流線位相構造の自動抽出の精度評価のための
テストデータセットの生成方法

2020SE051 齋藤智海

指導教員：横山哲郎

1　はじめに

流体力学は，川や大気の流れ，血流などの自然界や生体内で観察される気体，液体の流れの運動を研究する学問分野である．現実の流体の挙動は予測が困難であり，そのデータは膨大で解析が困難であるが，数値解析や離散解析の進歩により，実験的アプローチが難しい事故や災害などの問題も計算手法によって解析可能となっている．トポロジカルデータ解析（TDA: Topological Data Analysis）とは，データの形状や連結性をトポロジーの視点から分析する手法であり，パーシステントホモロジーの理論を用いて，複雑な多次元データの分類を行う．TDA は，データ科学や機械学習の分野での応用がされている（例えば，[2]）．流体力学の分野において，トポロジカル流れデータ解析（TFDA: Topological Flow Data Analysis）が注目されている．TFDA は，流体を対象に TDA を応用するものであり，流体の構造変化をそのトポロジーの変化として捉える．特に，TFDA のアプローチの一つとして，流線構造と一対一に対応する木構造である COT を用いる方法が提案されている．この方法により，構造安定な流体の完全な分類が行われ，その構造の変化をより明確に理解し解析することができる．大量の流れに関するデータを入力として TFDA を行うには，流れのデータからトポロジーを効率的かつ高精度に自動抽出する前処理が必要である．

`psiclone`[1] は，TFDA を支援する Python ライブラリである．このライブラリは，パーシステントホモロジーの理論を用いて構築したレーブグラフを活用し，二次元配列で表現されるハミルトンベクトル場から，COT 表現で表されたトポロジーを抽出する機能をもつ．現在，生成される COT 表現の精度や生成–消滅ペアの寿命に基づくフィルタリングの閾値に応じた解析の実行時間に関する定量的評価は行われていない．精度評価を行うためには，適切なテストデータセットが必要である．

本研究では，パーシステントホモロジーとレーブグラフを用いた COT 表現の自動抽出手法の評価を行うことを目的とする．このために，ハミルトンベクトル場を系統的に生成する方法を提案し，局所部分構造を最大 1 つ含む流線位相構造をもちノイズのないテストデータセットを生成する．さらに，`psiclone` を用いて，これらのデータセットから一定のクラスの流線位相構造が正確に抽出できることを確認する．

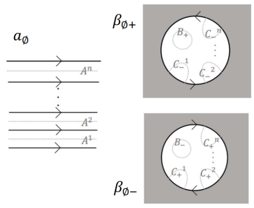

図 1　基本構造

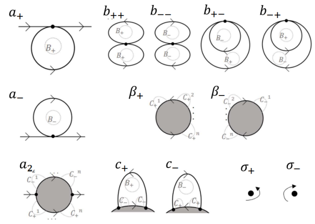

図 2　局所部分構造

2　COT 表現

ハミルトニアン $H\colon \mathbb{R}^2 \to \mathbb{R}$ およびハミルトンベクトル場 $u = (\partial_y H, -\partial_x H)$ は，非圧縮性・非粘性をもつ定常な流れを特徴づける [3, 1]．本稿では，流体力学の慣習に倣ってハミルトニアンの等高線を流線と呼ぶ．構造安定なハミルトンベクトル場の流線構造に一対一に対応する表現として，部分円順序木（COT: partially Cyclic Ordered Tree）表現 [1] が知られている．COT は，流線トポロジーを表す木であり，特異点や境界の周りの回転による同一視を同値関係とする同値類である．COT は，図 1にある基本構造のうちの 1 つをもち，再帰的に組み合わせた図 2の局所部分構造をもつ．

基本構造は，非有界領域内を横切る一様流 $a_\emptyset$ と有界領域内の回転流 $\beta_{\emptyset\pm}$ である．局所部分構造は 3 つの系列と回転流の中心である渦心点 $\sigma_\pm$ からなる．A 系列は，無限遠から無限遠への軌道上に一つの回転流が接続した構造 $a_\pm$ および穴（物体）が接続した構造 a_2 からなる．B 系列は，2 つの回転流が 8 の字に接続している構造 $b_{\pm\pm}$*1, 1 つの回転流の内部に 1 つの逆向きの回転流が 1 点で接続している構造 $b_{\pm\mp}$, 穴の境界周りを流れる構造 $\beta_\pm$ からなる．C 系列は，穴から軌道が接続する構造 $c_\pm$ からなる．

3　流線位相構造の抽出ライブラリ `psiclone`

`psiclone` は流体解析を支援する Python ライブラリである [1]．このライブラリは，ハミルトニアンを入力として受け取り，レーブグラフの生成や COT 表現のような流体解析を援用し結果を可視化することを支援する．ここでレーブグラフとは，根を最大値，鞍点を節，渦心点と穴を葉とする流線位相構造を表す二分木である．図 3に，流線位相構造をレーブグラフの構造によって $\mathrm{COT}b_{++}(b_{++}, b_{++}(\beta_+, b_{+-}))$ と決定する例を示す．

*1 本稿では符号は複号同順である．

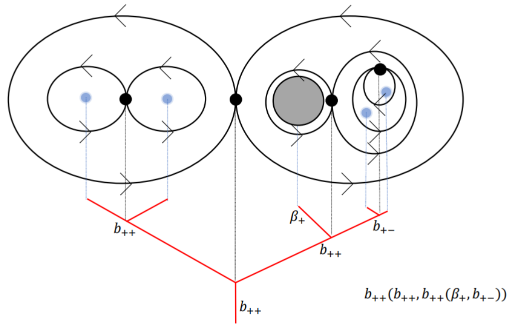

図 3　レーブグラフの構築手法と COT 表現

psiclone は入力データを高さ情報のみとしているところから，等高線の情報から位相構造の分類ができるだけでなく，画像の色などを高さ情報とした画像解析など，流線位相構造の分類の可能性を大きく広げるものとなる．

4　流線位相構造のデータセットの生成方法

本研究で扱う流線位相構造は，基本構造は一様流 $a_\emptyset$ および渦流 $\beta_{\emptyset+}$，局所部分構造を含む構造は一様流中に 1 つ渦流をもつ $a_\emptyset(a_+)$ および $a_\emptyset(a_-)$，さらにその渦流中に 8 の字の流線位相構造をもつ $a_\emptyset(a_+(b_{++}))$ および $a_\emptyset(a_-(b_{--}))$，ならびに渦流中に 8 の字の流線位相構造をもつ $\beta_{\emptyset+}(b_{++})$ とする．

関数定義において，ある程度の高さ，幅があること，連続かつ微分可能であること，極値が唯一であることが条件である．これらの条件のもと，基本的な流線位相構造をもつような滑らかな関数を次のように定義する：

$$f_a(x,y) = ay \quad \text{平面} \tag{1}$$

$$g_{b,c}(x,y) = (x-b)^2 + (y-c)^2 \quad \text{放物面} \tag{2}$$

$$h_{b,c}(x,y) = \exp(-g_{b,c}(x,y)) \quad \text{ベル曲面} \tag{3}$$

ここで，f は y 軸に沿って一定の勾配をもつ平面であり一様流を表し，g は中心 (b,c) の放物面であり渦流を表し，h は中心 (b,c) から離れると急激に減衰するベル曲線であり局所的な渦流を表す．一様流 f と回転流 g は基本構造であり，局所的な回転流 h は局所構造の構成要素である．一様流 f は極値は持たないが，ある程度の高さや幅を持ち，連続かつ微分可能な関数である．また，関数 g，h は唯一の極値が存在し，極値がある程度の高さをもち，極値周りはある程度の幅があり連続かつ微分可能である．

複数の記号で表される流線位相構造をもつ流線については，これらの関数の線型結合で定義する．例えば，8 の字の流線位相構造は中心をずらした 2 つの g や h を加算することで得られる．

5　データセットの生成方法の評価

これらの関数の線型結合によって，表 1の左の列にあるような穴を含まず 8 の字型の局所構造をたかだか 1 つもつ一様流と 8 の字型の局所構造をたかだか 1 つもつ渦流のすべての流線位相構造を生成することができた．

表 1　pscilone による COT の抽出結果

入力 (関数)	抽出された COT
$f_{0.1}$	$a_\emptyset$
$g_{0,0}$	β_+
$f_{0.1} + h_{0.5,2.4}$	$a_\emptyset(a_+)$
$f_{0.1} - h_{0.5,2.4}$	$a_\emptyset(a_-)$
$f_{0.1} + h_{1.0,2.4} + h_{-1.0,2.4}$	$a_\emptyset(a_+(b_{++}))$
$f_{0.1} - h_{1.0,2.4} - h_{-1.0,2.4}$	$a_\emptyset(a_-(b_{--}))$
$g_{2.0,0} + g_{-2.0,0}$	$b_{+-}(\beta_+, \sigma_-)$

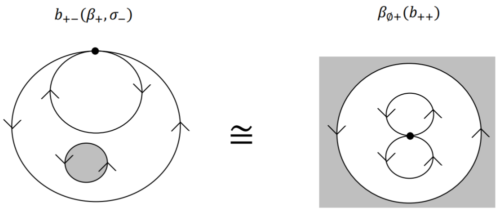

図 4　B 系列の局所部分構造を含む構造に穴 (物体) の構造がある場合の同相関係

これらの流線位相構造を psiclone によって抽出できるか実験を行った．これらの入力に対して右の列にあるような流線位相構造が抽出された．これらの流線位相構造は正しい流線構造である．3 つの基本構造に対応するいずれの記号も含んでいない β_+ と $b_{+-}(\beta_+, \sigma_-)$ についても，正しい流線位相構造である．なぜなら，図 4のように $b_{+-}(\beta_+, \sigma_-)$ は $\beta_\emptyset(b_{++})$ に連続変形できるため，両者は同相である．また，同様に β_+ は $\beta_\emptyset$ と同相である．

6　おわりに

本研究では，ノイズのない流れのデータセットの作成方法を提案した．この方法で生成されたデータセットを入力として，psiclone を用いて一定のクラスの流線位相構造を自動抽出できることを確認した．

任意の流線位相構造をもつデータセットの生成方法の検討，生成–消滅の寿命の閾値に応じたノイズありの流れ図に対する psiclone の精度評価が今後の課題である．

参考文献

[1] 宇田智紀，横山知郎，坂上貴之：パーシステントホモロジーとレーブグラフを用いた 2 次元ハミルトンベクトル場の流線位相構造の自動抽出アルゴリズム，日本応用数理学会論文誌， Vol. 29, No. 2, pp. 187–224 (2019).

[2] 梅田裕平，金児純司，菊池英幸：トポロジカルデータアナリシスと時系列データ解析への応用，*FUJITSU*, Vol. 69, No. 4, pp. 97–103 (2018).

[3] 坂上貴之，横山知郎，澤村陽一：二次元多重連結領域内における構造安定な非圧縮流れの文字列表現アルゴリズム，数理解析研究所講究録， Vol. 1900, pp. 11–25 (2014).