11-DA01

戸田幹人（奈良女子大学）

生体分子の大規模分子動力学計算に対する時系列解析の新展開

生体分子機能が、揺らぐ環境下、頑健に発現する動力学的機構を解明する

生体分子機能とは、揺らぐ外界の下、刺激に対する応答としての階層的集団運動が、分子自身の非線型動力学により頑健かつ自発的に形成され、それがさらに一連の構造変化・化学反応を動的に誘起すること。

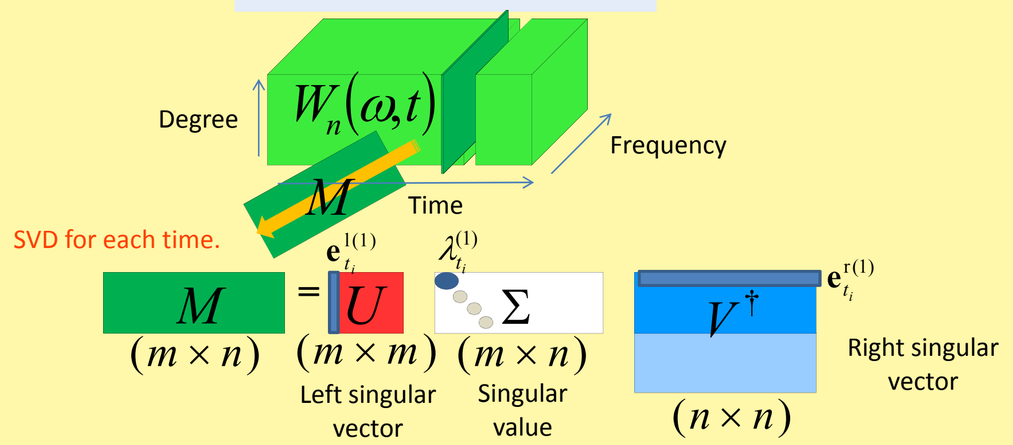
その動力学的機構を理解するため、分子動力学計算で得られる大規模時系列データから、様々な階層的集団運動を抽出し、それらの集団運動の動力学的な特性と機能との関連を明らかにする。

階層的集団運動を抽出する「データマイニング」

タンパク質分子の2次構造・3次構造と階層的集団運動との対応を明らかにし、特に近年新たに提唱されているfly-casting 機構の動力学的側面など、機能発現における「運動と構造」の関連を解明する。

そのために、ウエーブレット変換や特異値分解などを応用し、集団運動を抽出する「データマイニング」の新たな手法を開拓する。下図は富士香奈による。左特異ベクトルは空間、右特異ベクトルは周波数の情報を表す。

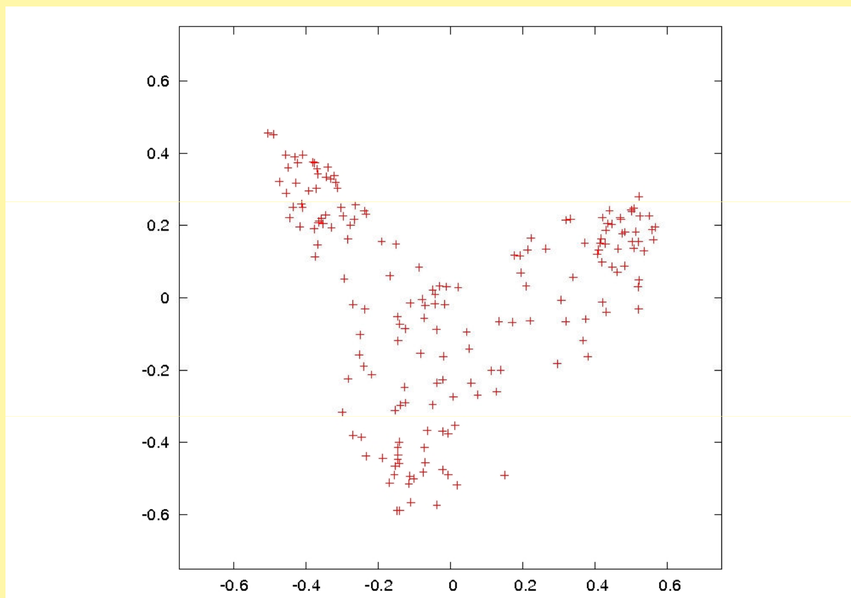
$$M = U \Sigma V^{\dagger}$$



相関・因果関係を解析する統計的因果推論

因果関係の厳密な定式化により、非線形相関の解析から因果関係を推論する。

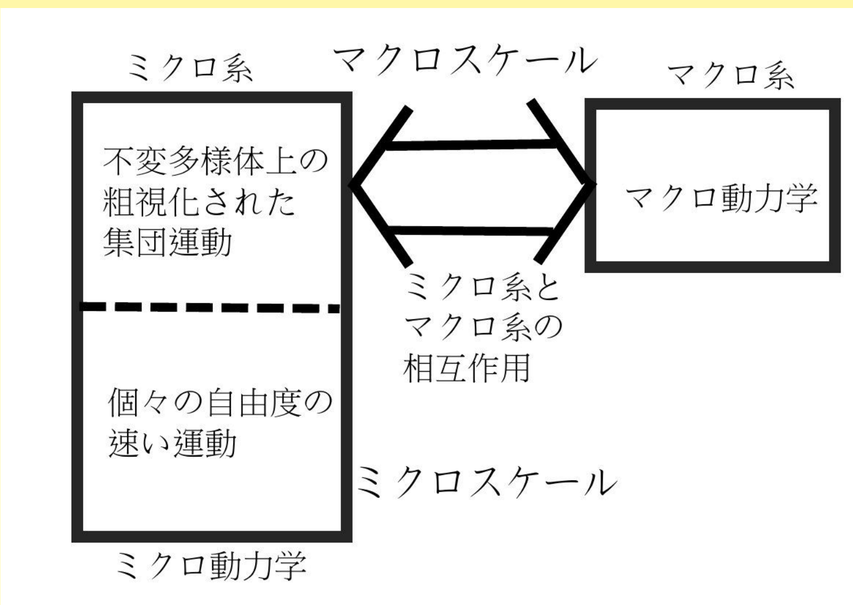
カーネル法をはじめとする近年の統計科学の成果を応用し、大規模な分子動力学シミュレーションで得られる時系列データに対して、階層的集団運動の相互間の非線形相関解析、さらにはそれらの因果関係の解析を遂行する。下図はカーネル主成分解析の一例。 第一と第二主成分への射影によってクラスター分析ができる。

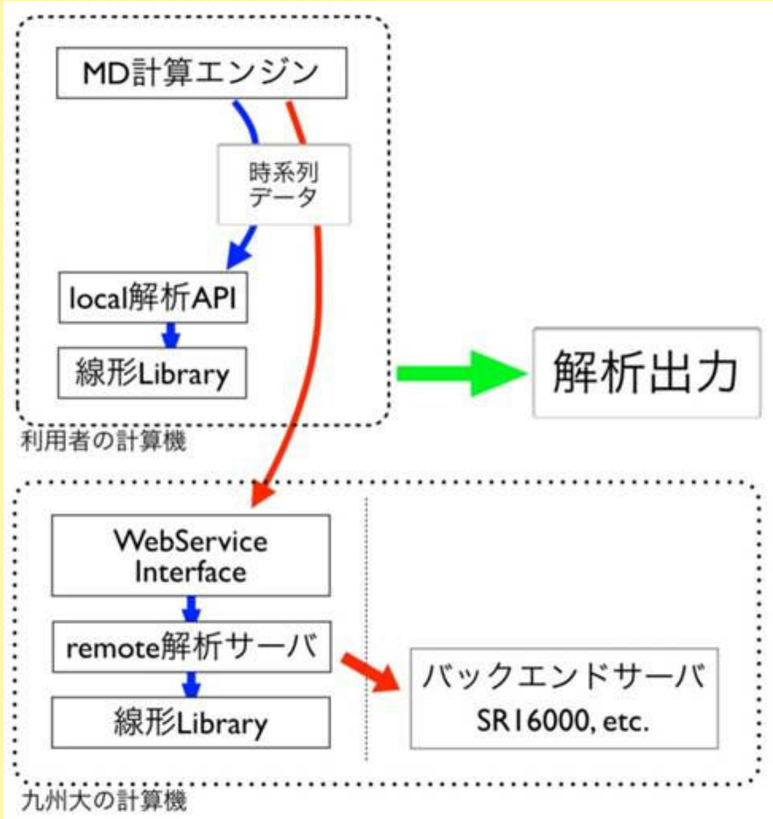


不変集合解析による「粗視化」とそれに基づく階層的アルゴリズム

大自由度系の時系列から相空間構造の不変集合を数値的に再構築する手法を開拓する。データ可視化において「多様体復元」と呼ばれる計算法を相空間に拡張し、特に「ゆっくりした集団運動」の動力学を記述する「不変集合」の構成を行う。

このように再構成された集団運動の動力学に依拠した、新たな「粗視化」モデルを提案する。



大規模時系列データに対する効率的な実装

大規模時系列データの「次元縮約」、「次元縮約」で得られた集団運動に基づく「粗視化モデル」構築、長時間時系列データに対するカーネル法の遂行において、システムの効率的な実装が必要であり、特に分子科学や統計科学の研究者にとって便利なシステム設計・構築を行う。