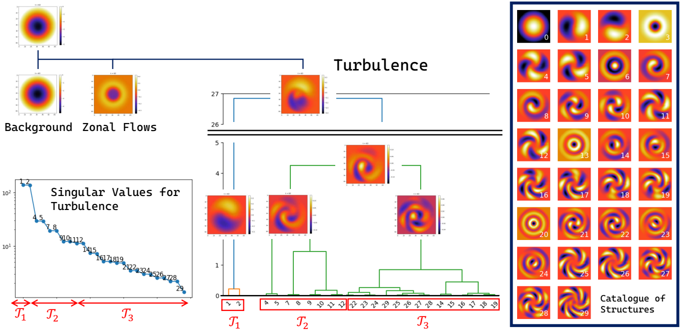
# 簡約流体モデルに基づく数値乱流場の構造クラスタリング

奥野彰文[1,2], 古田原拓実[3], 佐々木真[3]

[1]統計数理研究所　[2]理化学研究所　[3]日本大学

- 数値乱流場 (静電ポテンシャル$\phi(x, y, t)$) のモードのクラスタリング



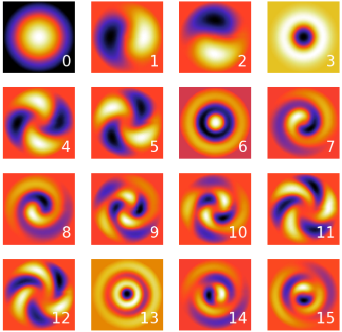

- Okuno, Kodahara, and Sasaki, “Hierarchical Clustering of Modes in Numerical Turbulence Fields”, PFR: Rapid Comunications, accepted.

# 背景 (1/2)

- 核融合プラズマの特性は乱流に支配されており，動的現象の理解が喫緊の課題.
  - 本研究ではKelvin-Helmholtz乱流シミュレーションにより得られた静電ポテンシャル$\phi(x, y, t)$を考える.
  - 乱流を直接的に扱うことは難しい．ので分解をしたい.
- 単純なフーリエ展開では多くの自由度(基底)が生じる.
  - 分解の粒度が細かすぎて，内部の挙動を理解することは困難.
- 動的モード分解 (Sasaki et al. PPCF2019) やニューラルネット (Jajima et al. PPCF2023) を利用した解析手法が報告されている.

# 背景 (2/2)

- 特異値分解 (SVD) を利用して乱流を分解し，その過程を解析するアプローチが提案されている (Sasaki et al. PPCF2019; Kodahara et al. PFR2023).
- SVDでも多くの自由度(基底)が残る.




- 本研究では「乱流の分解」を体系的に扱う方法を考える.

# 静電ポテンシャルの分解

- 乱流 (静電ポテンシャル) $\phi(x, y, t)$ をそのまま扱うのは難しいので，独立な要素$\phi_1, \phi_2, \ldots$に分解:

$$\phi(x, y, t) = \sum_{j \geq 0} \phi_j(x, y, t).$$

**特異値分解 (a.k.a., 主成分分析)**

$$\phi_j(x, y, t) = s_j \Psi_j(x, y) h_j(t).$$

$s_j$は特異値(要素$j$の重要度)，　$\Psi_j(x, y)$は空間構造，　$h_j(t)$は時間発展を表す．

- 実際にはシミュレーションの観測値から離散的に計算する．

# より詳細な構造$\Psi_j(x, y)$のリスト

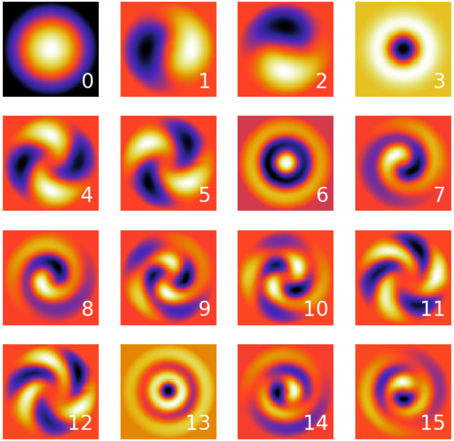
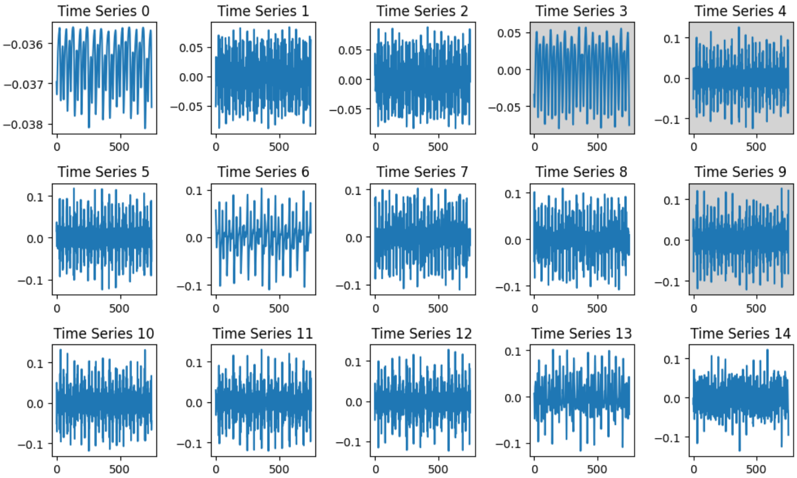



Figure: 背景構造 ($j = 0$), Zonal ($j = 3, 13, \ldots$), 乱流 (それ以外)

# 時間成分$h_j(t)$のリスト



(1) (とびぬけて寄与率が大きい)$j=0$を背景構造とする

(2) $v(r,\theta)=(r\cos\theta, r\sin\theta)$として，周方向の$\Psi_j$の変動の総和

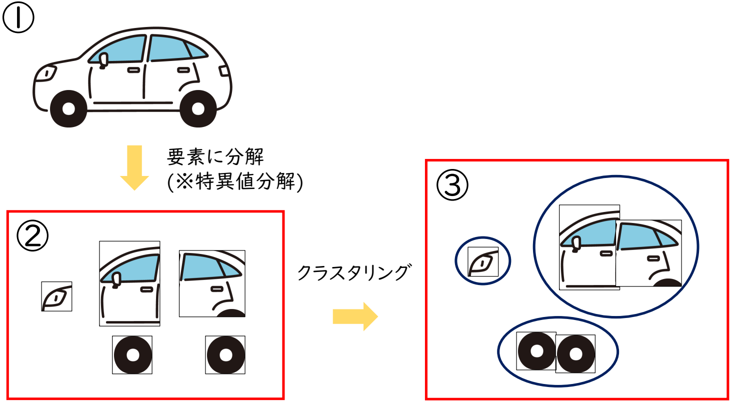
$$c_j := \iint \left\{ \frac{\partial \Psi_j(v(r,\theta))}{\partial \theta} \right\}^2 \mathrm{d}r\mathrm{d}\theta,\ (j=1,2,\ldots)$$

を計算する.

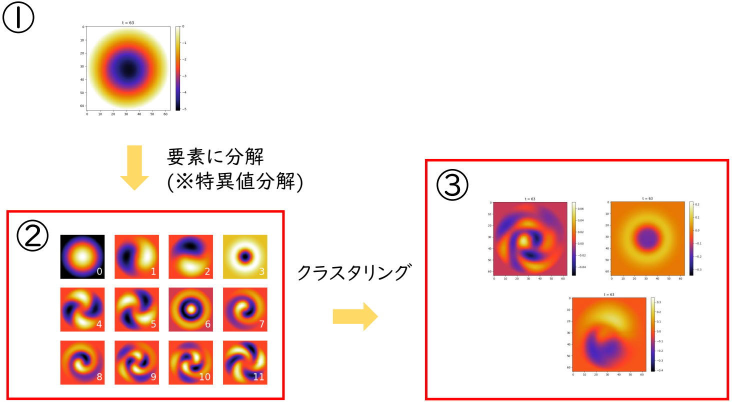
- $c_j$が特定の閾値$c_*$より小さければ(周方向に変動が小さいので)帯状流,
- そうでなければ乱流の成分とする.

(3) 帯状流，乱流の要素をそれぞれクラスタリングしてまとめる.

# クラスタリングとは?




- クラスタリング=似た要素をまとめ上げる手続き.




- 最終的に,「細かすぎない」分解が得られる.

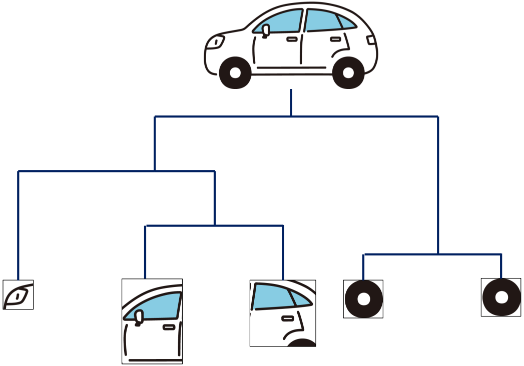
# 代表的なクラスタリング手法と要望

- Sasaki et al. (PoP2019) $\Rightarrow$ 手動でのクラスタリング.
- 自動化も可能.
  1. $k$-meansクラスタリング,
  2. mean-shiftクラスタリング,
  3. 混合正規クラスタリング,...

基本的には"いくつのクラスタにまとめるか"を指定してクラスタリングする.

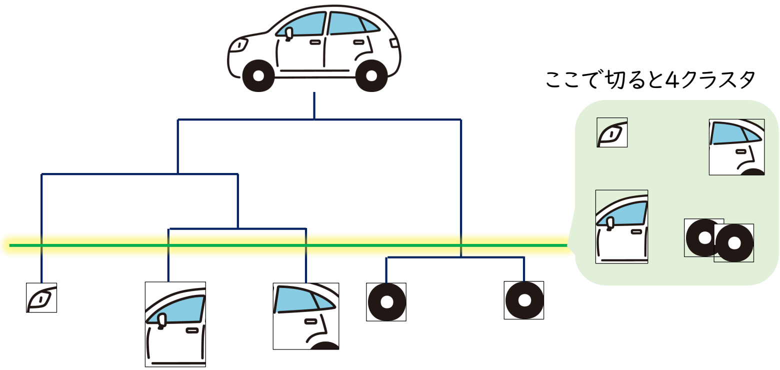
- 自動化しつつ，粒度の設定をフレキシブルにしたい.

- 全体を，徐々に要素に分解する手法.



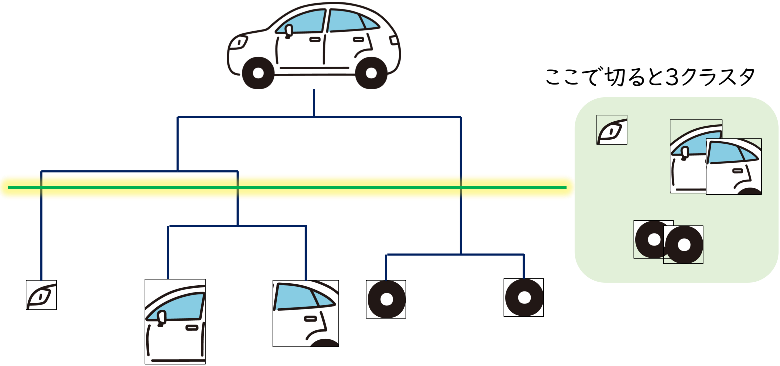
# 階層型クラスタリング

- 細かいまとめ方 (4クラスタ)



# 階層型クラスタリング: より荒い分類

- もっと大まかなまとめ方 (3クラスタ)



# どうやって計算するのか?

1. 要素$i, j$間の乖離度(距離)$d_{ij}$を計算.
2. パッケージに$D = (d_{ij})$を投げる⇒デンドログラムが生成される.
3. おわり.

```
linkage_D = linkage(D, method="ward")
dendrogram(linkage_D)
```

- 距離$d_{ij}$が近いものから徐々に合体させていくクラスタリング.

モード$\{\phi_j(x,y,t)\}_j$をまとめたいので，モード$\phi_i, \phi_j$の間の乖離度を定義する必要がある．
いくつかの候補があった．

- 回転を無視した構造の乖離度: $d_{ij} = \min_{\mathcal{O}:\text{回転行列}} \|\Psi_i - \mathcal{O}\Psi_j\|$,
- $h_i(t), h_j(t)$の周波数の乖離度,
- 特異値の乖離度 $d_{ij} = |s_i - s_j|$, …

- 特異値の乖離度で階層型クラスタリング(ward法)するのが一番素朴でよかった.

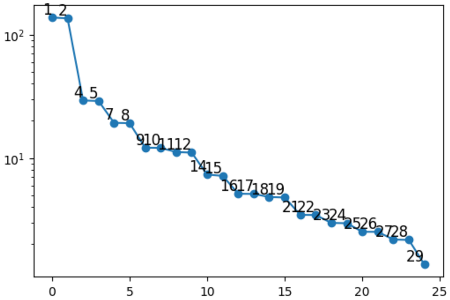



Figure: 第1特異値–第30特異値のプロット.

各クラスタ$\mathcal{J} \subset \{1, 2, \ldots\}$について, $\phi_{\mathcal{J}}(x, y, t) = \sum_{j \in \mathcal{J}} s_j \Psi_j(x, y) h_j(t)$を計算する.

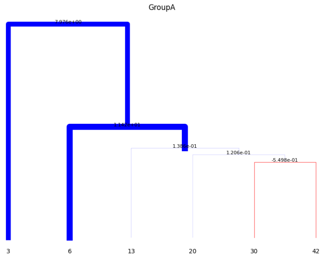
# 得られたデンドログラムと，その帰結としての分解

- 乱流がさらに線形不安定モード・非線形モードに分解された.
  - 非線形モードは突発性に重要な寄与をする可能性 (Sasaki et al. PPCF2020).
  - これまでの「背景・帯状流・乱流」の3分割より細かくできる.

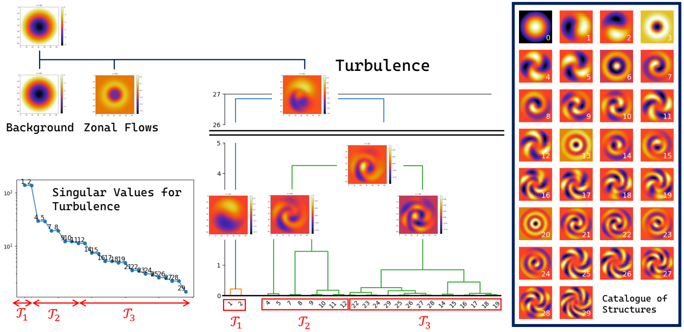
- テンソル型の分解：今回は$\phi_j(x, y, t) = s_j \Psi_j(x, y) h_j(t)$型だったが，

$$\phi_j(x, y, t) = \alpha_j(x) \beta_j(y) h_j(t)$$

  型などにも分解できる．いろいろ試してみたい．
- 各クラスタから背景へのエネルギー輸送



# まとめ




- Okuno, Kodahara, and Sasaki, “Hierarchical Clustering of Modes in Numerical Turbulence Fields”, PFR: Rapid Comunications, accepted.
- これまで手動でやっていた分解を体系的に自動化できた.
- 特異値が重複しないケースはより複雑で，研究中の課題.
- 連絡先：`okuno@ism.ac.jp`