Coherent sets in nonautonomous dynamics

Kathrin Padberg-Gehle

Technische Universität Dresden
Institut für Wissenschaftliches Rechnen

Madrid · July 7, 2016
Motivation

The Perceptual Ocean.
(Movie credit: NASA/Goddard Space Flight Center, Scientific Visualization Studio)
Mathematical challenges

• How to mathematically characterize structures that remain *coherent* for an extended time span (such as ocean gyres and eddies)?

• In other words: How to define what you can physically observe?

• How to systematically, reliably, and efficiently extract these structures?

• How to quantify the mass exchange between these structures and their surroundings?

• How to enhance/mitigate/control coherence?
Concepts for analyzing flow structures

- **Geometric approach**: detect barriers to particle transport
  - invariant manifolds [Rom-Kedar, Wiggins, Mancho, Balasuriya, ...]
  - Lagrangian coherent structures (LCS) [Haller, Shadden, Lekien, Marsden, Beron-Vera, ...]

Approximate set boundaries!

- **Probabilistic approach**: detect minimally dispersive regions
  - almost-invariant sets (spatially fixed) [Dellnitz, Junge, Schütte, Froyland, Koltai, P., ...]
  - finite-time coherent sets (moving) [Froyland, Santitissadeekorn, Monahan, Bollt, Junge, P., ...]

Approximate sets!
In this talk

**Probabilistic approach**

- Transfer operators and numerics
- Almost-invariant sets
- Functional analytic framework
- Finite-time coherent sets
- Clustering framework
- Examples and applications
- Conclusion and future research

Joint work with Gary Froyland, UNSW Australia
Notation

- **Discrete dynamical system**

\[ T : M \to M, \quad M \subset \mathbb{R}^d \text{ compact} \]

- Here: \( T \) diffeomorphism (not formally required!),
  
e.g. an autonomous flow map:

\[ T(\cdot) := x(\tau; \cdot), \]

where \( x(\tau; x_0) \) solves \( \dot{x} = f(x), \ x_0 = x(0) \) for fixed flow time \( \tau \in \mathbb{R} \)

- A set \( A \subset M \) is \( T \)-invariant if

\[ A = T^{-1}(A). \]

- \( \mathcal{M} \) space of finite signed measures on \( M \)

- Probability measure \( \mu \in \mathcal{M} \) is \( T \)-invariant if

\[ \mu(A) = \mu(T^{-1}(A)) \text{ for all } A \subset M. \]
Transfer operators in dynamics

- Define linear operator \( P : \mathcal{M} \to \mathcal{M} \)

\[
(P\nu)(A) = \nu(T^{-1}(A)), \quad A \subset M.
\]

More relevant: \( P : L^1(M, m) \otimes \) with

\[
\int_A Pf \; dm = \int_{T^{-1}(A)} f \; dm, \quad m \text{ Lebesgue measure}
\]

and for diffeomorphisms

\[
Pf(x) = \frac{f(T^{-1}x)}{|\det DT(T^{-1}x)|}.
\]

- \( P \) is the natural push-forward of densities under the action of \( T \)
- Invariant density corresponds to fixed point of \( P \), i.e. eigenfunction to eigenvalue 1
- **Transfer operator** or **Perron-Frobenius operator**
- \( P \) is a Markov operator
Numerical approximation of $P$

- Consider $T : M \to M$
- $\{B_1, \ldots, B_n\}$ partition of $M$
- Galerkin approximation of $P$ with indicator functions on $B_i$, $i = 1, \ldots, n$, as basis functions [Ulam 1960]
- $P$ represented by a sparse, stochastic matrix

$$P_{ij} = \frac{m(B_i \cap T^{-1}(B_j))}{m(B_i)} \approx \frac{\# \{k : T(x_{i,k}) \in B_j \}}{K},$$

with test points $x_{i,k}$, $k = 1, \ldots, K$ uniformly distributed in $B_i$ [Hunt 1993].

- Fixed points of $P$ converge to invariant density of $P$ as $n \to \infty$ for a very restricted class of systems [e.g. Li 1976]
- $P$ is in general assumed to be a good approximation of $P$
Almost-invariant sets

- μ preserved by $T : M \rightarrow M$
- $A \subset M$ is almost-invariant if $T(A) \approx A$, i.e.

$$\frac{\mu(A \cap T^{-1}(A))}{\mu(A)} \approx 1$$

- Application of transfer operator methods for almost-invariant sets:
  - Study eigenfunctions of transfer operator $P$ / eigenvectors of $P$ to real eigenvalues close to 1 [Dellnitz/Junge 1997/99, Deuflhard et al. 1998, Huisingsa, Schmidt 2006],
  - Consider eigenfunctions of the infinitesimal generator of $P$ and its discretization [Froyland/Junge/Koltai 2013]
  - Consider eigenvectors of a transition matrix $R$ of a reversible Markov chain constructed from $P$ [Froyland 2005; Froyland, P. 2009]
  - Consider optimal partitions of a directed graph induced by $P$ or $R$ [Froyland, Dellnitz 2003, Dellnitz et al. 2005]
  - Applications in physical oceanography, dynamical astronomy, fluid mixers, oil spills, epidemic spread,…

No uniform framework for deriving optimality criteria
Proposed construction  [Froyland, P. 2014]

- **Task**: find nontrivial sets $A$, $A^c$
  - that maximize
  \[
  \rho(A) := \frac{\mu(A \cap T^{-1}(A))}{\mu(A)} + \frac{\mu(A^c \cap T^{-1}(A^c))}{\mu(A^c)}
  \]
  - and are robust w.r.t. to perturbations
Proposed construction [Froyland, P. 2014]

- **Task:** find nontrivial sets \( A, A^c \)
  - that maximize
  \[
  \rho(A) := \frac{\mu(A \cap T^{-1}(A))}{\mu(A)} + \frac{\mu(A^c \cap T^{-1}(A^c))}{\mu(A^c)}
  \]
  - and are robust w.r.t. to perturbations

- **Construction:** Use operator \( L \) dynamically similar to \( P \), with \( L1 = 1 \)

- **Functional representation** of invariance condition \( A \approx T(A) \):
  \[
  L1_A \approx 1_A \quad \text{(solution to eigenequation)}
  \]

- When
  \[
  L := D_\epsilon P D_\epsilon / (D_\epsilon P D_\epsilon 1)
  \]
  where \( D_\epsilon \) is a diffusion operator then – under mild assumptions – \( L \) is compact on \( L^2(M, \mu) \) [Froyland 2013].
Optimization problem

**Goal**

Measurably partition

\[ M = A \cup A^c \]

s.t. \( L1_A \approx 1_A, \ L1_{A^c} \approx 1_{A^c} \) and \( \mu(A) \approx \mu(A^c) \)

- **Invariance ratio:**

\[
\rho(A) = \frac{\langle L1_A, 1_A \rangle}{\mu(A)} + \frac{\langle L1_{A^c}, 1_{A^c} \rangle}{\mu(A^c)} = \frac{\langle Q1_A, 1_A \rangle}{\mu(A)} + \frac{\langle Q1_{A^c}, 1_{A^c} \rangle}{\mu(A^c)}
\]

with compact, self-adjoint operator \( Q := (L + L^*)/2 \), where \( L^* \) dual

- **Q** describes mass transport in forward and backward time
Optimization problem

**Goal**
Measurably partition

\[ M = A \cup A^c \]

s.t. \( L^1_A \approx 1_A \), \( L^1_{A^c} \approx 1_{A^c} \) and \( \mu(A) \approx \mu(A^c) \)

- **Invariance ratio:**

\[
\rho(A) = \frac{\langle L^1_A, 1_A \rangle_\mu}{\mu(A)} + \frac{\langle L^1_{A^c}, 1_{A^c} \rangle_\mu}{\mu(A^c)} = \frac{\langle Q^1_A, 1_A \rangle_\mu}{\mu(A)} + \frac{\langle Q^1_{A^c}, 1_{A^c} \rangle_\mu}{\mu(A^c)}
\]

with compact, self-adjoint operator \( Q := (L + L^*)/2 \), where \( L^* \) dual

- **Q** describes mass transport in forward and backward time

- **Relaxed problem** of constrained maximization of \( \rho \) can be shown to be

\[
\max_{f \in L^2(M, \mu)} \left\{ \frac{\langle Qf, f \rangle_\mu}{\langle f, f \rangle_\mu} : \langle f, 1 \rangle_\mu = 0 \right\} \quad (*)
\]
Results

- $Q$ is self-adjoint and compact, $Q1 = 1$; i.e. $u_1 = 1$ is eigenfunction to eigenvalue $\lambda_1 = 1$
- $\lambda_1$ is simple [Froyland 2013]

Theorem

- Maximum in $(\ast)$ is $\lambda_2$ and maximizing $f = u_2$ [follows from min-max theorem],
- $2 - 2\sqrt{2(1 - \lambda_2)} \leq \sup_{A \subset X} \rho(A) \leq 1 + \lambda_2$

[Froyland, P. 2014]
Results

- Q is self-adjoint and compact, Q1 = 1; i.e. u1 = 1 is eigenfunction to eigenvalue λ1 = 1
- λ1 is simple [Froyland 2013]

Theorem

- Maximum in (⋆) is λ2 and maximizing f = u2 [follows from min-max theorem],
- 2 − 2\sqrt{2(1 − λ2)} ≤ \sup_{A \subset X} \rho(A) ≤ 1 + λ2

[Froyland, P. 2014]

- Note that f = f^+ − f^- is signed - with f^+ \approx 1_A and f^- \approx 1_{A^c}
- This suggest an extraction scheme.
- A priori bounds - verification of matrix based bounds
  [Froyland 2005, Froyland, P. 2009]
- A posteriori bounds also directly apply in this setting [Huisinga, Schmidt 2006]
- Influence of ϵ (e.g. spectral gaps, regularity of eigenvectors)
  [Froyland 2013, Froyland, P. 2014]
Numerical approximation of $L$

- $\{B_1, \ldots, B_n\}$ partition of $M$
- $P$ transition matrix obtained via Ulam’s method
- By $p = pP$ we obtain $\mu(B_i) \approx p_i$.
- Approximation to $L$, $L^*$:
  \[
  L_{ij} = \frac{p_i P_{ij}}{p_j}, \quad L_{ij}^* = P_{ji}
  \]

- Compute second left eigenvector of sparse $Q = (L + L^*)/2$
  (e.g. by iterative schemes)
- Carry out line search to find optimal sets [Froyland, P. 2009]
- Diffusion comes for free from numerical scheme but explicit incorporation is possible
- Set-oriented numerical approach implemented in software package GAIO
  [Dellnitz & Junge 2001]
Example: Lorenz system

Second eigenvector of $Q$ and almost-invariant sets in the Lorenz system for flow time $\tau = 0.4$ [Froyland, P. 2009]
Ridges in FTLE field bound almost-invariant sets. [Froyland, P. 2009]
Major gyres in Southern Ocean extracted as almost-invariant sets ($\tau = 2$ months)

Let’s move: coherent sets

- **Goal:** find optimal slow mixing time-dependent structures
- Flow map $T : X \rightarrow Y$ of a nonautonomous system $\dot{x} = f(x, t)$ on $[t, t + \tau]$, $X, Y \subset M$ compact
- Probability measure $\mu$ at $t$ (not invariant)
- **Finite-time coherent pairs:** $A_t, A_{t+\tau}$ satisfying $T(A_t) \approx A_{t+\tau}$, i.e. maximizing
  \[
  \rho(A_t, A_{t+\tau}) = \frac{\mu(A_t \cap T^{-1}(A_{t+\tau}))}{\mu(A_t)} + \frac{\mu(A_t^c \cap T^{-1}(A_{t+\tau}^c))}{\mu(A_t^c)}
  \]
  (plus robustness w.r.t. perturbations and mass constraints)
- Results for matrix-based setting [Froyland, Santitissadeekorn, Monahan 2010]
- Our optimization framework with compact self-adjoint operator
  \[
  Q = L^*L
  \]
  applies [Froyland 2013, Froyland, P. 2014]
Consider $T : X \to Y$ (i.e. $Y := T(X)$)

- $\{B_1, \ldots, B_m\}$ partition of $X$, $\{C_1, \ldots, C_n\}$ partition of $Y$.
- $P$ represented by

$$ P_{ij} = \frac{m(B_i \cap T^{-1}(C_j))}{m(B_i)} \approx \frac{\# \{k : T(x_{i,k}) \in C_j \}}{K}, $$

with test points $x_{i,k}$, $k = 1, \ldots, K$ uniformly distributed in $B_i$.

- Given a probability measure $\mu$ (not invariant!), set $p_i = \mu(B_i)$ and $q = pP$.

- Approximation to $L$, $L^*$:

$$ L_{ij} = \frac{p_i P_{ij}}{q_j}, \quad L^*_{ij} = P_{ji} $$

- Compute second left and right singular vectors of $L$ and do line search to find optimal sets [Froyland et al 2010]
Transitory double gyre \cite{Mosovsky & Meiss 2011}

\[ \dot{x} = -\frac{\partial}{\partial y} \Psi, \quad \dot{y} = \frac{\partial}{\partial x} \Psi \]

with stream function \( \Psi(x, y, t) = (1 - s(t))\Psi_P + s(t)\Psi_F \) where

\[ \Psi_P(x, y) = \sin(2\pi x) \sin(\pi y), \quad \Psi_F(x, y) = \sin(\pi x) \sin(2\pi y) \]

and transition function

\[ s(t) = \begin{cases} 
0, & t < 0, \\
t^2(3 - 2t), & 0 \leq t \leq 1, \\
1, & t > 1. 
\end{cases} \]
Coherent sets

Left and right singular vectors in the transitory double gyre flow w.r.t. time span $[0, 1]$ and extracted finite-time coherent sets at $t = 0$ and $t + \tau = 1$. 
Evolution of coherent sets

Particles evolved by the flow - red particles remain in coherent set.
Back to the ocean

Tracking of an Agulhas ring over two years. [Froyland, Horenkamp, Rossi, Sen Gupta 2015]

Clustering framework

- Transfer operator-based framework is very powerful but involves considerable computational effort

**Wish-list for new method:**
- Computationally more efficient
- Work with relatively small number of trajectories
- Respect entire trajectory not just end-points
- Deal with sparse and incomplete trajectory information
- Provide “quick and dirty” coherent sets diagnostics

**Our simple solution:** use geometric clustering algorithms on trajectory data

[Froyland, P. 2015]
Trajectory-based coherent sets

- $n$ trajectories given at discrete time instants:

$$x_{i,t} \in \mathbb{R}^d, \ i = 1, \ldots, n, \ t = 0, \ldots, T$$

- Extract bundles of trajectories that make up coherent sets, i.e. that minimally spread out over time

Something like that!
General clustering framework

- **Discrete dynamic metric:**

\[
D(x_{i,0}, x_{j,0}) = \sum_{t=0}^{T} \rho(x_{i,t}, x_{j,t})^2
\]

for \( x_{i,0}, x_{j,0} \in \mathbb{R}^d \), \( 1 \leq i, j \leq n \), based on some metric \( \rho \) on \( \mathbb{R}^d \) (Euclidean metric in the following)

- Gives \( n(n - 1)/2 \) interpoint distances \( D(x_{i,0}, x_{j,0}) \).
General clustering framework

- **Discrete dynamic metric:**

\[ D(x_{i,0}, x_{j,0}) = \sum_{t=0}^{T} \rho(x_{i,t}, x_{j,t})^2 \]

for \( x_{i,0}, x_{j,0} \in \mathbb{R}^d \), \( 1 \leq i, j \leq n \), based on some metric \( \rho \) on \( \mathbb{R}^d \) (Euclidean metric in the following)

- Gives \( n(n-1)/2 \) interpoint distances \( D(x_{i,0}, x_{j,0}) \).
- \( x_{i,0} \) and \( x_{j,0} \) are close if they stay close under time-evolution
General clustering framework

- **Discrete dynamic metric:**

  \[ D(x_i,0, x_j,0) = \sum_{t=0}^{T} \rho(x_i,t, x_j,t)^2 \]

  for \( x_i,0, x_j,0 \in \mathbb{R}^d, 1 \leq i,j \leq n \), based on some metric \( \rho \) on \( \mathbb{R}^d \) (Euclidean metric in the following).

- Gives \( n(n-1)/2 \) interpoint distances \( D(x_i,0, x_j,0) \).

- \( x_i,0 \) and \( x_j,0 \) are close if they stay close under time-evolution.

- **Cluster** the initial points \( x_i,0 \in \mathbb{R}^d \) according to \( D \).
General clustering framework

- **Discrete dynamic metric:**

\[
D(x_{i,0}, x_{j,0}) = \sum_{t=0}^{T} \rho(x_{i,t}, x_{j,t})^2
\]

for \( x_{i,0}, x_{j,0} \in \mathbb{R}^d \), \( 1 \leq i, j \leq n \), based on some metric \( \rho \) on \( \mathbb{R}^d \) (Euclidean metric in the following)

- Gives \( n(n-1)/2 \) interpoint distances \( D(x_{i,0}, x_{j,0}) \).
- \( x_{i,0} \) and \( x_{j,0} \) are close if they stay close under time-evolution
- **Cluster** the initial points \( x_{i,0} \in \mathbb{R}^d \) according \( D \)
- Any clustering method on \( \mathbb{R}^d \) could be employed at this point, e.g. k-means, fuzzy c-means, or density-based clustering approaches.
Possible simple clustering strategy

- Interpret \( \{x_i, t\}_{0 \leq t \leq T} \) as a point

\[
X_i = (x_i, 0, x_{i,1}, \ldots, x_{i,T}) \in \mathbb{R}^{d(T+1)}
\]

- Apply fuzzy \textit{c-means} [Bezdek 1981–] on the \( n \) data points in \( \mathbb{R}^{d(T+1)} \)

- For fixed number of clusters \( K \in \mathbb{N} \) fuzzy \textit{c-means} computes
  - a centre \( C_k \in \mathbb{R}^{d(T+1)} \) for each cluster \( k = 1, \ldots, K \) and
  - a likelihood of membership \( u_{k,i} \) of each \( X_i, i = 1, \ldots, n \) to each \( C_k \).
Possible simple clustering strategy

- Interpret \( \{x_{i,t}\}_{0 \leq t \leq T} \) as a point

\[
X_i = (x_{i,0}, x_{i,1}, \ldots, x_{i,T}) \in \mathbb{R}^{d(T+1)}
\]

- Apply fuzzy \( c \)-means [Bezdek 1981–] on the \( n \) data points in \( \mathbb{R}^{d(T+1)} \)

- For fixed number of clusters \( K \in \mathbb{N} \) fuzzy \( c \)-means computes
  - a centre \( C_k \in \mathbb{R}^{d(T+1)} \) for each cluster \( k = 1, \ldots, K \) and
  - a likelihood of membership \( u_{k,i} \) of each \( X_i, i = 1, \ldots, n \) to each \( C_k \).

- **Objective:** minimise

\[
\sum_{k=1}^{K} \sum_{i=1}^{n} u_{k,i}^m \|X_i - C_k\|^2 = \sum_{k=1}^{K} \sum_{i=1}^{n} u_{k,i}^m \sum_{t=0}^{T} \|x_{i,t} - c_{k,t}\|^2 \quad \text{(obj)}
\]

subject to \( \sum_{k=1}^{K} u_{k,i} = 1 \) and \( \sum_{i=1}^{n} u_{k,i} > 0 \)
Possible simple clustering strategy

- Interpret \( \{x_{i,t}\}_{0 \leq t \leq T} \) as a point
  \[
  X_i = (x_{i,0}, x_{i,1}, \ldots, x_{i,T}) \in \mathbb{R}^{d(T+1)}
  \]

- Apply **fuzzy c-means** [Bezdek 1981—] on the \( n \) data points in \( \mathbb{R}^{d(T+1)} \)

- For fixed number of clusters \( K \in \mathbb{N} \) fuzzy c-means computes
  - a centre \( C_k \in \mathbb{R}^{d(T+1)} \) for each cluster \( k = 1, \ldots, K \) and
  - a likelihood of membership \( u_{k,i} \) of each \( X_i, i = 1, \ldots, n \) to each \( C_k \).

- **Objective:** minimise
  \[
  \sum_{k=1}^{K} \sum_{i=1}^{n} u_{k,i}^m \|X_i - C_k\|^2 = \sum_{k=1}^{K} \sum_{i=1}^{n} u_{k,i}^m \sum_{t=0}^{T} \|x_{i,t} - c_{k,t}\|^2 \quad \text{(obj)}
  \]

  subject to \( \sum_{k=1}^{K} u_{k,i} = 1 \) and \( \sum_{i=1}^{n} u_{k,i} > 0 \)

- Fuzziness exponent \( m > 1 \) (increasing \( m \) means softer clusters)

- **Iterative scheme**, implemented e.g. as `fcm` in MATLAB.
Approximation of coherent sets for transitory double gyre flow at $t_0 = 0$, computed on $[0, 1]$. $2^{14}$ initial conditions, trajectory output in 0.1 time instants. (a) Membership functions for 2-clustering using entire trajectories and for (b) 2-clustering using trajectory endpoints. (c) Second singular vector of transition matrix.
Properties and extensions

- Clustering into spheres is preferred by Euclidean norm but other distance functions could be used.
- Interpretations for continuous time and continuous space available.
- Isotropic scaling of space and time has no effect.
- **Clustering results are frame-independent.**
- Weights can be included as coefficients for $\|x_{i,t} - c_{k,t}\|^2$, which could depend on $i$, $t$, or $k$ (e.g. discount distances far in the future).
- **Missing data can be easily handled** (restrict computations of $C_k$ to available data per time-slice).
- Treatment of almost-invariant sets also possible.
- Normalized entropy can serve as a measure for classification uncertainty.
Application I: global ocean drifters

Application II: turbulent convection

Experimental set-up in Barrel of Ilmenau, Germany.

Cell dimensions: 2.5 m (H) x 2.5 m (W) x 0.65 m (D); temperature difference: 10K $T_{bot} = 35^\circ C$

$T_{top} = 25^\circ C$; Rayleigh number $Ra \approx 1.5e10$
Coherent structures

Coherent plumes visible in experimental data (du Puits et al, PRL 2014)
Trajectory-based coherent sets

Coherent set in convection flow (extracted from PIV velocity data) – using only the trajectories shown.

Joint work with Ronald du Puits, Ilmenau
Conclusion and future research

- Computational study of coherent flow structures
- Optimization problems within transfer operator and clustering framework
- Both coherent sets and boundaries can be extracted
- Many applications, mathematical and computational challenges

Automatic extraction of many coherent sets

[Karrasch, Huhn, Haller 2014; Ma, Bollt 2013]

Systematic comparison with other approaches

Active transport, inertial particles

Early warning signals for bifurcations

Thank you!

kathrin.padberg@tu-dresden.de
Conclusion and future research

- Computational study of coherent flow structures
- Optimization problems within transfer operator and clustering framework
- Both coherent sets and boundaries can be extracted
- Many applications, mathematical and computational challenges

- Automatic extraction of many coherent sets
  [Karrasch, Huhn, Haller 2014; Ma, Bollt 2013]
- Systematic comparison with other approaches
- Active transport, inertial particles
- Early warning signals for bifurcations
Conclusion and future research

- Computational study of coherent flow structures
- Optimization problems within transfer operator and clustering framework
- Both coherent sets and boundaries can be extracted
- Many applications, mathematical and computational challenges

- Automatic extraction of many coherent sets
  [Karrasch, Huhn, Haller 2014; Ma, Bollt 2013]
- Systematic comparison with other approaches
- Active transport, inertial particles
- Early warning signals for bifurcations

Thank you!

kathrin.padberg@tu-dresden.de