Eurographics Conference on Visualization (EuroVis) 2016
E. Bertini, N. Elmqvist, and T. Wischgoll
(Guest Editors)



# Error Estimates for Lagrangian Flow Field Representations

M. Hummel[1] and R. Bujack[1] and K. I. Joy[2] and C. Garth[1]

[1]University of Kaiserslautern, Germany
[2]University of California, Davis, USA

**Abstract**

*Computing power outpaces I/O bandwidth in modern high performance computers, which leads to temporal sparsity in flow simulation data. Experiments show that Lagrangian flow representations (where pathlines are retrieved from short-time flow maps using interpolation and concatenation) outperform their Eulerian counterparts in advection tasks under these circumstances.*

*Inspired by these results, we present the theoretical estimate of the Lagrangian error for individual pathlines, depending on the choice of temporal as well as spatial resolution. In-situ, this measure can be used to steer the output resolution and post-hoc, it can be used to visualize the uncertainty of the pathlines. To validate our theoretical bounds, we evaluate the measured and the estimated error for several example flow fields.*

Categories and Subject Descriptors (according to ACM CCS): Visualization [Human-centered computing]: Visualization application domains—Scientific visualization

## 1. Introduction

In hydrodynamics, two specifications of a flow field are typically used. The Eulerian specification describes the flow passing through a given spatial domain, which is usually stored by means of the velocity field. The Lagrangian specification describes a specific fluid parcel that travels through space. The usual way of storing a flow field in its Lagrangian specification is by means of its flow map that describes how a flow parcel at $(x_0, t_0) \in R^d \times R$ moves to $F^t_{t_0}(x_0) \in R^d$ in the time interval $[t_0, t]$.

In theory, both representations are equivalent because the trajectories can be derived from the vectors through integration and the vectors from the trajectories through differentiation. But in practice, they differ because the data is given at discrete positions and times only and the integration or differentiation have to be approximated numerically.

Flow data is often the result of a hydrodynamics simulation run on a high performance supercomputer with many small time steps. In modern HPC architecture, the computational capacity exceeds the I/O bandwidth [CPA*10], which is why only few time steps of the simulation are stored to disk for post-hoc analysis. Under this temporal sparsity, the Lagrangian representation has proven less prone to error for the post-hoc construction of pathlines [ACG*14].

In this paper, we present a method to calculate an upper bound for the error of individual pathlines that are constructed by sequential interpolation of short-time flow maps.

The paper is organized as follows: In Section 2, we will provide a short overview of related work. We present the theoretical definition of our error estimate in Section 3 and describe our experiments in Section 4, before we conclude our findings in Section 5.

## 2. Related Work

Lagrangian methods have been long used for efficient particle tracer computation [BKHJ01], the calculation of Lagrangian Coherent Structures, [HY00, Hal02], or to incorporate flow maps into the Eulerian representation of a flow field [SS06, SGSM08, JEH02].

They are particularly useful because the majority of current state-of-the-art flow visualization techniques utilize advection [LHD*04, LHZP07, SJWS08, MLP*09, PPF*11, BCP*12] and the particle paths are explicitly given in the Lagrangian flow representations.

Hlawatsch et al. employ a hierarchical scheme to construct longer integral lines from previously computed partial solutions [HSW11]. Chandler et al. [COJ15] construct visualizations of smoothed particle hydrodynamics (SPH) applications [Mon05]. Agranovsky et al. show empirically that the Lagrangian method is more accurate than Eulerian advection methods [ACG*14] and Bujack and Joy [BJ15] extend their idea. They give the first theoretical estimate of the overall error of these kinds of Lagrangian flow field representations.

In this paper, we refine their findings to derive a local error for each pathline, which can be used for the visualization of the uncertainty of the method as well as for the in-situ adjustment of the resolution chosen for storage.



DOI: 10.2312/eurovisshort.20161153

## 3. Theoretical Error Estimate

In this section, we will derive a regional error estimate for the Lagrangian flow representation using the method of concatenating interpolations of sectional flow maps. Bujack and Joy [BJ15] have shown that this method is a special kind of numerical one-step integration method and have given an overall error estimate. The overall measure is not sufficient for our means, because we want to show the error of the method for each reconstructed pathline individually.

In our application, we use multi-linear interpolation [HH12] in the spatial domain $\mathbb{R}^d$ to estimate the endpoints of the arbitrarily seeded pathlines within each of the stored time slices. In order to keep the notation short, we will perform the calculations assuming the one-dimensional case. The error will be a two- or three-dimensional vector for 2D or 3D flows and each component will behave like the one-dimensional case.

Trajectories are constructed from the Lagrangian flow representation by interpolating sequential flow maps. The sequence $x(t_0),\dots,x(t_n)$ of exact pathline points is approximated by the sequence $x_0,\dots,x_n$ with $x_0 = x(t_0)$. Each point $x_j$ is obtained by interpolating the flow map $F^{t_j}_{t_{j-1}}$ at the previous point $x_{j-1}$.

We will first show how this construction method can be viewed as a on-step integration method before we describe the error measure.

### 3.1. One-step Integration Method

In numerics of ordinary differential equations [Sch02, GH10], for a given initial value problem

$$x'(t) = f(t,x(t)), \quad x_0 = x(t_0), \tag{1}$$

a one-step integration method produces a sequence of points $x_0,\dots,x_n$ by the rule

$$x_j = S(x_{j-1}) = x_{j-1} + hA(t_{j-1},x_{j-1},f,h), \tag{2}$$

where $S$ is the rule to advance from one step to the next and $A$ is the increment function. It can be seen that patching together the interpolated trajectories is a special kind of numerical one-step integration method with increment function

$$A(t_{j-1},x_{j-1},f,h) = L_x \dot{F}^t_{t_{j-1}}(x_{j-1}) \tag{3}$$

from using the Taylor expansion [GL10] with respect to time

$$\begin{aligned} L_x F^{t_j}_{t_{j-1}}(x) &\overset{\text{Taylor}}{=} L_x F^{t_{j-1}}_{t_{j-1}}(x) + h_t L_x \dot{F}^t_{t_{j-1}}(x) \\ &= x + h_t L_x \dot{F}^t_{t_{j-1}}(x), \end{aligned} \tag{4}$$

where the dot represents the temporal derivative with respect to the end time and $t \in [t_{j-1},t_j]$.

### 3.2. Local Truncation Error

The local truncation error $\tau_j$ for a step $j$ and an exact position $x(t_{j-1})$ is defined by

$$\tau_j(x(t_{j-1})) := |x(t_j) - S(x(t_{j-1}))| \tag{5}$$

and describes the error that the one-step method makes advancing one step. For our method, the local truncation error is the error that we make within each of the stored time slices when we interpolate the endpoints of the spanning trajectories in space to generate trajectories starting at arbitrary spatial locations. It coincides with the linear interpolation error [HH12] of a point in the interval $x \in [x_0,x_1]$, which follows from the Taylor series

$$\begin{aligned} \tau_j(x) &= |L_x F^{t_j}_{t_{j-1}}(x) - F^{t_j}_{t_{j-1}}(x)| \\ &\overset{\text{Taylor}}{\leq} \frac{1}{2}(x_1 - x)(x - x_0) \max_{\zeta \in [x_0,x_1]} |(F^{t_j}_{t_{j-1}})''(\zeta)|. \end{aligned} \tag{6}$$

### 3.3. Global Truncation Error

The global truncation error $e_j$ up to the step $j$ is defined by

$$e_j(x_0) := |x(t_j) - x_j|. \tag{7}$$

The notation $e_n(x_0)$ was chosen to stress that it is valid for the sequence $x_0,\dots,x_n$ that starts at $x_0$. It suffices

$$\begin{aligned} e_j(x_0) &\overset{(7)}{=} |x(t_j) - x_j| \\ &\leq |x(t_j) - S(x(t_{j-1}))| + |S(x(t_{j-1})) - x_j| \\ &\overset{(5)}{=} \tau_j(x(t_{j-1})) + |S(x(t_{j-1})) - x_j| \\ &\overset{(2)}{=} \tau_j(x(t_{j-1})) + |S(x(t_{j-1})) - S(x_{j-1})| \\ &\overset{(2)}{=} \tau_j(x(t_{j-1})) + |x(t_{j-1}) + hA(t_{j-1},x(t_{j-1}),f,h) \\ &\quad - x_{j-1} - hA(t_{j-1},x_{j-1},f,h)| \\ &\overset{(7)}{=} \tau_j(x(t_{j-1})) + e_{j-1} \\ &\quad + h|A(t_{j-1},x(t_{j-1}),f,h) - A(t_{j-1},x_{j-1},f,h)|. \end{aligned} \tag{8}$$

At this point in numerics of ordinary differential equations, the overall maximal error is derived by using the globally maximal $\tau_j$ and the global Lipschitz constant $L$. But for our application, we do not need this very coarse approximation, but one tailored towards the specific integration sequence $x_0,\dots,x_n$. That means we look for an estimate that holds for one pathline. We get it from not looking at the maximal possible errors of the whole domain, but at the maximal possible errors for the area around our point in the sequence $x_{j-1}$ that is big enough to contain the true value $x(t_{j-1})$. Therefore, we make use of the regional Lipschitz constant $L(x_{j-1})$ sufficing

$$\begin{aligned} &\forall x,y \in [x_{j-1} - e_{j-1}, x_{j-1} + e_{j-1}] : \\ &|A(t_{j-1},x,f,h) - A(t_{j-1},y,f,h)| \leq L(x_{j-1})|x - y|, \end{aligned} \tag{9}$$

as well as the regional value $\bar{\tau}_j(x_{j-1})$ that bounds $\tau_j(x(t_{j-1})$ from above

$$\bar{\tau}_j(x_{j-1}) := \max_{x \in [x_{j-1} - e_{j-1}, x_{j-1} + e_{j-1}]} \tau_j(x) \tag{10}$$

without depending on the true value $x(t_{j-1})$. Note that this value depends on the global truncation error $e_{j-1}(x_0)$ calculated thus far. In contrast to (5), it can therefore not be calculated from only looking at one time step. That way, we get the global truncation error

for the pathline starting at $x_0$

$$\begin{aligned} e_j(x_0) &\overset{(8)}{\leq} \tau_j(x(t_{j-1})) + e_{j-1} \\ &\quad + h|A(t_{j-1}, x(t_{j-1}), f, h) - A(t_{j-1}, x_{j-1}, f, h)| \\ &\overset{(9),(10)}{\leq} \bar{\tau}_j(x_{j-1}) + e_{j-1} + hL(x_{j-1})|x(t_{j-1}) - x_{j-1}| \\ &\overset{(7)}{=} \bar{\tau}_j(x_{j-1}) + e_{j-1} + hL(x_{j-1})e_{j-1} \\ &= \bar{\tau}_j(x_{j-1}) + e_{j-1}(1 + hL(x_{j-1})). \end{aligned} \tag{11}$$

This is the main theoretical result of this paper.

### 3.4. Error Within the Sections

So far, we have calculated the error of the sequence $x_0, \ldots, x_n$. To get the error of the whole pathline, we have to also evaluate it at arbitrary time between the stored time steps $t_j$. To approximate these values, we interpolate linearly with respect to time $L_t$ between the one-step points $x_{j-1}$ and $x_j$ . Due to the uncertainty of the one-step points, the additional approximation error for arbitrary times $t \in [t_{j-1}, t_j]$ is then given by

$$\begin{aligned} s_j(x_0, t) &:= |L_t F^t_{t_{j-1}}(x_{j-1}) - F^t_{t_{j-1}}(x_{j-1})| \\ &\leq \max_{x \in [x_{j-1} - e_{j-1}, x_{j-1} + e_{j-1}]} |L_t F^t_{t_{j-1}}(x) - F^t_{t_{j-1}}(x)| \\ &\overset{\text{Taylor}}{\leq} \max_{x \in [x_{j-1} - e_{j-1}, x_{j-1} + e_{j-1}]} \\ &\quad \frac{1}{2}(t - t_{j-1})(t_j - t) \max_{\theta \in [t_{j-1}, t_j]} |\dot{F}^{\theta}_{t_{j-1}}(x)|. \end{aligned} \tag{12}$$

It is added to the error (11) to approximate the absolute error of the Lagrangian representation

$$a_j(x_0, t) := e_j(x_0) + s_j(x_0, t). \tag{13}$$

## 4. Experimental Results

Multiple data sets were chosen to evaluate the theoretical error measures presented in this paper. We used a Runge-Kutta integration scheme to compute sets of consecutive flow maps in the high in-situ resolution for each scenario and stored their endpoints on a regular grid, as explained in [ACG*14]. Then, a number of pathlines were randomly seeded in the domain of the data sets. One set of pathlines was generated through consecutive interpolation of the stored flow maps for each interval. In addition, a set of ground truth pathlines was obtained by applying the Runge-Kutta integration method to the original flow field in its high resolution. From these two sets of pathlines, the actual reconstruction errors were calculated for each seed point as the Euclidean distance between the integrated and the interpolated position. Applying our error estimate to the reconstructed pathlines from integration allowed us to directly compare the actual errors to the theoretically predicted errors.

**Double Gyre** One of the data sets used for the evaluation of the theoretical error measure in this paper is the well known double gyre. It is a time-varying, two-dimensional vector field with a closed boundary. We use the analytical definition from [SLM05]. Example pathlines can be found in Figure 1. The global truncation error (11) plotted to densely spread seedpoints all over the domain for the 8-th time interval of 16 can be seen in Figure 4.

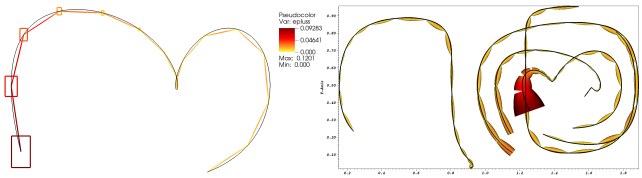

Figure 1: Double gyre. Left: A reconstructed pathline is shown with the estimated global truncation error $e$ from (11) represented by rectangles. Right: Five reconstructed pathlines are represented as tubes with a radius equal to the absolute estimated error $a$ from (13), which also determines the color. On both sides, the corresponding correct pathlines are plotted in black.

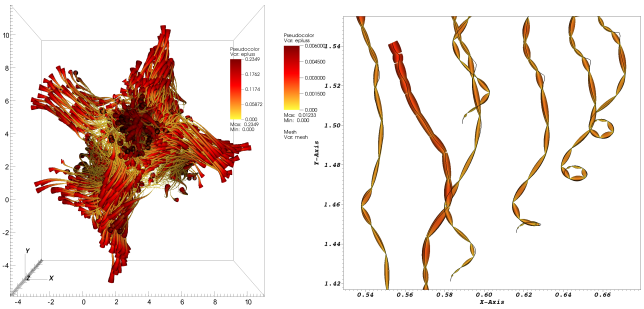

Figure 2: As in Figure 1, the estimated absolute error (13) is represented by the thickness and color of the tubes. Left: 1000 pathlines in the ABC data set reconstructed from 16 flow map intervals. Right: Several pathlines in the convection flow field with 32 intervals.

**ABC** The Arnold Beltrami Childress (ABC) vector field, as defined in [Hal01], is a time-varying, three-dimensional vector field that features periodic boundary conditions. Figure 2 shows a set of 1000 pathlines that were obtained from 16 consecutive flow map intervals.

**Convection** The convection dataset stems from a simulation of a two-dimensional flow around a small, heated cylinder. As heat is transferred from the cylinder, the surrounding air begins to rise. Figure 1 shows a selection of pathlines. The distribution of the global truncation error over the domain for the 12-th time interval of 16 can be found in Figure 5.

### 4.1. Evaluation

The behavior of the average and the maximal estimated as well as measured global truncation errors (11) for the different data sets can be found in Figure 3. It shows the errors for 100,000 randomly seeded pathlines as they progress through 16 time intervals. As theoretically predicted, the error increases exponentially over time, which is why we depict the results using a logarithmic scale. Our theoretical error estimate is constructed to bound the actual error from above, which can be well perceived in the charts. Since in applications, errors from consecutive time steps are likely to cancel

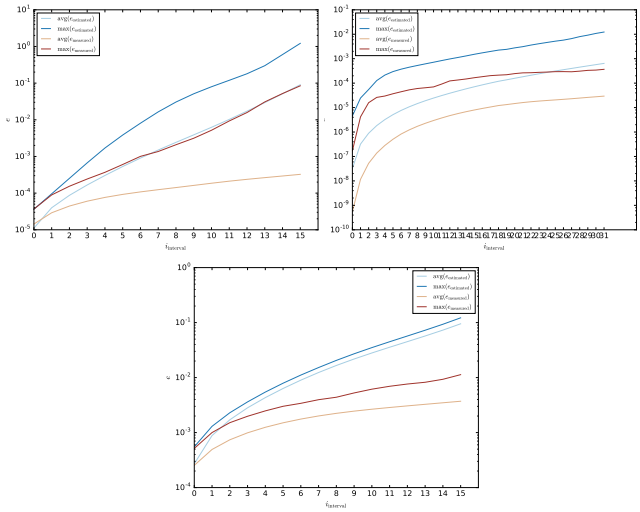

Figure 3: Averages and maxima of estimated and measured error over time: double gyre (top left), convection (top right), ABC (bottom).

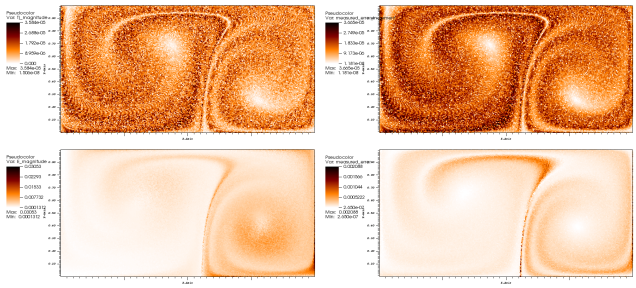

Figure 4: Double gyre: Estimation and measurement of the local truncation error $\tau$ (top row) and of the global truncation error $e$ (bottom row).

each other out, the real error is usually smaller than the estimated one. This difference increases with time.

The relationship between measured and estimated errors becomes clearer in Figure 4. Estimated errors (left column) and measured errors (right column) clearly exhibit the same features, which shows that the estimate does not only provide an upper bound, but also correlates with the measured error.

This can also be seen in the two supplemental videos, which show how the error estimates $\tau$ and $e$ and the actual errors evolve over time.

### 4.2. Visualization

In a $d$-dimensional flow field, the truncation errors are $d$-dimensional vectors. The global truncation error (11) at the endpoints of the sections per component mapped to a rectangle can be found in Figure 1. The real positions of the pathline can not be located outside the rectangles. It shows how the global truncation error increases exponentially over time.

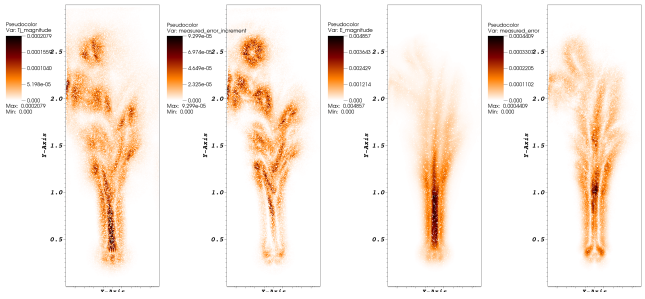

Figure 5: Convection: Estimation and measurement of the local truncation error $\tau$ (left pair) and of the global truncation error $e$ (right pair).

In order to also visualize the error between the stored time steps, we decided to encode the Euclidean norm of the absolute error vector (13) in the radius and color of a tube that is centered around the reconstructed pathline. This representation of the pathlines as tubes resembles the uncertainty visualization method *UFLOW* as presented by Lodha et al. [LPSW96]. The tubes have a very intuitive interpretation, because the actual pathline has to lie inside of it. For long advection times, the radii of the tubes may become very large, which can cause visual clutter. This is why we offer to only encode the error in the color and draw a line instead of a tube to get a more sparse representation of the reliability of the reconstructed pathlines. Figure 1 shows five pathlines in the double gyre data set that were reconstructed using a set of 16 consecutive flow map intervals. The image shows well how the error does not only increase over time, but also grows between each two consecutive stored time intervals to become smaller as we reach the next stored position.

## 5. Conclusion

In this paper, we have presented a method to estimate the error of individual pathlines that are reconstructed from sets of flow maps using interpolation and concatenation. We have compared the theoretical estimate to the actual errors for pathlines in a number of example flow data sets. The results show that using our method, a reasonably accurate upper bound of the reconstruction error can be computed. Further, our experiments have confirmed that the error estimates correlate with the actual errors and reveal similar features. We therefore believe that the error measure may become useful to control refinement in multi-resolution settings or lossy compression.

In the future, we will examine how this error estimate can be used in-situ to steer the accuracy of simulation output data.

## Acknowledgements

This research was funded in part by the German Research Foundation (DFG) within the IRTG 2057 "Physical Modeling for Virtual Manufacturing Systems and Processes" and under EU Career Integration Grant #304099.

## References


[ACG*14] AGRANOVSKY A., CAMP D., GARTH C., BETHEL E., JOY K., CHILDS H.: Improved post hoc flow analysis via lagrangian representations. In *Large Data Analysis and Visualization (LDAV), 2014 IEEE 4th Symposium on* (Nov 2014), pp. 67–75. 1, 3

[BCP*12] BRAMBILLA A., CARNECKY R., PEIKERT R., VIOLA I., HAUSER H.: Illustrative Flow Visualization: State of the Art, Trends and Challenges. *EG 2012 State of the Art Reports* (2012), 75–94. 1

[BJ15] BUJACK R., JOY K. I.: Lagrangian Representations of Flow Fields with Parameter Curves. In *Large Data Analysis and Visualization (LDAV), 2015 IEEE 4th Symposium on* (2015), IEEE. 1, 2

[BKHJ01] BRUCKSCHEN R., KUESTER F., HAMANN B., JOY K. I.: Real-time out-of-core visualization of particle traces. In *Proceedings of the IEEE 2001 symposium on parallel and large-data visualization and graphics* (2001), IEEE Press, pp. 45–50. 1

[COJ15] CHANDLER J., OBERMAIER H., JOY K.: Interpolation-based pathline tracing in particle-based flow visualization. *Visualization and Computer Graphics, IEEE Transactions on 21*, 1 (Jan 2015), 68–80. 1

[CPA*10] CHILDS H., PUGMIRE D., AHERN S., WHITLOCK B., HOWISON M., PRABHAT, WEBER G., BETHEL E. W.: Extreme Scaling of Production Visualization Software on Diverse Architectures. *IEEE Computer Graphics and Applications (CG&A) 30*, 3 (May/June 2010), 22–31. 1

[GH10] GRIFFITHS D. F., HIGHAM D. J.: *Numerical Methods for Ordinary Differential Equations: Initial Value Problems*. Springer, London, UK, 2010. 2

[GL10] GHORPADE S. R., LIMAYE B. V.: *A Course in Multivariable Calculus and Analysis*. Springer, New York, USA, 2010. 2

[Hal01] HALLER G.: Distinguished material surfaces and coherent structures in three-dimensional fluid flows. *Physica D 149*, 4 (2001), 248–277. 3

[Hal02] HALLER G.: Lagrangian Coherent Structures from Approximate Velocity Data. *Physics of Fluids 14* (2002), 1851–1861. 1

[HH12] HÄMMERLIN G., HOFFMANN K.-H.: *Numerical mathematics*. Springer Science & Business Media, 2012. 2

[HSW11] HLAWATSCH M., SADLO F., WEISKOPF D.: Hierarchical line integration. *Visualization and Computer Graphics, IEEE Transactions on 17*, 8 (Aug 2011), 1148–1163. 1

[HY00] HALLER G., YUAN G.: Lagrangian coherent structures and mixing in two-dimensional turbulence. *Phys. D 147*, 3-4 (Dec. 2000), 352–370. 1

[JEH02] JOBARD B., ERLEBACKER G., HUSSAINI H.: Lagrangian-Eulerian advection of noise and dye textures for unsteady flow visualization. *IEEE Transactions on Visualization and Computer Graphics 8*, 3 (2002), 211–222. 1

[LHD*04] LARAMEE R. S., HAUSER H., DOLEISCH H., VROLIJK B., POST F. H., WEISKOPF D.: The State of the Art in Flow Visualization: Dense and Texture-Based Techniques. *Computer Graphics Forum 23* (2004), 2004. 1

[LHZP07] LARAMEE R. S., HAUSER H., ZHAO L., POST F. H.: Topology-Based Flow Visualization, The State of the Art. In *Topology-based Methods in Visualization* (2007), pp. 1–19. 1

[LPSW96] LODHA S. K., PANG A., SHEEHAN R. E., WITTENBRINK C. M.: Uflow: Visualizing uncertainty in fluid flow. In *Visualization'96. Proceedings.* (1996), IEEE, pp. 249–254. 4

[MLP*09] MCLOUGHLIN T., LARAMEE R. S., PEIKERT R., POST F. H., CHEN M.: Over Two Decades of Integration-Based, Geometric Flow Visualization. In *EG 2009 - State of the Art Reports* (2009), pp. 73–92. 1

[Mon05] MONAGHAN J. J.: Smoothed particle hydrodynamics. *Reports on Progress in Physics 68*, 8 (2005), 1703. 1

[PPF*11] POBITZER A., PEIKERT R., FUCHS R., SCHINDLER B., KUHN A., THEISEL H., MATKOVIC K., HAUSER H.: The State of the Art in Topology-based Visualization of Unsteady Flow. *Computer Graphics Forum 30*, 6 (September 2011), 1789–1811. 1

[Sch02] SCHATZMANN M.: *Numerical Analysis: A Mathematical Introduction*. Oxford University Press, New York, USA, 2002. 2

[SGSM08] SALZBRUNN T., GARTH C., SCHEUERMANN G., MEYER J.: Pathline predicates and unsteady flow structures. *The Visual Computer 24*, 12 (2008), 1039–1051. 1

[SJWS08] SALZBRUNN T., JÄNICKE H., WISCHGOLL T., SCHEUERMANN. G.: The State of the Art in Flow Visualization: Partition-based Techniques. In *Simulation and Visualization 2008 Proceedings* (2008). 1

[SLM05] SHADDEN S. C., LEKIEN F., MARSDEN J. E.: Definition and properties of lagrangian coherent structures from finite-time lyapunov exponents in two-dimensional aperiodicăflows. *Physica D: Nonlinear Phenomena 212*, 3-4 (2005), 271 – 304. 3

[SS06] SALZBRUNN T., SCHEUERMANN G.: Streamline predicates. *IEEE Transactions on Visualization and Computer Graphics 12*, 6 (2006), 1601–1612. 1