P-spline Smoothing on Difficult Domains

Paul Eilers

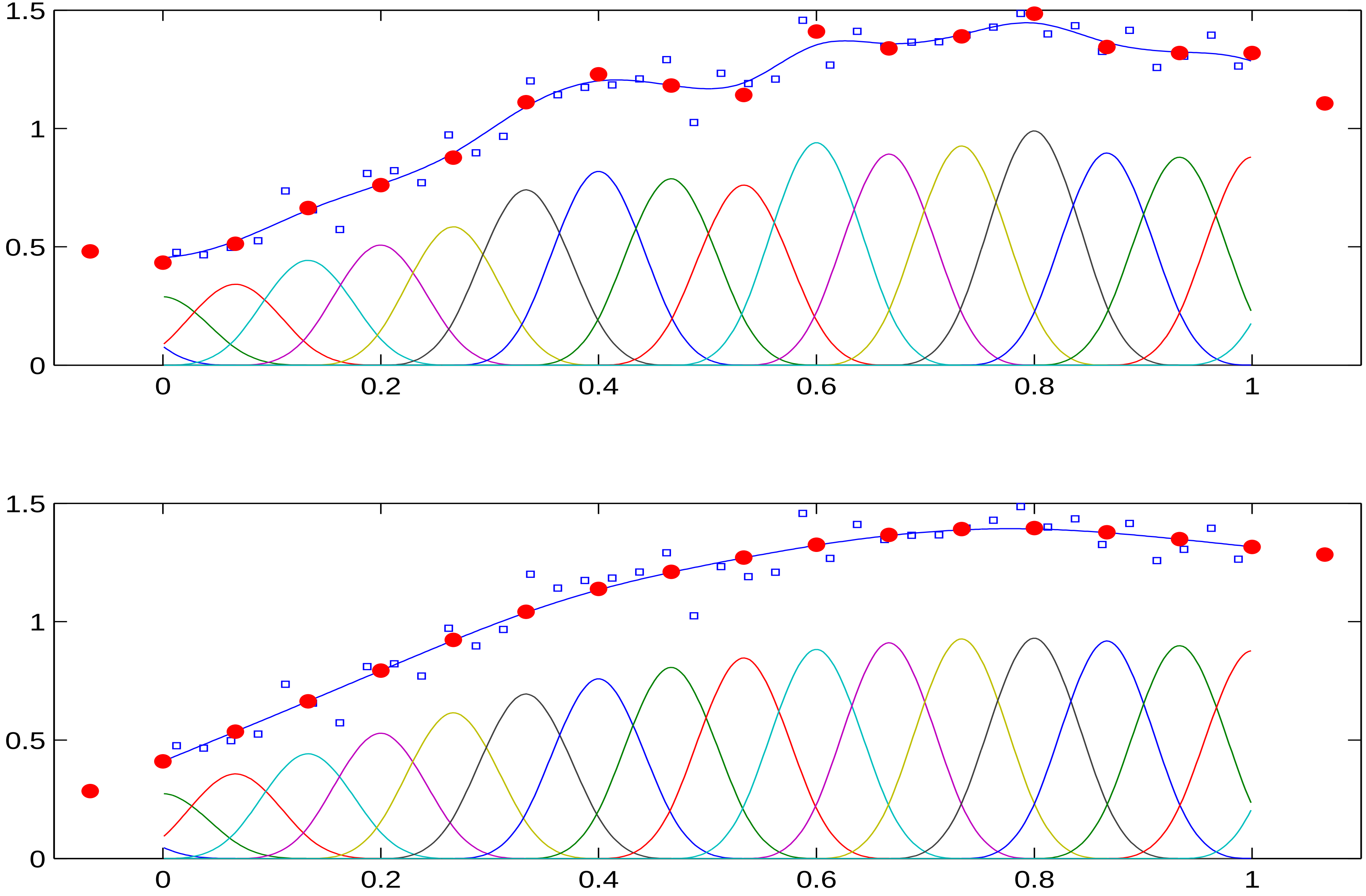
Leiden University Medical Centre

SFB386, October 2006

P-splines on one page

- Data in vectors x and y
- Compute a basis $B = B(x)$ of B-splines, equally spaced
- Minimize $S = ||y - B\alpha||^2 + \lambda||D\alpha||^2$
- Matrix D forms differences (order 2 or 3)
- First term: fit to data; second term: roughness penalty
- The larger λ, the smoother $\hat{\alpha}$ (and thus $\hat{y} = B\hat{\alpha}$)
- Use many B-splines, let λ do the work
- Don't worry about optimal knots

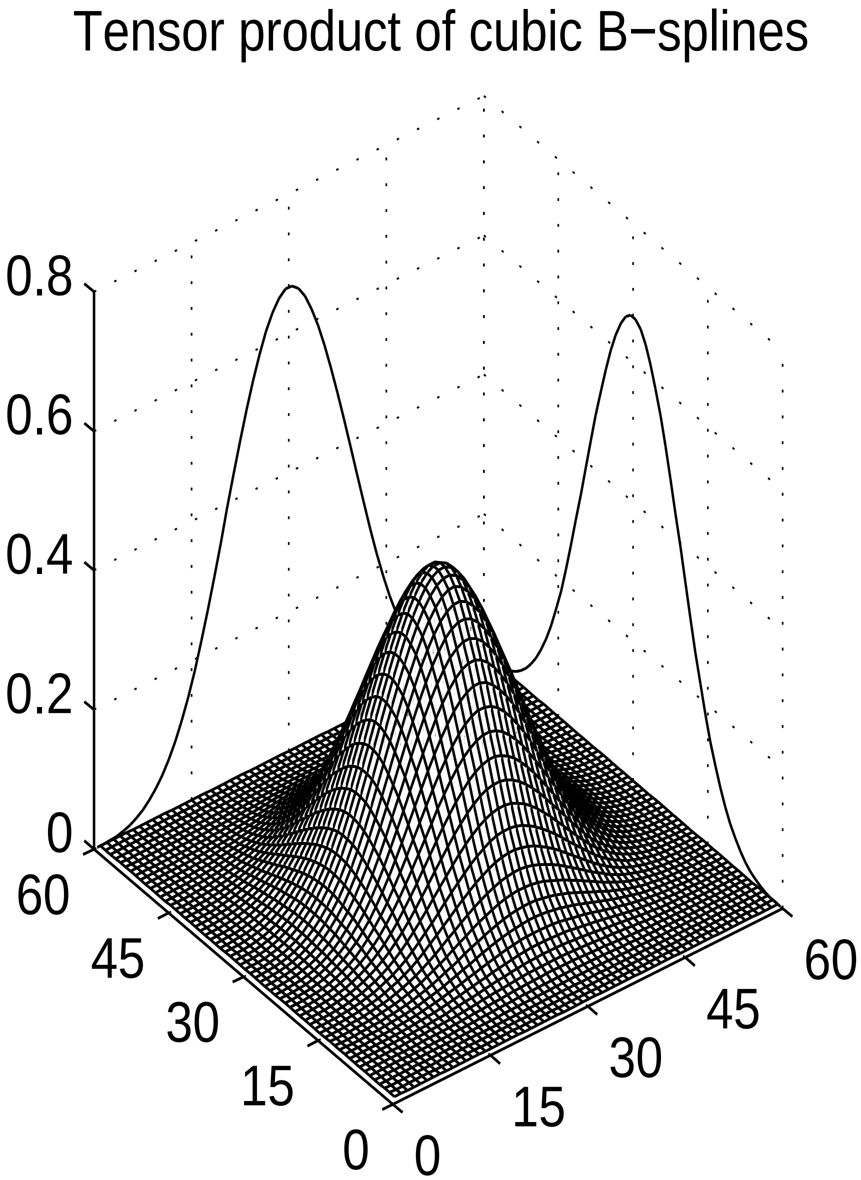
P-splines illustrated

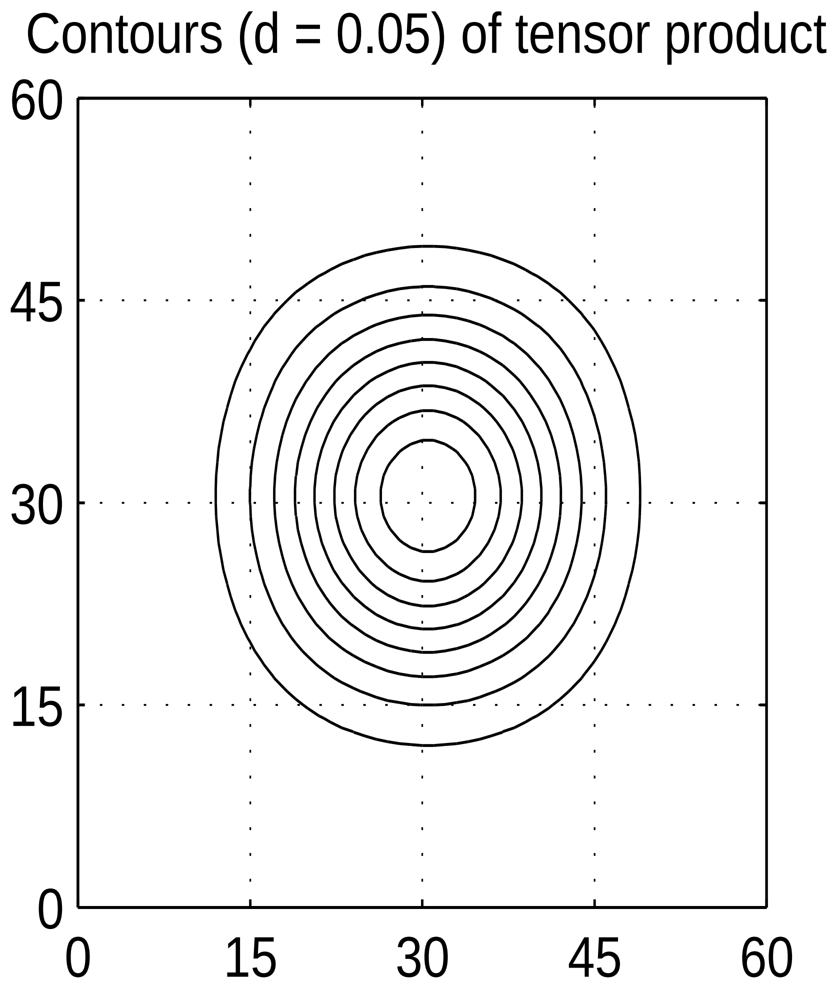


2-D grid P-splines on one page

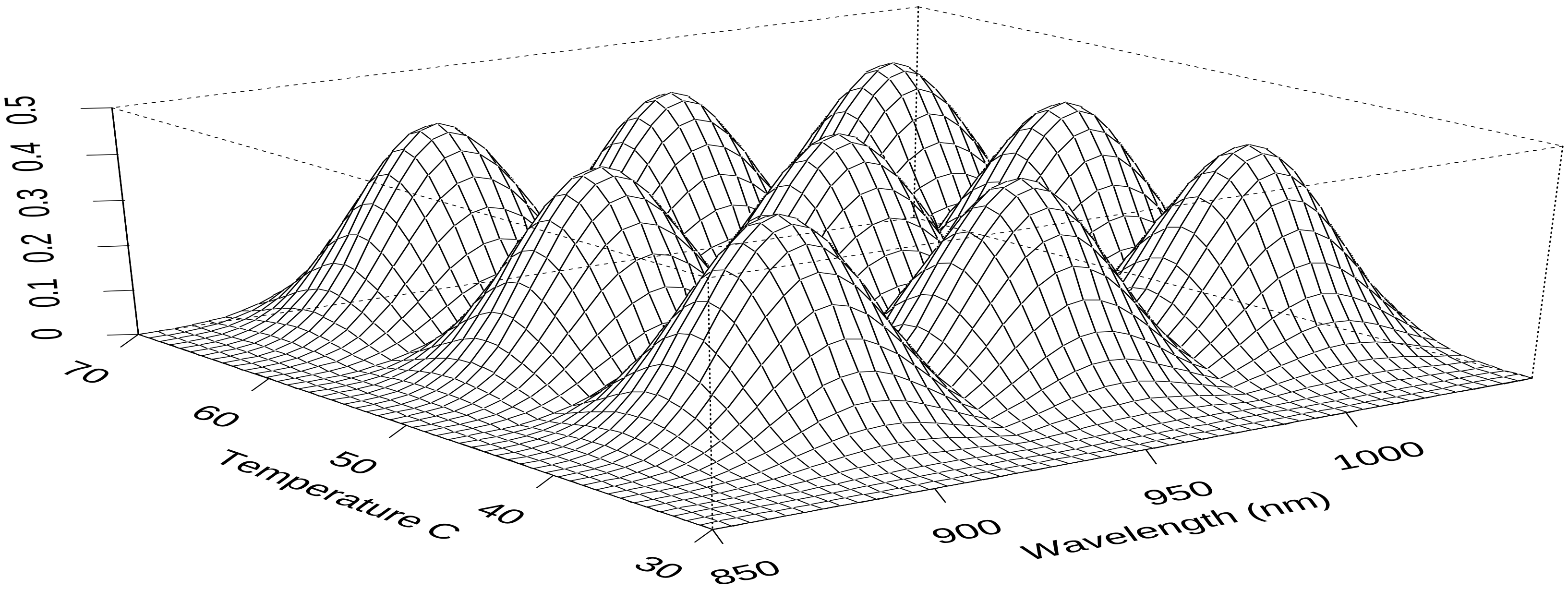
- Tensor products of bases B for rows, $\breve{B}$ for columns
- Knots on grid, data Y on (finer) rectangular grid
- Minimize $||Y - BA\breve{B}||_F^2 + \lambda||DA||_F^2 + \breve{\lambda}||AD'||_F^2$
- Frobenius norm: $||X||_F^2 = \sum_i \sum_j x_{ij}^2$
- Penalty down columns $\lambda||DA||_F^2$, along rows $\breve{\lambda}||AD'||_F^2$
- Fast algorithms are available, for weighted smoothing (GLM)
- Zero weights nice for missing points; automatic interpolation
- Iain Currie, Maria Durbàn & PE (CSDA, 2006; JRSS-B, 2006)

One tensor product





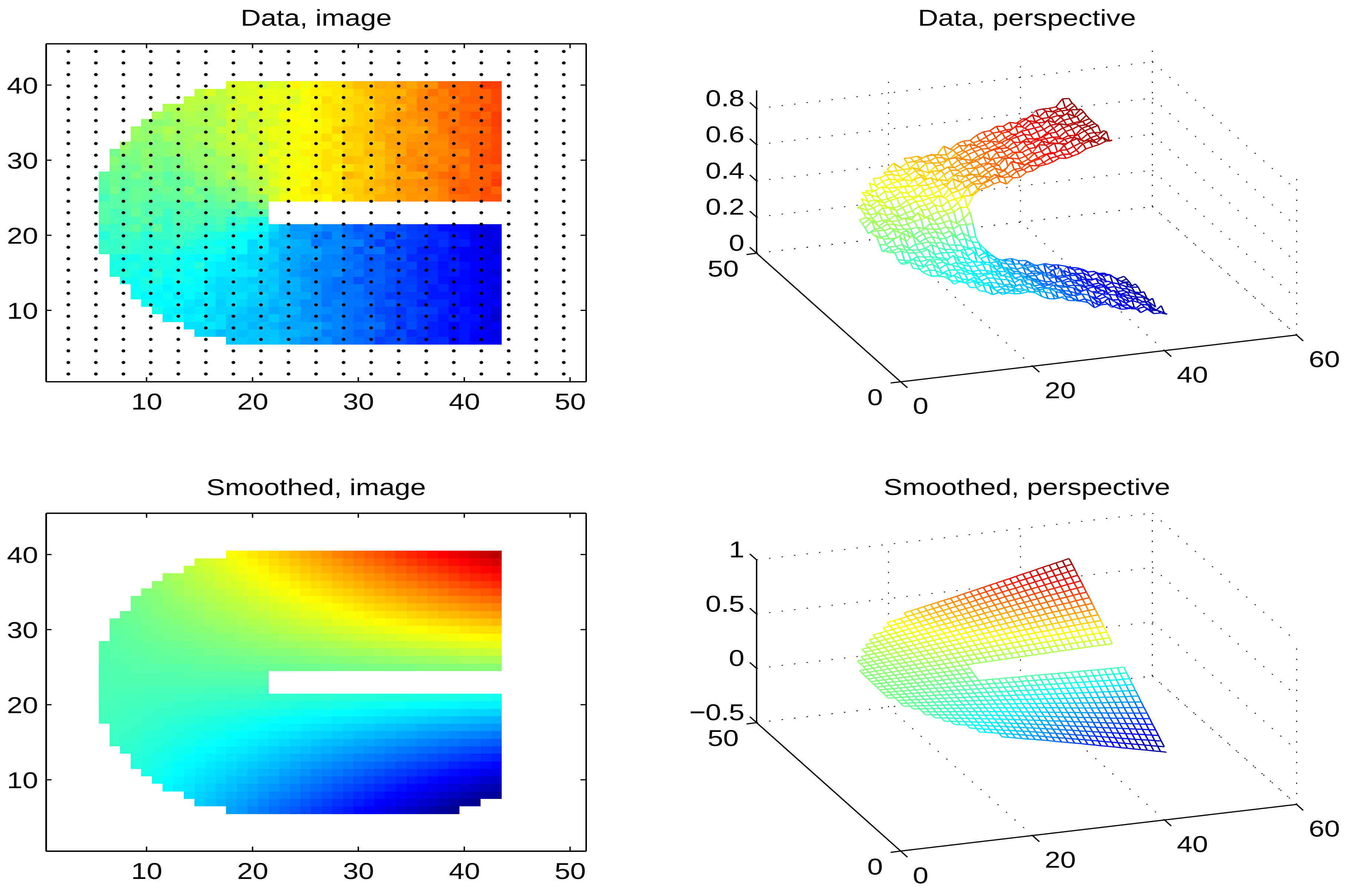
Selected tensor products



Tim Ramsay's problem: "difficult domains"

- Paper Ramsay in JRSS-B (2002)
- Natural (or artificial) barriers between regions may exist
- Rivers, lakes, mountains, industrial areas, ...
- Then smoothness has to be ″interrupted″
- Ramsay's application: survey from Montreal
- Also a simple example "tuning fork" (U shape)
- Arms physically near but independent

Standard smoothing bridges the gap



Ways to handle difficult domains

- Change the penalties
  - make some terms in $\lambda||DA||_F^2 + \breve{\lambda}||AD'||_F^2$ zero
  - using appropriate weights

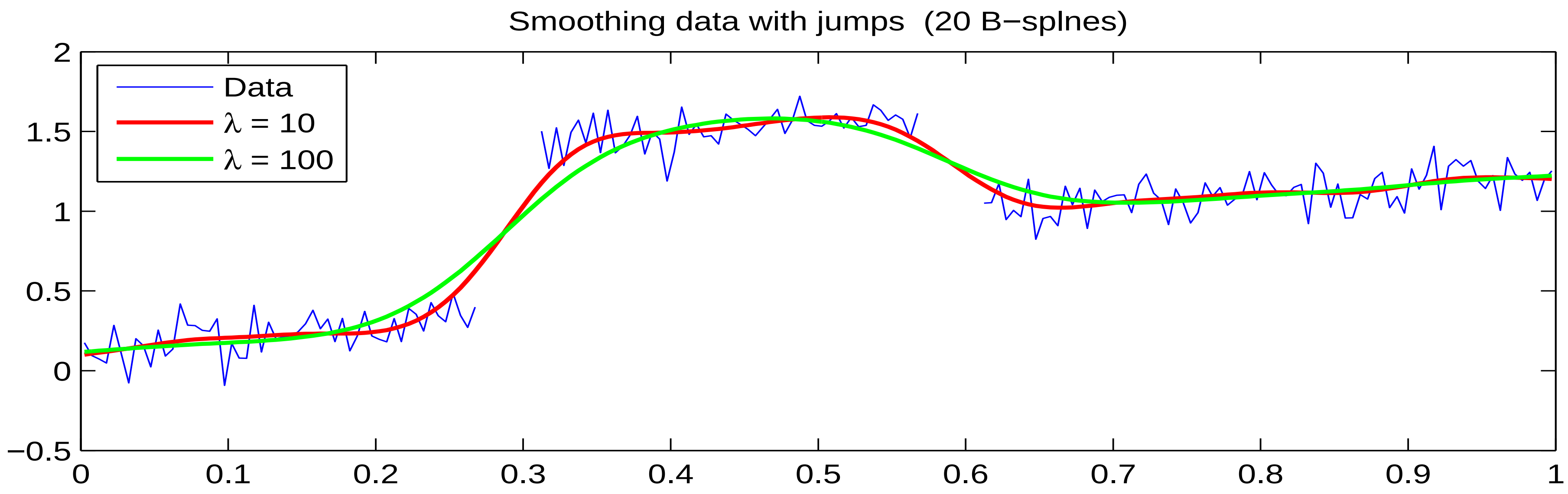
- Reshape the domain, using insight
  - bend the tuning fork to a rectangle
  - use polar coordinates for half ring

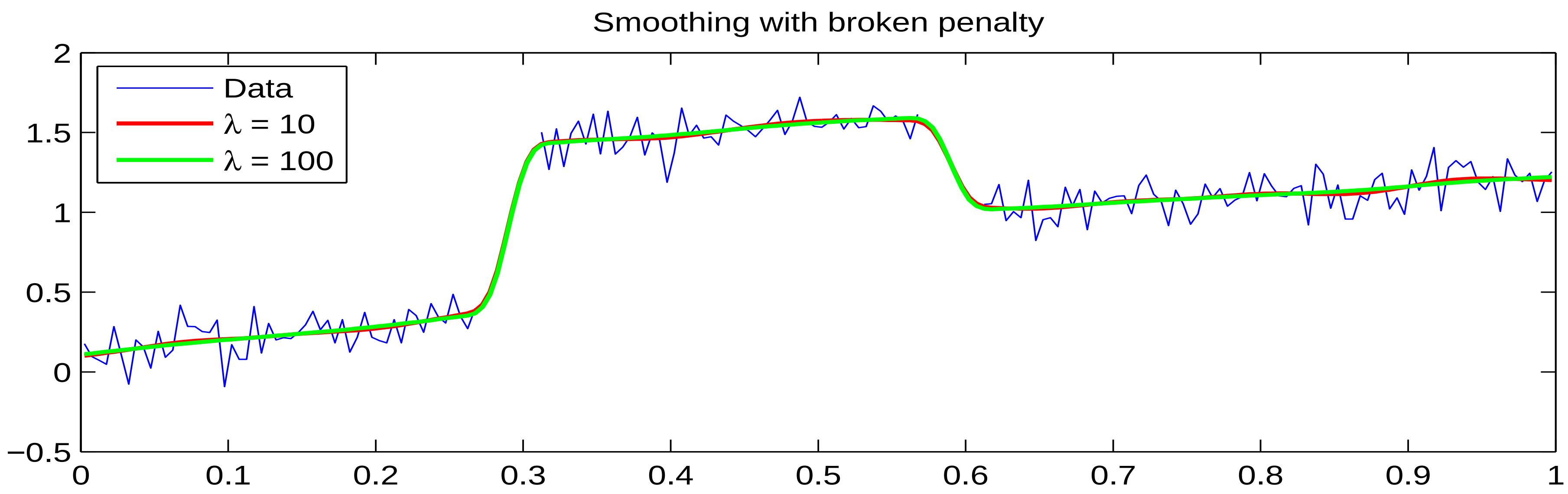
- Reshape the domain using conformal mapping
  - numerical Schwartz-Christoffel transform
  - based on complex function theory

Breaking the penalty

- Data: smooth curve with jumps
- Gaps in x at jumps
- Roughness penalty smoothly interpolates gaps
- Often a nice property, but not what we want now
- Solution: eliminate some terms in penalty $\sum_j (\Delta^2 \alpha_j)^2$
- For α coefficients in the middle of the gap
- Elegant implementation with weights v: $\sum_j v_j (\Delta^2 \alpha_j)^2$
- Equivalently: $\lambda \alpha' D' V D \alpha$

P-spline smoothing in 1-D with jumps

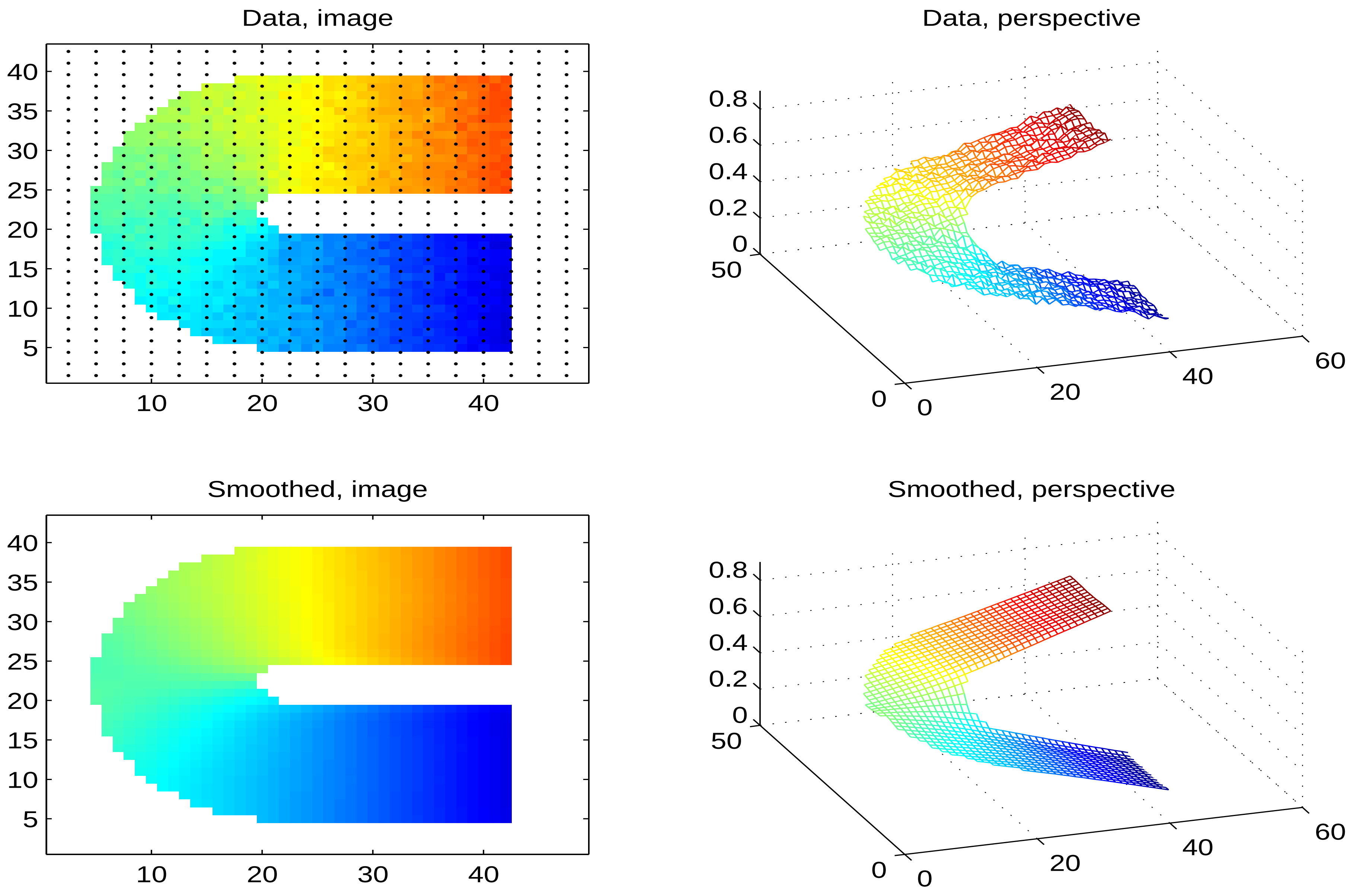




Breaking the penalty in 2 dimensions

- We have a penalty on each row of A and on each column
- Introduce weighting vectors for differences
- One for each row and one for each column
- Nice $||DA||_F^2$ and $||AD'||_F^2$ formulas no longer work
- Instead: $\alpha = \text{vec}(A)$, block-diagonal V and Kronecker products
- Details skipped

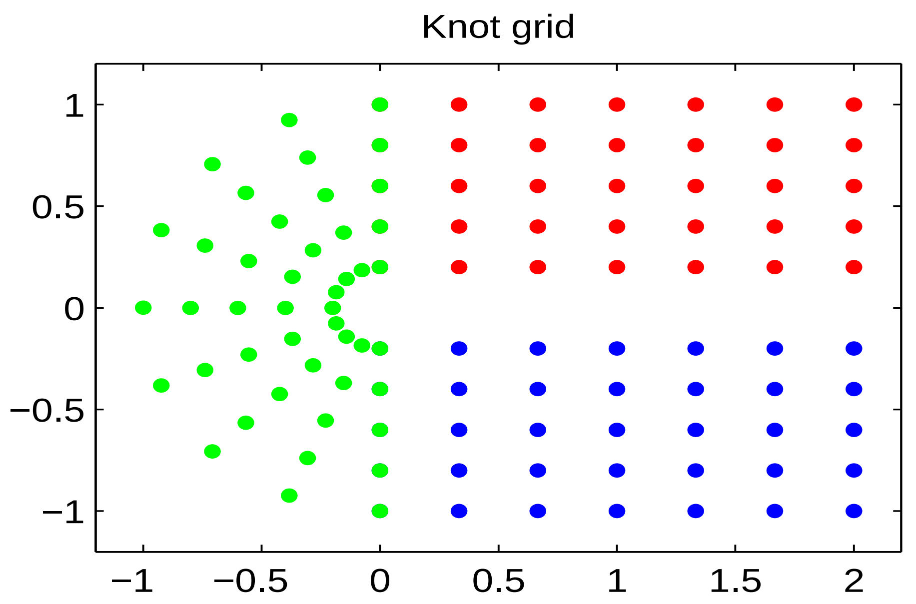
Smoothing with broken penalty

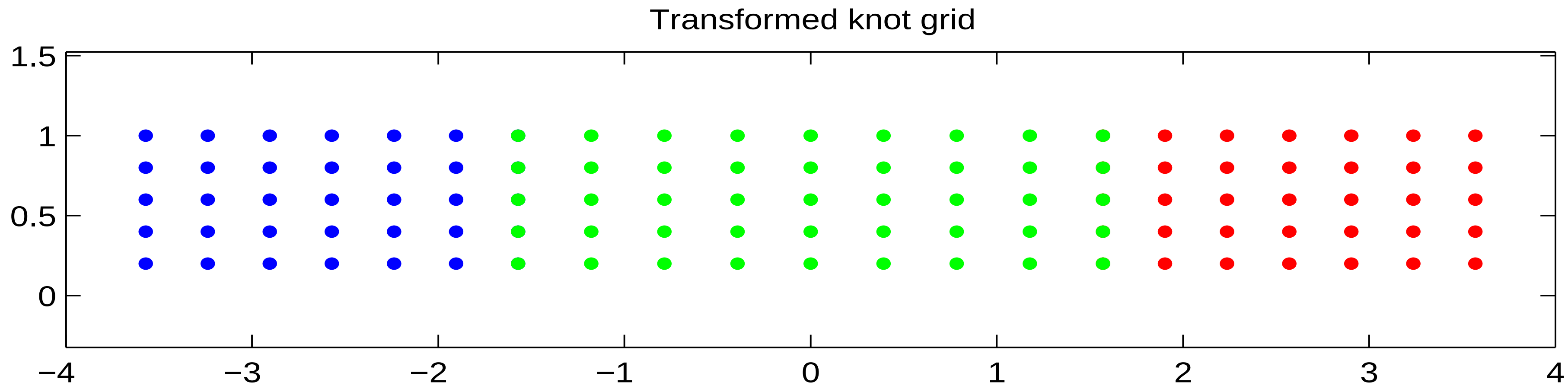


Domain transformation using insight

- The tuning fork has three parts
- Two rectangles and one semi-ring
- Introduce polar coordinates on ring: ϕ, ρ
- In polar domain half ring becomes rectangle
- Glue original rectangles to both sides
- Smooth and transform back

Domain transformation





Domain transformation issues

- Uniform grid is maintained on rectangles
- But not on the semi-ring: inner bend is stretched
- Data grid after transformation not uniform
- Fast grid algorithm no longer applies (on ring)
- Isotropy on transformed domain is not maintained
- This might be desirable or not
- Less smoothing in sharp turns natural?
- Polar data grid on semi-ring for data collection?

Transformation of really difficult domains

- Tuning fork is easy
- Mathematical insight quickly leads to transformation
- We will not always be that lucky
- Some domains might have really difficult shapes
- Solution: use numerical conformal mapping
- Schwartz-Christoffel transform

The Schwartz-Christoffel transform

- One of the beauties of complex function theory
- Commonly used to map one (polygonal) region into another
- Popular for solving partial differential equations
- Laplace's equation: $\partial^2 z/\partial x^2 + \partial^2 z/\partial y^2 = 0$
- Transformation $z = x + iy = f(w) = f(u + iv)$
- Laplace holds again: $\partial^2 w/\partial u^2 + \partial^2 w/\partial v^2 = 0$
- Grid lines follow transformation nicely

CS transform details

- Transformation $z = x + iy = f(w) = f(u + iv)$
- In problem domain: z; in other domain: w
- Polygon in z is transform of chosen shape in w

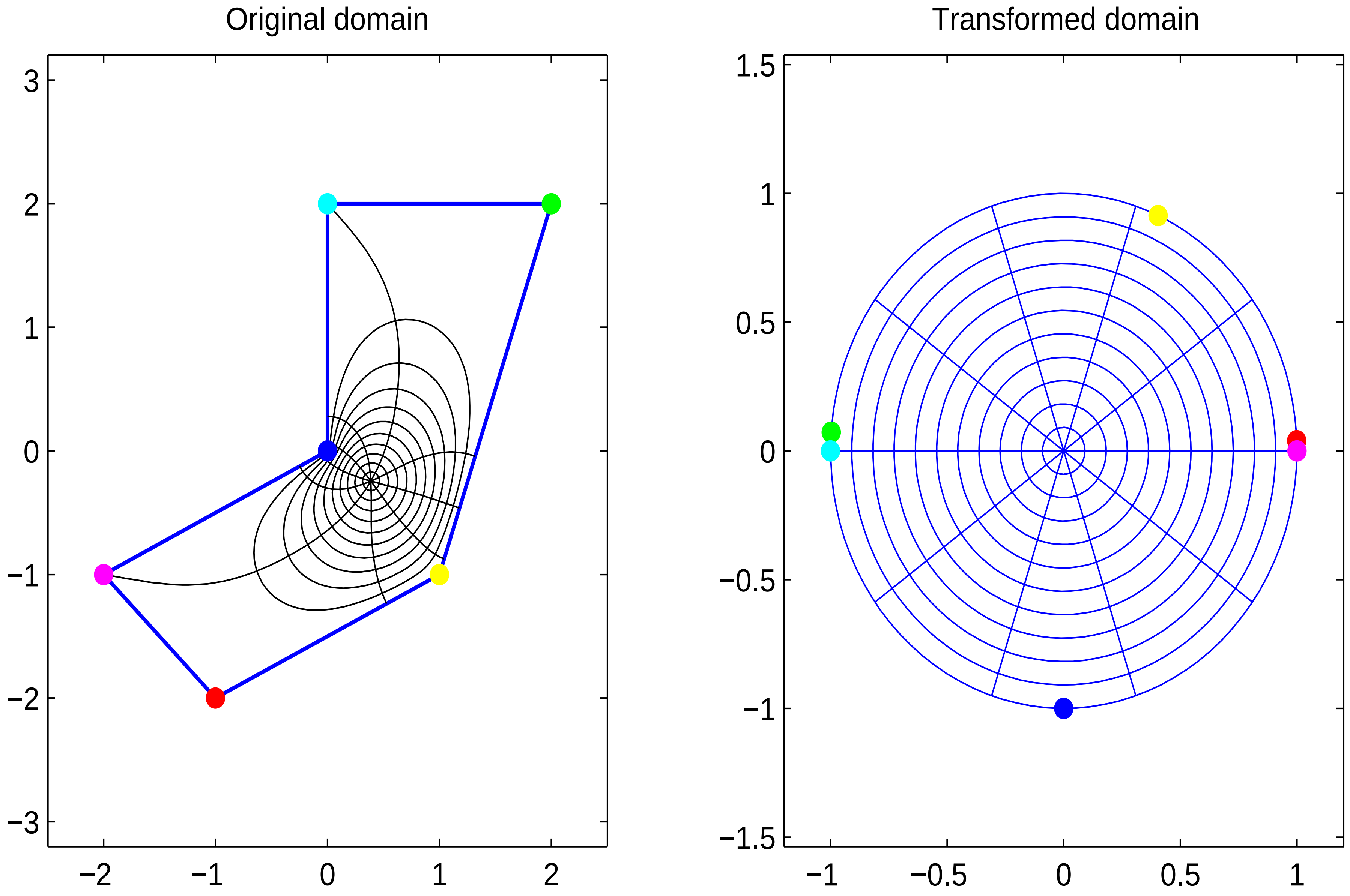
$$f'(w) = A \prod_i (w - t_i)^{k_i}$$

- Powers k_i determined by (external angles of) z-polygon
- Characteristic points t (and coefficient A) to be determined

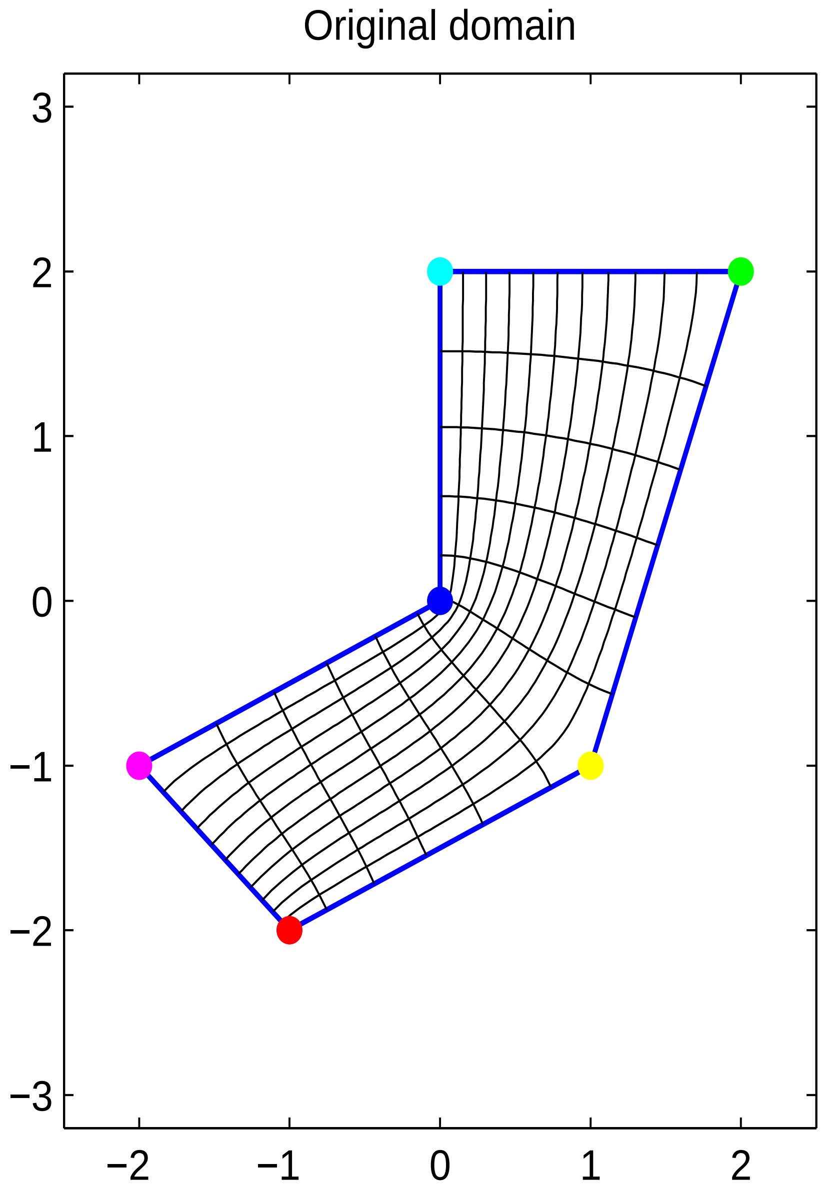
Numerical Schwartz-Christoffel transform

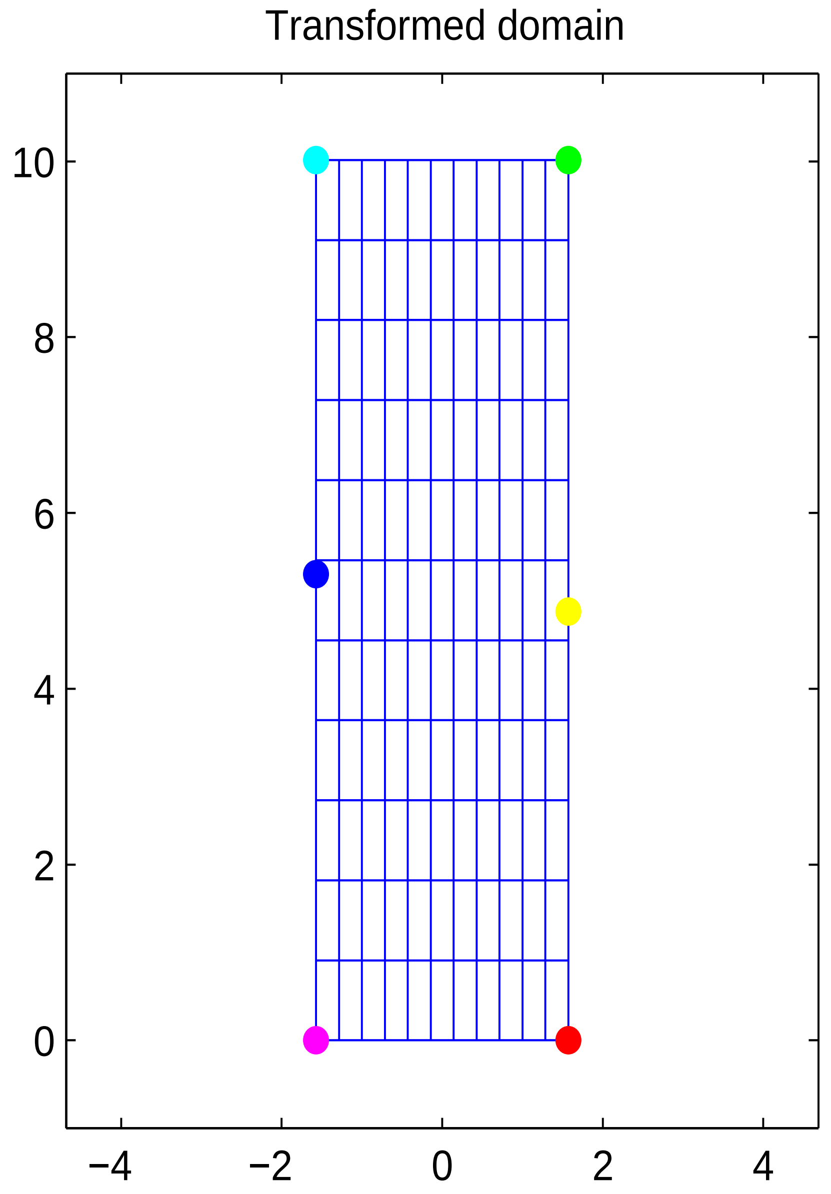
- Analytical results are limited
- But powerful numerical approaches exist
- I use Matlab toolbox by Toby Driscoll
- User specifies a closed polygon in z-plane
- Chooses target shape (rectangle, disk, ...) in w-plane
- Toolbox does the hard work

Example: map polygon to disk

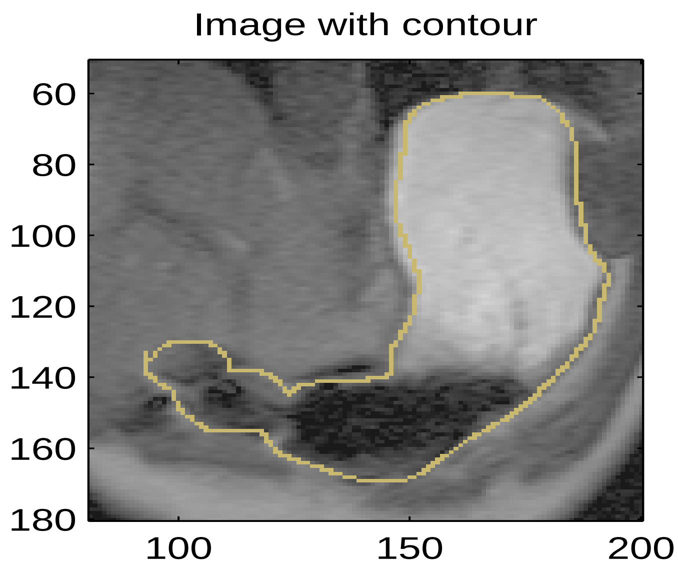


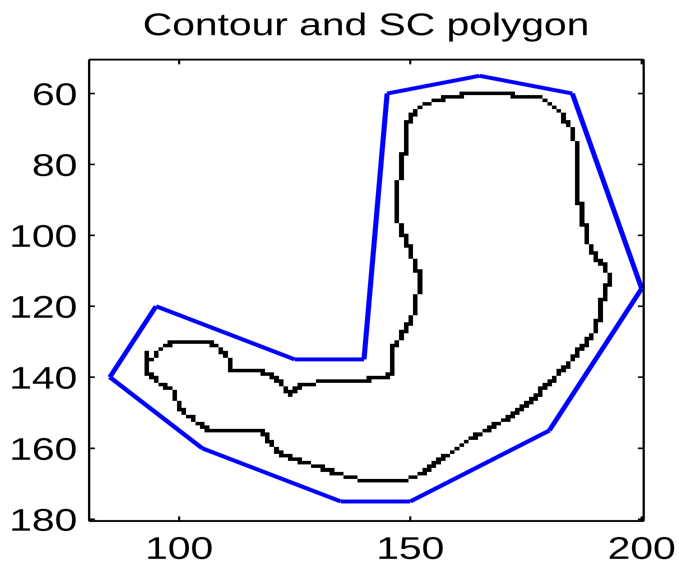
Example: map polygon to rectangle

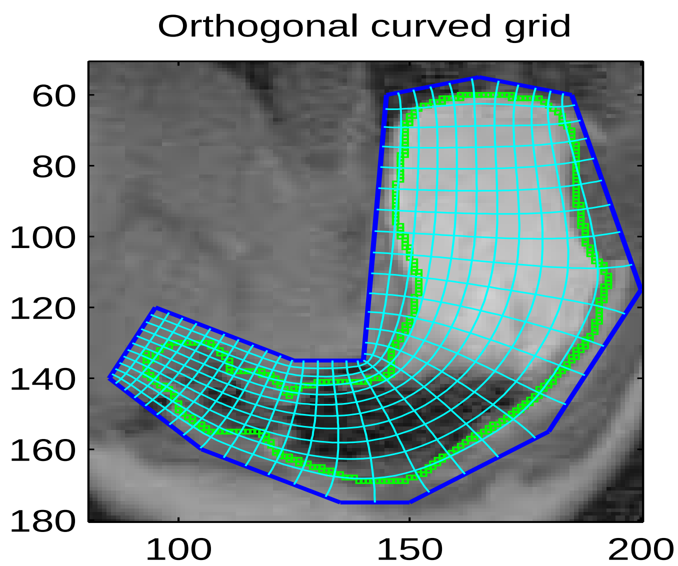


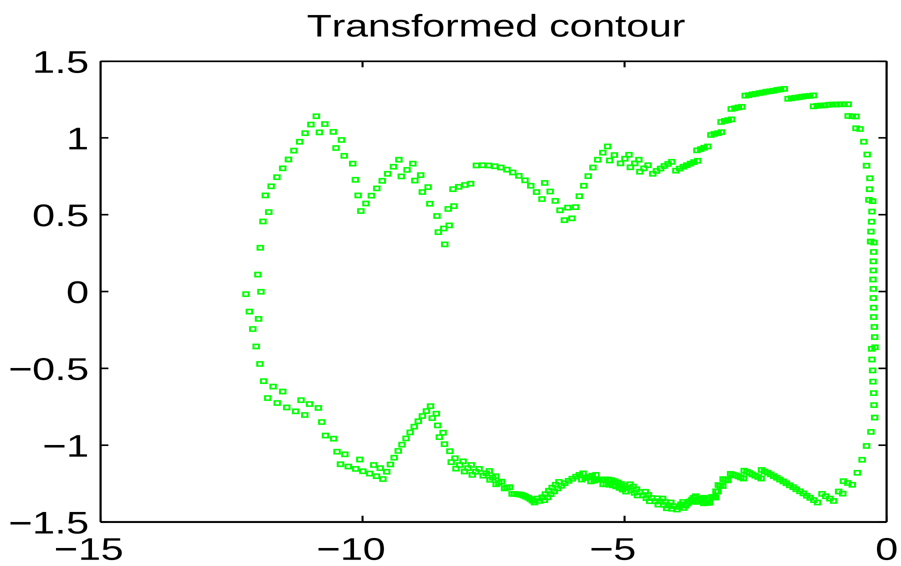


A real-life example





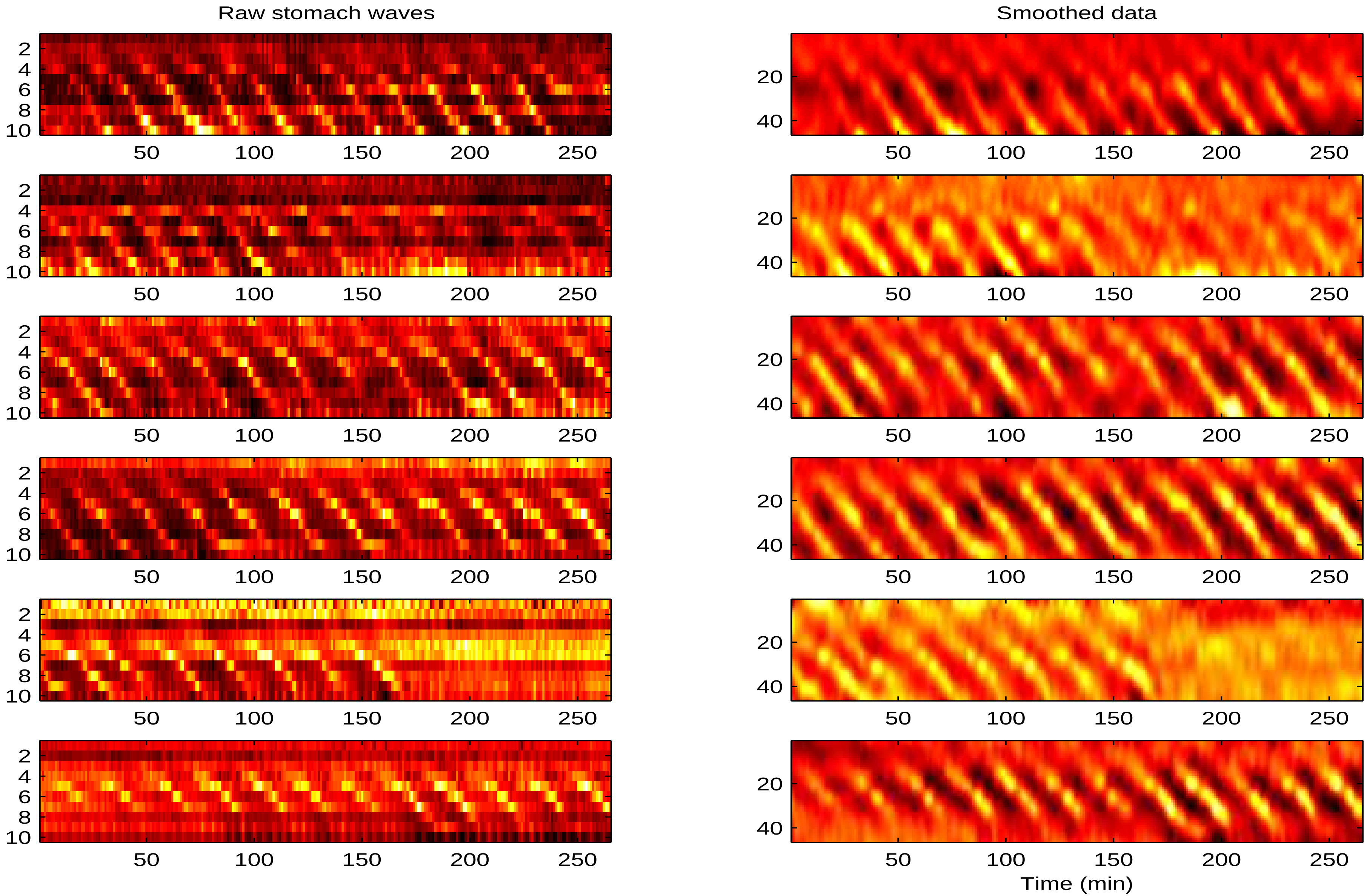




Where does the example come from?

- Consulting with colleagues working on stomach diseases
- MRI scans at regular intervals, after a meal
- To study emptying process
- Interesting phenomenon: local contractions
- Like eating Bavarian *Weiszwurst* (mit Brezel und süszem Senf)
- Contractions travel like waves down the stomach
- Can we measure amplitude, period and travel speed?
- Now a very time-consuming process (manual digitization)

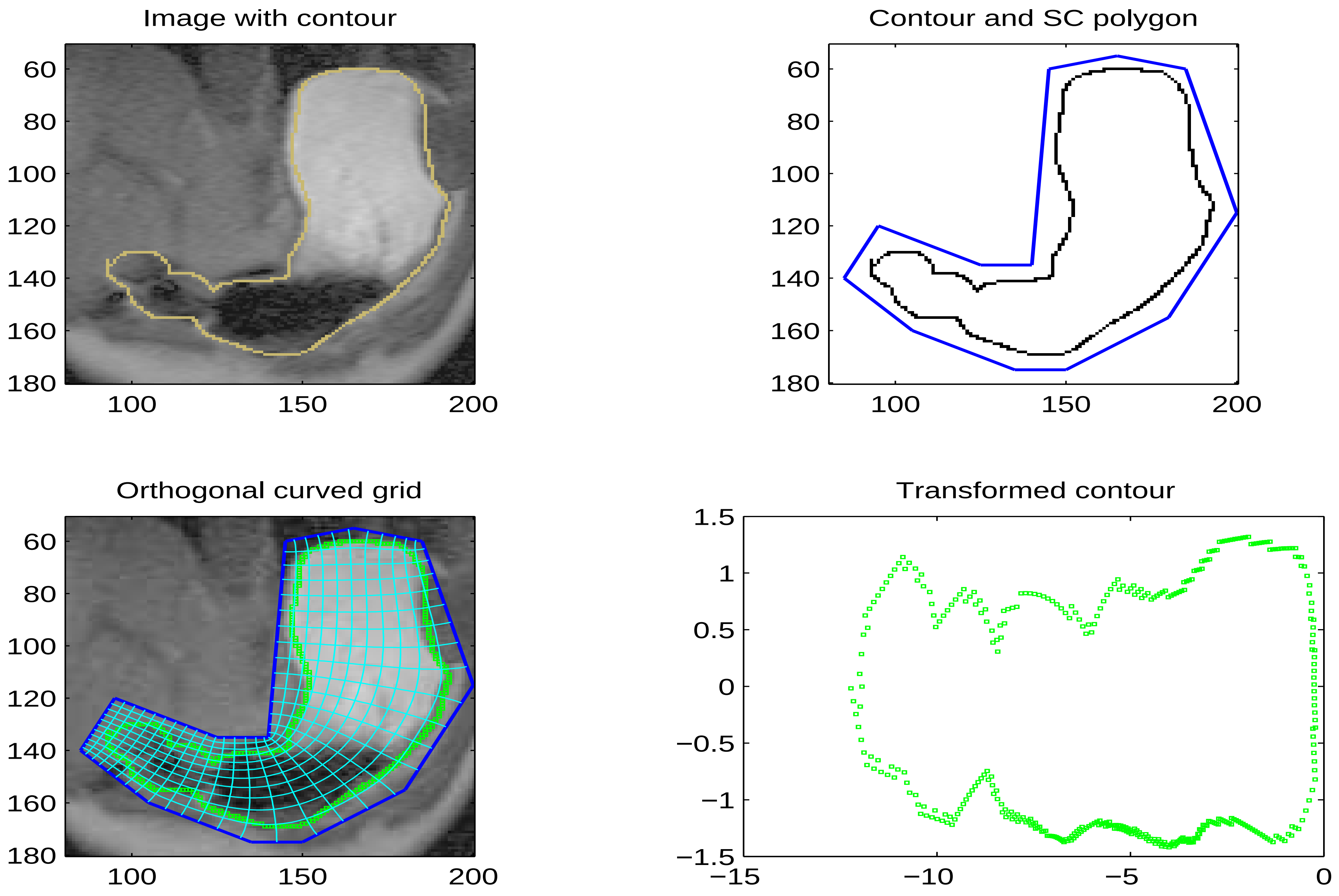
Stomach contractions



The plan

- Try to find a good grid on stomach
- One dimension following main “axis”
- The other locally orthogonal
- SC maps stomach neighborhood (polygon) to rectangle
- Standard curve fitting can be applied
- Working on the transformed stomach wall

A real-life example



Discussion

- Transformation depends on specified polygon
- Quite arbitrary
- How much does it matter?
- Specified polygon has variable “width”
- Can we connect to “principal curves”?
- Tim Ramsay connects to partial differential equations
- Smoothing spline: strip (membrane) with springs to data points
- Are second derivatives natural?