

Scientific Visualisation Summary

FS 2013

Pascal Spörri
pascal@spoerri.io

August 4, 2013

This summary is based on the course slides of the Scientific Visualisation course at ETH Zürich¹ taught by Prof. Ronald Peikert from spring semester 2013.

¹http://www.scivis.ethz.ch/education/scivis_course/notes

Contents

1	Introduction	7
1.0.1	SciVis and InfoVis	7
1.1	Visualisation Scenarios	7
1.1.1	Video/Movie	8
1.1.2	Tracking	8
1.1.3	Interactive Post Processing/Visualisation	9
1.1.4	Interactive Steering/Computational Steering	9
1.2	Data Discretisations	10
1.3	Unstructured Grids	10
1.3.1	2D Unstructured Grids	10
1.3.2	3D Unstructured Grids	10
1.4	Structured Grids	11
1.4.1	Point-Sampled Data/Scattered Data	12
1.5	Elementary Visualisation Methods	12
2	Contouring and Isosurfaces	14
2.1	Contours	14
2.1.1	Contours in a quadrangle cell	15
2.1.2	Basic contouring algorithms	16
2.1.3	Marching Squares	16
2.2	The Marching Cubes Algorithm	17
2.3	The Asymptotic Decider Algorithm	20
2.4	The Dividing Cubes Algorithm	20
2.5	Optimised Isosurface Algorithms	20
2.5.1	The Span-Space Algorithm	21
2.6	Selecting Contour Levels	21
2.7	Limitations of Isosurfaces	22
3	Raycasting	23
3.1	Direct Volume Rendering	23
3.2	Raycasting	23
3.2.1	Compositing	24
3.2.2	α -compositing	24
3.3	Transfer Functions	26
3.3.1	Pre- vs. Post-classification	27
3.4	Preintegration	28
3.4.1	Extinction-based volume rendering	29
4	Object Space Volume Rendering	30
4.1	Texture-based volume rendering	30
4.1.1	Volume rendering with 2D texturemapping	30
4.1.2	Volume rendering with 3D texture mapping	30

4.2	Shear-Warp Factorisation	31
4.3	Perspective shear warp	32
4.4	Object Space versus Image Space	34
4.5	Splatting	35
4.6	Cell Projection	36
4.6.1	Drawing	38
4.6.2	Example: Visualisation of smoke propagation	38
4.6.3	Opacities	39
4.6.4	Hardware-Assisted Visibility Sorting	39
5	Vector Field Visualisation	40
5.1	Visualisation	40
5.2	Vector fields as ODEs	41
5.2.1	Pathlines	41
5.2.2	Streamlines	42
5.2.3	Streaklines	44
5.2.4	Timelines	44
5.3	The Stencil Walk algorithm	45
5.3.1	Problems	46
5.4	Global Point Location	47
5.5	Computational Space Streamline Integration	48
5.6	Skin Friction Lines	49
5.7	Streamline Placement	50
5.8	Streamsurfaces	50
5.9	Streak Surfaces	52
6	Texture Advection	53
6.1	Line Integral Convolution	53
6.2	Multi-Layer Flow Textures	54
6.3	Lagrangian-Eulerian Advection	55
6.3.1	Lagrangian-Eulerian Advection for Vector Field Visualisation	55
6.4	Image-Based Flow Visualisation	57
6.5	Texture advection in surfaces	59
6.6	IBFV for Surfaces	59
6.7	Image-Space Advection	59
7	Feature Extraction	61
7.1	Region-Type Features	61
7.2	Point Features in Scalar Fields	62
7.3	Line-Like Features in Scalar Fields	63
7.3.1	Circular Gutter	65
7.3.2	Blended Height Fields	65
7.3.3	Watersheds vs. Height Ridges	66
7.4	Geometric Features of Surfaces	66

7.5	Line-Like Features in Vector Fields	67
7.6	Tracking of Features	68
7.7	Post-Filtering of Features	68
8	Vector Field Topology	69
8.1	Vector Fields as ODEs	69
8.2	Special Streamlines	69
8.3	Critical Points	70
8.3.1	Hyperbolic Critical Points	70
8.4	Critical Points in 2D	70
8.4.1	Node Source	71
8.4.2	Node Sink	72
8.4.3	Saddle	72
8.4.4	Focus Source	73
8.4.5	Focus Sink	73
8.4.6	Node Focus Source	74
8.4.7	Star Source	74
8.5	Nonhyperbolic Critical Points	75
8.5.1	Center	75
8.6	Other Stationary Points	75
8.7	The Topological Skeleton	76
8.8	Critical Points in 3D	76
8.9	Periodic Orbits	77
8.10	Saddle Connectors	78
9	Visualisation of Lagrangian Coherent Structures	80
9.1	The Finite-Time Lyapunov Exponent	80
9.2	Separatrices	81
9.2.1	LCS in 3D	81
9.3	Computation of FTLE Ridges	81
9.4	Lagrangian Vector Field Topology	82
9.4.1	Double Gyre Example	83
9.4.2	Petri Dish Example	83
9.4.3	Vortex Street Example	84
10	Tensor Field Visualisation	86
10.1	Tensors	86
10.2	Tensor Glyphs	86
10.3	Tensor Field Lines	88
10.4	Diffusion Tensor Fiber Bundle Tracking	90
10.4.1	Algorithmic Steps	91
11	Illustrative Visualisation	92
11.1	Smart Visibility	92

11.2	Silhouette Algorithms	92
11.2.1	Hatching	93
11.3	Illuminated Streamlines	93
11.3.1	Maximum Principle	93
11.3.2	Cylinder Averaging	94
11.4	Tone Shading	94
11.5	Context-Preserving Volume Rendering	95
11.6	Illustrative Visualisation of (Flow) Surfaces	96
12	Information Visualisation	97
12.1	Scatter Plots	97
12.2	Dimension Reduction	98
12.3	Parallel Coordinates	98
12.4	Pixel-Oriented Techniques	99
12.4.1	Space-Filling Curves	99
12.4.2	Spiral Technique	99
12.4.3	Axes Technique	99
12.5	Icon-Based Techniques	99
12.5.1	Chernoff Faces	99
12.5.2	Stick Figures	100
12.6	Hierarchical and Network Data	101
12.6.1	Cone Trees	101
12.6.2	Tree Maps	102
12.7	Voronoi Tree Maps	103
12.7.1	Additively Weighted (AW) Voronoi Tessellation	104
12.7.2	Power Weighted (PW) Voronoi Tessellation	104
12.7.3	Voronoi Tree Map	104
12.8	Clustering Techniques	105
12.9	Distortion Techniques	106
12.9.1	Hyperbolic trees	106

1 Introduction

SciVis is interdisciplinary the fields of application include engineering, natural sciences and medical sciences. There's a common application to all fields: There are *numerical datasets* providing an abstraction from the particular application. The characteristics of such datasets include:

Dimension of domain: Number of coordinates or parameters

Dimension of values: Scalar, vector or tensor fields

Type of data: Discrete values versus discretised data

Type of discretisation: (Un-)structured grid, scattered data

Time dependencies: Static versus time-dependent.

1.0.1 SciVis and InfoVis

Scientific Visualisation is mostly concerned with

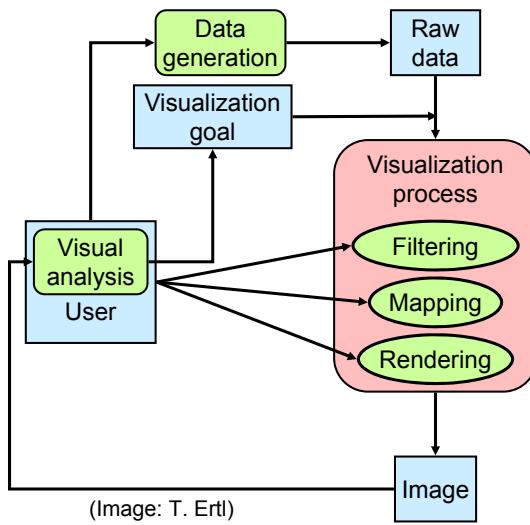
- 2,3,4 dimension spatial or spatio-temporal data
- discretised data

Information Visualisation focuses on:

- High-dimensional, abstract data
- Discrete data
- Financial, statistical, etc.
- Visualisation of large trees, networks, graphs
- Data mining:
 - Finding patterns
 - Clusters
 - Voids
 - Outliers

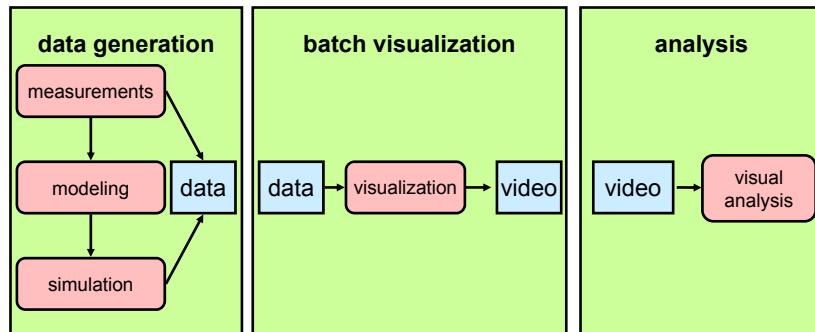
1.1 Visualisation Scenarios

The reference model for visualisation:



1.1.1 Video/Movie

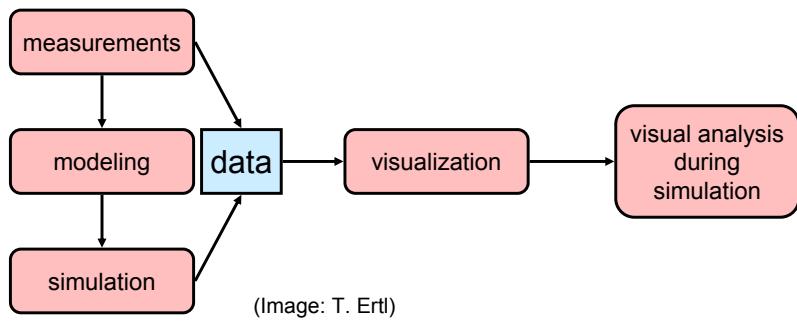
In a first step the data is generated. Then the data is visualised during a batch visualisation step and in the end the video is analized.



(Image: T. Ertl)

1.1.2 Tracking

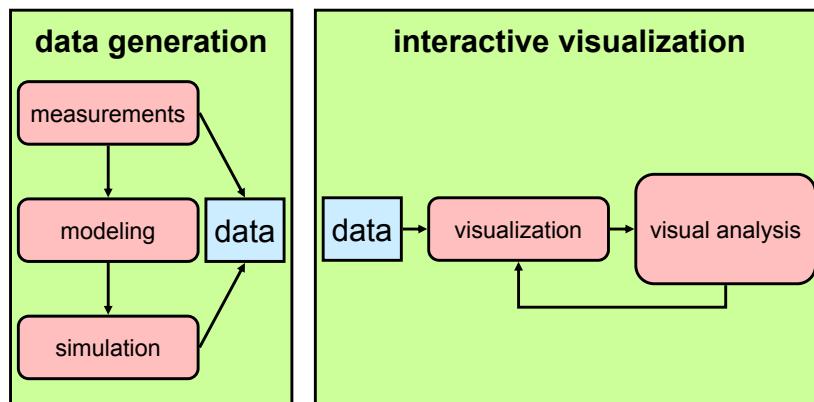
The gathered data is directly visualised and analysed.



(Image: T. Ertl)

1.1.3 Interactive Post Processing/Visualisation

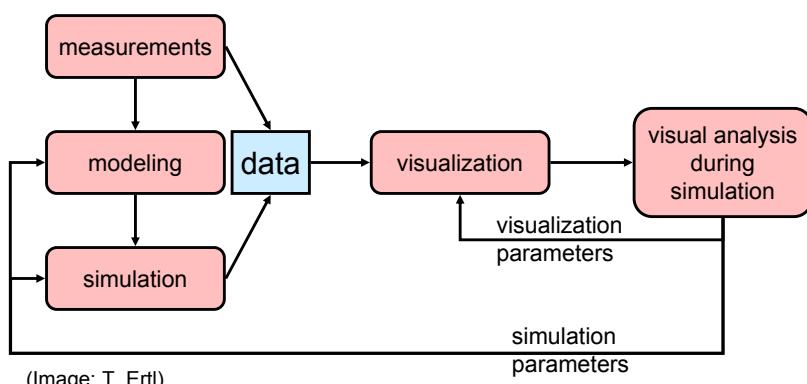
The data generation step is split from the visualisation step.



(Image: T. Ertl)

1.1.4 Interactive Steering/Computational Steering

The visualisation has a direct impact on the simulation and the visualisation.



1.2 Data Discretisations

Types of data sources have typical types of discretisations:

Measurement Data Typically scattered ("mesh-less", no grid)

Numerical Simulation Data

- Structured, block-structured or unstructured grids
- Adaptively refined meshes
- Multi-Zone grids with relative motion
- ...

Imaging Methods Uniform grids

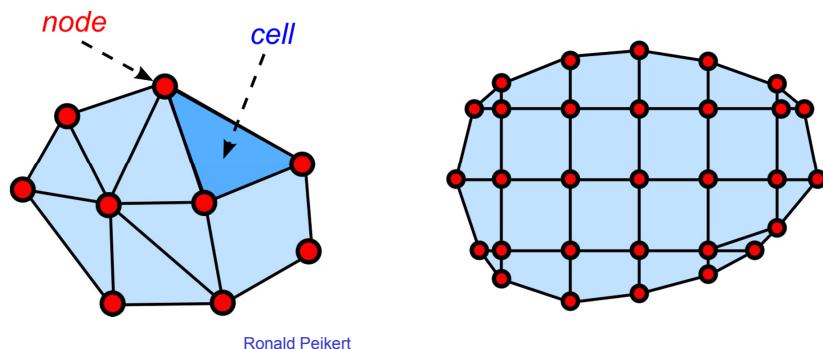
Mathematical Functions (and functionally represented data) can be sampled by demand:

- Uniform
- Adaptive

1.3 Unstructured Grids

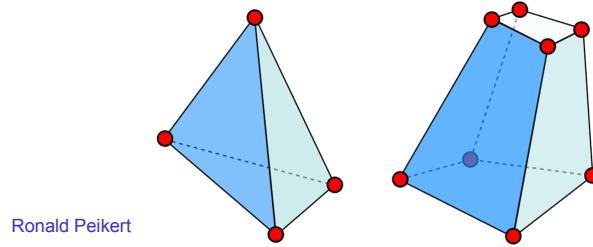
1.3.1 2D Unstructured Grids

Cells are *triangles* and/or quadrangles. The domain can be a surface embedded in 3-space (distinguish n -dimensional from n -space).

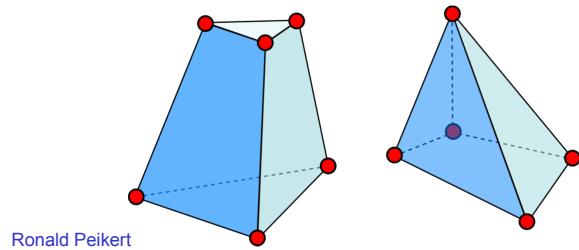


1.3.2 3D Unstructured Grids

Cells are *tetrahedra* or *hexahedra*.



Mixed grids ("zoo meshes") require additional types: *wedge* (3 sided prism), and a *pyramid* (4-sided).



1.4 Structured Grids

Curvilinear Grid (general case) Nodes are given in an array $N_i \times N_j \times N_k$ and the cells are implicit.

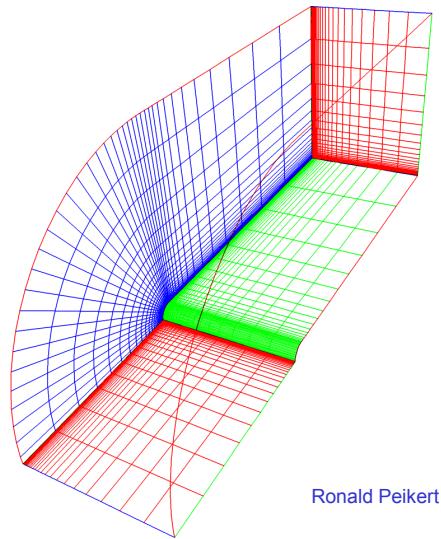


Figure 1: Curvilinear Grid

Rectilinear Grid (special case) The coordinate functions are simpler:

$$x = x(i) \quad y = y(j) \quad z = z(k)$$

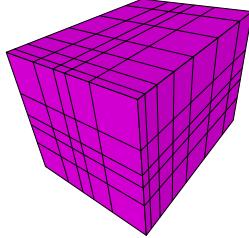


Figure 2: Rectilinear Grid, Source: Wikipedia

Uniform Grid (more special) The coordinates are defined by an *axis-aligned* bounding box.

1.4.1 Point-Sampled Data/Scattered Data

Point sampled data returns only nodes and no cells. Typical data sources are measurement data for example meteorological data.

Options for visualisation include:

Point-Based Methods (relatively few algorithms)

Triangulation for example constrained Delaunay (difficult in 3D)

Resampling onto uniform grid.

1.5 Elementary Visualisation Methods

Scalar Fields can be visualised by plotting its *function graphs*:

1D Field: The Graph is a curve:

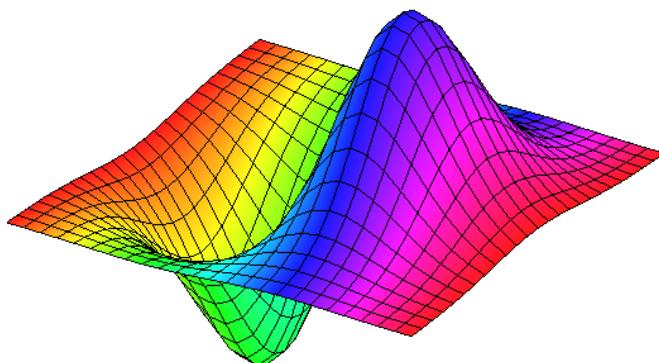
$$y = f(x)$$

2D Field: The Graph is a *height field*:

$$z = f(x, y)$$

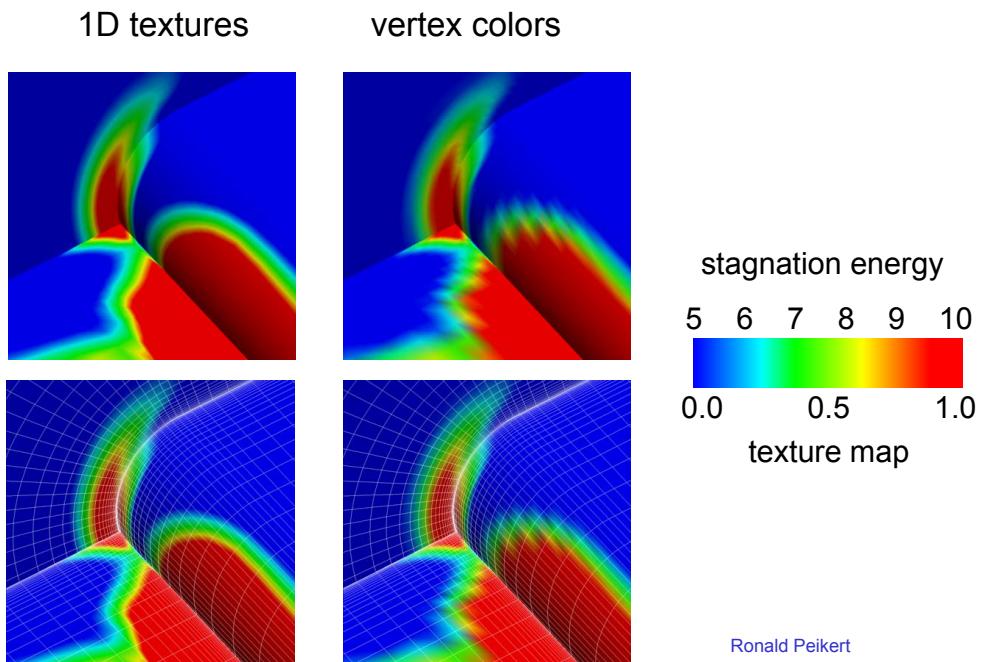
Visualisation is easy for rectilinear grids:

Painter's algorithm (hidden surface removal in software): Draw cells row by row from back to front.



Ronald Peikert

Scalar fields can also be visualised using *color coding* using *1D texture mapping*. Don't use *vertex colors* and Gouraud shading!



Ronald Peikert

- Problem of RGB colouring mode: The Interpolation is in the wrong colour space (RGB vs. colour table).
- Problem of Colour Index mode: Lighting is not possible.

2 Contouring and Isosurfaces

2.1 Contours

Contours are a set of points where the scalar field s has a given value c :

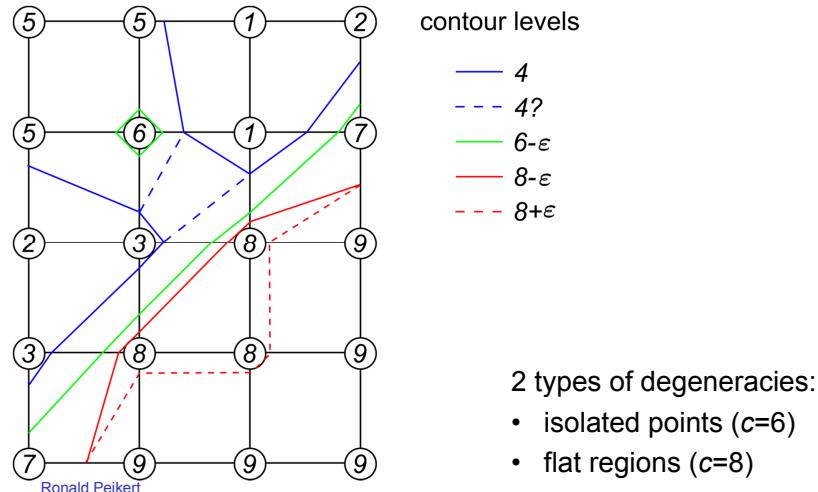
$$\{x \in \mathbb{R}^n : s(x) = c\}$$

Examples in 2D:

- Height contours on maps
- Isobars on weathermaps

Contouring algorithm:

- Find intersection with grid edges
- Connect points in each cell



Topological consistency To avoid degeneracies, use *symbolic perturbations*:

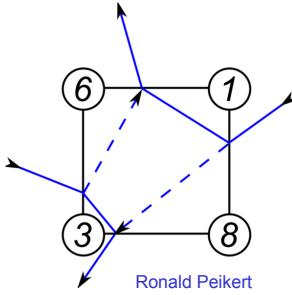
If level c is found as a node value, set the level to $c + \epsilon$ where ϵ is an infinitesimal, i.e., $\epsilon > 0$ and $\epsilon < x \forall x \in \mathbb{R}$.

Then:

- Contours intersect edges at some (possibly infinitesimal) distance from end points.
- Flat regions can be visualised by a pair of contours at $c - \epsilon$ and $c + \epsilon$.
- Contours are *topologically consistent*, meaning:

Contours are *closed, orientable, nonintersecting lines*.

Ambiguities of contours What is the *correct* contour of $c = 4$?



Ronald Peikert

Two possibilities (from which both are orientable):

- Connecting the high values _____
- Connecting the low values - - - - -

2.1.1 Contours in a quadrangle cell

Local Coordinates $(0, 0)$ $(1, 0)$ $(0, 1)$ $(1, 1)$

Function Values s_{00} s_{10} s_{01} s_{11}

Bilinear Interpolant

$$\begin{aligned} s(x, y) &= (1-x)(1-y) s_{00} + x(1-y) s_{10} + (1-x)y s_{01} + xy s_{11} \\ &= Axy + Bx + Cy + D \end{aligned}$$

with $A = s_{11} - s_{01} - s_{10} - s_{00}$

$B = s_{10} - s_{00}$

$C = s_{01} - s_{00}$

$D = s_{00}$

- If $A = 0$, then the contour equation is $c = Bx + Cy + D$ and the contours are *straight lines*, all parallel.
- If $A \neq 0$, then the contour equation is

$$c = A \left(x + \frac{C}{A} \right) \left(y + \frac{B}{A} \right) + D - \frac{BC}{A}$$

and the contours are *hyperbola* except for the level

$$c = D - \frac{BC}{A}.$$

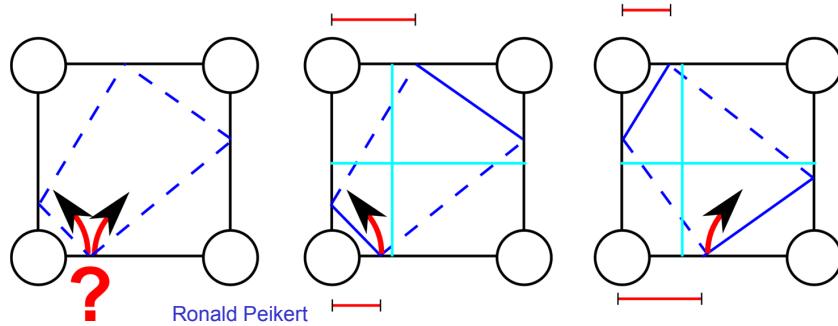
For the special level the contour equation is

$$0 = A \left(x + \frac{C}{A} \right) \left(y + \frac{B}{A} \right),$$

and the contour is a pair of axis-aligned straight lines with

$$\begin{aligned} x &= -\frac{C}{A}, \\ y &= -\frac{B}{A}. \end{aligned}$$

Decision can be made without computing special level or saddle points by just comparing fractions of edges:



By using local coordinates, this works also for curvilinear and unstructured grids.
Note that drawing hyperbola instead of straight lines does not lead to better contours:
The piecewise bilinear function is not in C^1 .

2.1.2 Basic contouring algorithms

Cell-By-Cell algorithms: Simple structure, but generate disconnected segments and require post-processing.

Contour Propagation methods: Complicated, but generate connected contours.

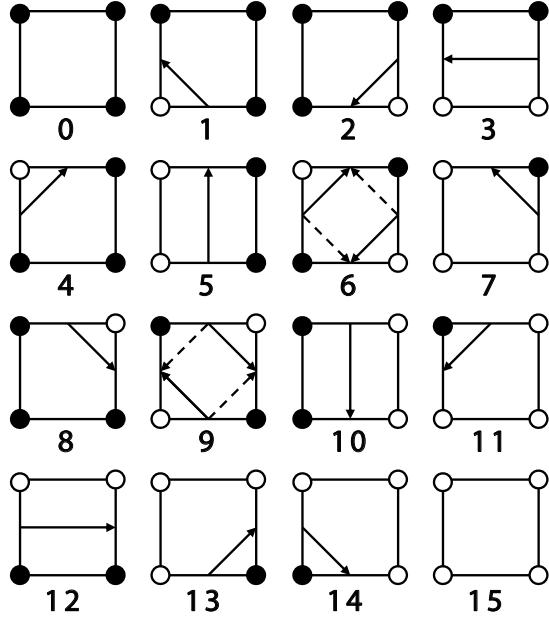
Marching Squares algorithm: Systematic cell-by-cell algorithm. (See below)

2.1.3 Marching Squares

Process nodes in ccw (counter-clockwise order), denoted here as x_0, x_1, x_2 and x_3 . At each node x_i compute the reduced field

$$\tilde{s}(x_i) = s(x_i) - (c - \varepsilon) \quad (\text{which is forced to be nonzero}).$$

Take it's signe as the i^{th} bit of a 4-bit integer. Use this as an index for the lookup table containing the connectivity information.



- $\tilde{s}(x_i) < 0$
- $\tilde{s}(x_i) > 0$

Alternating signs exist in cases 6 and 9. Chose the solid or the dashed line based on topological *consistency*.

Contours in triangle/tetrahedral cells Linear interpolation of cells implies piece-wise linear contours. Since contours are unambiguous let us introduce a "marching triangles" method. This however introduces periodic artefacts.

2.2 The Marching Cubes Algorithm

Contours of 3D scalar fields are known as *isosurfaces*. Before 1987, isosurfaces were computed as contours on planar *slices*, followed by "contour stitching".

The *marching cubes* algorithm computes contours *directly in 3D*:

- Pieces of the isosurfaces are generated on a cell-by-cell basis.
- Similar to marching squares, an 8-bit number is computed from the 8 signs of $\tilde{s}(x_i)$ on the corners of a hexahedral cell.
- The isosurface piece is looked up in a table with 256 entries.

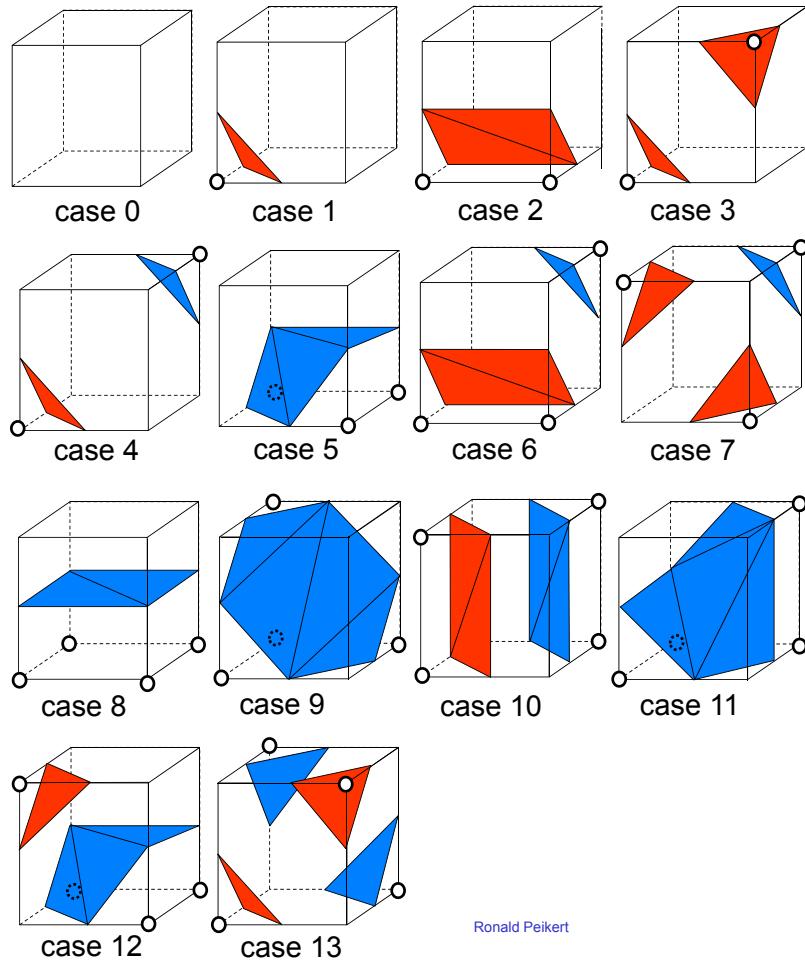
How to build up the table of 256 cases?

Lorensen and Cline (1987) exploited 3 types of symmetries:

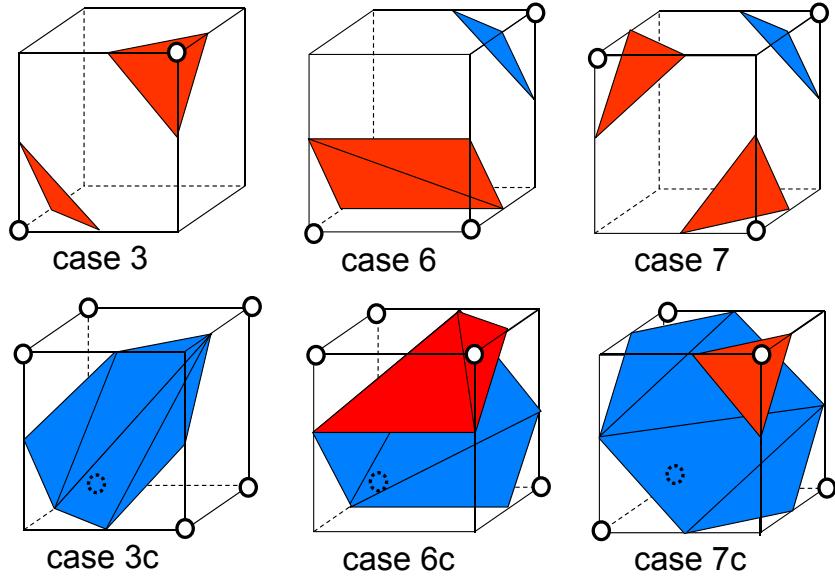
- Rotational symmetries of the cube
- Reflective symmetries of the cube
- Sign changes of $\tilde{s}(x)$.

They published a reduced set of 14 cases:

- White circle indicate positive signs of $\tilde{s}(x)$,
- The positive side of the isosurface is drawn in red, the negative side in blue.



Unfortunately not all pieces fit together (same problem as with marching squares) and additional cases need to be introduced.



Ronald Peikert

The remaining complementary cases are simply obtained by changing the orientation. Based on these 28 cases, the full 256 cases are obtained by rotations of the cube.

Summary of the algorithm:

1. Pre-processing:
 - Build a table of the 28 cases.
 - Derive a table of the 256 cases, containing info on
 - Intersected cell edges
 - Triangles based on these points
2. Loop over cells:
 - Find sign of $\tilde{s}(x)$ for the 8 corner nodes, giving 8-bit integer.
 - Use as index into lookup table.
 - Find intersection points on edges listed in table, use linear interpolation.
 - Generated triangles according to table.
3. Post-processing steps:
 - Connect triangles (share vertices)
 - Compute normal vectors:
 - By averaging triangle normals (problem: thin triangles!).
 - By estimating the gradient of the field $s(x)$ (better).

2.3 The Asymptotic Decider Algorithm

Motivation for a different isosurface algorithm: Marching cubes can produce "bad" topology.

Asymptotic decider algorithm (Nielson and Hamann 1991):

- Generate topologically *correct* contours (as oriented straight line segments) on the cell surfaces.
- Connect these around the cell, resulting in one or more polygons.
- Triangulate the polygons.

In general, the AD algorithm generates better isosurfaces. However,

- It cannot be easily implemented with a table like Marching Cubes (too many cases).
- It generates polygons with up to 12 sides (Marching Cubes: up to 7).
- The topology is correct with respect to the trilinear interpolant, but the geometry can deviate.
- Some polygons cannot be "cleanly" triangulated.

2.4 The Dividing Cubes Algorithm

An early *point-based* algorithm (Crawford et al. '87): For each cell

- Check whether it is intersected by the isosurface:

$$\min_{i \in \text{cell}} s_i < c < \max_{i \in \text{cell}} s_i$$

- Subdivide the intersected cell into $m \times m \times m$ subcells using trilinear interpolation.
- Draw the centers of all intersected subcells.
- To light the points estimate the gradient and use it as the normal vector.

2.5 Optimised Isosurface Algorithms

Approaches to speeding up isosurface computation:

View Dependent algorithms:

- Occluded triangles are not computed
- GPU-based isosurface computation and redenring

Data Processing for fast computation of *multiple* isosurfaces (multiple levels), for example for interactive exploration of the data.

- Many methods: Octree, Extrema Graph, Span Space.
- Common Goal: Avoid computation in non-intersected cells.

2.5.1 The Span-Space Algorithm

Method by Livnat (1996).

1. Pre-Processing. For each element

- Compute min and max ,
- Treat (min, max) as a point in the *span space* (Euclidean plane).
- Store points in boxes, non-empty boxes are stored in a linked list.

2. Computation of the isosurface level c :

Find the intersected cells in the *quadrant* $\min < c, \max > c$.

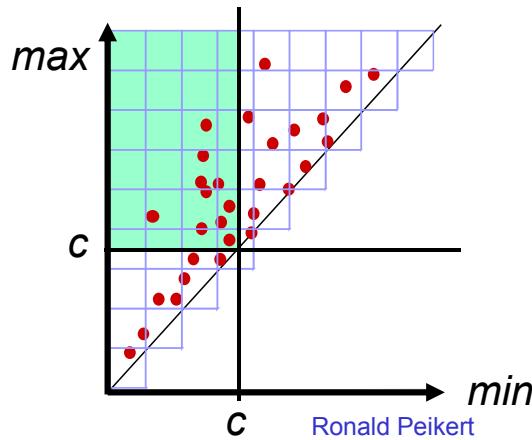


Figure 3: Node distribution in span space.

This algorithm yields a performance gain for datasets with a small local variation, i.e. points in the span space are distributed on the diagonal.

2.6 Selecting Contour Levels

Several types of isosurface statistics can help with level selection.

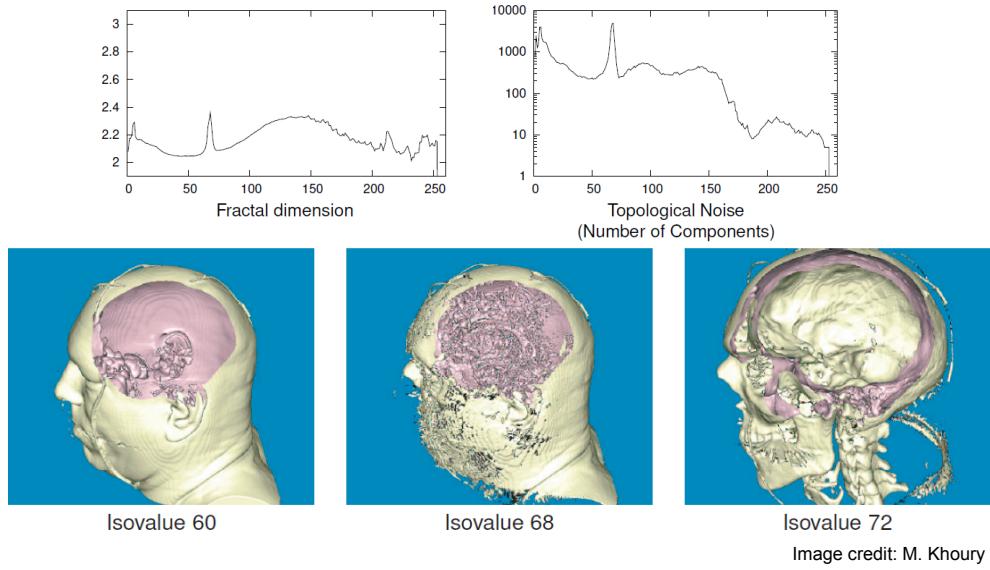


Image credit: M. Khoury

2.7 Limitations of Isosurfaces

Isosurfaces represent only a single level within the data range. In practical data, there is often not a single "interesting" level.

Transparent rendering of multiple isosurfaces is possible, but

- Limited to a small number of surfaces by visibility
- Alpha-blending require depth sorting.

Alternatives:

- *Feature Extraction methods*, e.g. detecting "blobs" (maximal ellipse-like contours)
- *Volume rendering* can show ranges of "interesting" levels of the field and/or its gradient.

3 Raycasting

3.1 Direct Volume Rendering

Volume rendering (sometimes called *direct volume rendering*) stands for methods that generate images directly from 3D scalar data. "Directly" means: *no intermediate geometry* (such as an isosurface) is generated.

Volume rendering techniques

- Depend strongly on the grid type.
- Exist for structured and unstructured grids.
- Are predominantly applied to uniform grids with 2D or 3D image data with *cell-centered* data. Cell-centered data
 - are attributed to cells (pixels, voxels) rather than nodes,
 - can also occur in (finite volume) CFD datasets,
 - are converted to node data:
 - * By taking the *dual grid* (easy for uniform grids: n cells $\rightarrow n - 1$ cells!),
 - * or by interpolating.

3.2 Raycasting

Raycasting is historically the first volume rendering technique. It is very similar to *raytracing*:

- *image-space* method: The main loop iterates over the pixels of the output image,
- a *view ray* per pixel (or per subpixel) is traced backward,
- and samples are taken along the ray and *composited* to a single color.

The differences are

- No secondary (reflected, shadow) rays,
- the transmitted ray is not refracted,
- more elaborate compositing functions,
- and samples are taken at intervals (not at object intersections).

3.2.1 Compositing

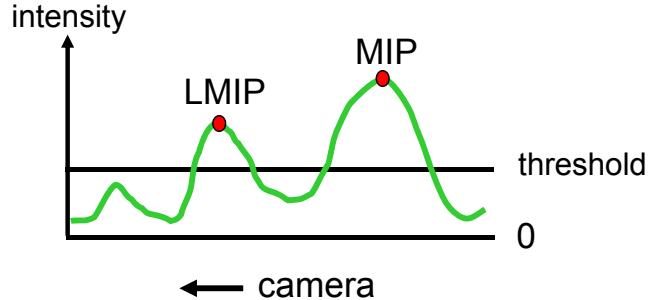
Two simple compositing functions can be used for previewing:

Maximum intensity projection (MIP):

- Maximum of sampled values
- Result resembles X-ray image.

Local maximum Intensity projection (LMIP):

- Choose first local maximum which is above a prescribed threshold.
- Process approximates occlusion.
- It is faster and better than MIP.



3.2.2 α -compositing

Assume that each sample on a view ray has a *color* and *opacity*.

$$(C_0, \alpha_0), \dots, (C_N, \alpha_N) \quad C_i \in [0, 1]^3, \quad \alpha_i \in [0, 1]$$

where the 0^{th} sample is next to the camera and the N^{th} one is a (fully opaque) background sample:

$$\begin{aligned} C_N &= (r, g, b)_{\text{background}} \\ \alpha_N &= 1. \end{aligned}$$

α -compositing can be defined recursively:

Let C_f^b denote the *composite color* of samples $f, f + 1, \dots, b$. The recursion formula for *back-to-front* compositing yields:

$$\begin{aligned} C_b^b &= \alpha_b C_b \\ C_f^b &= \alpha_f C_f + (1 - \alpha_f) C_{f+1}^b. \end{aligned}$$

With *transparency* set to $T_i = 1 - \alpha_i$ we get a *closed formula* for α -compositing:

$$C_f^b = \sum_{i=f}^b \alpha_i C_i \prod_{j=f}^{i-1} T_j$$

The *front-to-back* compositing can be derived from the closed formula: Let T_f^b denote the *composite transparency* of samples $f, f+1, \dots, b$:

$$T_f^b = \prod_{j=f}^b T_j.$$

Then the *simultaneous recursion* for front-to-back composition is:

$$\begin{aligned} C_f^f &= \alpha_f C_f \\ T_f^f &= 1 - \alpha_f \\ C_f^{b+1} &= C_f^b + \alpha_{b+1} C_{b+1} T_f^b \\ T_f^{b+1} &= (1 - \alpha_{b+1}) T_f^b. \end{aligned}$$

The emission-absorption model How realistic is α compositing? The *emission-absorption* model (Sabella 1988) yields a basic *volume rendering equation*

$$L(x) = \int_x^{x_b} \varepsilon(x') \exp \left(- \int_x^{x'} \tau(x'') dx'' \right) dx'$$

The equation describes the *radiance* arriving along a ray at the position x on this ray. The *emission* function $\varepsilon(x)$ describes the photons "emitted" by the volume along the ray. The *absorption* function $\tau(x)$ is the probability that that photon traveling over a unit distance is lost by absorption.

The emission-absorption model is based on the Boltzmann transport equation in statistical physics, but completely *ignores scattering*. In more general models $\tau(x)$ is an *extinction function* having both an absorption term and a scattering term.

Instead, in the emission-absorption model:

- *Incident scattering* is modelled by the emission function
- *Loss by scattering* can be thought to be part of the absorption.

The discrete version of the emission absorption model:

$$L(x) = \sum_{i=0}^n \varepsilon_i \Delta x \exp \left(- \sum_{j=0}^{i-1} \tau_j \Delta x \right) = \sum_{i=0}^n \varepsilon_i \Delta x \prod_{j=0}^{i-1} e^{-\tau_j \Delta x},$$

matches the α -compositing formula

$$C_f^b = \sum_{i=f}^b \alpha_i C_i \prod_{j=f}^{i-1} (1 - \alpha_j)$$

and gives interpretations of "opacity" and "color":

$$\begin{aligned} \alpha_i &= 1 - e^{\tau_j \Delta x} \approx \tau_j \Delta x && \text{if } \Delta x \ll 1 \\ \alpha_i C_i &= \varepsilon_i \Delta x. \end{aligned}$$

The product $\tilde{C} = \alpha_i C_i$ is called a *premultiplied* or *associated* colour.

3.3 Transfer Functions

Transfer functions map raw voxel data to opacities and colours as needed for the α -compositing. Inputs of the TF (one or more):

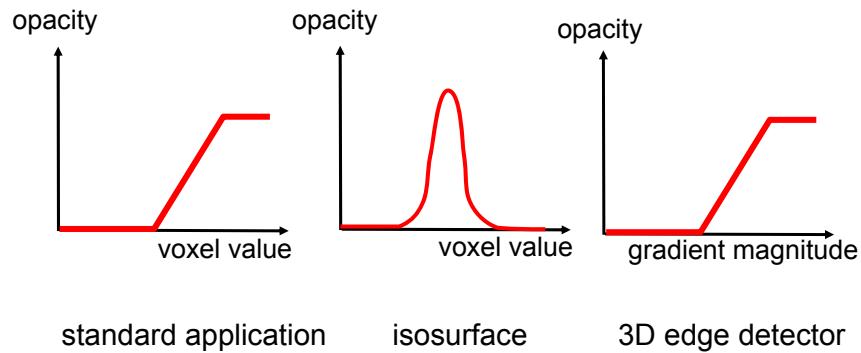
- Voxel value $s(x)$
- Gradient magnitude $\|\nabla s(x)\|$
- Higher derivatives of $s(x)$.

Opacity transfer function $\alpha(s(x), \|\nabla s(x)\|, \dots)$

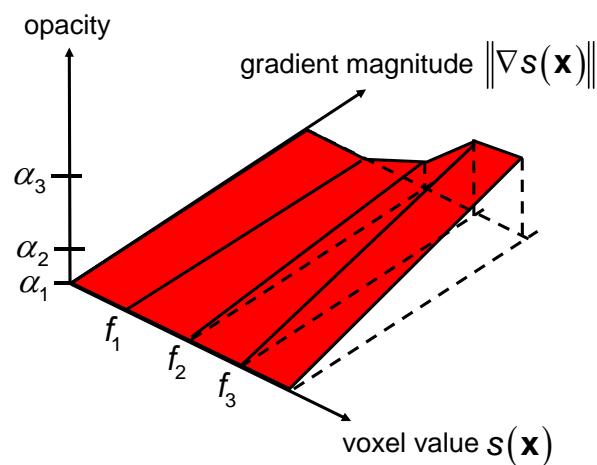
Color transfer function $C(s(x), \|\nabla s(x)\|, \dots)$

In general TF don't depend on *spatial location*, exception for *focus and context* techniques.

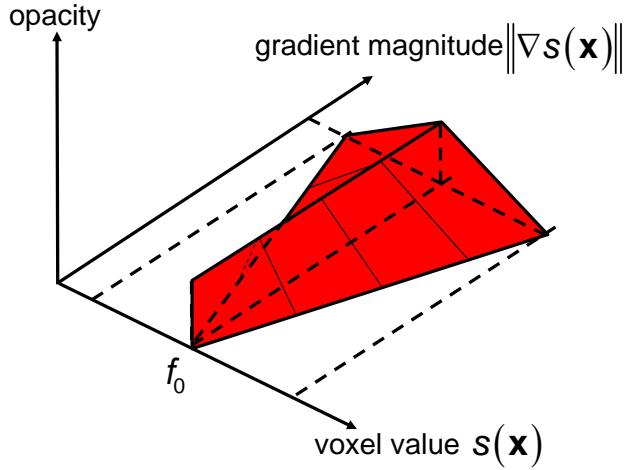
By choosing different opacity transfer functions different types of applications can be achieved.



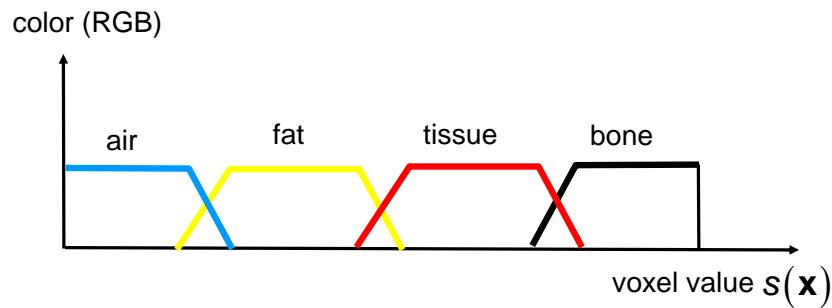
Example of a *bivariate* (=2D) transfer function:



Example of a bivariate transfer function for an isosurface of constant thickness.



The color transfer function allows to make a simple *classification*.



Pre-classification In pre-classification, the voxels can also be lit:

- The gradient is perpendicular to the local isosurface. It can be used as a normal vector for *Phong lighting* (without rendering the isosurface itself).
- *Reflection coefficients* can be assigned by a separate transfer function ("materials" instead of only colors).
- *Diffuse lighting* can be applied to the entire volume dataset as a pre-processing since it's independent of the viewing direction.

3.3.1 Pre- vs. Post-classification

For quality reasons, current volume rendering implementations often use *post-classification*.

Pre-Classification

1. Transfer functions are applied to voxels.
2. Results are interpolated to sample locations.

Post-Classification

1. Raw data are interpolated to sample locations.
2. Transfer functions are applied to sampled data.

3.4 Preintegration

Idea (Engel 2001): Simulate *infinitely many* interpolated samples between two successive samples $s_i = s(x_i)$ and $s_{i+1} = s(x_{i+1})$. Assuming that

- Field $s(x)$ varies *linearly* between samples
- and that the transfer functions don't depend on derivatives.

The discrete formula for opacity at a sample was

$$\alpha_i = 1 - e^{-\tau_i \Delta x}.$$

The continuous version for a sample interval $[x_i, x_{i+1}]$ is

$$\alpha_i = 1 - e^{-\int_{x_i}^{x_{i+1}} \tau(s(x)) dx}.$$

Assuming now $s(x)$ to be linear between samples, we get:

$$\alpha_i = 1 - \exp\left(-\frac{d}{s_{i+1} - s_i} \int_{s_i}^{s_{i+1}} \tau(s) ds\right) \quad \text{with } d = \|x_{i+1} - x_i\|,$$

which is called a *preintegrated opacity transfer function*.

The integral

$$\int_{s_i}^{s_{i+1}} \tau(s) ds = \int_0^{s_{i+1}} \tau(s) ds - \int_0^{s_i} \tau(s) ds$$

can be evaluated by two lookups in a precomputed table of

$$\int_0^s \tau(s') ds'.$$

The composite colour of the same interval

$$C_i = \int_{x_i}^{x_{i+1}} \varepsilon(s(x)) \exp\left(-\int_{x_i}^x \tau(s(x')) dx'\right) dx,$$

simplifies for linear $s(x)$ to:

$$C_i = \frac{d}{s_{i+1} - s_i} \int_{s_i}^{s_{i+1}} \varepsilon(s) \exp\left(-\frac{d}{s_{i+1} - s_i} \int_{s_i}^s \tau(s') ds'\right) ds,$$

which can be precomputed for all combinations of s_i , s_{i+1} and d .

3.4.1 Extinction-based volume rendering

Instead of an opacity Transfer function, use an extinction TF [Schlegel 2011]. Advantage: Extinction is *additive*.

- Riemann sums for numerical integration give better accuracy.
- Additivity can be used for efficient lighting.

Screen-Space ambient occlusion Approximate the fraction of ambient light that is occluded. Method: Compute the total extinction per shell (boxes approximating spheres) b using a *summed area table*.

4 Object Space Volume Rendering

In object space rendering methods the main loop is not over the pixels but over the objects in 3-space. In case of direct volume rendering "objects" can mean:

- Layers of voxels: Use *Image compositing* methods:
 - 2D texture based
 - 3D texture based
- Voxels: Use *splatting* methods.
- Cells: Use *cell projection* methods.

4.1 Texture-based volume rendering

4.1.1 Volume rendering with 2D texturemapping

- Use planes parallel to the *base plane*. The base plane is the front face of the volume which is "most orthogonal" to the view ray.
- Draw the texture rectangles using a bilinear interpolation filter.
- Render back-to-front using α -blending for the α -composition.

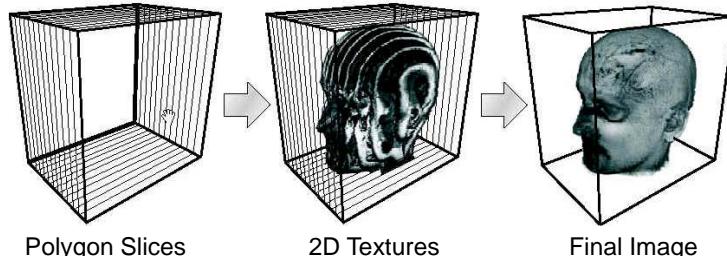


Image credit: H.W.Shen, Ohio State U.

4.1.2 Volume rendering with 3D texture mapping

Cabral 1994.

- Use the voxel data as the 3D texture.
- Render an arbitrary number of slices (eg. 100 or 1000) parallel to the image plane (3- to 6- sided polygons).
- Back-to-front compositing as in the 2D texture method.

This method is limited by the size of the texture memory.

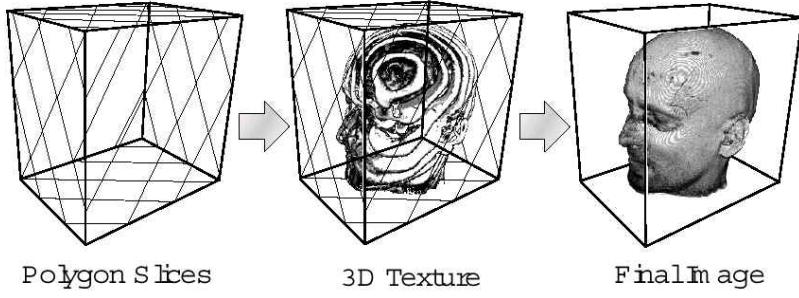
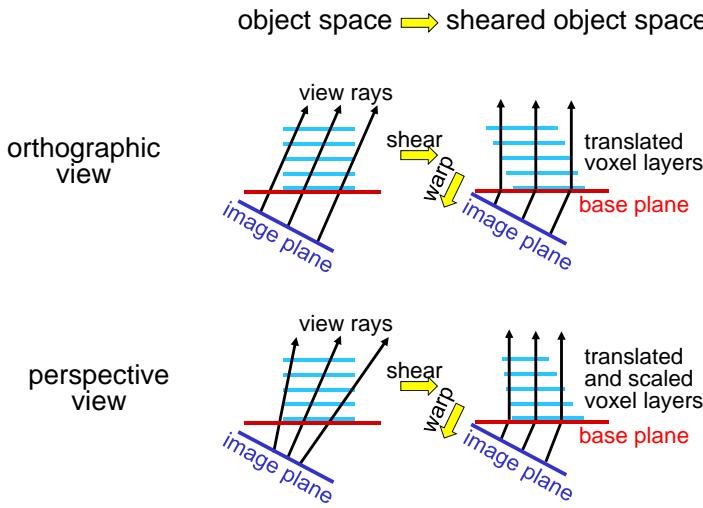


Image credit: H.W.Shen, Ohio State U.

4.2 Shear-Warp Factorisation

In general the image plane is not parallel to a volume face. The *shear-warp* method allows to render an intermediate image in the base plane:

1. Transform the *sheared object space* by translating (and possibly scaling) the voxel layers.
2. *Render* the intermediate image in the base plane.
3. *Warp* the intermediate image.



The *view transformation* is an affine transformation consisting of a rotation and a translation. Ignoring the translation, the 3×3 matrix can be factorised:

$$M_{\text{view}} = W \cdot S \cdot P$$

where

- P is a permutation matrix mapping the base plane to the xy plane.
- S is the shear matrix. The *shear* is of the form

$$S \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} + z \begin{pmatrix} s_x \\ s_y \\ 0 \end{pmatrix}.$$

With S being:

$$S = \begin{pmatrix} 1 & 0 & s_x \\ 0 & 1 & s_y \\ 0 & 0 & 1 \end{pmatrix}$$

where s_x and s_y have to be solved for from M_{view} .

- W is the warp matrix. The *warp* is a 3×3 matrix and effectively an affine transformation of the xy -plane. The third row of W is irrelevant while two zeros in the third column are required to make the warp independent of z .

$$W = \begin{pmatrix} w_{00} & w_{01} & 0 \\ w_{10} & w_{11} & 0 \\ w_{20} & w_{21} & w_{22} \end{pmatrix}$$

Assuming for simplicity that P is the identity, we get:

$$M_{\text{view}} = \begin{pmatrix} v_{00} & v_{01} & v_{02} \\ v_{10} & v_{11} & v_{12} \\ v_{20} & v_{21} & v_{22} \end{pmatrix} = W \cdot S = \begin{pmatrix} w_{00} & w_{01} & s_x w_{00} + s_y w_{01} \\ w_{10} & w_{11} & s_x w_{10} + s_y w_{11} \\ w_{20} & w_{21} & s_x w_{20} + s_y w_{21} + w_{22} \end{pmatrix}.$$

It follows for the warp coefficients $w_{ij} = v_{ij}$ ($j \neq 2$) and for the shear coefficients:

$$\begin{pmatrix} s_x \\ s_y \end{pmatrix} = \begin{pmatrix} v_{00} & v_{01} \\ v_{10} & v_{11} \end{pmatrix}^{-1} \begin{pmatrix} v_{02} \\ v_{12} \end{pmatrix}$$

and for w_{22} (not needed):

$$w_{22} = -s_x v_{20} - s_y v_{21} + v_{22}.$$

If P is not the identity, permuted versions of S and W can be used.

4.3 Perspective shear warp

The same factorisation as before can be used. But now *homogeneous coordinates* are used:

$$M_{\text{view}} = W \cdot S \cdot P.$$

The *shear and scaling* matrix S gets the form

$$S = \begin{pmatrix} 1 & 0 & s_x & 0 \\ 0 & 1 & s_y & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & s_w & 1 \end{pmatrix}.$$

It does

- a translation of x by $s_x z$ and of y by $s_y z$ followed by
- a scaling with $\frac{1}{1+s_w z}$.

The *warp* matrix now is

$$W = \begin{pmatrix} w_{00} & w_{01} & 0 & w_{03} \\ w_{10} & w_{11} & 0 & w_{13} \\ w_{20} & w_{21} & w_{22} & w_{23} \\ w_{30} & w_{31} & 0 & w_{33} \end{pmatrix}.$$

Assuming again that P is the identity we get:

$$M_{\text{view}} = \begin{pmatrix} v_{00} & v_{01} & v_{02} & v_{03} \\ v_{10} & v_{11} & v_{12} & v_{13} \\ v_{20} & v_{21} & v_{22} & v_{23} \\ v_{30} & v_{31} & v_{32} & v_{33} \end{pmatrix} = W \cdot S = \begin{pmatrix} w_{00} & w_{01} & s_x w_{00} + s_y w_{01} + s_w w_{03} & w_{03} \\ w_{10} & w_{11} & s_x w_{10} + s_y w_{11} + s_w w_{13} & w_{13} \\ w_{20} & w_{21} & s_x w_{20} + s_y w_{21} + w_{22} + s_w w_{23} & w_{23} \\ w_{30} & w_{31} & s_x w_{30} + s_y w_{31} + s_w w_{33} & w_{33} \end{pmatrix}.$$

It follows that for the warp conditions $w_{ij} = v_{ij}$ with $(j \neq 2)$ holds. For the shear coefficients:

$$\begin{pmatrix} s_x \\ s_y \\ s_w \end{pmatrix} = \begin{pmatrix} v_{00} & v_{01} & v_{03} \\ v_{10} & v_{11} & v_{13} \\ v_{30} & v_{31} & v_{33} \end{pmatrix}^{-1} \begin{pmatrix} v_{02} \\ v_{12} \\ v_{32} \end{pmatrix}$$

and for w_{22} (not needed):

$$w_{22} = -s_x v_{20} - s_y v_{21} - s_w v_{23} + v_{22}.$$

For the shear-warp volume rendering algorithm now works as follows:

1. For each voxel layer (parallel to base plane):
 - Shear and scale the layer image by multiplying with S .
 - Apply the transfer functions.
2. Generate intermediate image with α compositing.
3. Warp the image by multiplying with W .

An advantage of this algorithm is that an aliasing filter can be used to prevent undersampling when scaling the image.

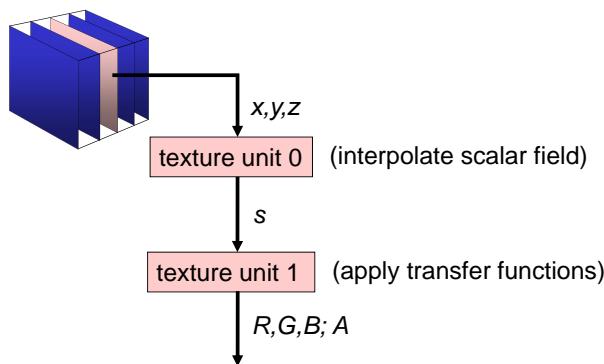
4.4 Object Space versus Image Space

Comparison of the object space methods introduced and image space methods such as *raycasting*.

Formally both methods are equivalent only the nesting order of the loops is different.
Practical differences:

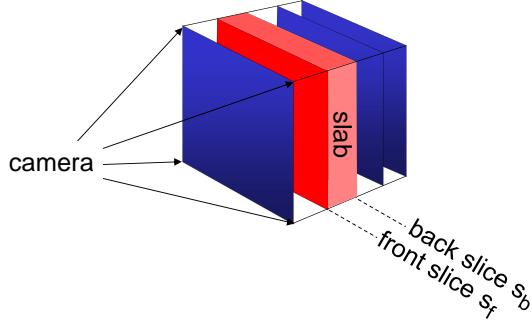
- Image space methods with FTB compositing allow early termination.
- Object space methods using a framebuffer for intermediate results suffer from quantisation artefacts.
- Object space methods can exploit texture mapping hardware and MIPmap textures for antialiasing.
- Image space methods would need supersampling in x and y to get antialiasing.

Post-classification The post-classification can be done directly in the graphics hardware. Using (OpenGL) *dependent texture* (two texture mapping stages):



Pre-integration It's also possible to pre-integrate in object space:

- Use *slabs* (space between two slices) and
- Dependent textures:
 - 1st stage: Interpolate scalar filed in front and back slice.
 - 2nd stage: Look up integrated transfer function.



4.5 Splatting

Raycasting "What does **each** voxel contribute to a given pixel?"

Splatting "What does **a given** voxel contribute to each pixel?"

Algorithm:

- Pre-processing:
 - For each voxel x_i render (raycast) a field $s(x_i) = \delta_{ij}$.
 - Store the resulting *footprint* images.
- Main loop:
 - For each voxel x_i adjust the footprint image to effective TF value.
 - Do α -compositing of all footprint images.

Advantages of splatting:

- Applicable to structured and unstructured grids.
- Other reconstruction filters than trilinear interpolation are possible (for example a sinc filter).

Original algorithm Westover 1990: Orthographic view and uniform grids. All footprints are translations of a template.

Sheet buffer method Westover 1991:

- Blend all footprint images of a voxel layer ("sheet buffer")
- Do α -compositing of sheet buffers.

Elliptical weighted average splatting (EWA), Zwicker et al. 2001:

- Ellipsoidal Gaussians are used as footprints
- Perspective view, low-pass filter for antialiasing.

4.6 Cell Projection

Projected tetrahedra (PT) is an object space method for tetrahedral grids [Shirley, Tuchmann 1990]. Each (tetrahedral) cell is decomposed into 3 or 4 tetrahedra along those edges which are not part of the silhouette.

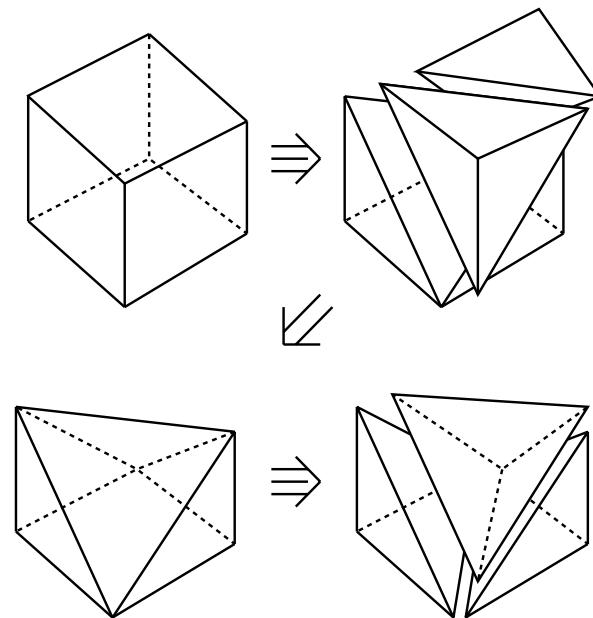


Figure 4: Decomposition of a cube into five tetrahedra. Source: A Polygonal Approximation to Direct Scalar Volume Rendering - [Shirley, Tuchmann 1990]

Cells are projected to *triangle fans* consisting of

- 1 *thick vertex* (projection of the common edge of the tetrahedra)
- 3 or 4 *thin vertices* (on the silhouette)

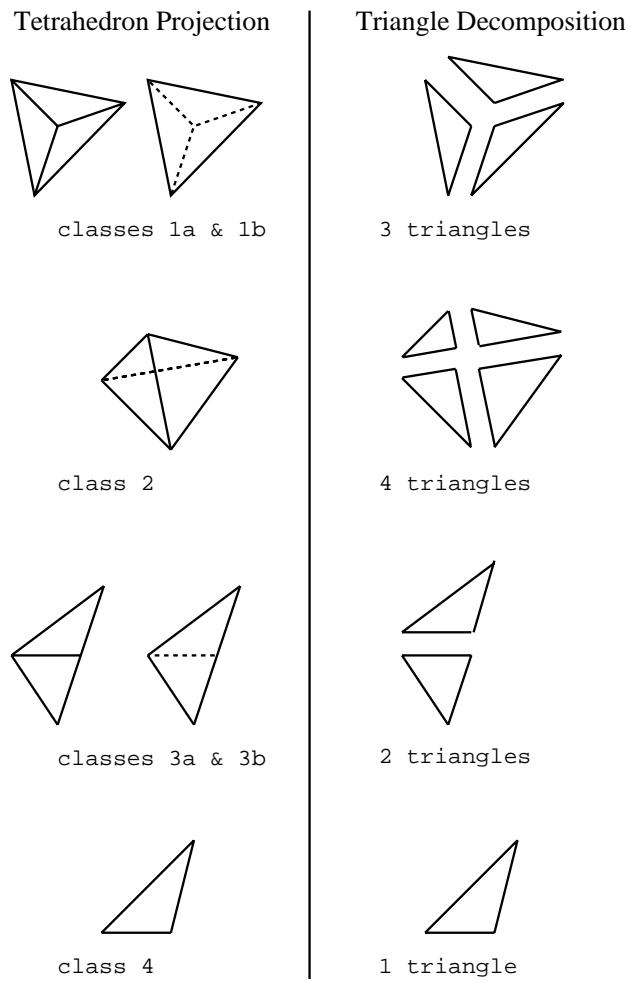


Figure 5: Splitting tetrahedra. Source: A Polygonal Approximation to Direct Scalar Volume Rendering - [Shirley, Tuchmann 1990]

Original algorithm: Store *triangle fan* in *space*:

- Thin vertices are kept in the *original position*.
- The thick vertex is set to the *midpoint* of the projected edge.

Advantages:

- *Depth test* can be used (allows volume rendering into a scene).
- *Viewing direction* and field-of-view can be changed (for a fixed camera position) by keeping the projection.

Computation of the thick vertex

- Compute the determinants

$$d_i = \det(x_j, x_k, x_l) \quad i \in [0, 3]$$

where x_j , x_k and x_l are the vertices of the i^{th} face relative to camera position, ordered counter clockwise on the outside of the face.

- If the number of positive determinants is

Odd Class 1

Even Class 2

- Interpolate weights

Painting The thick and thin vertices are then sent to the graphics card (with supplied texture coordinates).

4.6.1 Drawing

Back-to-front compositing

- Cells must be depth-sorted \Rightarrow visibility sorting.
- Drawing possible without re-sorting on camera turn and zoom.
- Depth-test (z-buffer) must be enabled.
- Additional (opaque) objects must be rendered before the volumes.

Visibility sorting (MPVO algorithm, Williams 1992):

- Generate *partial ordering* of cells based on adjacent pairs.
- Break *cycles* (rare, small rendering error, alternative: Split a cell).
- Sort the list of *front cells* by distance to centroid (heuristic)

4.6.2 Example: Visualisation of smoke propagation

For this model we use a simple smoke model which is mostly used in fire protection engineering. The *absorption* τ is proportional to $s(x)$ (particle concentration). Which leads to a simple preintegrated opacity transfer function:

$$\alpha = 1 - \exp\left(-c \frac{\tau_f + \tau_b}{2} \|x_b - x_f\|\right)$$

4.6.3 Opacities

When compositing cells with low opacity then opacities are essentially *added*. Adding many very small opacities (e.g. between $[0, 1/255]$) leads to *quantisation artefacts*. Options to reduce artefacts:

- Compositing with 16 bits
- α -dithering use *randomised rounding* instead of standard rounding:

$$x \rightarrow \lfloor x \rfloor + (x - \lfloor x \rfloor \geq \text{rand}) \quad \text{with } \text{rand} \in [0, 1]$$

4.6.4 Hardware-Assisted Visibility Sorting

HAVS, Silva et al. 2005 is a faster cell projection algorithm:

- Requires 4 RGBA float buffers for storing 7 pairs of (s, d) per pixel:
 - With a scalar field value s ,
 - and distance d to camera.
- The initial cell sorting is done on the CPU based on the centroids. Which results in a *k-nearly sorted sequence* with $k \leq 7$.
- Main loop: Draw all cell faces from back to front.
- Fragment shader:
 - Insert (s, d) into buffer
 - If the buffer is full:
 - * Take out furthes pair of (s, d)
 - * Compute thickness of cell behind the pixel:

$$\Delta d = d - d_{\text{old}}$$

- * Do a preintegrated TF lookup with $s, s_{\text{old}}, \Delta d$
- * Apply α compositing.

5 Vector Field Visualisation

- A *static vector field* $v(x)$ is a vector-valued function of space.
- A *time-dependent vector field* $v(x, t)$ also depends on time.

In the case of *velocity fields*, the terms *steady* and *unsteady flow* are used.

The dimensions of x and v are equal, often 2 or 3 and we denote the components by x, y, z and u, v, w :

$$x = (x, y, z), \quad v = (u, v, w).$$

Sometimes a vector field is defined on a surface $x(i, j)$. The vector field is then a function of parameters and time:

$$v(i, j, t).$$

5.1 Visualisation

An elementary visualisation of a vector field is to *draw arrows*:

- At the data points (grid nodes or cell centers), or
- at a new (uniform) grid. For 3D fields it's often a 2D slice.

Arrows can visualise:

- Direction
- Relative magnitude (when appropriately scaled)
- Time dependency (when animated)

Problems

- It is not clear whether arrows represent vector values at the start point or at the midpoint of the arrow.
- Often there exist no satisfactory scaling factors:
 - Large scaling: Arrows occlude each other
 - Small scaling: Direction is not recognisable in some regions
 - Fixed length: Magnitude information is lost.

5.2 Vector fields as ODEs

For simplicity, the vector field is now interpreted as *velocity field*. The field $v(x, t)$ describes the connection between location and velocity of a (massless) particle. Which can equivalently be expressed as an *ordinary differential equation*:

$$\dot{x}(t) = v(x(t), t).$$

This ODE, together with an *initial condition*

$$x(t_0) = x_0,$$

is a so-called *initial value problem* (IVP). Its solution is the *integral curve* (or *trajectory*)

$$x(t) = x_0 + \int_{t_0}^t v(x(\tau), \tau) d\tau.$$

The integral curve is a *pathline* describing the *path* of a massless *particle* which was released at time t_0 at position x_0 .

Remark: $t < t_0$ is allowed.

For static fields the ODE is *autonomous*:

$$\dot{x}(t) = v(x(t))$$

and its integral curves

$$x(t) = x_0 + \int_{t_0}^t v(x(\tau)) d\tau$$

are called *field lines* or in the case of velocity fields *streamlines*.

- In *static vector fields* pathlines and streamlines are *identical*.
- In *time-dependent* vector fields *instantaneous streamlines* can be computed from a "snapshot" at a fixed time T (which is a static vector field):

$$v_T(x) = v(x, T).$$

In practice time-dependent fields are often given as a dataset per time step. Each dataset is then a snapshot.

Computing streaklines or timelines is more expensive than solving a single IVP.

5.2.1 Pathlines

Pathlines are the trajectories that individual fluid particles follow. These can be thought of as "recording" the path of a fluid element in the flow over a certain period. The direction the path takes will be determined by the streamlines of the fluid at each moment in time.²

²https://en.wikipedia.org/wiki/Streamlines,_streaklines,_and_pathlines

- Are physically meaningful
- Allow comparison with experiment (observe marked particles)
- Are well suited for dynamic visualisation (of particles).

5.2.2 Streamlines

Streamlines are a family of curves that are instantaneously tangent to the velocity vector of the flow. These show the direction a fluid element will travel in at any point in time.³

- Are only geometrically but not physically meaningful,
- Are easiest to compute (no temporal interpolation, single IVP),
- Are better suited for static visualisation (prints),
- Don't intersect (under reasonable assumptions)

Inputs

- Static vector field $v(x)$
- Seed points with time of release (x_0, t_0)
- Control Parameters:
 - Step size (temporal, spatial or in local coordinates)
 - Step count limit, time limit, etc...
 - Order of integration scheme

Output Streamlines as "polylines", with possible attributes.

Preprocessing

- Set up search structure for point location
- For each seed point:
 - *Global point location:*
Given a point x , find the cell containing x and the local coordinates (ξ, ν, ζ) . Alternatively if the grid is structured find the computational space coordinates

$$(i + \xi, j + \nu, k + \zeta)$$
 - If x is not found in a cell then remove the seed point.

³https://en.wikipedia.org/wiki/Streamlines,_streaklines,_and_pathlines

Integration For each seed point x :

- Interpolate v trilinearly to local coordinates (ξ, ν, ζ)
- Do an integration step producing a new point x'
- *Incremental point location*: For position x' find cell and local coordinates (ξ', ν', ζ') making use of information (coordinates, local coordinates, cell) of old point x .

Integration step Widely used methods:

- *Euler*, inaccurate, used only in special speed-optimised techniques:

$$x_{\text{new}} = x + v(x, t) \cdot \Delta t$$

- *Runge-Kutta*, 2nd or 4th order. Higher order schemes are often too slow for visualisation. Yeung/Pope 1987 showed that when using standard trilinear interpolation *interpolation errors* dominate *integration errors*.

- There are several options for choosing a step size:
 - Fixed time stepping Δt
 - Fixed spatial step Δs :

The time step is derived from the spatial step

$$\Delta t = \frac{\Delta s}{\|v(x)\|}$$

which needs to be corrected iteratively and is used for methods such as LIC.

- Adaptive:
 - * Adapting to grid resolution (such as cell size)
 - * Adapting to data variation (Runge-Kutta-Fehlberg method)
 - * Used for interactive viewing (with zooming)

Termination criteria

- Grid boundary reached
- Step count limit reached
- Optional: Velocity close to zero
- Optional: Time limit reached
- Optional: Arc length limit reached

5.2.3 Streaklines

Streaklines are the locus of points of all the fluid particles that have passed continuously through a particular spatial point in the past. Dye steadily injected into the fluid at a fixed point extends along a streakline.

- Are physically meaningful
- Allow a direct comparison with the experiment (i.e. dye injection)
- Are well suited for static and dynamic visualisation
- Are a good choice for fast moving vortices
- Can be approximated by a set of disconnected particles.

Algorithm:

1. For time samples t_0, t_1, \dots, t_n solve the IVP

$$\begin{aligned}\dot{x}_i(t) &= v(x_i(t), t) \\ x_i(t_i) &= y\end{aligned}$$

2. Extract the point $x_i(t_n)$

In numerical computation the temporal interval must be *adaptively refined* if two successive particles diverge too much.

5.2.4 Timelines

Timelines are the lines formed by a set of fluid particles that were marked at a previous instant in time, creating a line or a curve that is displaced in time as the particles move.

- Are physically meaningful
- Are well suited for static and dynamic visualisation
- Can be approximated by a set of disconnected particles.

Algorithm:

1. For point samples y_0, y_1, \dots, y_n on the seed curve solve the IVP:

$$\begin{aligned}\dot{x}_i(t) &= v(x_i(t), t) \\ x_i(t_0) &= y_i\end{aligned}$$

2. From the integral curve $x_i(t)$ extract the point $x_i(T)$.
3. Connect these points.

The result is a timeline for time T .

In the numerical computation the spatial interval must be *adaptively refined* if two neighbour particles diverge too much.

5.3 The Stencil Walk algorithm

Computing the *incremental point location* is nontrivial for curvilinear and unstructured grids.

Buning's *stencil walk* algorithm solves this problem.

Given

- Point with coordinates x
- Cell (as three parameters (i, j, k) or as index c respectively)
- Local coordinates (ξ, ν, ζ)
- Coordinates of a new point x'

Wanted

- New cell as (i', j', k') or as index c' respectively
- New local coordinates (ξ, ν, ζ)

Algorithm In a first phase the algorithm finds the cell containing x' by iteratively doing:

- Take the difference vector $\Delta x = x' - x$
- Intersecting the ray $x + t\Delta x$ with the cell boundary giving a t value.
 - Linearise the coordinate transform

$$\varphi : (\xi, \nu, \zeta) \mapsto (x, y, z)$$

in the point $x = (x, y, z)$ by computing the *Jacobian*

$$J = \frac{\delta(x, y, z)}{\delta(\xi, \nu, \zeta)} = \begin{bmatrix} \frac{\delta\varphi}{\delta\xi} & \frac{\delta\varphi}{\delta\nu} & \frac{\delta\varphi}{\delta\zeta} \end{bmatrix}.$$

- Using the *Jacobian* to transform the difference vector $\Delta x = x' - x$ into the local coordinate frame of the cell:

$$(\Delta\xi, \Delta\nu, \Delta\zeta) = J^{-1}\Delta x.$$

- Find the intersection of the ray

$$(\xi, \nu, \zeta) + t(\Delta\xi, \Delta\nu, \Delta\zeta)$$

with the cell boundary having the equations:

$$\xi, \nu, \zeta = 0 \tag{1}$$

$$\xi, \nu, \zeta = 1 \quad \text{for hex cell} \tag{2}$$

$$\xi + \nu + \zeta = 1 \quad \text{for tet cell} \tag{3}$$

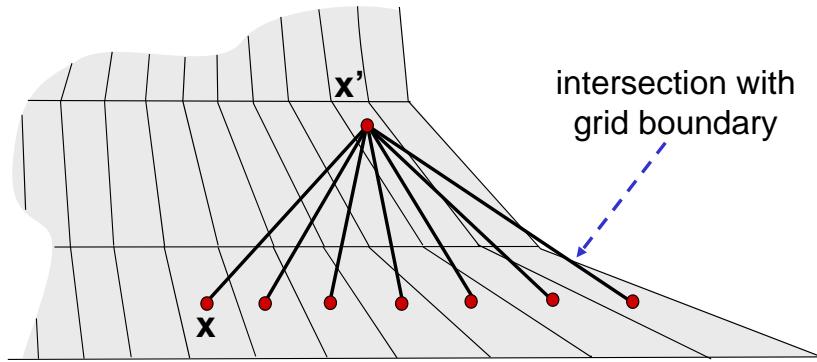
$$0 \leq \xi, \nu, \zeta \leq 1 \quad \text{inequalities} \tag{4}$$

Due to linearisation the point is not exact and in most cases the correct neighbour cell is found.

- If $t \geq 1$ the point x' lies in the current cell and iteration can be stopped.
- Otherwise move to the neighbour cell adjacent at the intersection point
- If no such cell exists terminate with failure
- Set the *cell centroid* as the new x for the next iteration.

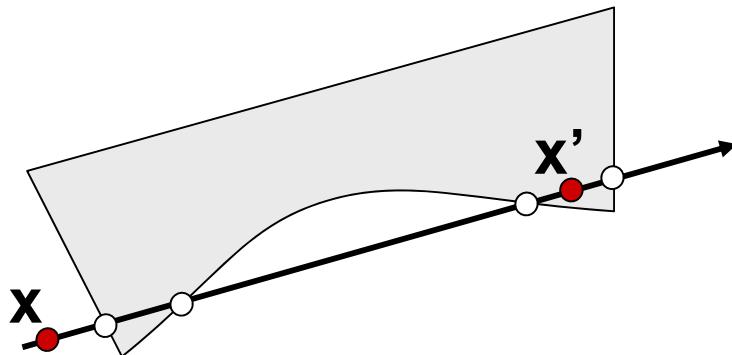
5.3.1 Problems

If the cells are sufficiently skewed the algorithm can walk away from the target cell.



This problem can be solved with a *modification* of the algorithm:

- Keep ray (x, x') unchanged
- New Problem: Cell faces of the type *quadrangle* (nonplanar!) can be intersected *twice!*



Therefore *exact* intersection points of the ray with bilinear surface patches must be calculated:

Exact intersection calculation If the quadrangle is *nonplanar* the four corners P_0, P_1, P_2, P_3 can be mapped to:

$$\begin{aligned}\tilde{P}_0 &= (0, 0, 0) \\ \tilde{P}_1 &= (1, 0, 0) \\ \tilde{P}_2 &= (1, 1, 1) \\ \tilde{P}_3 &= (0, 1, 0)\end{aligned}$$

by the *affine transformation*

$$\tilde{x} = (\tilde{P}_1 | \tilde{P}_2 | \tilde{P}_3)(P_1 | P_2 | P_3)^{-1}(x - P_0).$$

The bilinear surface containing P'_0, P'_1, P'_2 and P'_3 is the *hyperbolic paraboloid*

$$z = xy.$$

Inserting the transformed view ray $\tilde{x} + t\Delta\tilde{x}$ leads to a quadratic equation for t :

$$(\tilde{z} + t\Delta\tilde{z}) = (\tilde{x} + t\Delta\tilde{x})(\tilde{y} + t\Delta\tilde{y}).$$

If real solutions with $0 < t < 1$ exist the intersection points $(\tilde{x}_i, \tilde{y}_i, \tilde{z}_i)$ are computed and transformed back.

In the second phase the stencil walk algorithm computes the local coordinates of the point in the cell known to contain it. The local coordinates are the inverse of the coordinate function:

$$\varphi : (\xi, \nu, \zeta) \mapsto (x, y, z)$$

evaluate at the given point

$$x = (x, y, z).$$

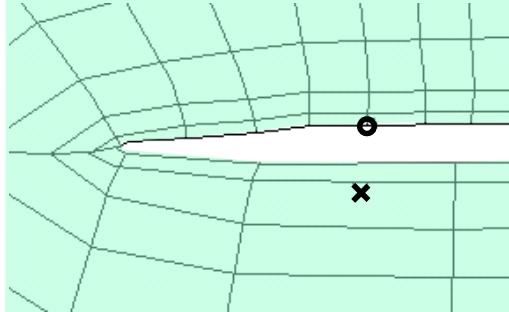
However, the trilinear function π is a cubic polynomial and its inverse is a sixth-degree polynomial. Hence the problem needs to be solved with Newton's method.

5.4 Global Point Location

Global point location is more expensive than the stencil walk algorithm. Many methods trade in safety for efficiency. A few methods are:

1. Search for the point in every grid cell using Newton's Method.
Hypothetic "brute force" method. Safe.
2. Buning's method.
Do incremental point location starting from a boundary cell. **Problem:**

Node (O) nearest to given point (X) is not necessarily adjacent to the cell containing it. Furthermore the straight line between the two points can leave the grid.



Bunings method is safe if incremental search is repeated with a different boundary cell as long as the point is not found. Instead of using all boundary cells, precompute a subset of boundary cells to guarantee to find all points within the grid.

3. Incremental point location starting from a node near the grid center.

Simple method. Only safe for *star-shaped* grids.

4. Search structure based.

Efficient methods use a search structure such as a uniform grid, octrees or kd-trees for nodes or cell centers:

- When a point query is not sufficient a *range query* is need with a range determined by the cell size.
- Problem: Cells can have extreme aspect ratios (especially from CFD).

5. Bounding box hierarchy

Use a bounding box hierarchy for a recursively subdivided grid. Efficient and safe method and easy for structured grids.

More preprocessing is required for unstructured grids (cell tree, Garth 2010)

5.5 Computational Space Streamline Integration

In structured grids *point location* can be *avoided* by using a different approach: Integration can be done in computational space \mathcal{C} instead of physical space \mathcal{P} . This requires a modification of the integration algorithm:

- *Before* the integration step:

Transform the *velocity* $v(x)$ to \mathcal{C} by multiplying with J^{-1}

- *After* the integration step:

This step is only required if graphical output of this step is needed. Transform the new *position*

$$x = (i + \xi, j + \nu, k + \zeta)$$

to \mathcal{P} by trilinear interpolation.

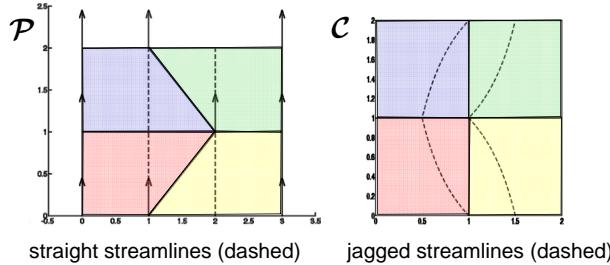
Main problem of the integration in \mathcal{C} :

- The ordinate function

$$\varphi : (i + \xi, j + \nu, k + \zeta) \mapsto (x, y, z)$$

is only C^0 continuous at cell boundaries.

- Therefore J is discontinuous.
- Example (Sadarjoen 1994): Four cells with a constant velocity field.



Two sources of error:

- Integration steps across cell boundaries.

This can be avoided by shortening such steps.

- Use of a (precomputed) single transformed field vector by node.

Can be fixed by transforming all eight vectors of a cell on the fly when entering a new cell.

The main advantage of integration in \mathcal{C} is its algorithmic simplicity. But if it's done correctly (avoiding the errors above) it can be slower than an integration in \mathcal{P} .

5.6 Skin Friction Lines

Velocity fields of fluid flow can have grid boundaries which are walls (i.e. solid material surfaces). At walls the velocity vector is usually zero as a result of the no-slip boundary condition. Therefore a derived vector field is often used: The *wall shear stress*:

$$\tau_w = \mu \frac{\delta v_w}{\delta s},$$

can be obtained as the limit of the wall-parallel velocity component v_w divided by the wall distances s and multiplied with the *dynamic viscosity* μ (material constant). This limit is typically nonzero except at isolated points.

Streamlines of τ_w are called *skin friction lines*. They are an example of a vector field defined on a surface in 3-space.

5.7 Streamline Placement

The problems of visualisation by streamlines:

- Dependency on seed points,
- and the density of streamlines can be largely inhomogeneous.

Solution Automatically optimise choice of seed points.

1. *Streamlets* (short streamline segments)

The length is proportional to the velocity magnitude (obtained automatically by using a fixed integration time).

Start with a uniform grid and make spacing roughly even by locally adapting (displacing, inserting, removing) seeds.

Typical use: Displaying weather maps

2. Algorithm by Turk and Bank (longer streamlines):

Objective: Create a streamline image which when low-pass filtered has a uniform grey level.

Optimise seed positions and integration lengths.

Operations:

- Insert
- Delete
- Move
- Lengthen
- Shorten

Apply operations either randomly or based on oracles.

5.8 Streamsurfaces

Definition: Union of streamlines seeded densely on a curve.

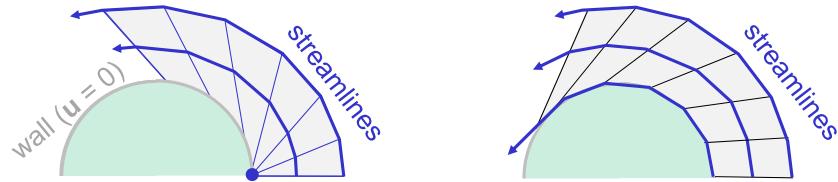
Advantage for visualisation: Structured, better spatial perception.

Naive algorithm:

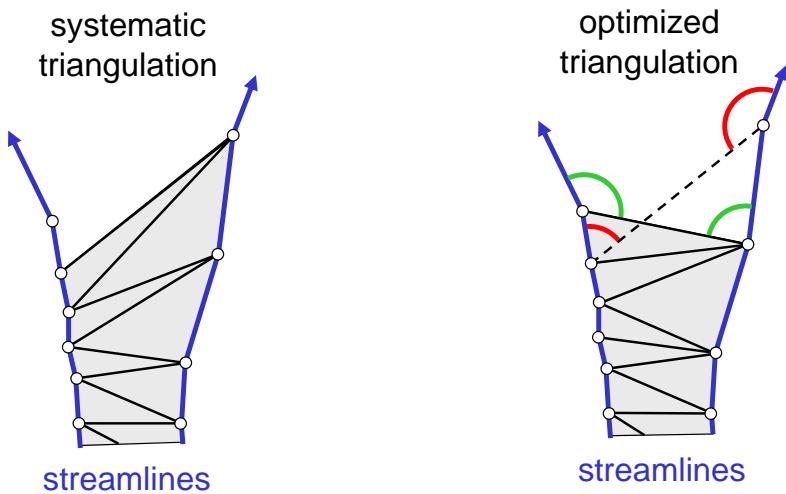
- Start integration at discrete samples on the seed curve.

- Connect points of equal integration time resulting in a quad mesh.

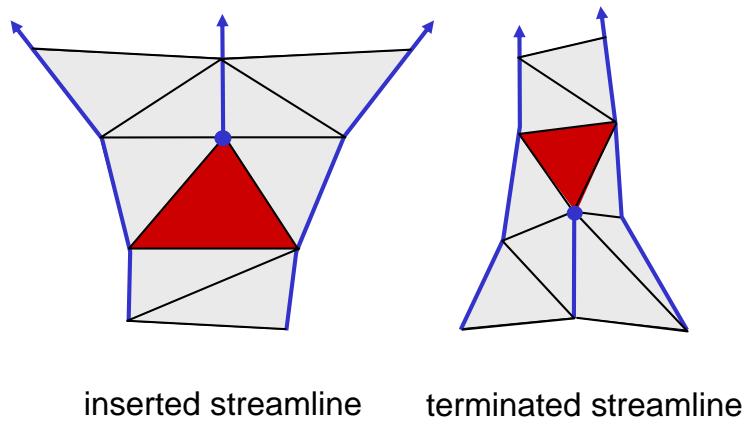
This fails if streamlines diverge or grow at largely different speeds.



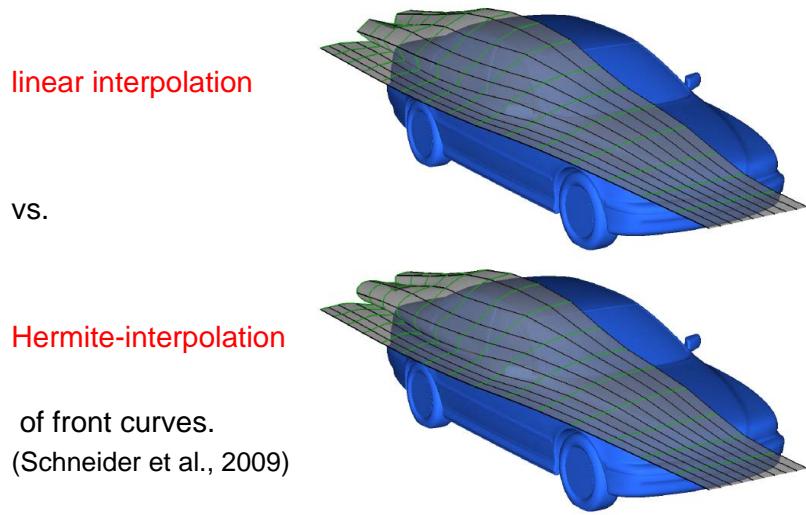
Hultquist's algorithm solves this problem of speed differences by *optimised triangulation*: Of two possible connections choose the one which is closer to orthogonal (sic) to both streamlines.



The problem of divergence or convergence is solved by *inserting* or *terminating streamlines*.



Also use hermite interpolation for streamsurface patches:



5.9 Streak Surfaces

Stream lines in 3D on a fully adaptive triangle mesh.

6 Texture Advection

Motivation: *Dense* visualisation of vector fields without any seed points.

Methods for *static* fields:

LIC - Line integral convolution (Cabral/Leedom 1993)

Methods for *time-dependent* fields:

- LEA - Lagrangian-Eulerian Advection (Jobard et al. 2001)
- IBFV - Image-Based Flow VIs (van Wijk 2002)

Methods for vector fields *on surfaces*

- IBFVS - IBV for Surfaces
- ISA - Image-Space Advection (Laramee 2003)

6.1 Line Integral Convolution

Line integral convolution (LIC) is a family of 10+ variants. The original method by Cabral and Leedom assumes 2D vector fields on rectilinear grids.

Basic idea:

1. Generate a grey level image of random pixels at the desired resolution.
2. Compute forward and backward streamline segments of fixed arc length for all pixels.
3. Sample the random image along the streamline and compute the average (i.e. convolve with a box filter).
4. Use the computed values as the pixels of the output image.
5. Stretch the range of the output image.

LIC images can be combined with a colour coding of a scalar field:

in *HSV* space:

hue: Scalar field

saturation: 1

value: LIC

in *HLS* space:

hue: scalar field

lightness: LIC

saturation: 1

The Fast LIC method (Stalling) is an order of magnitude faster by reusing parts of streamlines where possible. Fast LIC is the basis of most the newer LIC methods.

LIC for unstructured grids (Battke) uses a *procedural* 3D random image. For each triangle a LIC image is computed separately and displayed as a separate texture map. This method can also be used for *vector fields on surfaces*.

LIC method for curvilinear grids (Forsell). Generate a LIC in computational space \mathcal{C} and use it as a texture map for the grid in physical space \mathcal{P} .

Problems of this approach: If parameters lines are not smooth and cell sizes have a large variation the then resulting image might show artefacts.

Animated LIC The LIC of static vector fields can be easily animated to show the relative velocity magnitudes:

- Use samples at constant *time steps*
- Replace the box filter by *sinusodial* filter with exactly one period.
- Shift the kernel *backward* in steps of one sample.

This results in the texture moving forward.

3D LIC LIC can be computed easily in 3D but the result is a 3D image. Rendering options:

Isosurfaces: NO

Sensitive to noise, LIC is a near worst case!

Direct volume rendering: Yes

6.2 Multi-Layer Flow Textures

Multi-Layer LIC-like textures (Carneky 2012) require:

- Multi-Layer screen-space data structure ("illustration buffer")
- Procedural 3D noise texture with constant screen-space frequency
- Anisotropic diffusion to obtain flow aligned texture.

6.3 Lagrangian-Eulerian Advection

Dynamic behaviour can be expressed in either Eulerian or Lagrangian formulation:

Eulerian or *grid-based*:

Fields are given at grid nodes

Lagrangian or *particle-based*:

A set of particles is advected by the velocity field $v(x, t)$, other fields are given at particle positions.

The temporal change of a function $f(x, t)$ (*while following a particle*) is expressed by the *material derivative* (or convective derivative):

$$\frac{Df(x, t)}{Dt} = \frac{\delta f(x, t)}{\delta t} + \Delta f(x, t) \cdot v(x, t)$$

6.3.1 Lagrangian-Eulerian Advection for Vector Field Visualisation

The LEA method for vector field visualisation (Jobard 2011) uses *one particle per cell*: Initialise a white noise texture (same as for LIC).

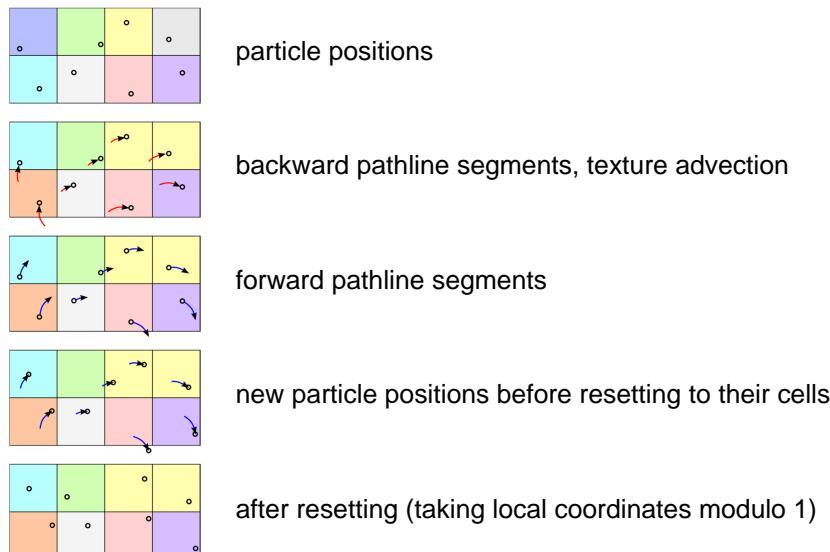
For each time step do:

For each particle do:

- Integrate *backward* pathline segment giving a new *texel value* for the cell.
- Integrate *forward* pathline segment giving the new *particle position* in the same cell (local coordinates modulo 1).

Why backward integration? The texture advection can be done as *forward mapping* (Lagrangian scheme) or *backward mapping* (Eulerian scheme).

With backward mapping being the better choice.



Special choices made by LEA

- 1st order integration
- Simplification: Forward segment = –backward segment
Better: Backward segment = –previous forward segment
- Add buffer cells at grid boundaries:
 - Contain texture but no particles
 - Allow texture advection at inflow boundaries
 - Random texture is refreshed after each time step to avoid artefacts
- Post-Processing: Apply a LIC filter to each image before outputting.

Interpolation Backward mapping scheme allows 2 interpolation choices:

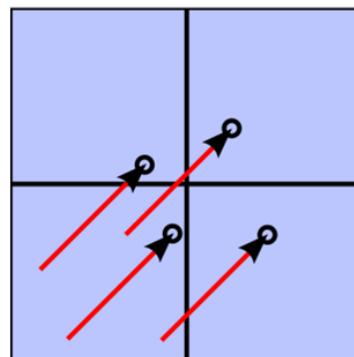
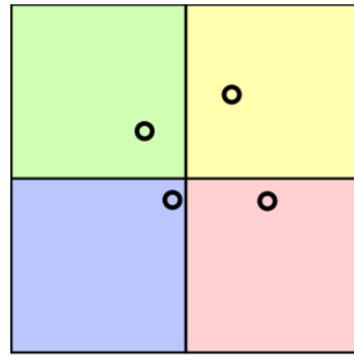
- *nearest-neighbour*
- Bilinear

LEA uses *both*:

- Nearest-neighbour is used for updating the *stored* texture.
- Bilinear interpolation is used for *displayed* texture

Noise injection Backward mapping can have a duplication effect. Causes are:

- Divergence of the vector field.
- Nearest-neighbour interpolation

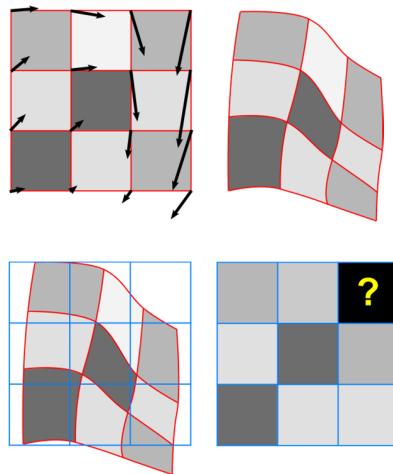


Solution: *Noise injection*: A small percentage of noise is added after each step.
 Trade-off: Keep high frequencies, but also temporal correlation.

6.4 Image-Based Flow Visualisation

IBFV algorithm (van Wijk 2002). Main idea:

- Initialise a noise texture image.
- For each time step do:
 - *Advect nodes* of the texture image resulting in a warped grid.
 - *Render* the *warped grid*, texture mapped.
 - Resample the image to original mesh:
Read back the rendered image to texture memory.
 - Use as next texture image.



Detailed algorithm:

- Initialise a noise texture image.
- For each time step do:
 - Advect nodes of the texture image, resulting in a warped grid.
 - Render the warped grid, texture mapped.
 - *Blend with noise image.*
 - *Apply dye injection.*
 - Resample image to original mesh.
 - Use as next texture image
 - *Draw overlaid graphics*

Noise image

Static results in static image for steady flow

Temporally coherent, using *spot noise* texture.

Boundary areas For the boundary areas a special solution is needed. Simple solution:
Don't clear the screen before redrawing.

Comparison with LEA IBFV is a much faster algorithm than LEA. The coherence is not as good.

6.5 Texture advection in surfaces

Texture advection on surfaces can be used for:

- Boundary flow (wall shear stress)
- Flow on streamsurfaces
- Less meaningful: Project flow on other surfaces (isosurfaces)

Possible but expensive:

- Work in object space
- Use 3D texture

Alternatives:

- Work in image space:
 - IBFV for surfaces (van Wijk)
 - Image-space advection (Laramee)

6.6 IBFV for Surfaces

Idea for IBFVS:

- Use screen coordinates from previous rendering as texture coordinates.
- Advect in object space.
 - I.e. Distort the surface mesh.
- Render the distorted mesh, keeping the texture coordinates
- Apply noise injection and blending
- Overlay the image

6.7 Image-Space Advection

Idea for ISA:

- Project the velocity field to image space.
- Do IBFV within boundary silhouette.
 - I.e. Advect rectangles.
- Apply noise injection and blending.
- Overlay image.

Comparison with IBFVS:

Advantages:

- Projected velocity field simplifies advection
- No computation time is spent for hidden polygons and polygons smaller than a pixel

Problems:

- Artificial continuity across interior silhouettes:
ISA uses edge detection (depth discontinuities)
- The texture is not attached to the surface when the camera is moving.

7 Feature Extraction

Features are inherent properties of data, independent of coordinate frames.
Dimensions of a feature:

- 0: *Point Feature* (often defined by n equations of n coordinates)
- 1: *Line-like feature* ($n-1$ equations)
- 2: *Surface-like feature*
- n : *region-type feature* (typically defined by a single inequality).

7.1 Region-Type Features

A feature is often indicated by high or low values of a *derived field*.
Example: *vortical regions* in a flow field have been defined by:

- Large magnitude of *vorticity*

$$\omega(x) = \Delta \times v(x).$$

- High absolute *helicity*

$$\omega(x) \cdot v(x),$$

or *normalised helicity*:

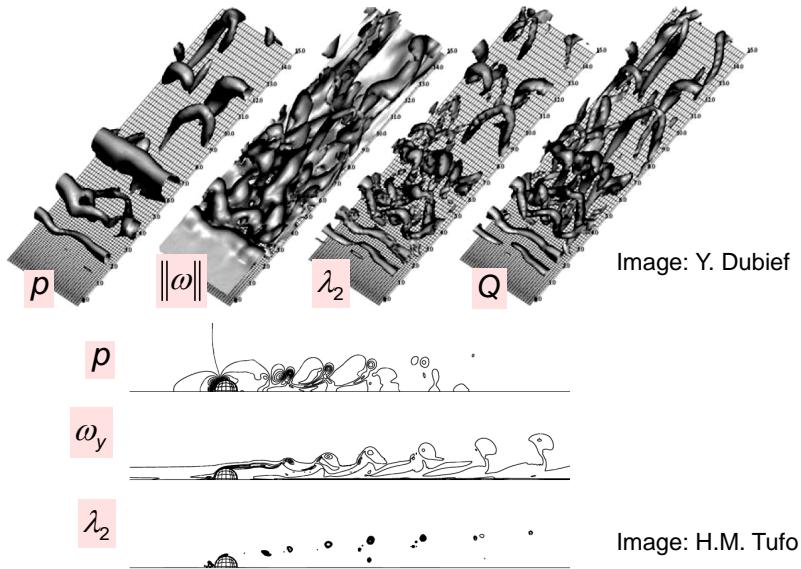
$$\frac{\omega(x)}{\|\omega(x)\|} \cdot \frac{v(x)}{\|v(x)\|}.$$

- Positive *pressure Laplacian*

$$\Delta \cdot \Delta p(x).$$

- Positive *second invariante* Q of the velocity gradient $\Delta v(x)$.
- Two negative *eigenvalues* of

$$\frac{\Delta v(x)^2 + (\Delta v(x)^T)^2}{2}.$$

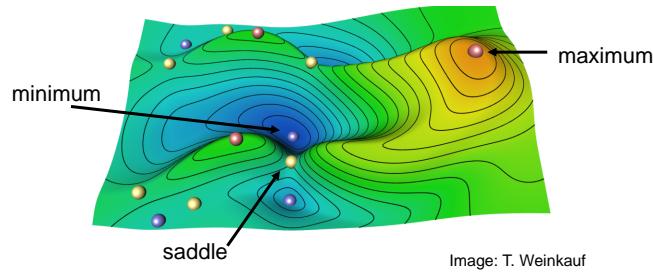


7.2 Point Features in Scalar Fields

Point features in scalar fields:

- Local minima/maxima
- Saddle points

occur at places where the height field is horizontal ($\Delta s(x) = 0$)



These point features are the places where the isoline or isosurface changes its topology when the level is varied.

The *contour tree* or *Reeb graph* describes the *split*, *join*, *creation* and *deletion* events.

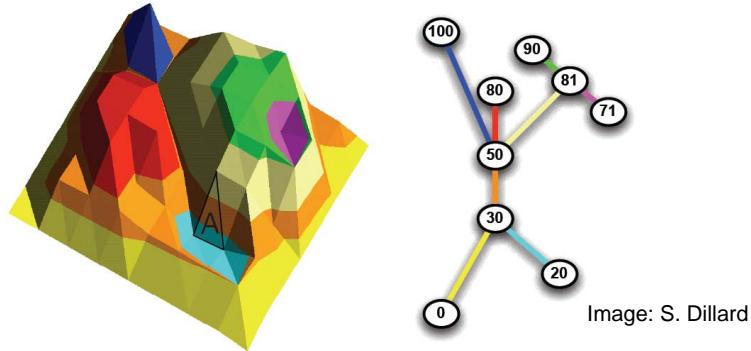
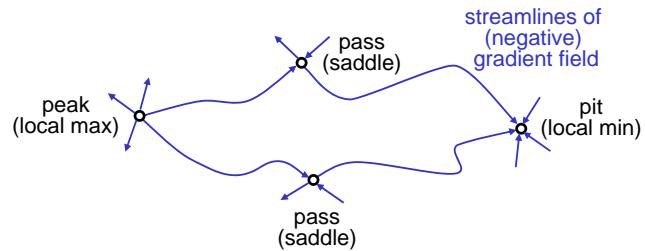


Figure 6: Contour tree

7.3 Line-Like Features in Scalar Fields

Watersheds describe ridges/valleys of a height field $s(x)$: Integrate the *gradient field* $\Delta s(x)$ (backward/forward) starting from *saddle points*.

The watersheds provide a *segmentation* of the domain into so-called *Morse-Smale cells*.



Watersheds require integration and are therefore not locally detectable.
Often used concepts:

- Section-based ridges
- Curvature extrema on height contours.

Height Ridges "Most natural extensions of local maxima to 1D features".

A formal definition of height ridges was given only in the 1990s (Eberly, Lindeberg) based on Haralicks's definition (1983).

At a given point x_0 the scalar field has the Taylor approximation:

$$s(x_0 + x) = s(x_0) + \Delta s \cdot x + x^T H x + \mathcal{O}(|x|^3),$$

where H is the *Hessian* matrix of second derivatives

$$H = \left(\frac{\delta^2 s(x)}{\delta x_i \delta x_j} \right).$$

With H having real eigenvalues and orthogonal eigenvectors. By taking the eigenvectors as the coordinate frame H becomes the diagonal matrix:

$$H = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

A point $x \in \mathbb{R}^n$ is a *local maximum* of $s(x)$ if *for all n* axes if:

- The first derivatives are zero:

$$s_{x_1} = \dots = s_{x_n} = 0$$

- The second derivatives are negative:

$$s_{x_1 x_1}, \dots, s_{x_n x_n} < 0.$$

In the appropriate coordinate frame this generalises to:

A point $x \in \mathbb{R}^n$ is on a d -dimensional *height ridge* of $s(x)$ if *for the first $n - d$* axes:

- The first derivatives are zero:

$$s_{x_1} = \dots = s_{x_n} = 0$$

- The second derivatives are negative:

$$s_{x_1 x_1}, \dots, s_{x_n x_n} < 0.$$

”Appropriate coordinate frame” means that axes are:

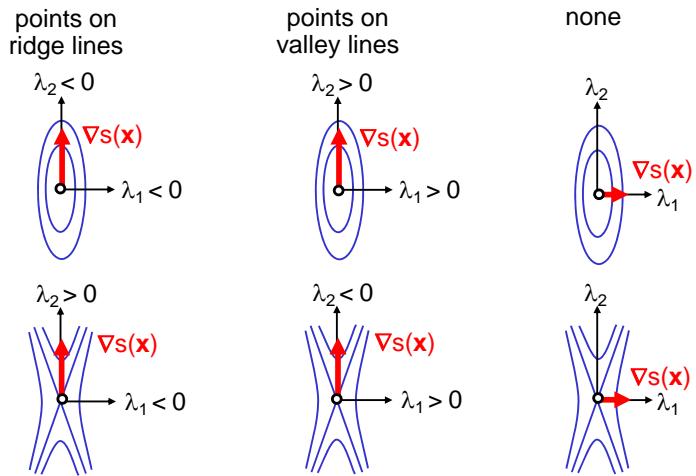
- Aligned with eigenvectors of H ,
- Ordered by absolute eigenvalues

$$|\lambda_1| \geq \dots \geq |\lambda_n|$$

(Lindeberg’s version), or

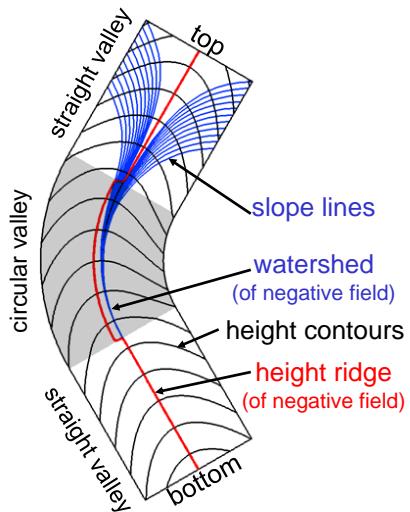
Ordered by signed eigenvalues:

$$\lambda_1 \leq \dots \leq \lambda_n$$



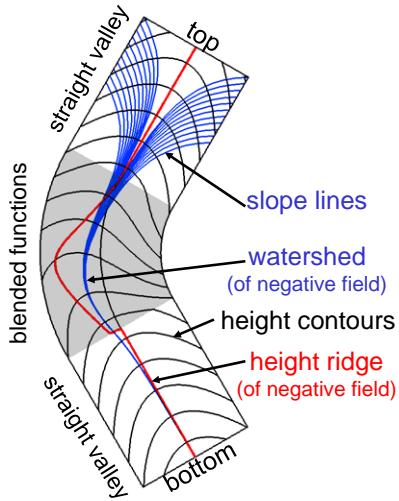
7.3.1 Circular Gutter

Example where slope lines and height ridge don't match. A "counter-example" for height ridges.



7.3.2 Blended Height Fields

By replacing the circular part of the circular gutter by a blend of the two height fields we observe that the watershed deviates (in the lowerpart) from the obvious symmetric valley line. A "counter-example" for watersheds.



7.3.3 Watersheds vs. Height Ridges

The discussion of watersheds vs. height ridges started in computer vision (Koendrick/van Doorn 1993) and is still ongoing.

Watersheds:

- + Are *slope lines* of the height field (=streamlines of the gradient field).
- Depend on boundaries.
- Require existence of a saddle point.

Height ridges in 3D scalar fields can be used for defining/detecting *vortex core lines*. These are

- Height ridges (valley lines) of *pressure* (Kida and Miura).
- Height ridges of *vorticity magnitude* (Ahmad/Kenwright/Strawn).

7.4 Geometric Features of Surfaces

On *surfaces* in 3-space, 0- and 1-dimensional features can be defined by the (differential) geometry alone.

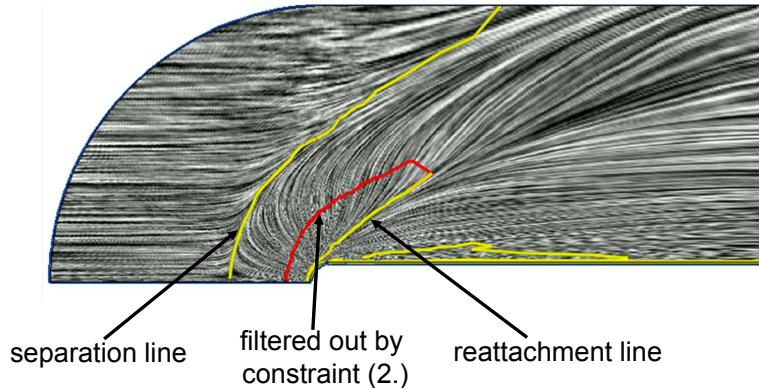
The term "ridge" can refer to either height ridges or curvature ridges. Curvature ridges are not appropriate as features of a scalar field (height field). Reasons:

1. Invariance under rotation (tilting the terrain)
2. No invariance under scaling.

7.5 Line-Like Features in Vector Fields

Separation lines in boundary shear flow (Kenwright): A point x lies on a *separation* or *reattachment line* if

1. At x the field vector $v(x)$ is an eigenvector of $\nabla v(x)$,
2. the corresponding eigenvalue is the one with smaller absolute value.



Vortex core lines (Sujudi/Haimes) A point x lies on a *vortex core line* if

1. At x the field vector $v(x)$ is an eigenvector of $\Delta v(x)$
2. the two other eigenvalues of $\Delta v(x)$ are complex.

Alternative definitions:

- According to Levy et al., longitudinal vortices have a high normalised helicity (or small angles between velocity and vorticity).

Vortex core line criterion: $v(x)$ is (anti-) parallel to $\omega(x)$.

- Singer and Banks' method:

- Find a first point on the core line
- Repeat
 - * Predict the next point along $\omega(x)$
 - * Correct to pressure minimum in normal plane of $\omega(x)$
- Compute vortex hull.

Computing Line-Like Features Instead of computing eigenvectors for height ridges, Sujud-Haines core lines etc. Make use of the following observation:

v is an eigenvector of A iff Av is *parallel* to v .

Recipe:

- Compute $w = Av$ as a *derived field*.
- Find places where v and w are parallel (or one of them is 0).
- Apply constraints
- Apply post-filtering.

Parallel vectors operator: Given v, w : Returns points where v and w are parallel.

Implementation:

2D: $v \times w = 0$ is just a contour line problem.

3D: $v \times w = 0$ is 3 equations for 3 unknowns:

- Equations are linearly dependent
- Can be solved with Marching Cubes like method.

7.6 Tracking of Features

In time-dependent data, features are usually extracted for single time steps. Howto recognise a feature in a different time step? Some methods are:

- Decide on *spatial overlap* (Silver et al.)
 - Appropriate for region-type features
 - Detects motion and events (split, merge, birth, death)
- Decide on *feature attributes* (Reinders)
 - Use attributes such as position, shape (fitted ellipsoid), orientation, spin, data values, etc...
 - Combine with motion prediction.

7.7 Post-Filtering of Features

Specific for ridges:

Height > Threshold

Specific for vortex axes:

Vortex strength > Threshold

General line-type features:

Angle (line, vector) < Threshold

Length (line) > Threshold

8 Vector Field Topology

8.1 Vector Fields as ODEs

What are conditions for existence and uniqueness of streamlines?

For the initial value problem

$$\dot{x}(t) = v(x(t)) \quad x(t_0) = x_0$$

a solution *exists* if the velocity field $v(x)$ is *continuous*.

The solution is *unique* if the field is *Lipschitz*-continuous, i.e. if there is a constant M such that

$$\|v(x) - v(x')\| \leq M\|x - x'\| \quad \forall x, x' \in \mathbb{R}.$$

Lipschitz-continuous is stronger than continuous (C^0) but weaker than continuously differentiable (C^1).

Important for scientific visualisation:

- *Piecewise multilinear* functions are Lipschitz-continuous,
- *cellwise* bi- or trilinear interpolation is Lipschitz-continuous.

Consequence: *Numerical* vector fields do have unique streamlines, but *analytic* vector fields don't necessarily.

8.2 Special Streamlines

It is possible that a streamline $x(t)$ maps two different times t and t' to the same point:

$$x(t) = x(t') = x_1.$$

There are two types of such special streamlines:

Stationary points If $v(x_1) = 0$, the streamline degenerates to a single point

$$x(t) = x_1 \quad (t \in \mathbb{R}).$$

Periodic points If $v(x_1) \neq 0$, the streamline is periodic:

$$x(t + kT) = x(t) \quad (t \in \mathbb{R}, k \in \mathbb{Z}).$$

Regular streamlines can *converge* to stationary points or periodic orbits in either positive or negative time. However (because of this uniqueness) a regular streamline cannot *contain* a stationary point or *periodic orbit*.

8.3 Critical Points

A stationary point x_c is called a *critical point* if the velocity gradient $J = \nabla v(x)$ at x_c is regular (is a non-singular matrix with a nonzero determinant). Near a critical point, the field can be approximated by its linearisation:

$$v(x_c + x) = Jx + \mathcal{O}(x^2).$$

Properties of critical points:

- In a neighbourhood, the field takes all possible directions.
- Critical points are *isolated* (as opposed to general stationary points).

Critical can point can have different *types* depending on the eigenvalues of J .

8.3.1 Hyperbolic Critical Points

A critical point is called *hyperbolic* if all eigenvalues of J have *nonzero real parts*.

The main property of hyperbolic critical points is *structural stability*:

Adding a small perturbation to $v(x)$ does not change the topology of the nearby streamlines.

Hyperbolic critical points in 3D can be classified as follows:

Two real eigenvalues:

- Both positive: *Node source*
- Both negative: *Node sink*
- Different signs: *Saddle*

Two conjugate complex eigenvalues:

- Positive real parts: *Focus source*
- Negative real parts: *Focus sink*

8.4 Critical Points in 2D

In 2D the eigenvalues are the zeros of:

$$x^2 + px + q = 0,$$

where p and q are the two *invariants*:

$$\begin{aligned} p &= -\text{trace}(J) = -(\lambda_1 + \lambda_2), \\ q &= \det(J) = \lambda_1 \lambda_2. \end{aligned}$$

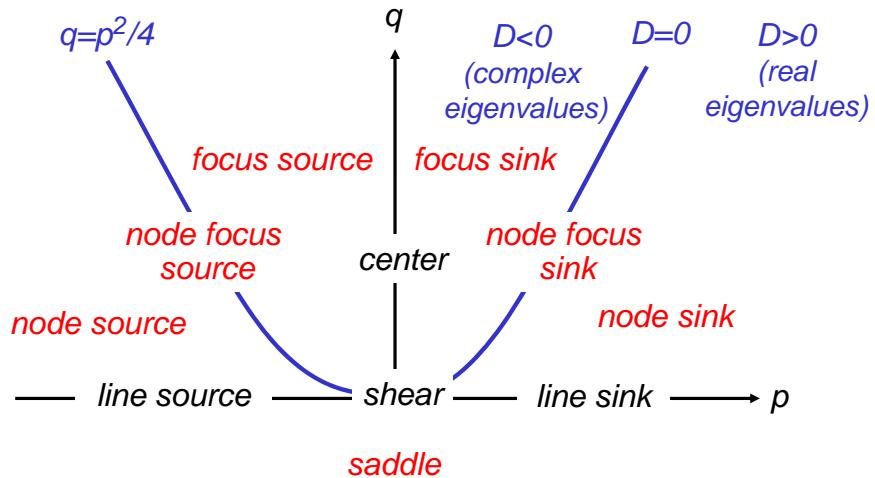
The eigenvalues are complex when the *discriminant*

$$D = p^2 - 4q,$$

is negative.

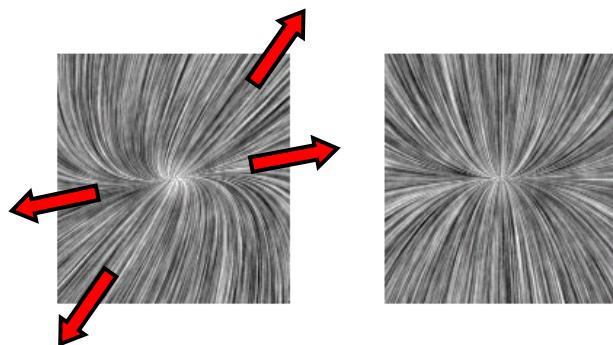
It follows:

- Critical point types depend on signs of p , q and D .
- Hyperbolic points have either $q < 0$, or $q > 0$ and $p \neq 0$.



8.4.1 Node Source

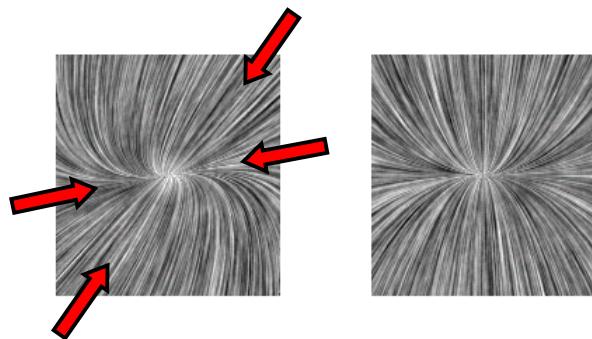
- Positive trace
- Positive determinant
- Positive discriminant



$$J = \begin{pmatrix} 0.425 & 0.431 \\ -0.1 & 1.075 \end{pmatrix} = A^{-1} \begin{pmatrix} 0.5 & 0 \\ 0 & 1 \end{pmatrix} A$$

8.4.2 Node Sink

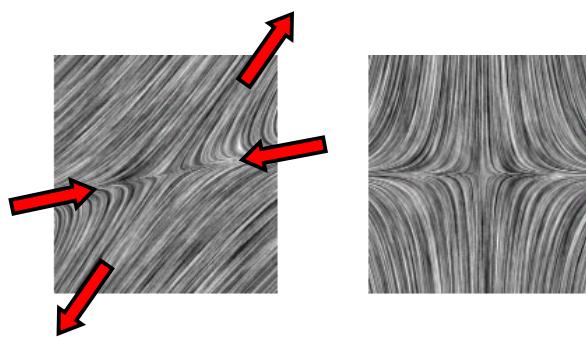
- Negative trace
- Positive determinant
- Positive discriminant



$$J = \begin{pmatrix} -0.425 & -0.431 \\ 0.1 & -1.075 \end{pmatrix} = A^{-1} \begin{pmatrix} -0.5 & 0 \\ 0 & -1 \end{pmatrix} A$$

8.4.3 Saddle

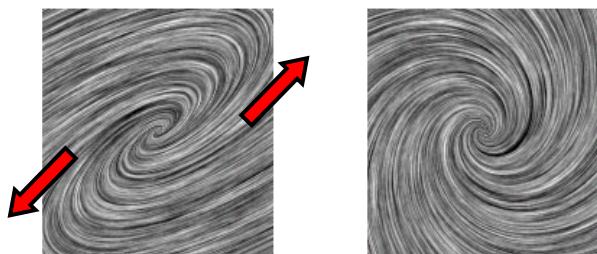
- Any trace
- Negative determinant
- Positive discriminant



$$J = \begin{pmatrix} -0.434 & 1.078 \\ -0.25 & 1.15 \end{pmatrix} = A^{-1} \begin{pmatrix} -0.25 & 0 \\ 0 & 1 \end{pmatrix} A$$

8.4.4 Focus Source

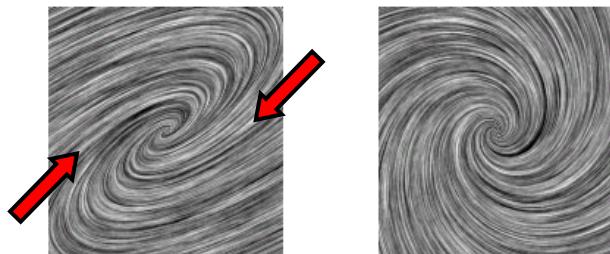
- Positive trace
- Positive determinant
- Negative discriminant



$$J = \begin{pmatrix} 1.48 & -1.89 \\ 1.04 & -0.48 \end{pmatrix} = A^{-1} \begin{pmatrix} 0.5 & -1 \\ 1 & 0.5 \end{pmatrix} A$$

8.4.5 Focus Sink

- Negative trace
- Positive determinant
- Negative discriminant

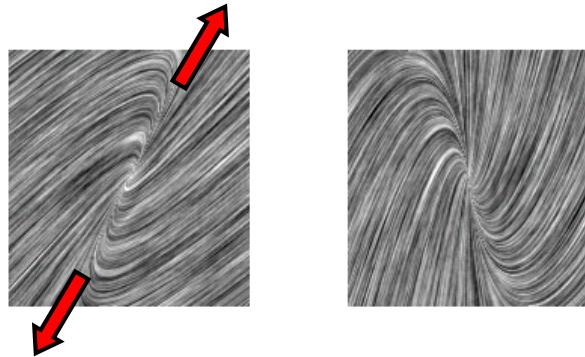


$$J = \begin{pmatrix} -1.48 & 1.89 \\ -1.04 & 0.48 \end{pmatrix} = A^{-1} \begin{pmatrix} -0.5 & 1 \\ -1 & -0.5 \end{pmatrix} A$$

8.4.6 Node Focus Source

- Positive trace
- Positive determinant
- Zero discriminant

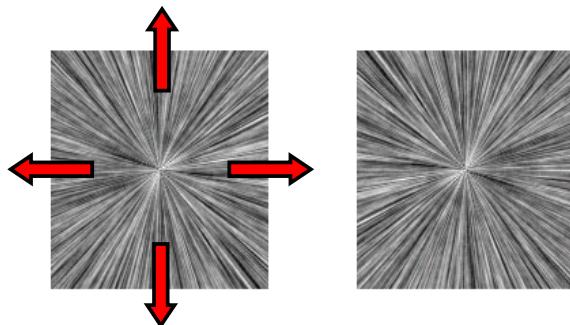
Between node source and focus source (double real eigenvalue).



$$J = \begin{pmatrix} 1.25 & -0.56 \\ 1 & -0.25 \end{pmatrix} = A^{-1} \begin{pmatrix} 0.5 & 0 \\ 1 & 0.5 \end{pmatrix} A$$

8.4.7 Star Source

Special case of the node focus source: Diagonal matrix.



$$J = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

8.5 Nonhyperbolic Critical Points

If the eigenvalues are purely imaginary (i.e. the real parts are zero), then the critical point is the boundary case between focus source and focus sink.

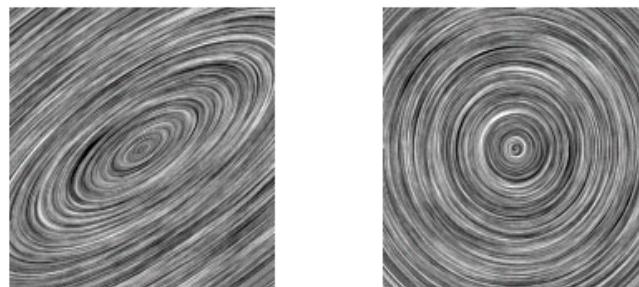
This type of critical point is called a *center*. Depending on the higher derivatives it can behave as a source or as a sink. Because a center is nonhyperbolic, it is not structurally stable in general:



but in the case of a *divergence-free* field it is structurally stable.

8.5.1 Center

- Zero trace
- Positive determinant
- Negative discriminant



$$J = \begin{pmatrix} 0.98 & -1.885 \\ 1.04 & -0.98 \end{pmatrix} = A^{-1} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} A$$

8.6 Other Stationary Points

If J is a singular matrix the following stationary (but not critical!) points are possible:

- A single eigenvalue is zero:
Line source, *Line sink*.
- If both eigenvalues are zero:
Pure Shear.

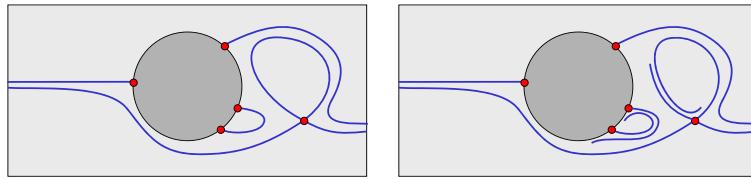
8.7 The Topological Skeleton

The *topological skeleton* consists of all periodic orbits and all streamlines converging (in either direction of time) to

- A saddle point (*separatrix* of the saddle),
- A critical point on a no-slip boundary.

It provides a kind of *segmentation* of the 2D vector field.

Example:



Irrational Vector Fields An irrotational (conservative) vector field is the gradient of a scalar field (its potential). The skeleton of an irrotational vector field: Watershed image of its potential field.

Discussion:

- Watersheds are topologically defined and an integration is required to compute them.
- Height ridges are geometrically defined and locally detectable.

8.8 Critical Points in 3D

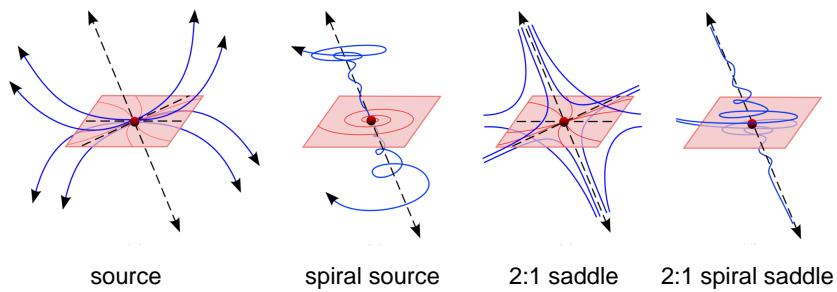
Hyperbolic critical points in 3D can be classified as follows:

Three real eigenvalues:

- All positive: *Source*
- Two positive, one negative: *1 : 2 saddle* (1 in, 2 out)
- One positive, two negative: *2 : 1 saddle* (2 in, 1 out)
- All negative: *Sink*

One real, two complex eigenvalues:

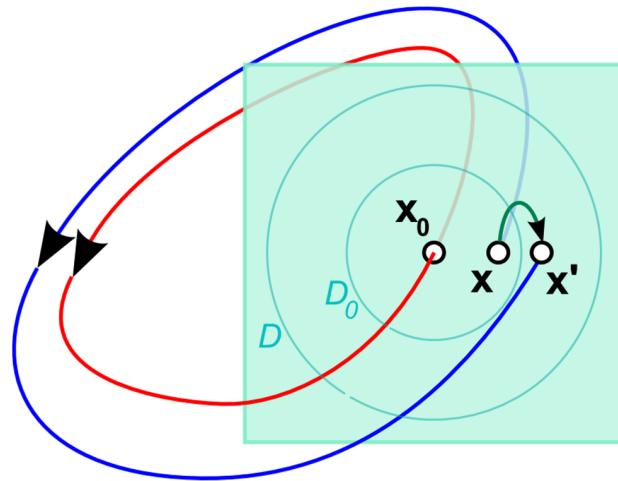
- Positive real eigenvalue, positive real parts: *Spiral source*
- Positive real eigenvalue, negative real parts: *2 : 1 Spiral saddle*
- Negative real eigenvalue, positive real parts: *2 : 1 Spiral saddle*
- Negative real eigenvalue, negative real parts: *Spiral sink*



8.9 Periodic Orbits

The *Poincaré map* of a periodic orbit in 3D:

- Choose a point x_0 on the periodic orbit.
- Choose an open circular disk D centered at x_0 on a plane which is not tangential that is *not tangential* to the flow and small enough that the periodic orbit intersects D only in x_0 .
- Any streamline seeded at a point $x \in D$ intersects D the next time at a point $x' \in D$ defines a mapping from x to x' .
- There exists a smaller open disk $D_0 \subseteq D$ centered at x_0 such that this mapping is defined for all points $x \in D_0$.
- This is the Poincaré map.



Using the coordinates on the plane of D and with origin at x_0 the Poincaré map can now be linearised:

$$x \mapsto Px$$

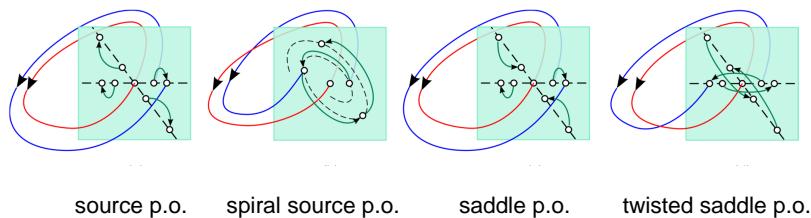
where P is a 2×2 matrix.

Important facts about Poincaré maps: The eigenvalues of P are independent of

- The choice of x_0 on the periodic orbit,
- The orientation of the plane of D ,
- The choice of coordinates for the plane.

A periodic orbit is called *hyperbolic* if its eigenvalues *lie off the complex unit circle*. Hyperbolic periodic orbits are structurally stable.

Types of hyperbolic orbits in 3D:



8.10 Saddle Connectors

The topological skeleton of 3D vector fields contains 1D and 2D separatrices of (spiral) saddles.

This is not directly usable for visualisation (too much occlusion). Alternative: Only show *intersection curves* of 2D separatrices.

Two types of saddle connectors:

Heteroclinic orbits: Connects two (spiral) saddles

Homoclinic orbits: Connects a (spiral) saddle with itself.

In rotational flow, a connected pair of spiral saddles can describe a *vortex breakdown bubble*.

Ideal case:

- $W_S(P_1)$ coincides with $W_u(P_2)$
- No saddle connector

Perturbed case:

- Transversal intersection of $W_S(P_1)$ and $W_U(P_2)$

- Saddle connector consists of two streamlines

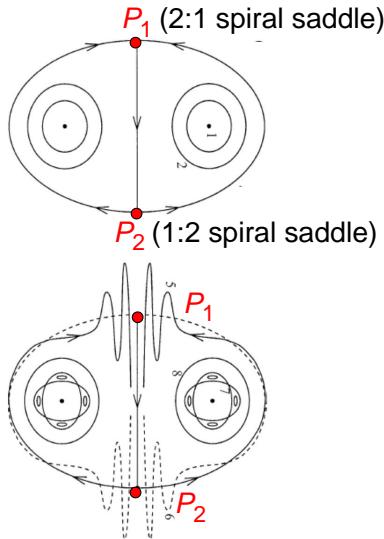


Image: Krasny/Nitsche

Figure 7: Top: Ideal case, Bottom: Perturbed case

Pair of spiral saddles, 3D view:

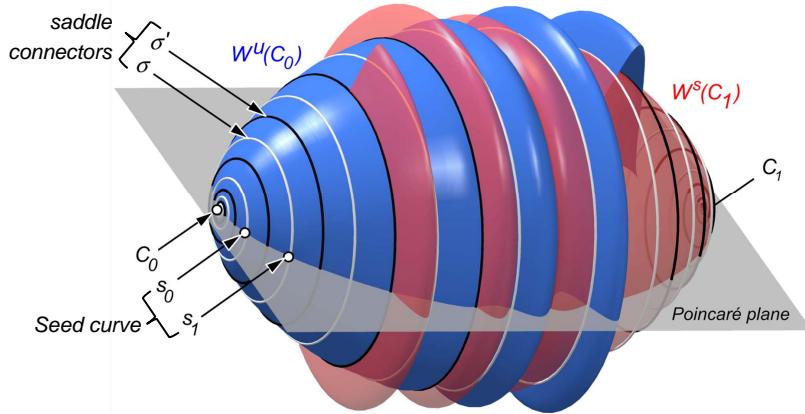


Figure 8: Perturbed saddle connector connecting C_0 and C_1

If v is the velocity field of a fluid: Folds must have a constant mass flux. Close to P_1 or P_2 this is approximately

$$\int \rho v \, dn \approx \rho \omega Ar \quad (\text{density} \cdot \text{angular velocity} \cdot \text{cross section area} \cdot \text{radius})$$

9 Visualisation of Lagrangian Coherent Structures

Motivation: Vector field topology does not well describe the topology of a "strongly" time-dependent vector field.

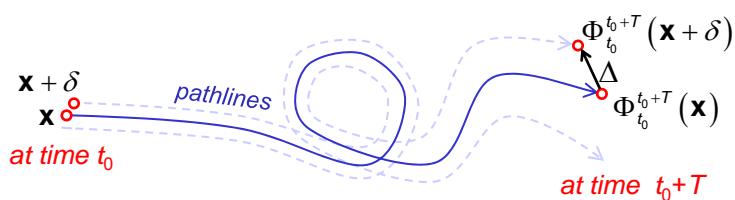
Separatrices are defined in terms of streamlines, not pathlines.

Critical points of a saddle types are not the places where flow separation happens.

9.1 The Finite-Time Lyapunov Exponent

The FTLE describes the amount of separation (stretching) after a finite advection time T .

Principle:



Definition:

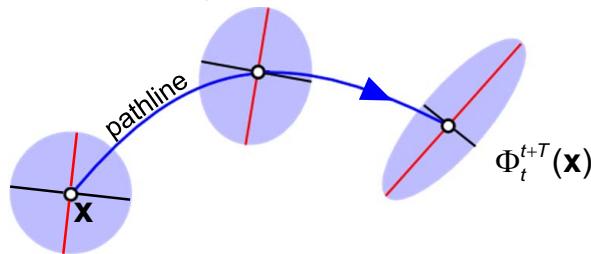
$$\text{FTLE}(x, t_0, T) = \lim_{\delta \rightarrow 0} \max_{\text{direction of } \delta} \frac{1}{|T|} \log \left(\frac{\|\Delta\|}{\|\delta\|} \right) = \frac{1}{|T|} \log \left(\sqrt{\lambda_{\max}(F^* F)} \right)$$

with $F = \nabla \Phi_{t_0}^{t_0+T}(x)$ and $F^* F$ being the Cauchy-Green deformation tensor.

Application

1. Compute the flow map:

$$x \mapsto \Phi_t^{t+T}(x) \quad (\text{for fixed times } t, t+T)$$



2. Compute Cauchy-Green tensor:

$$C = F^* F$$

$$F = \nabla \Phi_t^{t+T}(x)$$

3. FTLE:

$$\sigma = \frac{1}{|T|} \log \sqrt{\lambda_1(C)}$$

9.2 Separatrices

Separatrix: A streamline converging to a saddle point (interesting because it acts as a flow barrier).

- Convergence in positive time: *Attracting separatrix*.
- Convergence in negative time: *Repelling separatrix*.

9.2.1 LCS in 3D

Visualise using *direct volume rendering*. Use a transfer function which maps high FTLE values to high opacity: Approximates ridges.

LCSs have replaced VFT in feature-based visualisation of *unsteady* flow. Two approaches are commonly used:

- Geometric extraction of ridges
- Visualisation focusing on high FTLE values:
 - Isosurfaces
 - DVR

Advantages of *geometric* ridges:

- Follows Haller's definition more strictly.
- Allows for quantitative checking of cross-flux.

9.3 Computation of FTLE Ridges

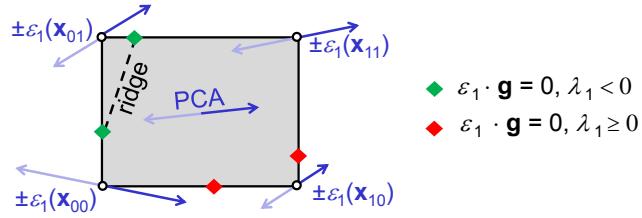
Efficient computation of height ridges (of a scalar field $s(x)$ in n -space):

- Compute derived fields $g = \nabla s$ and $H = \nabla g$
- For ridges of *dimension 1* use the *Parallel Vectors method*:
 - Find the places where g and Hg are parallel vectors.
 - Compute the eigenvalues of H , test if those in the directions of $\perp g$ are all negative and less than the one in direction g
- For ridges of *co-dimension 1* (i.e. of dimension $n-1$) use *Marching Ridges* method (Furst et al. 2001):
 - Compute eigenvalues of H : $\lambda_1 \leq \dots \leq \lambda_n$

- ε_1 : Eigenvector for λ_1 ($\varepsilon_1 \perp$ ridge)
- Solve for $\varepsilon_1 \cdot g = 0$ (single scalar equation!)

Problem: ε_1 is not a vector field (ambiguous directions). Marching Ridges does the following per cell:

- Orient ε_1 at nodes of the cell by PCA.
- Evaluate $\varepsilon_1 \cdot g$ at node
- Interpolate zero crossings on edges
- Use zero crossings with $\lambda_1 < 0$
- Generate triangles for *Marching Cubes* case:

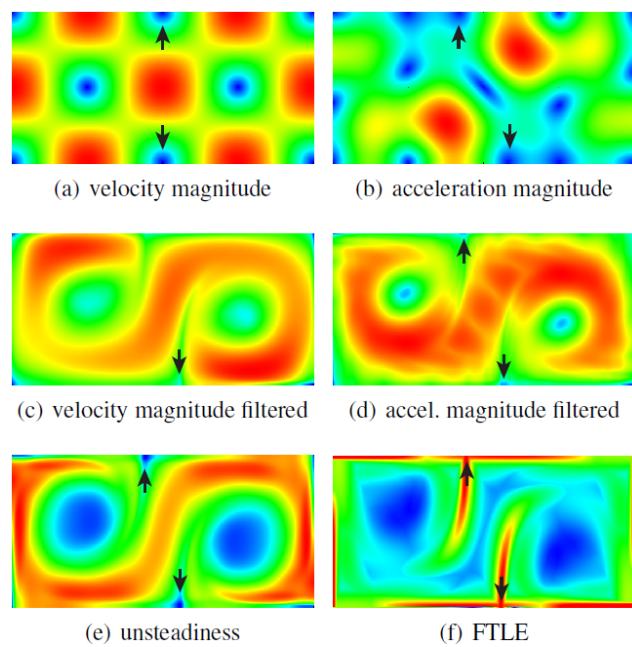


9.4 Lagrangian Vector Field Topology

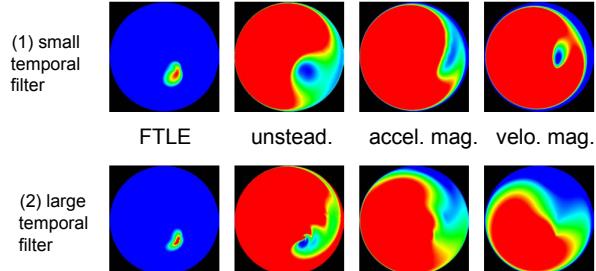
Alternatives for critical points:

- FTLE maxima
- Acceleration minima [Kasten 2009]
- Unsteadiness minima [Fuchs 2010]

9.4.1 Double Gyre Example



9.4.2 Petri Dish Example



9.4.3 Vortex Street Example

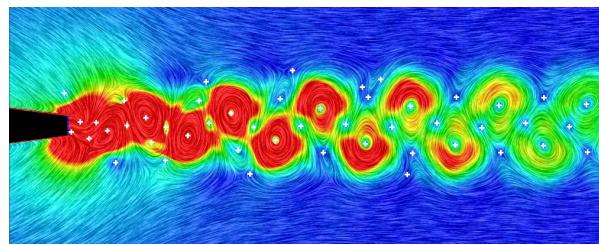


Figure 9: Acceleration minima, temporally filtered

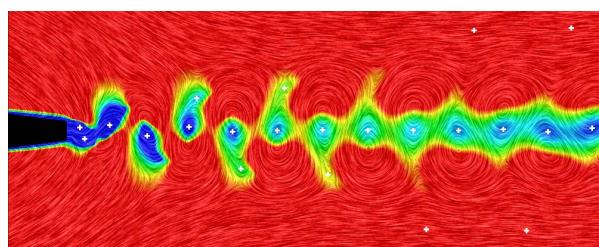


Figure 10: Velocity minima, temporally filtered

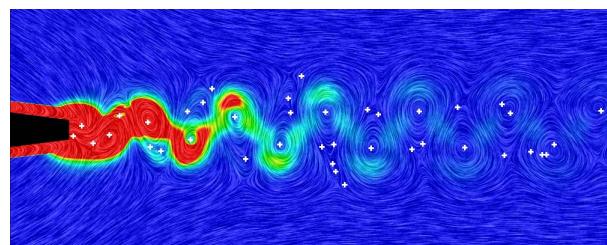


Figure 11: Unsteadiness minima, temporally filtered

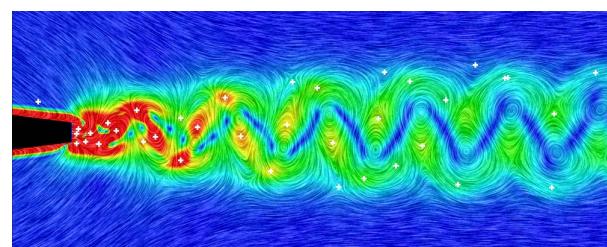


Figure 12: FTLE maxima

Comparison

	Double Gyre	Petri Dish	Vortex Street
FTLE max	(+)	+	-
Acceleration min	+	-	+
Unsteadiness min	+	+	(+)

10 Tensor Field Visualisation

10.1 Tensors

H. Hagen: Tensors are the language of mechanics.

Tensors of *order* (rank):

0 : Scalar

1 : Vector

2 : Matrix

Tensors can have "lower" and "upper" indices indicating different transformation rules for change of coordinates.

Tensor field visualisation almost always deals with *2nd order tensors*. Eigenvectors and eigenvalues contain full information.

Separate visualisation methods for *symmetric* and *nonsymmetric* tensors. Visualisation methods for tensor fields include:

- Tensor glyphs
- Tensor field lines, hyperstreamlines
- Tensor field topology
- Fiber bundle tracking.

10.2 Tensor Glyphs

In 3D, tensors are 3×3 matrices. The *velocity gradient tensor* is non-symmetric: 9 degrees of freedom for the local change of the velocity vector.

A *glyph* developed by de Leeuw and van Wijk can visualise all these 9 DOFs:

- Tangential acceleration: Green "membrane"
- Orthogonal acceleration: Curvature of arrow
- Twist: Candy stripes
- Shear: Orange ellipse
- Convergence/divergence: White "parabolic reflector"

Symmetric 3D tensors Real eigenvalues, orthogonal eigenvectors. If *positive definite* the data can be represented by *ellipsoids*.

Three types of anisotropy:

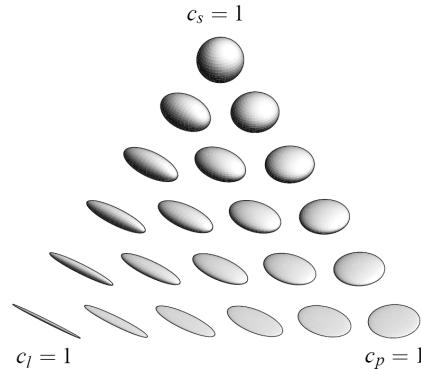
- Linear anisotropy
- Planar anisotropy
- Isotropy (spherical)

Anisotropy measure:

$$c_1 = \frac{\lambda_1 - \lambda_2}{\lambda_1 + \lambda_2 + \lambda_3}$$

$$c_p = 2 \frac{\lambda_2 - \lambda_3}{\lambda_1 + \lambda_2 + \lambda_3}$$

$$c_s = 3 \frac{\lambda_3}{\lambda_1 + \lambda_2 + \lambda_3}$$



Problem of *ellipsoid* glyphs: The shape is poorly recognised in projected views.

Problem of *cuboid* glyphs: Small differences in eigenvalues are overemphasised.

Problem of *cylinder* glyphs: Discontinuity at $c_l = c_p$ and artificial orientation at $c_s = 1$.

Superquadrics Combining advantages.

Superquadrics with z as primary axis:

$$q_z(\theta, \varphi) = \begin{pmatrix} \cos^\alpha \theta \sin^\beta \varphi \\ \sin^\alpha \theta \sin^\beta \varphi \\ \cos^\beta \varphi \end{pmatrix}$$

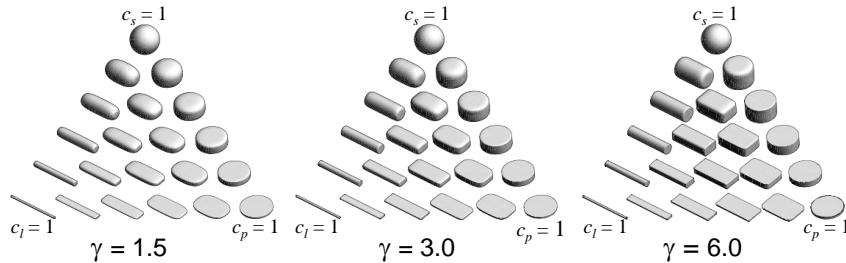
with $\cos^\alpha \theta$ used as shorthand for $|\cos \theta|^\alpha \operatorname{sgn}(\cos \theta)$.

Superquadric Glyphs (G. Kindlmann): Given c_l , c_p and c_s .

- Compute a base superquadric using a *sharpness* value γ :

$$q(\theta, \varphi) = \begin{cases} c_l \geq c_p : & q_z(\theta, \varphi) \text{ with } \alpha = (1 - c_p)^\gamma \text{ and } \beta = (1 - c_l)^\gamma \\ c_l < c_p : & q_x(\theta, \varphi) \text{ with } \alpha = (1 - c_l)^\gamma \text{ and } \beta = (1 - c_p)^\gamma \end{cases}$$

- Rotate into eigenvector frame and scale with λ_1 , λ_2 and λ_3 .



10.3 Tensor Field Lines

Let $T(x)$ be a 2^{nd} order symmetric tensor field (real eigenvalues, orthogonal eigenvectors).

Tensor field line Compute by integrating along one of the eigenvectors. Important: Eigenvector fields are *not* vector fields:

- Eigenvectors have *no magnitude* and *no orientation*
- The *choice* of the eigenvector (minor, major) can be made consistently only as long as all eigenvalues are all different.
- Tensor field lines can *intersect* (only) at points where two or more eigenvalues are equal, so-called *degenerate points*.

Tensor field lines can be rendered as *hyperstreamlines*: Tubes with an elliptic cross section and a radius proportional to the 2^{nd} and 3^{rd} eigenvalue.

Based on tensor field lines, a *tensor field topology* can be defined in analogy to the vector field topology.

Degenerate points play the role of critical points. At degenerate points infinitely many directions (of eigenvectors) exist.

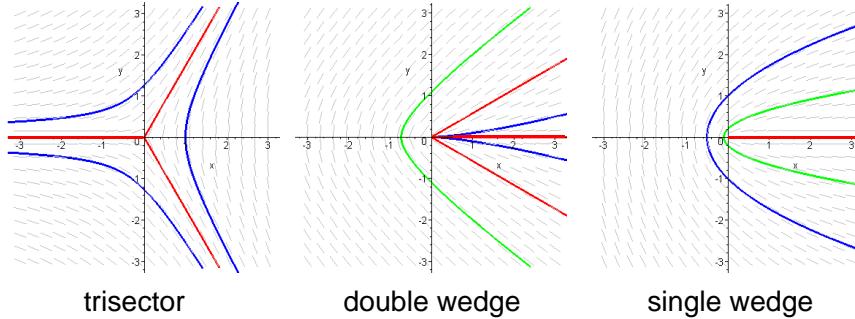
For simplicity we only study the *2D* case. A 2×2 tensor at a degenerate point has (in the given coordinate frame!) the form

$$T = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} = \lambda I.$$

Hence, degenerate points are found by solving the equations

$$\begin{aligned} T_{11}(x) - T_{22}(x) &= 0 \\ T_{12}(x) &= 0 \end{aligned}$$

Types of degenerate points illustrated with linear tensor fields:



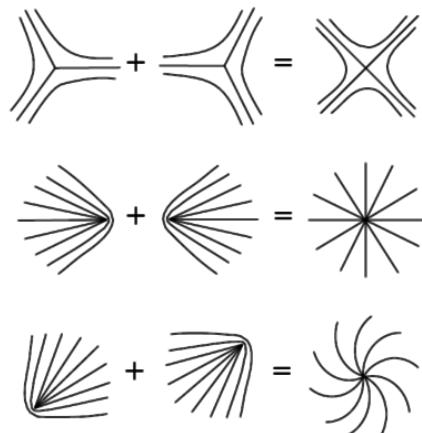
Separatrices are tensor field lines converging to the degenerate point with a radial tangent.

They are straight lines in the special case of a linear tensor field.

Double Wedges Have one "hidden separatrix" and two other separatrices which actually separate regions of different field line behaviour.

Single Wedges Have just one separatrix.

Saddles, nodes and foci can exist as non-elementary (higher-order) degenerate points. They are created by merging trisectors or wedges. They are not structurally stable and break up in their elements if perturbed.



Topological Skeleton The *topological skeleton* is defined as the set of separatrices of trisector points.

10.4 Diffusion Tensor Fiber Bundle Tracking

DTI is a newer magnetic resonance imaging (MRI) technique. DTI produces a tensor field of the anisotropy of the brain's white matter.

Most important application: *Tracking fiber bundles*.

Interpretation of anisotropy types:

- Isotropy: No white matter
- Linear anisotropy: Direction of the fiber bundle
- Planar anisotropy: Can have different meanings!

Fiber bundle tracking \neq tensor field line integration because bundles may *cross* each other.

Method 1 *Best neighbour* algorithm (Poupon), based on the idea of restricting the curvature:

- At each voxel compute the eigenvector of the dominant eigenvalue.
Get a "direction map".
- At each voxel M find the "best neighbour voxel" P according to angle criterion.
Get a "tracking map".
- Connect voxels (within a "white matter mask") with its best neighbour.

Method 2 Apply *moving least squares* filter which favors the current direction of the fiber bundle (Zhukov and Barr).

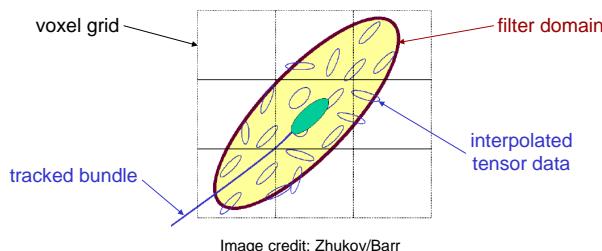


Image credit: Zhukov/Barr

Method 3 *Tensor deflection* (TEND, Lazar et al.)

Idea: If v is the incoming bundle direction use Tv as the direction of the next step.

Reasoning:

- Tv bends the curve towards the dominant eigenvector.

- Tv has the unchanged direction of v if v is
 - An eigenvector of T ,
 - Or a vector within the eigenvector plane if the two dominant eigenvalues are equal (rotationally symmetric T).

10.4.1 Algorithmic Steps

Clustering of fibers: The goal is to identify nerve tracts.

1. *Clustering* based on geometric attributes (centroid, variance, curvature).
Possible method: n -dimensional mass-spring model.
2. *Center line*: Take sets of "corresponding vertices" of each bundle and average them.
3. *Wrapping surface*: Compute the convex hull in orthogonal slices using Graham's scan algorithm.

11 Illustrative Visualisation

Computer supported interactive and expressive visualisations through abstractions as in traditional illustrations.

Illustrative visualisation uses server *non-photorealistic rendering* (NPR) techniques:

- Smart visibility
 - Silhouettes
 - Hatching (shading based on stroke density)
 - Tone Shading
 - Focus and Context techniques
- Context-reserving volume rendering.

11.1 Smart Visibility

- Cut-aways
- Ghosted views
- Section views
- Exploded views
- Browsing Deformations:
 - Leaffer
 - Peeler

11.2 Silhouette Algorithms

The *silhouette* of a surface consists of these points where the view vector V and surface normal N are orthogonal.

Silhouettes can be either *outlines* or *internal silhouettes*.

In contrast to other important feature lines such as curvature ridges/valleys and texture boundaries, silhouettes are view-dependent.

Object Space Algorithms exist for:

Polygonal surfaces

For each polygon

- Set Front-Facing flag to all edges if $N \cdot V \geq 0$
- Set Back-Facing flag to all edges if $N \cdot V < 0$

for each edge draw if both flags are set.

Implicit surfaces

NURBS surfaces

Image Space Algorithms exist for:

Polygonal Surfaces:

- Render polygons with depth buffer enabled
- Look for discontinuities in depth buffer:
 - Compute depth difference between two adjacent pixels, or the Laplacian on a 3×3 stencil.
 - If larger than the threshold, draw a silhouette pixel.

Volume Data (Ebert and Rheingans):

Idea: "Silhouette points" are where the gradient is orthogonal to the view vector.
Use an opacity transfer function depending on $|\nabla s \cdot V|$.

11.2.1 Hatching

Volume illustration with hatching:

- Compute an isosurface
- Compute *curvature fields* (1^{st} and 2^{nd} principal curvature directions on the isosurface), fast algorithm by Monga et al.
- Compute hatching as streamlines of both curvature fields using *streamline placement techniques*.

11.3 Illuminated Streamlines

Phong's local lighting model

$$I = I_{\text{ambient}} + I_{\text{diffuse}} + I_{\text{specular}} = k_a + k_d L \cdot N + k_s (V \cdot R)^n$$

requires a *normal vector* which for a curve in space is underspecified.

11.3.1 Maximum Principle

[Banks]

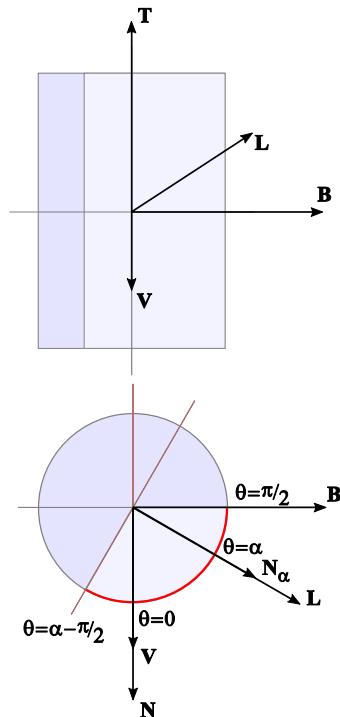
- Diffuse reflection is computed using the rotation angle which gives a maximum value.
- Specular reflection is computed using the rotation angle which gives a maximum value (a different angle in general!).
- Implemented with OpenGL texture mapping.

11.3.2 Cylinder Averaging

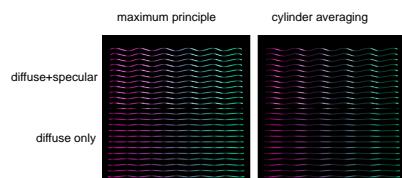
[Schüssmann]

Diffuse and specular reflection are computed as average over the part of a cylinder which is both *visible* and *lit*.

Implemented as a fragment shader.



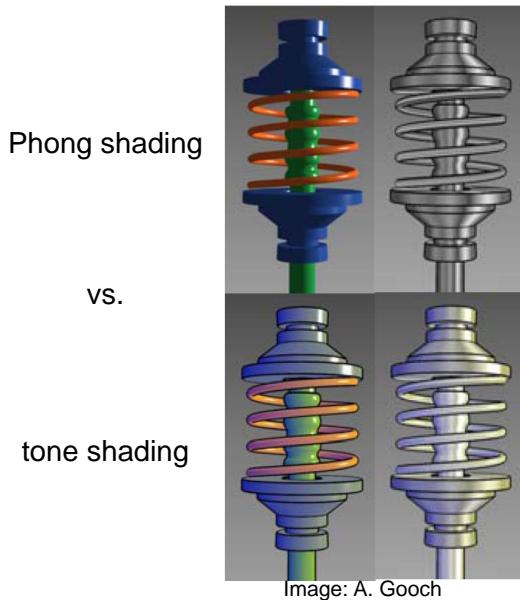
Example: Horizontal sine curves vertically stacked:



The comparison shows: Maximum principle works well only for specular diffusion.

11.4 Tone Shading

Tone shading or "toon shading" uses *tones* instead of *luminance* for shading.



- Use Warm to cool hue shift
- Depth cue: Warm colours advance while cool colours recede.

11.5 Context-Preserving Volume Rendering

Ghosted view: Surface transparency depends on the *grazing angle* (angle between view ray and surface).

More transparent for a large, more opaque for a small grazing angle.

Context-preserving volume rendering (Bruckner): Use of ghosted views in Volume rendering:

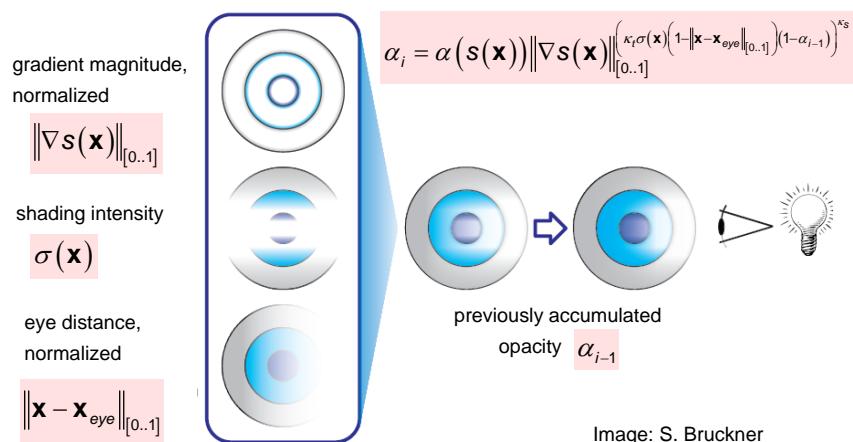


Image: S. Bruckner

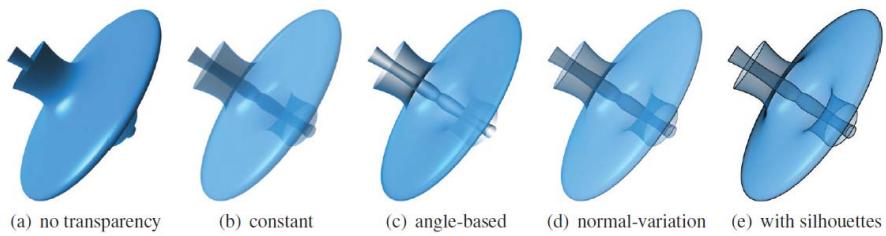
Parameters:

k_t corresponds roughly to the depth of a clipping plane.

k_s controls the sharpness of the transition between visible and clipped.

11.6 Illustrative Visualisation of (Flow) Surfaces

Transparency methods for surface rendering:



Angle based: $\alpha_{\text{view}} = \frac{2}{\pi} \arccos(N \cdot V)$

Normal variation: $\alpha_{\text{view}} = \left(\left(\frac{\delta N_z}{\delta i} \right)^2 + \left(\frac{\delta N_z}{\delta j} \right)^2 \right)^{\gamma/2}$

Normal variation approximates curvature and makes thinnest tubes most opaque.
Illustrators have numerous techniques for rendering of surfaces and textures.

12 Information Visualisation

Techniques for *high-dimensional* data:

- Scatter plots, PCA
- Parallel coordinates
- Link+Brush
- Pixel-oriented techniques
- Icon-based techniques

Techniques for *hierarchical* data and *networks*

- Trees: Treemaps
- Graph clustering
- Distortion, Focus and Context

”Dimension” often refers to *data channels* (attributes) and not to true *spatial dimension* (coordinates).

Roles of data and coordinates can be swapped: In the *data domain* the data becomes coordinates and vice versa.

12.1 Scatter Plots

Scatter plots or multi-dimensional histograms are projections of the attribute space to a 2D (or sometimes 3D) subspace.

Point density in scatter points is difficult to perceive. Colour coding of the density improves this somewhat especially in the high-density regions.

Continuous scatter plots (Bachthaler) work also for low densities if the data has a spatial domain!

Idea: Modify the *projected tetrahedra* algorithm. Instead of view projection transform to the data domain.

Density at the thick vertex: $\frac{1}{3}$ Volume/Area_{data domain}.

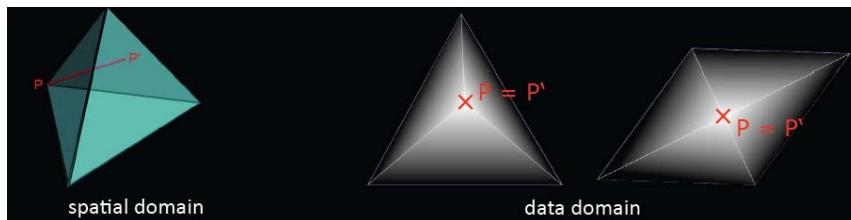


Image: S. Bachthaler

For $n > 3$ a single scatter plot cannot be used. n -dimensional data leads to a $n \times n$ **matrix** of 2D scatter plots.

For small n the matrix can directly serve as a visualisation:

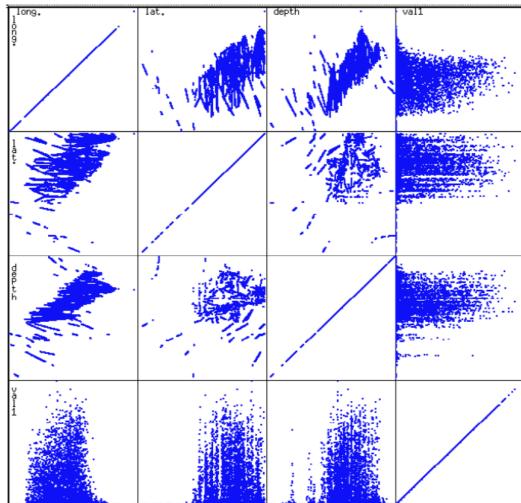


Image: M. Ward

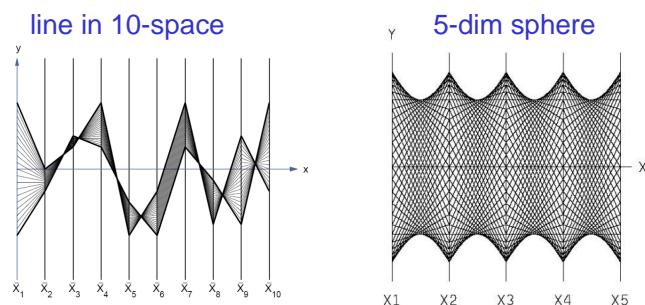
12.2 Dimension Reduction

See CIL.

12.3 Parallel Coordinates

Visualisation method of **parallel coordinates** (Inselberg 1985):

- n parallel and equidistant axes (one per attribute)
- Axes scaled to $[\min, \max]$ range of corresponding attribute.
- Every data item is represented by a polyline which intersects each of the axes at the point corresponding to its attribute.



Linear or spherical arrangement can be "seen" (Inselberg). Algorithm for testing if a point is in the convex hull of a set of points: Check if the polyline is within the two envelopes of the set of polylines.

12.4 Pixel-Oriented Techniques

12.4.1 Space-Filling Curves

Represent each record by a single pixel. Map one attribute to a colour, map sorting key to space-filling curve.

- Peano-Hilbert
- Z-Curve (Morton)

12.4.2 Spiral Technique

The spiral technique (Keim) for query dependent visualisation:

- Sort records (near a query point) by distance to query.
- Map sorted list to spiral.

12.4.3 Axes Technique

For query dependent visualisation (Keim):

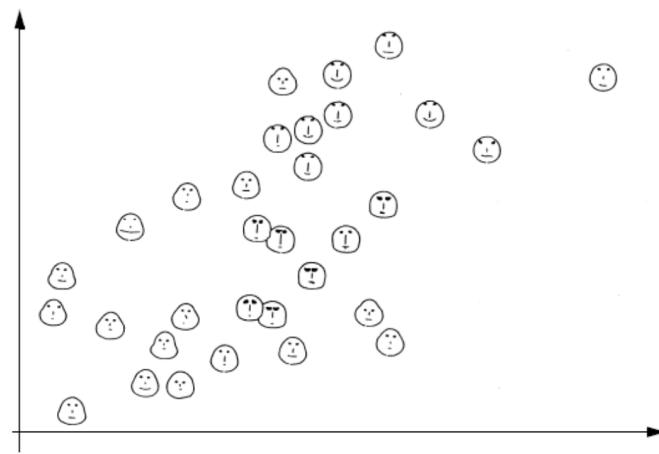
- For two selected attributes separate space into lower/higher attribute values.
- Draw spirals per quadrant.

12.5 Icon-Based Techniques

12.5.1 Chernoff Faces

- Two attributes are mapped to the display axes
- Remaining attributes are mapped to shape and size of hair, eyebrows, eyes, nose, mouth, etc...

Idea: Use the human ability to recognise and memorise faces.



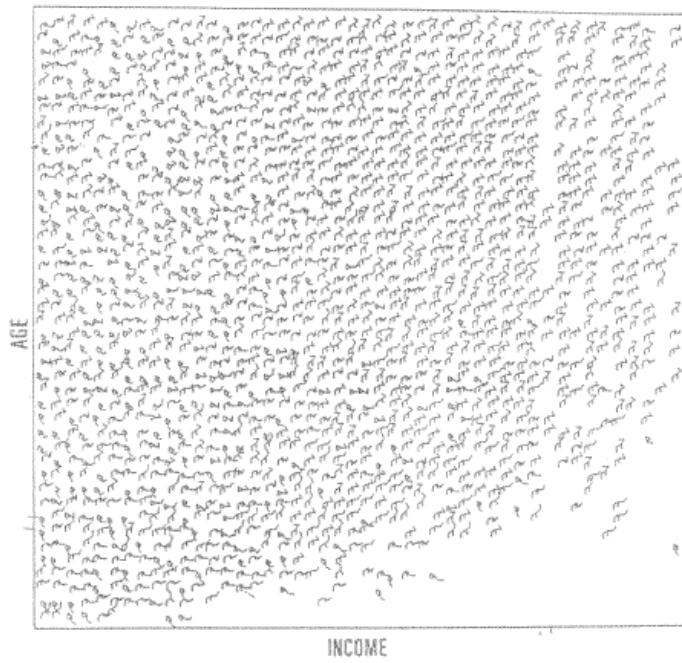
12.5.2 Stick Figures

(Grinstein)

- Two attributes are mapped to the display axes.
- The remaining attributes are mapped to lengths of limbs or angles between them.

Idea: Texture pattern in visualisation shows certain characteristics.

Example: Census data (age, income, sex, education, etc...)



It can be observed that the structure is more homogenous for higher incomes than for lower ones.

12.6 Hierarchical and Network Data

Mathematical description of hierarchies and networks: Graphs.

Some important special types of graphs:

- Undirected graphs
 - Directed graphs
 - Directed acyclic graphs (DAGs)
 - Rooted trees
 - Unrooted trees (i.e. every node can be chosen as the root)
 - Forests, etc..

12.6.1 Cone Trees

Cone trees (Robertson) are 3D embeddings of trees.

- Children are arranged on circular cones
 - Navigation by interactive rotation at all hierarchy levels.

Useful for trees with high branching (no binary trees!).

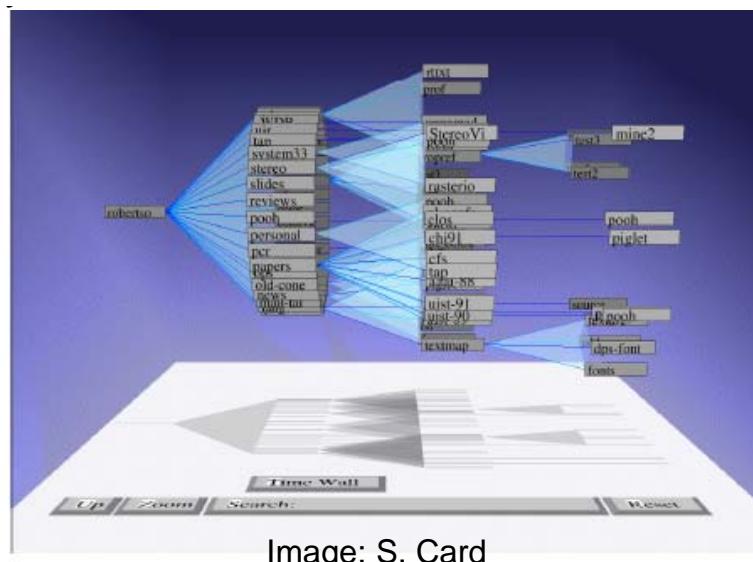


Image: S. Card

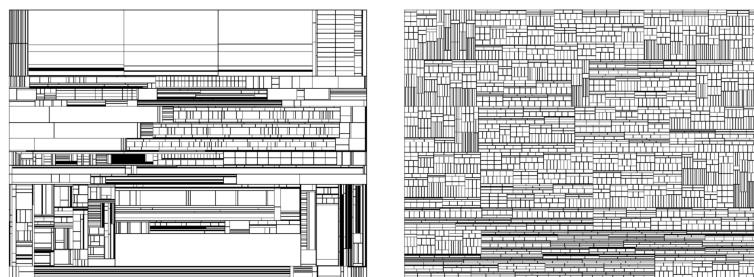
12.6.2 Tree Maps

Trees with weight attribute at nodes can be visualised using *tree maps* (Johnson and Shneiderman).

Tree maps are special Venn diagrams where

- Subtrees are represented by rectangles
- Rectangle area is proportional to total weight of the subtree
- Split direction is vertical/horizontal for odd/even hierarchy level
- Nodes can have colours, labels, tool-tip info, etc...

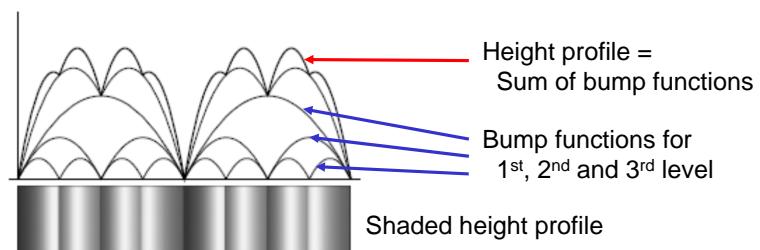
Problem of tree maps: In large trees, hierarchical levels can be hard to see:



Solution: *Cushion tree maps* (van Wijk, van de Wetering 99).

Idea: Give rectangles a height profile with height depending on the hierarchy level.

Example (1D): Height profile for a binary tree.



2nd problem of tree maps: Bad aspect ratios.

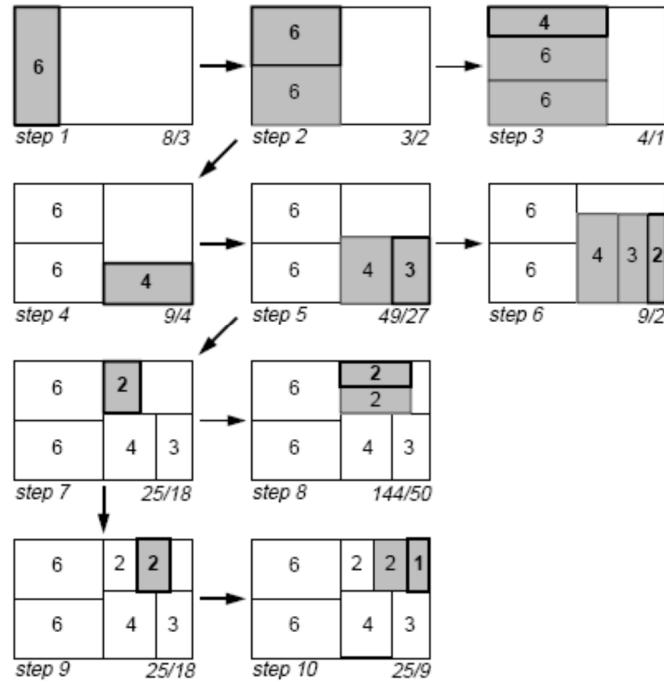
Solution: *Squarefied tree maps* (Bruls et al.).

Idea: Allow both vertical and horizontal splits within the same level of the tree.

Algorithm:

- Sort children by descending weight.
- While list of children not empty:
 - Insert the first child, splitting the larger edge

- Repeat:
 - * "Squeeze" the next child into the same "row" (along the shorter edge)
 - * If aspect ratio is worse than that of the previous step, undo the step and break the inner loop.

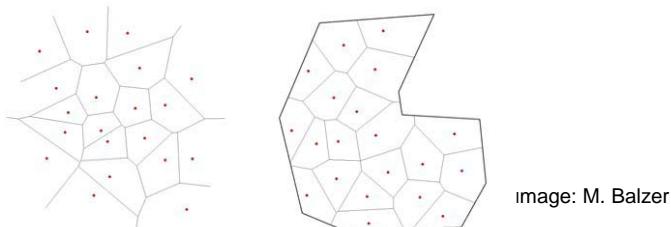


12.7 Voronoi Tree Maps

$S = \{s_1, \dots, s_n\}$ is a set of "sites" (points) in \mathbb{R}^2 . The cell generated by s_i is:

$$c_i = \{p \in \mathbb{R}^2 : \|s_i - p\| < \|s_j - p\| \ (j \neq i)\}$$

Example: Unbounded and bounded Voronoi tessellation:



12.7.1 Additively Weighted (AW) Voronoi Tessellation

Each site s_i has a *weight* w_i . The cell generated by $\langle s_i, w_i \rangle$ is:

$$c_i = \{p \in \mathbb{R}^2 : \|s_i - p\| - w_i < \|s_j - p\| - w_j \ (j \neq i)\}$$

AW Voronoi tessellation is a generalised Voronoi tessellation with circles as generators. The edges are now hyperbola.

12.7.2 Power Weighted (PW) Voronoi Tessellation

The cell generated by $\langle s_i, w_i \rangle$ is:

$$c_i = \left\{ p \in \mathbb{R}^2 : \|s_i - p\|^2 - w_i < \|s_j - p\|^2 - w_j \ (j \neq i) \right\}$$

Weight now corresponds to squared radii of circles. The edges are now straight lines.

12.7.3 Voronoi Tree Map

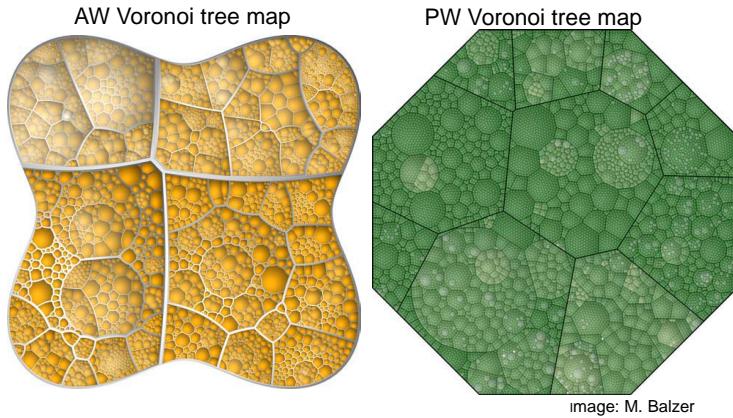
- Start with initial cell representing the full tree.
- Recursively subdivide the cell into AW or PW cells.
- Use weights for controlling the area of cells.

Problem: The area of the cells differs from the circle area. Controlling the cell area is difficult.

Solution: Use a *centroidal Voronoi tessellation* (CVT), where sites are at the centroids of cells. This results in "rounder" cells.

Subdivision of a cell now requires an *iteration*:

- Set initial weights according to tree attribute data.
 - Place sites "randomly" within a cell (Additional complication: empty cells must be avoided).
 - Repeat:
 - Compute AW or PW Voronoi tessellation
 - Compute cell areas
 - Compute errors (=deviation from intended areas)
 - Adjust weights
 - Replace sites by cell centroids
- until errors < tolerance.



12.8 Clustering Techniques

Motivation for *clustering* in visualisation of graphs: Multiple levels-of-detail are obtained by identifying "highly connected" subsets and representing them by glyphs.

Clustering techniques are often based on *force models*. Assume an undirected graph $G = (V, E)$ with a set of nodes V and a set of edges E .

Notation:

e_{ij} Edge connecting nodes i and j .

p_i Position of node i

$$p_{ij} = \|p_i - p_j\|$$

The *attractive force* is usually *Hooke's spring law*

$$f(x) = A \cdot (x - x_0),$$

where x_0 is the zero energy length of the spring.

The *repulsive force* generally follows an inverse square law inspired by *electrostatic fields*:

$$g(x) = \frac{B}{x^2}.$$

The total *potential energy* is then:

$$P = \frac{A}{2} \sum_{e_{ij} \in E} (p_{ij} - x_0)^2 - B \sum_{i \neq j} \frac{1}{p_{ij}}.$$

Difficult to visualise: *Small world* graphs (Watts and Strogatz).

Small world graphs are connect graphs having

- a small *average path length* (between a pair of nodes) and
- a high *clustering index*

both compared to a random graph with the same number of nodes and edges.

The *clustering index of a node* v is the ratio between

- the number of *existing* edges in the 1-neighbourhood $N(v)$ of v .
- the number of *possible* edges, which is $k \frac{k-1}{2}$ if $k = |N(v)|$.

The *clustering index of the graph* is the average of the clustering indices of its nodes.

Energy models suited for small-world problems: *r-PolyLog* energy models (Noack):

Potential Energy:

$$P = \sum_{e_{ij} \in E} (p_{ij} - x_0)^r - \sum_{j \neq i} \log(p_{ij}).$$

Attractive and repulsive forces are obtained by taking the derivative for 1-Polylog:

$$\begin{aligned} f(x) &= 1 \\ g(x) &= \frac{1}{x} \end{aligned}$$

Minimum energy configuration of 1-PolyLog has the following property:

Distance between two clusters C_1 and C_2 is inversely proportional to their *coupling*:

$$\frac{|\{e_{ij} : i \in C_1, j \in C_2\}|}{|C_1| |C_2|}$$

12.9 Distortion Techniques

Various techniques:

- Perspective wall (Robertson)
- Table lens (Rao and Card)

12.9.1 Hyperbolic trees

Hyperbolic trees trees are based on the *Poincaré Disk* model (projection) of the *hyperbolic space H_2* .

In the Poincaré disk, the role of *straight lines* is taken by:

- *Circles* which intersect the bounding circle $x^2 + y^2 = 1$ orthogonally
- *Diameters* of the bounding circle.

Properties

- Triangles have the sum of angles $< 180^\circ$.
- It has the *metric*

$$ds = \frac{\sqrt{dx^2 + dy^2}}{1 - x^2 - y^2}.$$

- The bounding circle is at infinity
- Circle perimeter grows exponentially with its radius.

As a consequence trees can be drawn *undistorted* in hyperbolic space:

- All edges having about the same length *and*
- All nodes having the same angle available for their children.

Rigid transformations of the Poincaré disk: *Möbius transformations* of complex numbers:

$$z' = T_{c\theta}(z) = \frac{\theta z + c}{\bar{c}\theta z + 1}, \quad |\theta| = 1, |c| < 1$$

These are:

for $c = 0$: *rotations* around 0.

for $\theta = 1$: *translations* (mapping 0 to c and $-c$ to 0).

Combinations:

$$T_{c_2\theta_2}(T_{c_1\theta_1}(z)) = T_{c\theta}(z)$$

with

$$c = \frac{\theta_2 c_1 + c_2}{\theta_2 c_1 \bar{c}_2 + 1}, \quad \theta = \frac{\theta_1 \theta_2 + \theta_1 \bar{c}_1 c_2}{\theta_2 c_1 \bar{c}_2 + 1}.$$

Technique (Lamping et al) Change of focus, i.e. moving a different node towards the center, is achieved by performing a *translation* in hyperbolic space.

Example: Visualisation of a large organisational hierarchy in hyperbolic space with different foci.

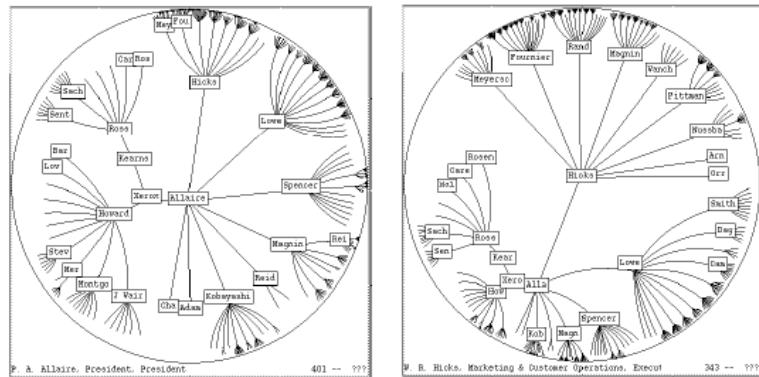


Image: R. Rao