano, J. Phys. Soc. Jpn. **66**, 3399 (1997). Copyright 1997, The Physical Society of Japan.

$$(4.1) \quad \mathbf{R} = {}^t(\mathbf{R}_1, \dots, \mathbf{R}_N) = {}^t({}^t\mathbf{r}_1 - {}^t\langle \mathbf{r}_1 \rangle, \dots, {}^t\mathbf{r}_N - {}^t\langle \mathbf{r}_N \rangle)$$

である. ここで, \mathbf{r}_i は i 番目の原子の重心からの相対位置座標であり, $\langle \mathbf{r}_i \rangle$ はその平均位置である. 平均位置で作られる構造を平均構造と呼ぶことにする. シミュレーションから得られた多数の構造のサンプルのそれぞれに対し, 1つの構造と平均構造の2乗変位が最小になるように, その構造を並進, 回転するという操作をしたものを用いる. さらに, 操作後の多数の構造を平均したものが平均構造と一致するように平均構造を選ぶ. 扱う原子は, 主鎖上の10個の C_{α}, C

原子とする。即ち $N = 10$ である。

まず、主成分分析 (PCA) について説明する。主成分分析では分散共分散行列

$$(4.2) \quad C = \langle \mathbf{R}^t \mathbf{R} \rangle$$

に対する固有値問題

$$(4.3) \quad C \mathbf{F}_p = \Lambda_p \mathbf{F}_p$$

を規格直交条件

$$(4.4) \quad {}^t \mathbf{F}_p \mathbf{F}_q = \delta_{p,q}$$

のもとで解く。 p 番目に大きい固有値を Λ_p とする。固有ベクトル \mathbf{F}_p は第 p 主成分軸と呼ばれ、この方向に $\mathbf{R}(Q)$ を射影した

$$(4.5) \quad \Phi_p(Q) = {}^t \mathbf{F}_p \mathbf{R}(Q)$$

は第 p 主成分と呼ばれる。 Λ_p は第 p 主成分の分散を与える。また、 $\mathbf{R}(Q)$ は \mathbf{F}_p を使って

$$(4.6) \quad \mathbf{R}(Q) = \sum_p \Phi_p(Q) \mathbf{F}_p$$

のように展開できる。

これに対し、緩和モード解析 (RMA) は時間相関行列

$$(4.7) \quad C(t) = \langle \mathbf{R}(t) {}^t \mathbf{R}(0) \rangle$$

に対する一般化固有値問題

$$(4.8) \quad C(t_0 + \tau) \mathbf{f}_p = e^{-\lambda_p \tau} C(t_0) \mathbf{f}_p$$

を規格直交条件

$$(4.9) \quad {}^t \mathbf{f}_p C(t_0) \mathbf{f}_q = \delta_{p,q}$$

のもとで解く。今はモンテカルロ・シミュレーションの場合を扱っているので、 $\phi_n(Q) = \hat{\phi}_n(Q)$ である。 p 番目に小さい緩和率を λ_p とする。対応する緩和モード $X_p(Q)$ と $\mathbf{R}(Q)$ の第 i 成分 $R_i(Q)$ の関係は

$$(4.10) \quad R_i(Q) \simeq \sum_p \tilde{g}_{i,p} X_p(Q)$$

で与えられる。これより、 $\mathbf{R}(Q)$ は $\tilde{g}_{i,p}$ を第 i 成分とするベクトル $\tilde{\mathbf{g}}_p$ を用いて

$$(4.11) \quad \mathbf{R}(Q) \simeq \sum_p X_p(Q) \tilde{\mathbf{g}}_p = \sum_p Y_p(Q) \tilde{\mathbf{g}}_p / |\tilde{\mathbf{g}}_p|$$

と展開できる。ここで、

$$(4.12) \quad Y_p(Q) = X_p(Q) |\tilde{\mathbf{g}}_p|$$

は PCA の $\Phi_p(Q)$ に対応する量である。

PCA で得られた分散 Λ_p と RMA で得られた緩和率 λ_p を図 3 に示す。PCA では $p = 1$ だけが特に大きい分散を与えるのに対し、RMA では $p = 1$ と 2 が特に小さい緩和率を与えている。PCA で得られた \mathbf{F}_p と RMA で得られた $\tilde{\mathbf{g}}_p / |\tilde{\mathbf{g}}_p|$ を $p = 1, 2, 3$ の場合に図 4 に示す。 \mathbf{F}_p の $p = 1$

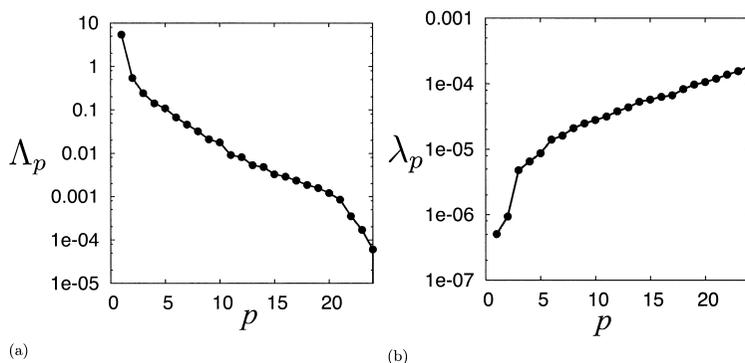


図 3. PCA で得られた分散 Λ_p の分布 (a) と RMA で得られた緩和率 λ_p の分布 (b). Λ_p の単位は \AA^2 , λ_p の単位は自由度あたりモンテカルロ・ステップの逆数. Reprinted with permission from A. Mitsutake, H. Iijima, and H. Takano, J. Chem. Phys. **135**, 164102 (2011). Copyright 2011, American Institute of Physics.

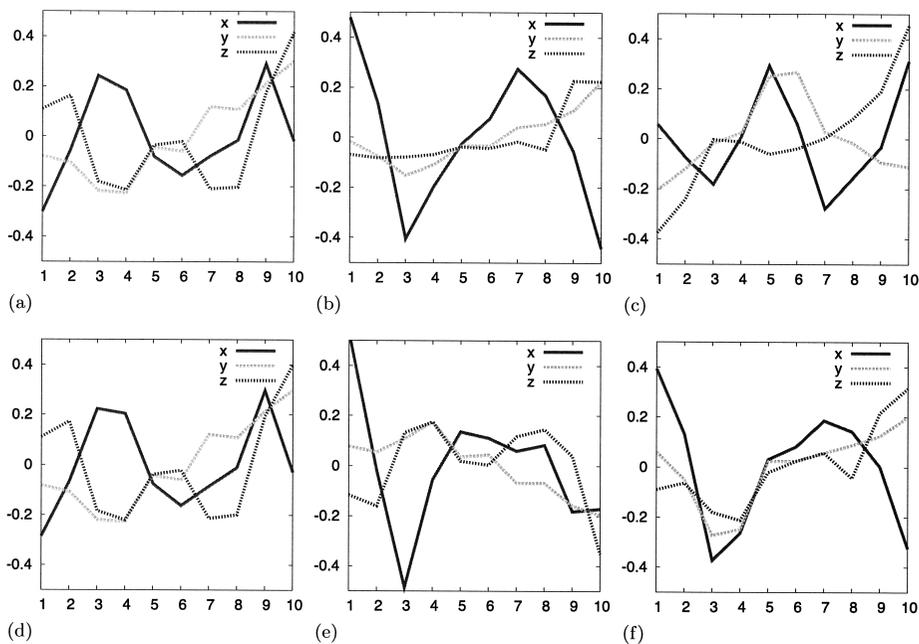


図 4. PCA で得た F_p と RMA で得た $\tilde{g}_p/|\tilde{g}_p|$. 横軸は主鎖上の原子の番号, 縦軸はその原子に対応した F_p と $\tilde{g}_p/|\tilde{g}_p|$ の x, y, z 成分の値. 濃い灰色, 中位の灰色, 薄い灰色がそれぞれ x, y, z 成分を表す. F_p の $p=1$ を (a) に, $p=2$ を (b) に, $p=3$ を (c) に, $\tilde{g}_p/|\tilde{g}_p|$ の $p=1$ を (d) に, $p=2$ を (e) に, $p=3$ を (f) に示した.

と $\tilde{g}_p/|\tilde{g}_p|$ の $p=1$ が似ていることがわかる. F_p の $p=2$ と $\tilde{g}_p/|\tilde{g}_p|$ の $p=3$ が似ているが, $\tilde{g}_p/|\tilde{g}_p|$ の $p=2$ に似ているものはない.

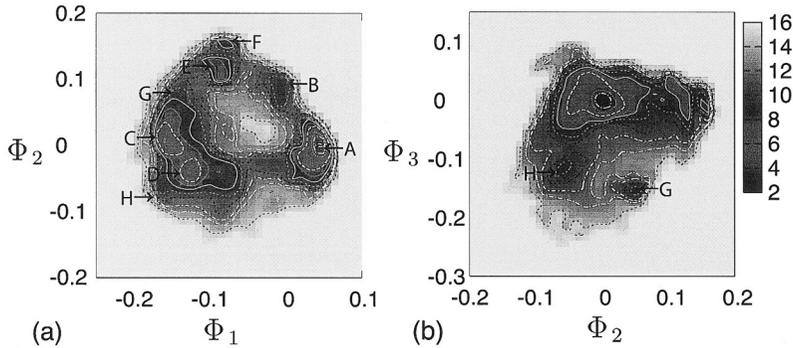


図 5. PCA で得られた自由エネルギー地形. 各グリッドに入る確率を P とするとき, $-\ln P$ の値を図の右に示した濃度で表している. Φ_2 対 Φ_1 を (a) に, Φ_3 対 Φ_2 を (b) に示した. 図中の A-H は極小点を示す. Reprinted with permission from A. Mitsutake, H. Iijima, and H. Takano, *J. Chem. Phys.* **135**, 164102 (2011). Copyright 2011, American Institute of Physics.

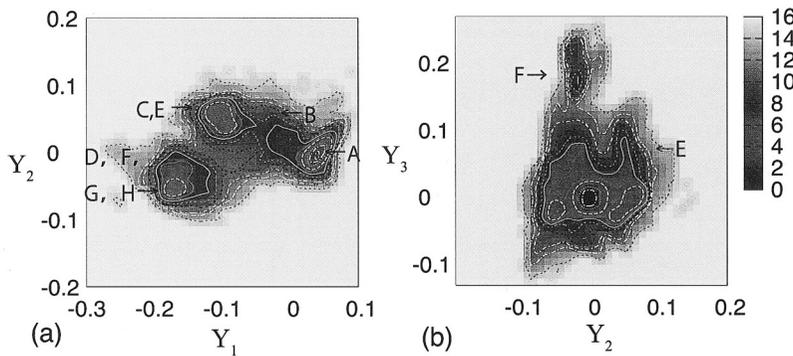


図 6. RMA で得られた自由エネルギー地形. 各グリッドに入る確率を P とするとき, $-\ln P$ の値を図の右に示した濃度で表している. Y_2 対 Y_1 を (a) に, Y_3 対 Y_2 を (b) に示した. 図中の A-H は, 図 5 中の極小点 A-H がこの図でどこに相当するかを示している. Reprinted with permission from A. Mitsutake, H. Iijima, and H. Takano, *J. Chem. Phys.* **135**, 164102 (2011). Copyright 2011, American Institute of Physics.

次に自由エネルギー地形について考える. $\Phi_p(Q)$ または $Y_p(Q)$ を X , $\Phi_q(Q)$ または $Y_q(Q)$ を Y として, シミュレーション中に X と Y のとる値から, その確率密度分布 $P(X, Y)$ を求め, 無次元化した自由エネルギー $F(X, Y) = -\ln P(X, Y)$ を求め, X - Y 平面上の等高線図を描いたものが自由エネルギー地形である. 図 5 に $p = 1, 2, 3$ の $\Phi_p(Q)$ で調べた自由エネルギー地形を, 図 6 に $p = 1, 2, 3$ の $Y_p(Q)$ で調べた自由エネルギー地形を示す. PCA で得られた自由エネルギー地形の極小点を図 5 の中に A-H で示した. これらの極小点が図 6 中のどこに相当するかを図 6 中に A-H で示した. 図 6(a) 中で, 深い極小は A, C, D で, B は A と C の間に位置し, E は C に属し, F, G, H は D の一部となっている. 第 1 緩和モード $\tilde{g}_1/|\tilde{g}_1|$ は第 1 主成分軸 F_1 と同様, A と C または A と D の間の遷移を表している. 第 2 緩和モード $\tilde{g}_2/|\tilde{g}_2|$ は C と D の間の

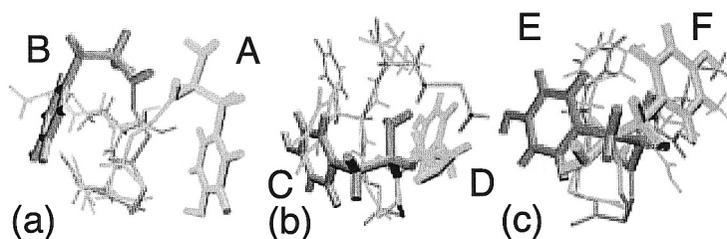


図7. PCA で得られた自由エネルギー極小の位置での Tyr1 の側鎖の配座の比較. B(濃灰色)とA(淡灰色)の比較を(a)に, C(濃灰色)とD(淡灰色)の比較を(b)に, E(濃灰色)とF(淡灰色)の比較を(c)に示した. それぞれの組について, Tyr1 以外の残基の主鎖部分の平均2乗変位が最小となるように配置してある. Reprinted with permission from A. Mitsutake, H. Iijima, and H. Takano, *J. Chem. Phys.* **135**, 164102 (2011). Copyright 2011, American Institute of Physics.

遷移を表している. 同時にこのモードは残基 Tyr1 の側鎖の配座の変化を記述している. 図7はAとB中の配座の比較, DとC中の配座の比較, FとE中の配座の比較を示している. それぞれの組で, Tyr1 以外の残基の主鎖部分間の平均2乗変位が最小となるように配置している. B, C, EではTyr1の側鎖は主鎖の左側に, A, D, Fでは右側にある. B, C, Eは Y_2 が正の領域に, A, D, Fは Y_2 が負の領域に存在している. 図4(e)に描かれた $\tilde{g}_2/|\tilde{g}_2|$ で, Tyr1に対応する第1成分の振幅が大きいことがわかる. PCAで得られた F_p には $\tilde{g}_2/|\tilde{g}_2|$ に似たものはなく, PCAでは抽出できなかった遅い運動, 即ち, Tyr1の側鎖の遅い運動をRMAにより抽出できたことがわかる.

また, 図5と図6を比較すると, PCAの自由エネルギー地形では, 極小点が原点を取り囲むように並ぶのに対し, RMAの自由エネルギー地形では, 極小点が軸に沿って並ぶ傾向がある. これは, 安定状態, 準安定状態間の遷移を正しく記述できているためと考えられる.

5. タンパク質系への応用

生体高分子系へのもう1つの応用例として, 水中のヒトリゾチームの分子動力学シミュレーションへ緩和モード解析を適用したNagai et al. (2013)の結果を紹介する. ヒトリゾチームは130残基, 2021原子からなるタンパク質で, 細菌の細胞壁にある多糖類の加水分解を行うという機能をもった酵素である. 天然構造で4個の α ヘリックスと1個の逆平行 β シートをもっている. 活性部位にはポケットがあり, このポケットの開閉運動で基質を捕捉することが知られている. ヒトリゾチームに7482個の水分子と8個の Cl^- を加えた24475個の全原子に対し, Langevin熱浴を使った等温分子動力学シミュレーションを温度300Kで100ns実行した.

解析に用いる変数は前節と同じ(4.1)式の R であるが, この節では, リゾチームの水素以外の全原子1029個を対象とする. 前節と同様に, 全原子の構造の多数のサンプルそれぞれを, 平均構造との2乗変位が最小になるように並進, 回転したものを用いる.

前節同様, 主成分分析を行う. 分散共分散行列(4.2)式に対して固有値問題(4.3)式を規格直交条件(4.4)のもとで解く. 固有値 Λ_p は大きい順に番号をつける. 第 p 主成分 $\Phi_p(Q)$ は(4.5)式で与えられる.

緩和モード解析を前節と同様に行うことを考える. 時間相関行列(4.7)式に対して一般化固有値問題(4.8)式を規格直交条件(4.9)式のもとで解く. 分子動力学シミュレーションではあるが,

位置のみを扱っているので、時間相関行列 $C(t)$ は対称行列となり、前節のモンテカルロ法の場合と同様の解析となる。これは、速度の緩和が位置に比べて非常に速いため、位置の遅い緩和だけに興味があるときは妥当な解析法である。

しかし、上記の方法はそのままでは使えない。それは、緩和モード解析の方法では、時間相関行列が比較的高い統計精度で評価されている必要があるからである。このことは、次のように理解できる。1成分の場合の緩和モード解析を考えてみる。1成分の確率過程を r とし、時間相関関数 $C(t) = \langle r(t)r(0) \rangle$ に対し、一般化固有値問題 $C(t_0 + \tau)f = \exp(-\lambda\tau)C(t_0)f$ を規格化条件 $fC(t_0)f = 1$ の下で解く。その結果は $\lambda = \{-\ln C(t_0 + \tau) + \ln C(t_0)\}/\tau$, $f = C(t_0)^{-1/2}$ で与えられる。緩和モードは $X = fr(t_0/2; r) = C(t_0)^{-1/2}r(t_0/2; r)$ で与えられ、 $g = C(t_0)f = C(t_0)^{1/2}$ を用いて逆変換 $r(t_0/2; r) = gX = C(t_0)^{1/2}X$ が与えられる。これより、 $\langle r(t_0)r(0) \rangle = C(t_0)$ および $\langle r(t_0 + \tau)r(0) \rangle = C(t_0 + \tau)$ が再現される。また、 $\tilde{g} = \exp(\lambda t_0/2)g = \exp(\lambda t_0/2)C(t_0)^{1/2}$ により $\langle r(t)r(0) \rangle \simeq \tilde{g}^2 \exp(-\lambda t)$ が与えられる。この式は $t = t_0$ と $t = t_0 + \tau$ では $C(t_0)$ と $C(t_0 + \tau)$ を正しく与える。 λ に関する結果は、 $C(t)$ 対 t の片対数プロットで、 $t = t_0$ と $t = t_0 + \tau$ に対する $C(t)$ の2点を通る直線の傾き $-\lambda$ を求めていることになる。 g に関する結果は、傾き $-\lambda$ の直線が $C(t_0)$ を通るように g を決めていくことになる。まとめると、この方法は $C(t) = \tilde{g}^2 \exp(-\lambda t)$ を仮定して、 $\ln C(t)$ 対 t の片対数プロットで、 $(t_0, \ln C(t_0))$ と $(t_0 + \tau, \ln C(t_0 + \tau))$ の2点を通るように直線を引き、傾き $-\lambda$ と振幅 \tilde{g}^2 を決めていくことになる。これより、 $C(t_0)$ と $C(t_0 + \tau)$ の値が比較的高い統計精度で求まっていなると、 λ と \tilde{g}^2 が良い精度で求まらないことになる。多成分の場合も同様で、緩和モード解析の方法は、 $C_{i,j}(t) = \sum_p \tilde{g}_{i,p} \tilde{g}_{j,p} \exp(-\lambda_p t)$ の形を仮定して、 $C_{i,j}(t_0)$, $C_{i,j}(t_0 + \tau)$ の値から λ_p と $\tilde{g}_{i,p}$ を決めることと等価である。従って、 $C_{i,j}(t_0)$, $C_{i,j}(t_0 + \tau)$ の値が比較的高い統計精度で求まっている必要がある。この要請のため、緩和モード解析ではタンパク質の全原子の巨大な自由度を直接扱うことはできず、何らかの方法で自由度を遁減する必要がある。

自由度遁減の方法として主成分分析の結果を用いる方法である主成分緩和モード解析(PCRMA)の方法について説明する。この方法では、まず全原子に対して主成分分析を行う。揺らぎの大きい上位の N_c 個の主成分によって張られる自由度に対して緩和モード解析を行う。このとき、試行関数は上位主成分

$$(5.1) \quad \Phi_c = {}^t(\Phi_1, \Phi_2, \dots, \Phi_{N_c}),$$

と変分パラメーター

$$(5.2) \quad \mathbf{f}_p = {}^t(f_{p,1}, f_{p,2}, \dots, f_{p,N_c}),$$

を用いて

$$(5.3) \quad X_p(Q) = {}^t \mathbf{f}_p \Phi_c(t_0/2; Q).$$

と与えられる。この試行関数より、相関行列

$$(5.4) \quad C(t) = (\Phi_c(t))^t \Phi_c(0)$$

に対する一般化固有値問題(4.8)式を規格直交条件(4.9)式のもとで解くことになる。ここでは、位置のみを扱っているので、 $\hat{\mathbf{f}}_p = \mathbf{f}_p$ が成立する。

上記の方法を試みたところ、主成分の自己相関関数をうまく再現できなかった。用いた発展時間 t_0 の値においては、与えた変分パラメーターでは表現できない速い緩和モードが、試行関数の中に含まれていたためと考えられる。言い換えると、上位の主成分には多くの速い緩和モードが含まれており、緩和モードの良い近似になっていない。そこで、2つの発展時間 t_1 と t_2 を

使った緩和モード解析を導入する。この場合、試行関数は変分パラメーター

$$(5.5) \quad \mathbf{f}_p^1 = {}^t(f_{p,1}^1, f_{p,2}^1, \dots, f_{p,N_c}^1), \quad \mathbf{f}_p^2 = {}^t(f_{p,1}^2, f_{p,2}^2, \dots, f_{p,N_c}^2),$$

を用いて

$$(5.6) \quad X_p(Q) = {}^t\mathbf{f}_p^1 \Phi_c(t_1/2; Q) + {}^t\mathbf{f}_p^2 e^{-\Gamma^+ t_2/2} \Phi_c(t_2/2; Q)$$

と与えられる。このように試行関数をとるのは、 $\Phi_c(t_1/2; Q)$ と $\Phi_c(t_2/2; Q)$ では、消去したい速い緩和モードの振幅が異なっているため、これらの線形結合で速い緩和モードを表現できると考えるからである。先程の1成分の例でいえば、単指数関数が良い近似になってない時間領域で、 $\langle r(t)r(0) \rangle = \tilde{g}_1^2 \exp(-\lambda_1 t) + \tilde{g}_2^2 \exp(-\lambda_2 t)$ のように2つの指数関数の和の形を仮定し、4つの t の値における $C(t)$ の値から4つのパラメーター $\tilde{g}_1^2, \tilde{g}_2^2, \lambda_1, \lambda_2$ を決めることに相当する。試行関数(5.6)式から一般化固有値問題

$$(5.7) \quad \tilde{C}(\tau) \mathbf{f}_p = e^{-\lambda_p \tau} \tilde{C}(0) \mathbf{f}_p$$

と、規格直交条件

$$(5.8) \quad {}^t\mathbf{f}_p \tilde{C}(0) \mathbf{f}_q = \delta_{p,q}$$

が得られる。ここで、

$$(5.9) \quad {}^t\mathbf{f}_p = ({}^t\mathbf{f}_p^1, {}^t\mathbf{f}_p^2)$$

であり、 $\tilde{C}(t)$ は(5.4)式で定義された $C(t)$ を用いて

$$(5.10) \quad \tilde{C}(t) = \begin{pmatrix} C(t_1 + t) & C(\frac{t_1+t_2}{2} + t) \\ C(\frac{t_2+t_1}{2} + t) & C(t_2 + t) \end{pmatrix}$$

と定義される。逆変換は

$$(5.11) \quad \mathbf{g}_p = {}^t({}^t\mathbf{g}_p^1, {}^t\mathbf{g}_p^2) = \tilde{C}(0) \mathbf{f}_p$$

を用いて

$$(5.12) \quad \begin{pmatrix} \Phi_c(t_1/2; Q) \\ \Phi_c(t_2/2; Q) \end{pmatrix} = \sum_{p=1}^{2N_c} \mathbf{g}_p X_p(Q)$$

と与えられる。このとき、 $\Phi_c(Q)$ の展開としては

$$(5.13) \quad \Phi_c(Q) \simeq \sum_{p=1}^{2N_c} \exp(\lambda_p t_1) \mathbf{g}_p^1 X_p(Q), \quad \Phi_c(Q) \simeq \sum_{p=1}^{2N_c} \exp(\lambda_p t_2) \mathbf{g}_p^2 X_p(Q)$$

という2種類が考えられる。ここでは2つをを平均した展開係数

$$(5.14) \quad \tilde{\mathbf{g}}_p^{\text{av}} = \{\exp(\lambda_p t_1) \mathbf{g}_p^1 + \exp(\lambda_p t_2) \mathbf{g}_p^2\} / 2$$

を用いた展開

$$(5.15) \quad \Phi_c(Q) \simeq \sum_{p=1}^{2N_c} \tilde{\mathbf{g}}_p^{\text{av}} X_p(Q)$$

を使うことにする。この展開により、(5.4)式で定義された主成分の相関行列は

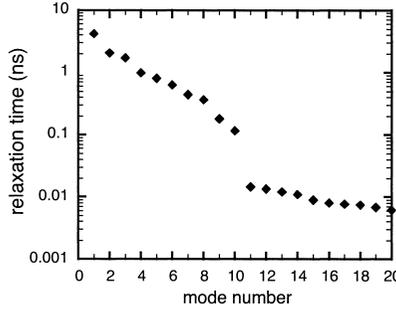


図 8. PCRMA で得られた緩和時間 λ_p^{-1} の分布. Reprinted with permission from T. Nagai, A. Mitsutake, and H. Takano, J. Phys. Soc. Jpn. **82**, 023803 (2013). Copyright 2013, The Physical Society of Japan.

$$(5.16) \quad C(t) \simeq \sum_{p=1}^{2N_c} \tilde{g}_p^{\text{av}} \tilde{g}_p^{\text{av}} \exp(-\lambda_p t)$$

のように再構成される.

位置の揺らぎを表す変数 \mathbf{R} は主成分を使って(4.6)式のように表される. 上位の主成分だけを使って, 小さい揺らぎを無視したという意味で粗視化された位置の揺らぎを表す変数 \mathbf{R}^c を

$$(5.17) \quad \mathbf{R}^c(Q) = \sum_{p=1}^{N_c} \Phi_p(Q) \mathbf{F}_p = \mathbf{F}^c \Phi_c(Q)$$

で定義する. ここで

$$(5.18) \quad \mathbf{F}^c = (\mathbf{F}_1, \dots, \mathbf{F}_{N_c})$$

である. (5.17)式に(5.15)式を代入することにより, \mathbf{R}^c は

$$(5.19) \quad \tilde{\mathbf{h}}_p = \mathbf{F}^c \tilde{g}_p^{\text{av}}$$

を用いて

$$(5.20) \quad \mathbf{R}^c \simeq \sum_{p=1}^{2N_c} X_p(Q) \tilde{\mathbf{h}}_p$$

と表される. (4.11)と同様, $\tilde{\mathbf{h}}_p/|\tilde{\mathbf{h}}_p|$ が第 p 緩和モードの運動の方向を表し, $X_p(Q)|\tilde{\mathbf{h}}_p|$ が PCA の $\Phi_p(Q)$ に相当する量であることがわかる.

実際の計算は各原子の位置座標に原子質量の平方根をかけた, 質量重み付けした座標を用いて行った. 上位主成分の数 N_c は 10 とした. これらの主成分の全分散への寄与は 38%であった. $t_1 = 0.8$ ps, $t_2 = 30.0$ ps, $\tau = 15.0$ ps と選び PCRMA を適用した. 得られた緩和率は小さい順に番号をつける. 図 8 は得られた緩和時間 λ_p^{-1} の分布を示している. $p = 1 - 10$ の緩和時間は $p = 11 - 20$ の緩和時間と比べて約 1 桁長い. 2つの発展時間を用いた PCRMA を適用したのは, 速い緩和モードの寄与を減らすためであり, $p = 11 - 20$ の緩和モードはそのような速い緩和モードで, ここでは重要ではない.

(5.16)式を用いて, 主成分の自己相関関数 $\langle \Phi_p(t) \Phi_p(0) \rangle$ を計算し, シミュレーションで直接計算した結果と比較した. 図 9 に $p = 1, 2$ の場合の結果を示す. ここで示したのは短時間領域

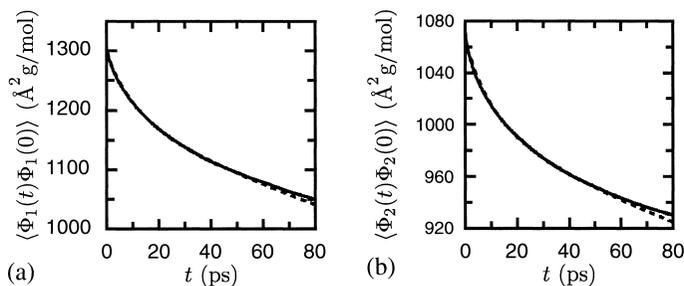


図9. 主成分の自己相関関数 $\langle \Phi_p(t)\Phi_p(0) \rangle$ の t 依存性. $p=1$ の場合を (a) に, $p=2$ の場合を (b) に示した. 実線がシミュレーションから直接計算した結果, 破線が(5.16)式を用いて PCRMA から計算した結果を表す. Reprinted with permission from T. Nagai, A. Mitsutake, and H. Takano, *J. Phys. Soc. Jpn.* **82**, 023803 (2013). Copyright 2013, The Physical Society of Japan.

であるが, PCRMA の結果がシミュレーションによる直接計算の結果と良く一致していることがわかる. 同様の良い一致は $p=3-10$ でも見られた. このことから, PCRMA の結果は妥当なものと考えられる.

図10に第1, 第2, 第3モードベクトルをPCAの場合とPCRMAの場合に示した. 図10(a)でわかるように, PCAの第1モードは外に出ているターンの部分のPro103とGln104の残基の揺らぎが局所的に大きい. 一方, RMAの第1モードにはそのような大きな揺らぎはPro103とGln104の近傍には無く, 協同的なモードになっている. Pro103とGln104の大きな揺らぎはRMAの第3モードに見ることができる.

ここでの結果が例示するように, PCAは速くて局所的な大きな揺らぎを上位の主成分として取り出すことがある. PCAとPCRMAのモードベクトルの類似性を調べてみると, 揺らぎの大きさが必ずしも運動の遅さに対応しておらず, PCRMAが遅い緩和を取り出すのに有効であることがわかる.

6. まとめ

詳細釣り合いの成り立つシミュレーション法に対して, 時間発展演算子の固有値, 固有関数から, 扱っている系の緩和率, 緩和モードをそれぞれ定義することができることを説明した. 計算機シミュレーションの時系列データから緩和率分布と緩和モードを変分問題に基づいて評価する方法である緩和モード解析の方法について説明し, 孤立線状高分子系の分子動力学シミュレーションへの応用例を示した. 生体高分子系への応用として, 真空中のペプチド系のモンテカルロ・シミュレーションと水中のタンパク質系の分子動力学シミュレーションへ適用した例を紹介した. ペプチド系の例では, 主成分分析では抽出できない遅い運動を抽出できることを示した. また, 緩和モード解析による自由エネルギー地形では, 極小点が軸に沿って並ぶ傾向があり, 安定・準安定状態間の遷移を良く記述できていると考えられることを見た. タンパク質系の例では, 全原子の大きな自由度を系統的に削減する方法として, 主成分分析を先に行い, 上位の主成分だけを用いて緩和モード解析を行う, 主成分緩和モード解析の方法を紹介した. さらに, 1つの発展時間だけでは残ってしまう遅い緩和モードの影響を除くための方法として, 2つの発展時間を用いる緩和モード解析の方法を紹介した. 2つの発展時間を用いる主成分緩和モード解析により, 全原子を扱いながら遅い緩和モードを系統的に抽出できることを示した.

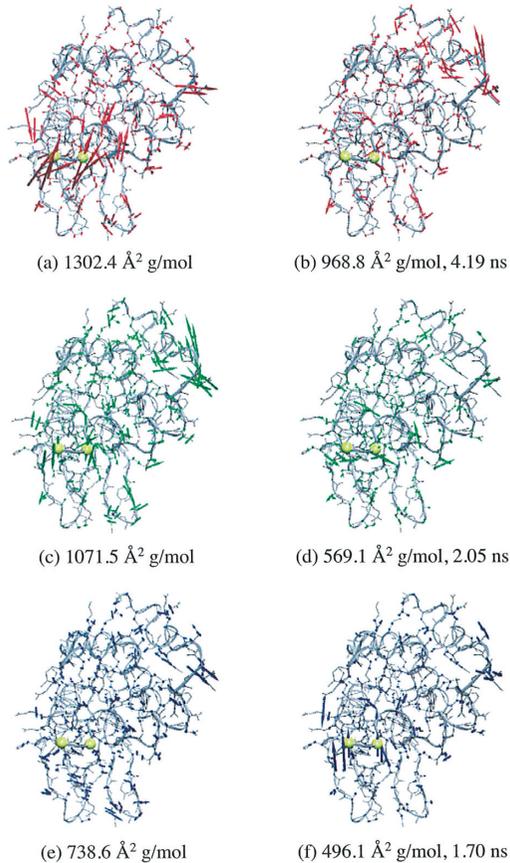


図 10. 平均構造上の C_α , C_β , C_γ 原子の位置でのモードベクトル. PCA で得られた $\Lambda_1^{1/2} \mathbf{F}_1$ を (a) に, $\Lambda_2^{1/2} \mathbf{F}_2$ を (c) に, $\Lambda_3^{1/2} \mathbf{F}_3$ を (e) に, PCRMA で得られた \mathbf{h}_1 を (b) に, \mathbf{h}_2 を (d) に, \mathbf{h}_3 を (f) に示した. 質量重み付けを外し, 振幅を 8 倍している. PCA の結果には分散 Λ_n を, PCRMA の結果には振幅 $|\mathbf{h}_p|^2$ と緩和時間 λ_p^{-1} を数値として示してある. Pro103 と Gln104 の C_α 原子を黄色い球で示した. Reprinted with permission from T. Nagai, A. Mitsutake, and H. Takano, J. Phys. Soc. Jpn. **82**, 023803 (2013). Copyright 2013, The Physical Society of Japan.

以上, 生体高分子系のシミュレーションから得られる時系列データに対して, 緩和モード解析の方法が適用できることを示した. 近年, タンパク質の分子動力学シミュレーションが盛んに行われてきているため, その時系列データの動的解析方法として, 2つの発展時間を用いる主成分緩和モード解析は特に有用であると考えられる.

付 録

A. 固有値方程式と変分問題

(3.3)式に対し, $\delta \mathcal{R} = 0$ が固有値方程式(2.13)式, (2.14)式を与えることを見る. $\delta \mathcal{R}$ は次の

ように計算される.

$$(A.1) \quad \delta\mathcal{R} = \frac{1}{\sum_Q \phi_n(Q)\psi_n(Q)} \left[\sum_Q \delta\phi_n(Q) \left\{ \sum_{Q'} T_\tau(Q|Q')\psi_n(Q') - \mathcal{R}\psi_n(Q) \right\} \right. \\ \left. + \sum_{Q'} \left\{ \sum_Q \phi_n(Q)T_\tau(Q|Q') - \mathcal{R}\phi_n(Q') \right\} \delta\psi_n(Q') \right].$$

ここで, $\psi_n(Q) = \hat{\phi}_n(Q)P_{\text{eq}}(Q) = \phi_n(\epsilon Q)P_{\text{eq}}(Q)$ であるから, 上式は

$$(A.2) \quad \delta\mathcal{R} = \frac{1}{\sum_Q \phi_n(Q)\psi_n(Q)} \left[\sum_Q \delta\phi_n(Q) \left\{ \sum_{Q'} T_\tau(Q|Q')\phi_n(\epsilon Q')P_{\text{eq}}(Q') - \mathcal{R}\phi_n(\epsilon Q)P_{\text{eq}}(Q) \right\} \right. \\ \left. + \sum_{Q'} \left\{ \sum_Q \phi_n(Q)T_\tau(Q|Q') - \mathcal{R}\phi_n(Q') \right\} \delta\phi_n(\epsilon Q')P_{\text{eq}}(Q') \right] \\ = \frac{1}{\sum_Q \phi_n(Q)\psi_n(Q)} \left[\sum_Q \delta\phi_n(Q) \left\{ \sum_{Q'} T_\tau(Q|Q')\phi_n(\epsilon Q')P_{\text{eq}}(Q') - \mathcal{R}\phi_n(\epsilon Q)P_{\text{eq}}(Q) \right\} \right. \\ \left. + \sum_{Q'} \left\{ \sum_Q \phi_n(Q)T_\tau(\epsilon Q'|\epsilon Q)P_{\text{eq}}(\epsilon Q) - \mathcal{R}\phi_n(Q')P_{\text{eq}}(Q') \right\} \delta\phi_n(\epsilon Q') \right] \\ = \frac{1}{\sum_Q \phi_n(Q)\psi_n(Q)} \left[\sum_Q \delta\phi_n(Q) \left\{ \sum_{Q'} T_\tau(Q|Q')\phi_n(\epsilon Q')P_{\text{eq}}(Q') - \mathcal{R}\phi_n(\epsilon Q)P_{\text{eq}}(Q) \right\} \right. \\ \left. + \sum_{\epsilon Q'} \delta\phi_n(Q') \left\{ \sum_{\epsilon Q} T_\tau(Q'|Q)\phi_n(\epsilon Q)P_{\text{eq}}(Q) - \mathcal{R}\phi_n(\epsilon Q')P_{\text{eq}}(Q') \right\} \right] \\ = \frac{2}{\sum_Q \phi_n(Q)\psi_n(Q)} \sum_Q \delta\phi_n(Q) \left\{ \sum_{Q'} T_\tau(Q|Q')\phi_n(\epsilon Q')P_{\text{eq}}(Q') - \mathcal{R}\phi_n(\epsilon Q)P_{\text{eq}}(Q) \right\}$$

と変形され, $\delta\mathcal{R} = 0$ より固有値方程式 (2.14) 式

$$\sum_{Q'} T_\tau(Q|Q')\phi_n(\epsilon Q')P_{\text{eq}}(Q') - \mathcal{R}\phi_n(\epsilon Q)P_{\text{eq}}(Q) = 0$$

が得られ, このときの \mathcal{R} の値が固有値 $\exp(-\lambda_n)$ となる. ここで, (3.16) 式と (2.12) 式を用いた. 同様に, (A.2) 式は

$$(A.3) \quad \delta\mathcal{R} = \frac{1}{\sum_Q \phi_n(Q)\psi_n(Q)} \left[\sum_Q \delta\phi_n(Q) \left\{ \sum_{Q'} T_\tau(\epsilon Q'|\epsilon Q)\phi_n(\epsilon Q')P_{\text{eq}}(\epsilon Q) - \mathcal{R}\phi_n(\epsilon Q)P_{\text{eq}}(Q) \right\} \right. \\ \left. + \sum_{Q'} \left\{ \sum_Q \phi_n(Q)T_\tau(Q|Q') - \mathcal{R}\phi_n(Q') \right\} \delta\phi_n(\epsilon Q')P_{\text{eq}}(Q') \right] \\ = \frac{1}{\sum_Q \phi_n(Q)\psi_n(Q)} \left[\sum_{\epsilon Q} \left\{ \sum_{\epsilon Q'} \phi_n(Q')T_\tau(Q'|Q) - \mathcal{R}\phi_n(Q) \right\} \delta\phi_n(\epsilon Q)P_{\text{eq}}(Q) \right]$$

$$\begin{aligned}
 & + \sum_{Q'} \left\{ \sum_Q \phi_n(Q) T_\tau(Q|Q') - \mathcal{R} \phi_n(Q') \right\} \delta \phi_n(\epsilon Q') P_{\text{eq}}(Q') \Big] \\
 & = \frac{2}{\sum_Q \phi_n(Q) \psi_n(Q)} \sum_{Q'} \left\{ \sum_Q \phi_n(Q) T_\tau(Q|Q') - \mathcal{R} \phi_n(Q') \right\} \delta \phi_n(\epsilon Q') P_{\text{eq}}(Q')
 \end{aligned}$$

と変形され、 $\delta \mathcal{R} = 0$ より固有値方程式 (2.13) 式

$$\sum_Q \phi_n(Q) T_\tau(Q|Q') - \mathcal{R} \phi_n(Q') = 0$$

が得られ、このときの \mathcal{R} の値が固有値 $\exp(-\lambda_n)$ となる。

$\phi_n(Q) = \hat{\phi}_n(Q)$ となるモンテカルロ法や Brown 動力学法では、緩和率 λ_n は全て 0 以上の実数で、 $\delta \mathcal{R} = 0$ となる \mathcal{R} の値は極値となり、変分問題 $\delta \mathcal{R} = 0$ は通常の変分法としての意味をもつ。これに対し、分子動力学法では、緩和率 λ_n が複素数となるため、 $\delta \mathcal{R} = 0$ となる \mathcal{R} の値は一般に極値とはならず、通常の変分法としての意味はもたない。1 次の変分 $\delta \mathcal{R} = 0$ が固有値方程式を与え、そのときの \mathcal{R} の値が固有値になるという意味で、 \mathcal{R} は有用な量である。

Hirao et al. (1997) では、(A.1) 式の $\delta \phi_n(Q)$ と $\delta \psi_n(Q)$ があたかも独立であるかのように扱っている。得られる結果は正しいが、詳細釣り合いの条件 (2.11) 式が成立しているので、 $\delta \phi_n(Q)$ と $\delta \psi_n(Q)$ は独立ではなく、上記のように扱わなくてはならない。

B. 試行関数の選び方

試行関数 (3.5) 式の選び方について考える。緩和モード解析により、複数の自由度がどのように協力して遅い緩和を実現しているかを明らかにすることができる。従って、対象としている運動を構成している自由度のうち、重要と考えられる自由度の線形結合を試行関数とする。例えば、ランダムスピン系でスピンの協力的な運動を取り出すには、各スピンの値の線形結合を用いる (Takano and Miyashita, 1995)。第 3 節の孤立高分子系の例では、高分子の重心周りでの変形運動を問題にしたので、各セグメントの重心からの相対座標の線形結合を用いた。第 4 節、第 5 節で出てくる生体高分子系の例では、並進と回転を除いた後の平均的な構造のまわりの遅い揺らぎを調べるため、平均的な構造からの原子の変位の線形結合を用いている。遅い緩和に重要だと思われる自由度が予想できていれば、それらの線形結合を用いれば、より良い試行関数を構成することができる。

次に、試行関数の中の t_0 の役割について考える (Mitsutake et al., 2011)。簡単のため、純粋緩和だけの場合を考え、第 4 節の変数 (4.1) 式を用いる。(3.5) 式に相当する試行関数 (近似的緩和モード) は

$$(B.1) \quad X_p(Q) = \sum_i f_{p,i} R_i(t_0/2; Q)$$

となる。 $R_i(Q)$ が厳密な緩和モード $\{\phi_n(Q)\}$ で

$$(B.2) \quad R_i(Q) = \sum_n r_{i,n} \phi_n(Q)$$

と展開されるとすると、(3.6) 式より、

$$(B.3) \quad R_i(t; Q) = \sum_{Q'} \sum_n r_{i,n} \phi_n(Q') T_t(Q'|Q) = \sum_n r_{i,n} \exp(-\lambda_n t) \phi_n(Q)$$

となる．ここで， λ_n は $\phi_n(Q)$ の緩和率で，番号は λ_n の小さい順につけるものとする．第4節の例では， \mathbf{R} が $3N - 6$ 個の自由度をもち，一般化固有値問題(4.8)式および(4.9)式を解くことで， \mathbf{R} に含まれる遅い方から $3N - 6$ 個の緩和モードを近似的に推定する．もし， \mathbf{R} が $3N - 6$ 個の緩和モードしか含んでいなければ， $t_0 = 0$ としても緩和モードを精度良く推定することができる．しかし， \mathbf{R} には最も遅い $3N - 6$ 個の緩和モード以外に，より速い緩和モードが多数含まれている．(B.3)式を見ると，速い緩和モードの遅い緩和モードに対する相対的寄与の大きさ，例えば， $1 \leq m \leq 3N - 6 < n$ として， n 番目の緩和モードの m 番目の緩和モードに対する相対的寄与の大きさは $|r_{i,n}/r_{j,m}| \exp[-(\lambda_n - \lambda_m)t]$ で与えられ， $\lambda_n > \lambda_m$ であることから， t を大きくすることで，速い緩和モードの相対的寄与が小さくなることがわかる．(B.1)式の中の t_0 は，試行関数に含まれる推定範囲外の速い緩和モードの寄与を減らすことで，緩和モード推定の精度を上げる役割をしている．

謝 辞

共同研究者である宮下精二氏，平尾秀友氏，古関幸子氏，光武亜代理氏，飯島寛允氏，長井俊樹氏，山本卓氏，小泉祐太氏，そして，この報告を書く機会を与えて下さいました伊庭幸人氏に深く感謝致します．

参 考 文 献

- Doi, M. and Edwards, S. F. (1986). *The Theory of Polymer Dynamics*, Oxford University Press, Oxford.
- Drozhdov, A. N. (2000). Variational theory of activated rate processes for an arbitrary barrier, *Physical Review E*, **62**, 1879–1892.
- Hirao, H., Koseki, S. and Takano, H. (1997). Molecular dynamics study of relaxation modes of a single polymer chain, *Journal of the Physical Society of Japan*, **66**, 3399–3405.
- Koseki, S., Hirao, H. and Takano, H. (1997). Monte Carlo study of relaxation modes of a single polymer chain, *Journal of the Physical Society of Japan*, **66**, 1631–1637.
- Kremer, K. and Grest, G. S. (1990). Dynamics of entangled linear polymer melts: A molecular-dynamics simulation, *The Journal of Chemical Physics*, **92**, 5057–5086.
- Mitsutake, A., Hansmann, U. H. E. and Okamoto, Y. (1998). Temperature dependence of distributions of conformations of a small peptide, *Journal of Molecular Graphics and Modelling*, **16**, 226–238.
- Mitsutake, A., Iijima, H. and Takano, H. (2011). Relaxation mode analysis of a peptide system: Comparison with principal component analysis, *The Journal of Chemical Physics*, **135**, 164102-1-15.
- Nagai, T., Mitsutake, A. and Takano, H. (2013). Principal component relaxation mode analysis of an all-atom molecular dynamics simulation of human lysozyme, *Journal of the Physical Society of Japan*, **82**, 023803-1-4.
- Risken, H. (1996). *The Fokker-Planck Equation: Methods of Solution and Applications*, Springer-Verlag, Berlin, Heidelberg.
- Takano, H. and Miyashita, S. (1995). Relaxation modes in random spin systems, *Journal of the Physical Society of Japan*, **64**, 3688–3698.
- Talkner, P. (1994). Finite barrier corrections for the Kramers rate problem in the spatial diffusion regime, *Chemical Physics*, **180**, 199–213.

Relaxation Mode Analysis of Biopolymer Systems

Hiroshi Takano

Faculty of Science and Technology, Keio University

For simulation methods that satisfy the detailed balance condition, relaxation rates and modes of a system concerned can be defined from the eigenvalues and eigenvectors of the time-evolution operator, respectively. The relaxation mode analysis method, which estimates the relaxation rate distribution and relaxation modes on the basis of the variational method, is explained. Applications of the method to Monte Carlo simulation of a peptide system and molecular dynamics simulation of a protein system are reviewed.