University of Basel

Covariance functions

Defining a Gaussian process

University of Basel

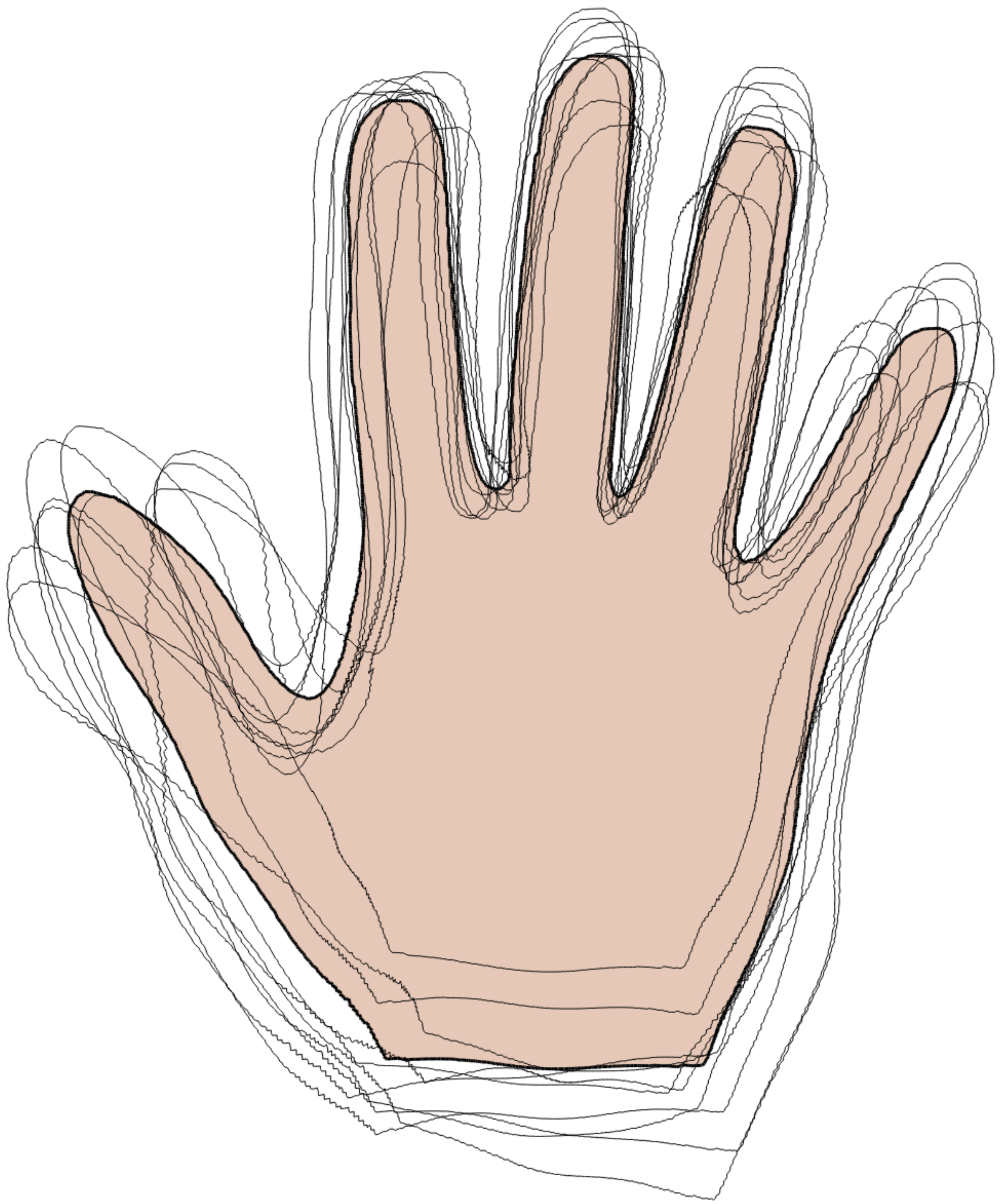
A Gaussian process

$$GP(\mu, k)$$

is completely specified by a mean function μ and covariance function (or kernel) k.

- $\mu: \mathcal{X} \to \mathbb{R}^d$ defines how the average deformation looks like
- $k: \mathcal{X} \times \mathcal{X} \to \mathbb{R}^{d \times d}$ defines how it can deviate from the mean

The mean function

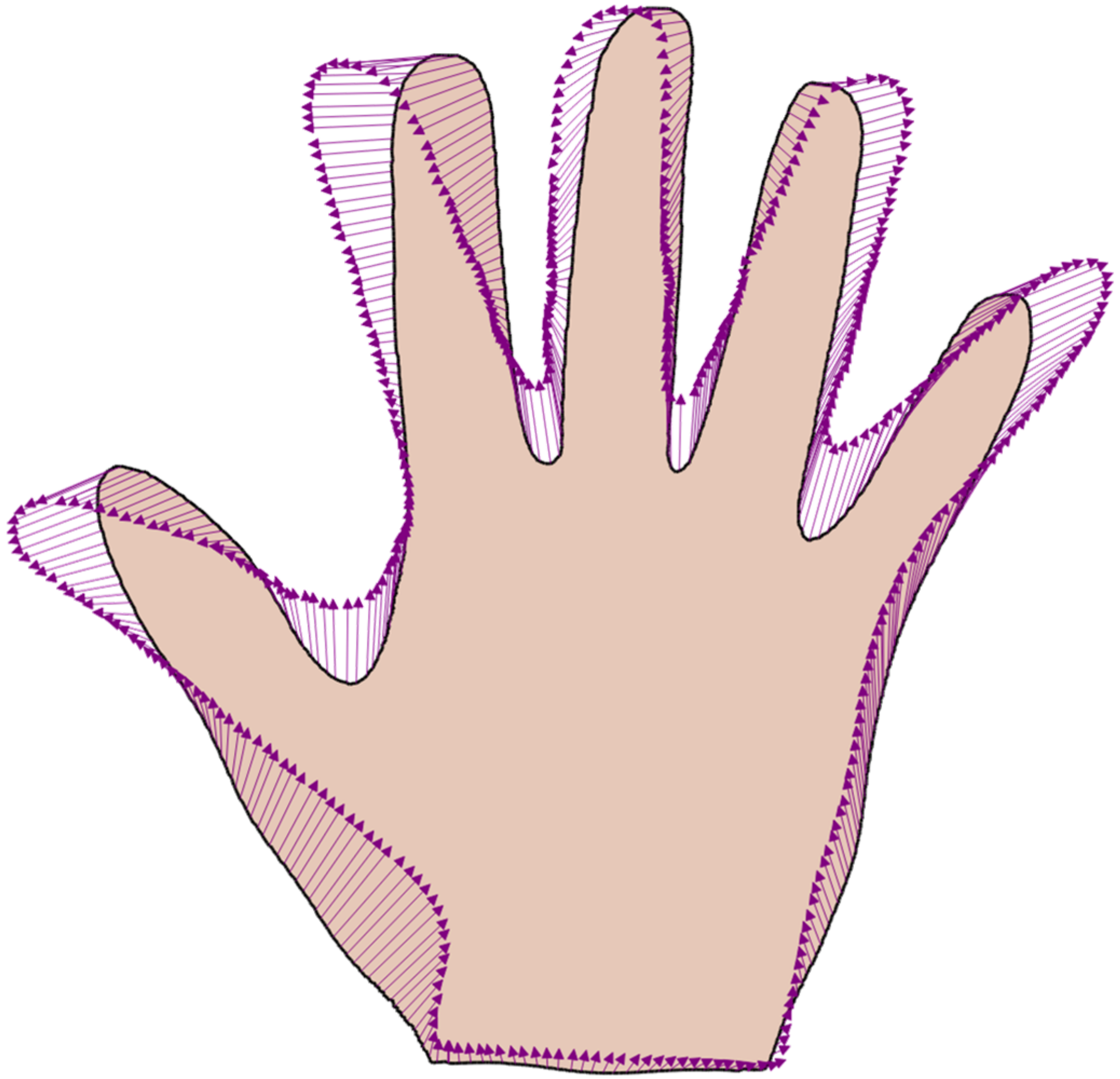


- Usual assumption:

$$\mu(x) = \begin{pmatrix} \mu_1(x) \\ \mu_2(x) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

  – The reference shape is an average shape.

The covariance function



- Defines characteristics of the deformations fields
  - Main assumption: deformation fields are smooth
- Mathematical requirement
  - The covariance function $k(x, x')$ should be a symmetric, positive semi-definite kernel.

Positive semi-definite kernels

Positive semi-definite matrix

A real $n \times n$ matrix K which satisfies

$$v^T K v \geq 0$$

for all vectors $v \in \mathbb{R}^n$ is called positive semi-definite.

Positive semi-definite kernels

University of Basel

Positive semi-definite kernel

A kernel $k : \mathcal{X} \times \mathcal{X} \to \mathbb{R}^{d \times d}$ is called positive semi-definite, if it gives rise to a positive-semi-definite kernel matrix K

with $K_{ij} = k(x_i, x_j), \quad i, j = 1, \ldots, n$

for any choice of n and $X = (x_1, \ldots, x_n)$

Positive semi-definite kernels

Positive semi-definite kernel

A kernel $k : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}^{d \times d}$ is called positive semi-definite, if it gives rise to a positive-semi-definite kernel matrix K

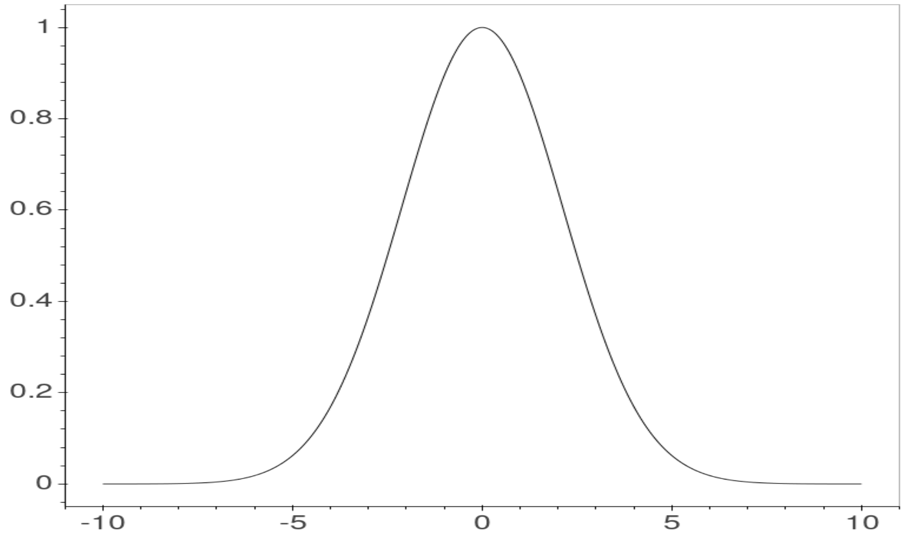
with $K_{ij} = k\left(x_i, x_j\right), \quad i, j = 1, \dots, n$

for any choice of n and $X = (x_1, \dots, x_n)$

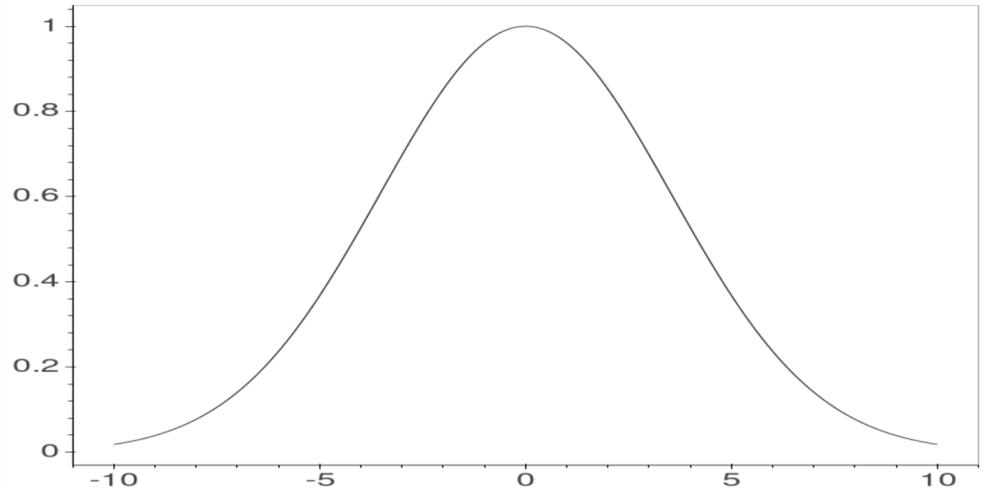
- Ensures that the GP defines a valid $N(\mu, K)$ for any discretization.

Scalar-valued Gaussian kernel

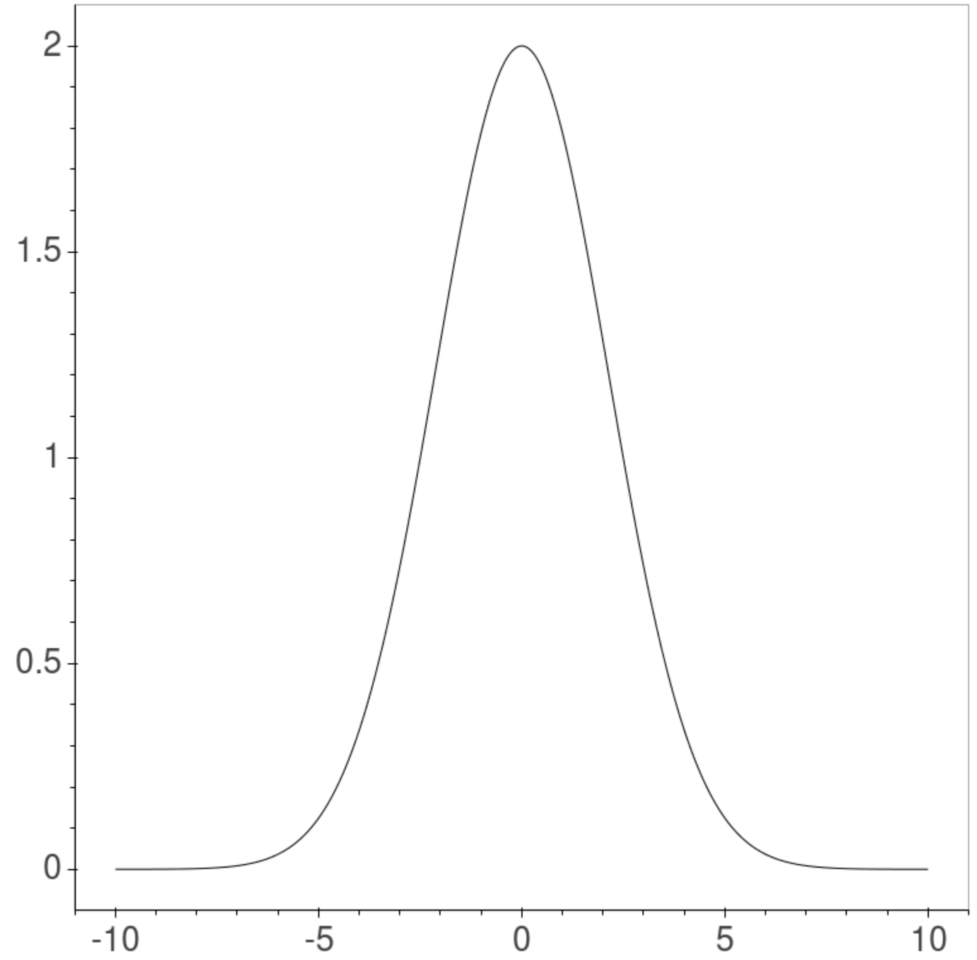
$$k(x, x') = s \exp(-\frac{\|x - x'\|^2}{\sigma^2})$$



$s = 1, \sigma = 3$



$s = 1, \sigma = 5$



$s = 2, \sigma = 3$

Diagonal kernels

$$k(x, x') = \begin{pmatrix} k^{(1)}(x, x') & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & k^{(d)}(x, x') \end{pmatrix}$$

- $k^{(1)}, \dots, k^{(d)}: \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ are scalar-valued kernels
- $k : \mathcal{X} \times \mathcal{X} \to \mathbb{R}^{d \times d}$ becomes a matrix valued kernel.

Assumption: Each dimension is modelled independently.

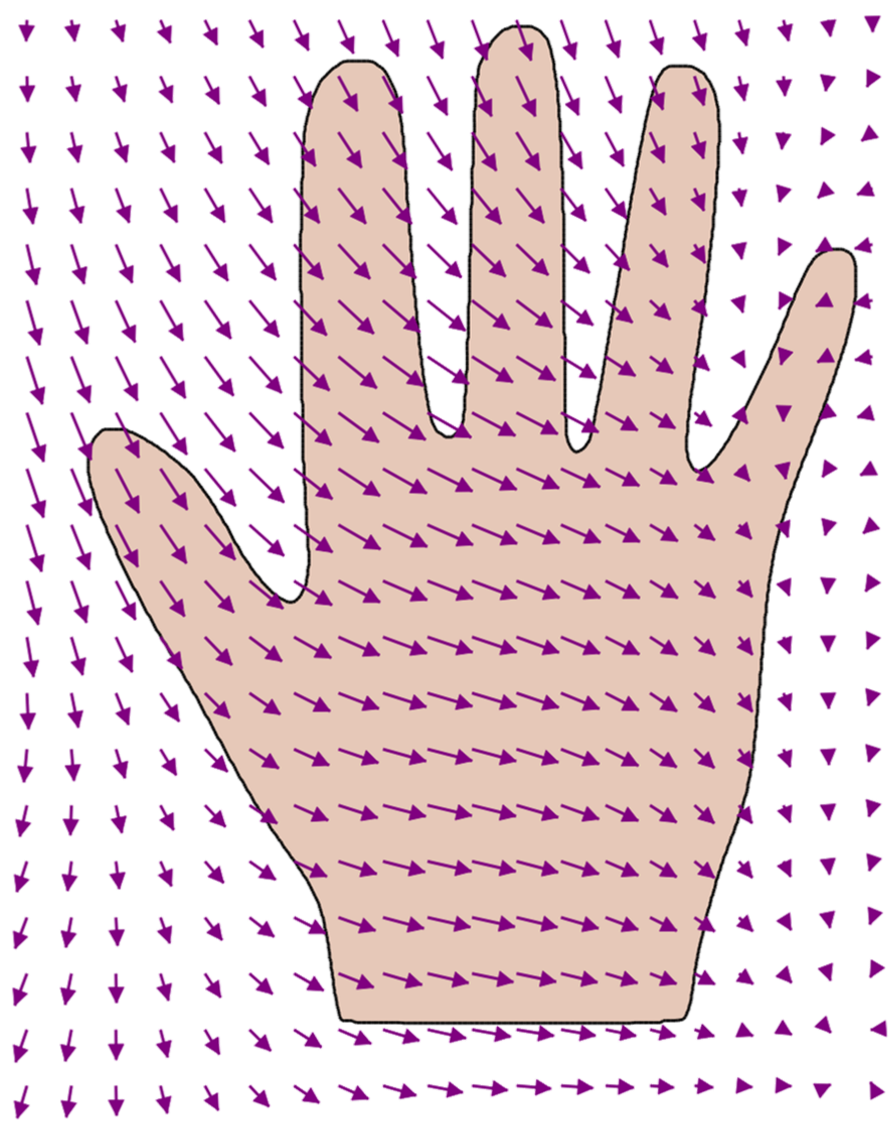
- the output-dimensions are uncorrelated.

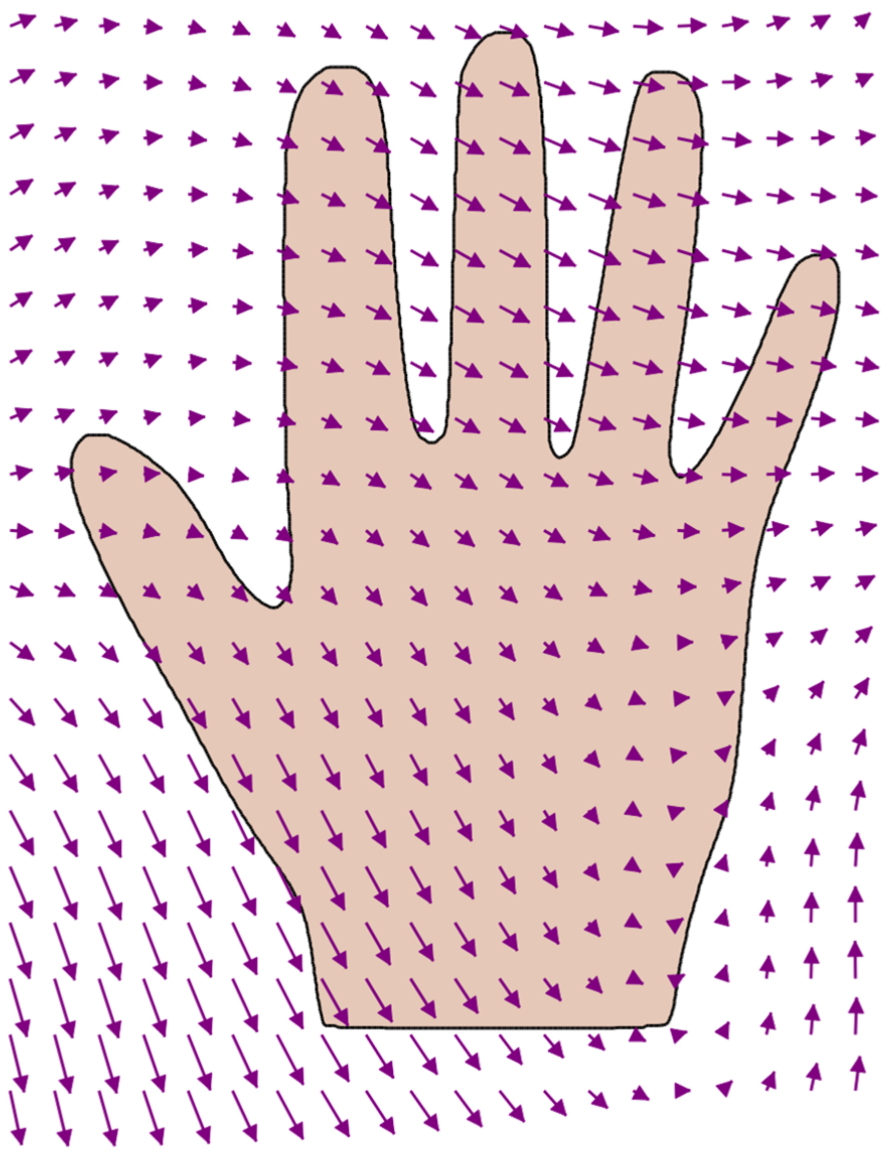
A model for smooth deformations

$$k(x, x') = \begin{pmatrix} s_1 \exp\left(-\frac{\|x - x'\|^2}{\sigma_1^2}\right) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & s_d \exp\left(-\frac{\|x - x'\|^2}{\sigma_d^2}\right) \end{pmatrix}$$

A model for smooth deformations

University of Basel

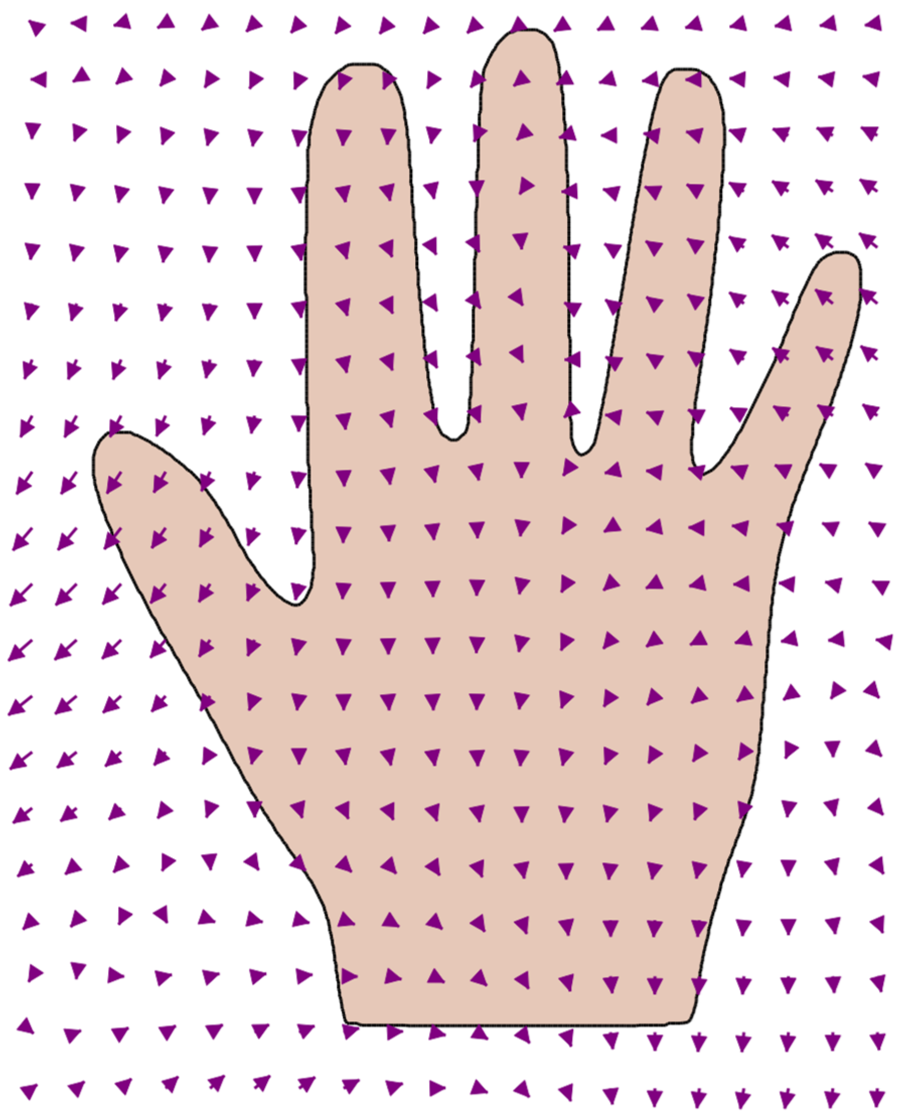


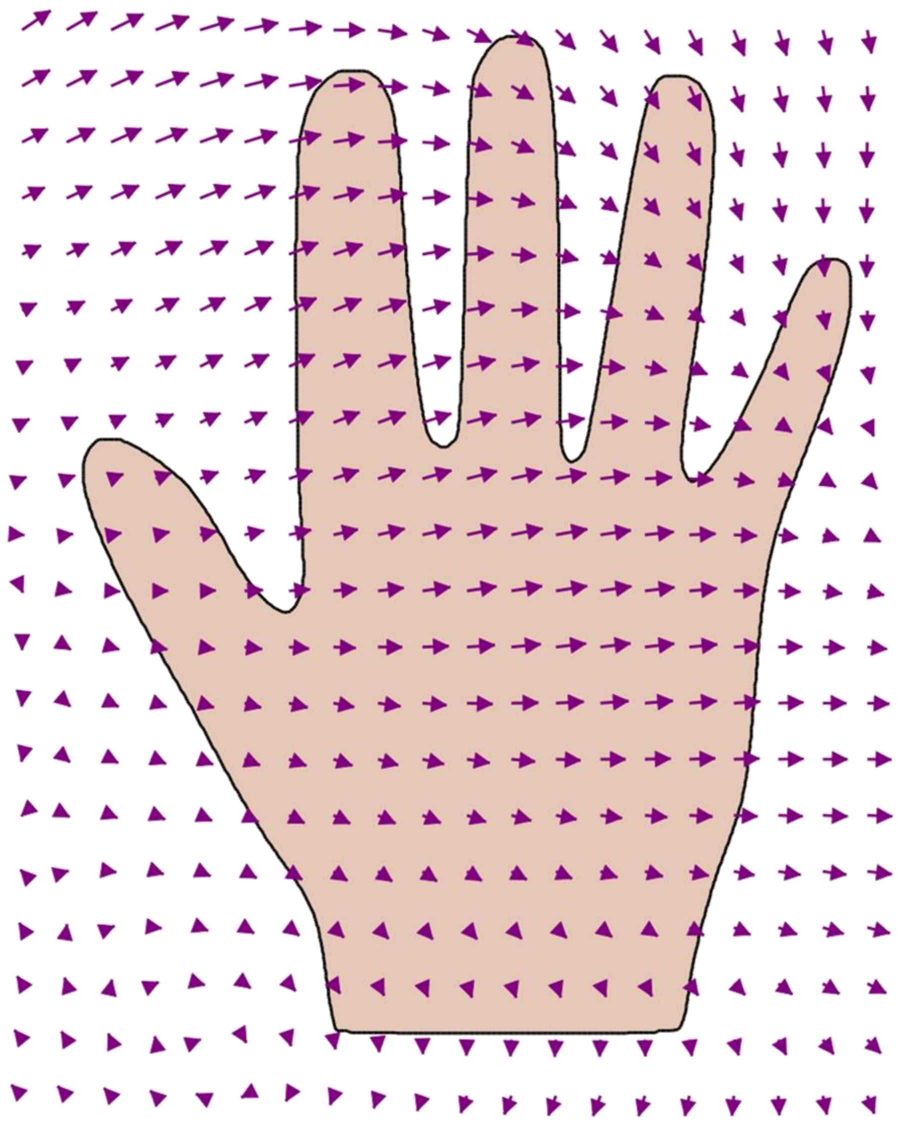


$s_1 = s_2$ large, $\sigma_1 = \sigma_2$ large

A model for smooth deformations

University of Basel

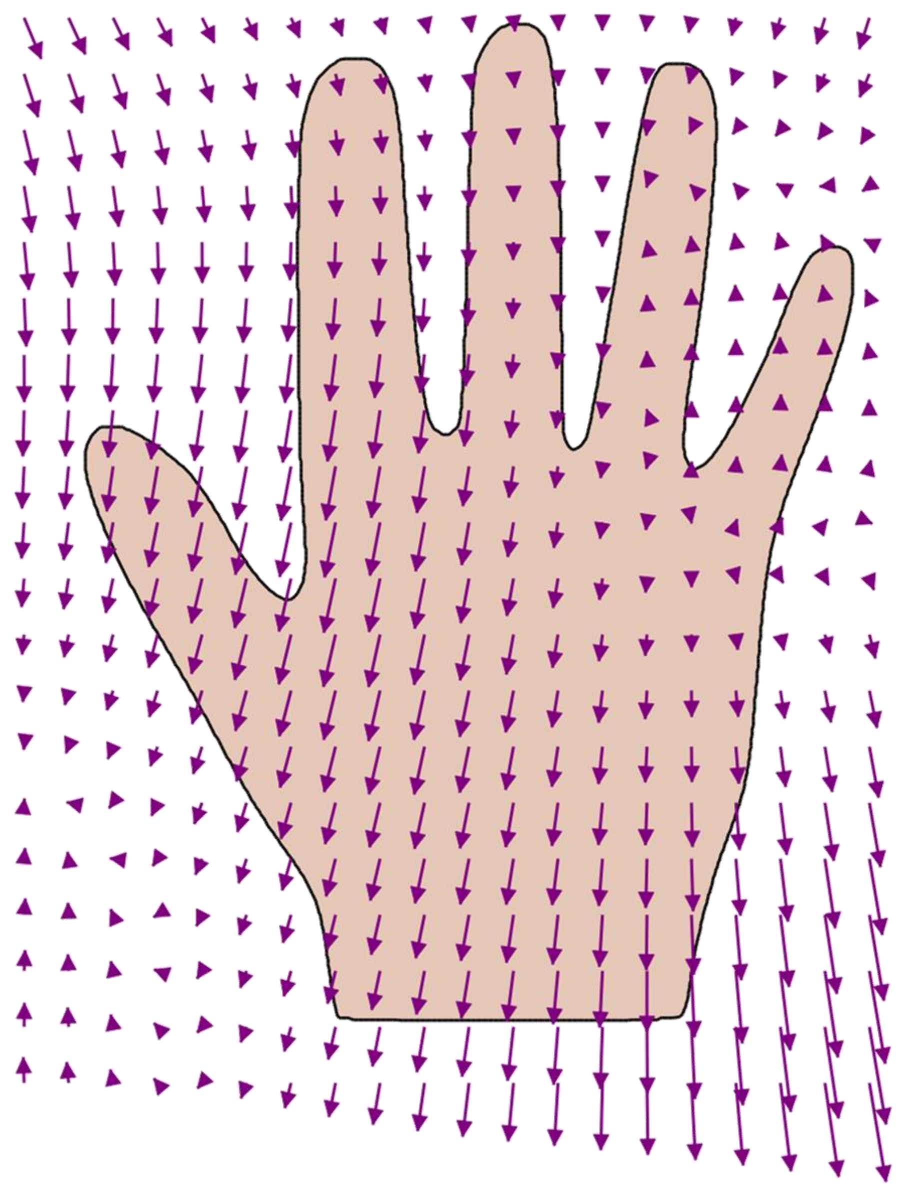


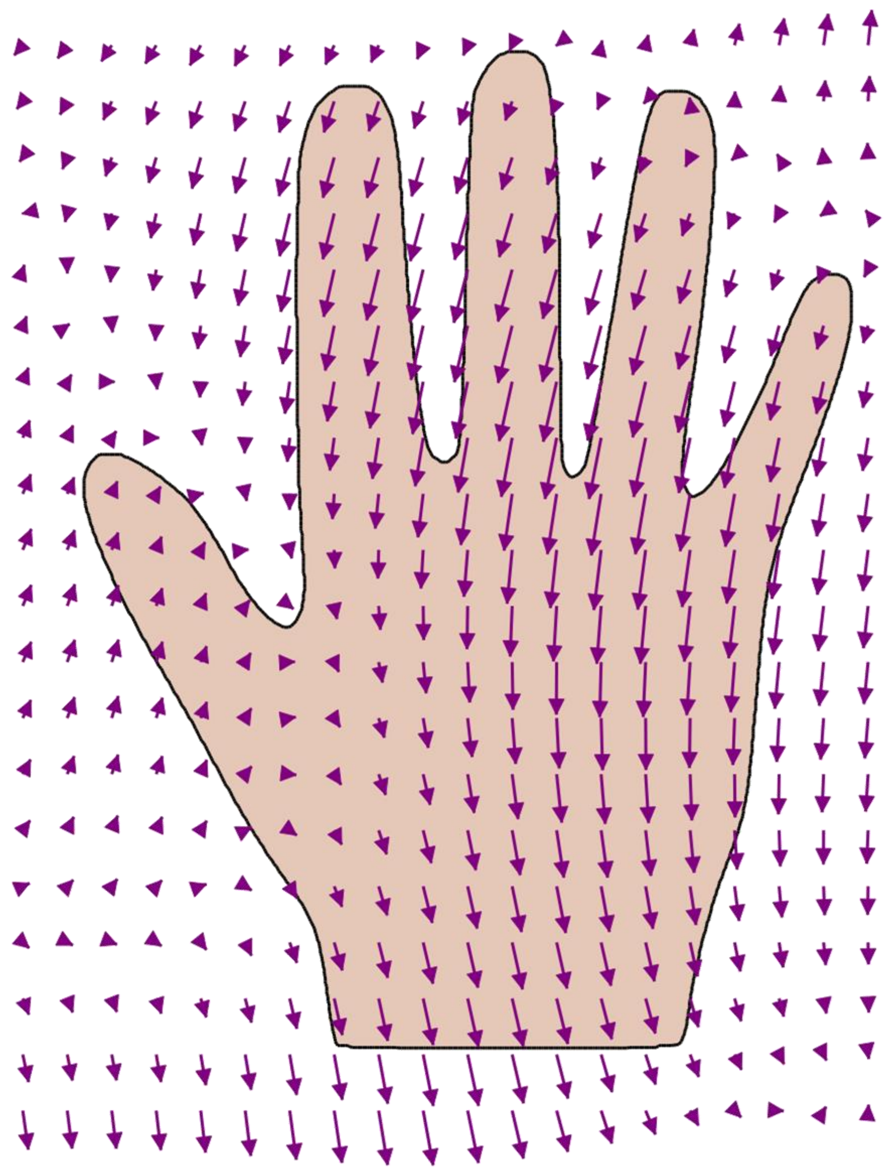


$s_1 = s_2$ small, $\sigma_1 = \sigma_2$ large

A model for smooth deformations

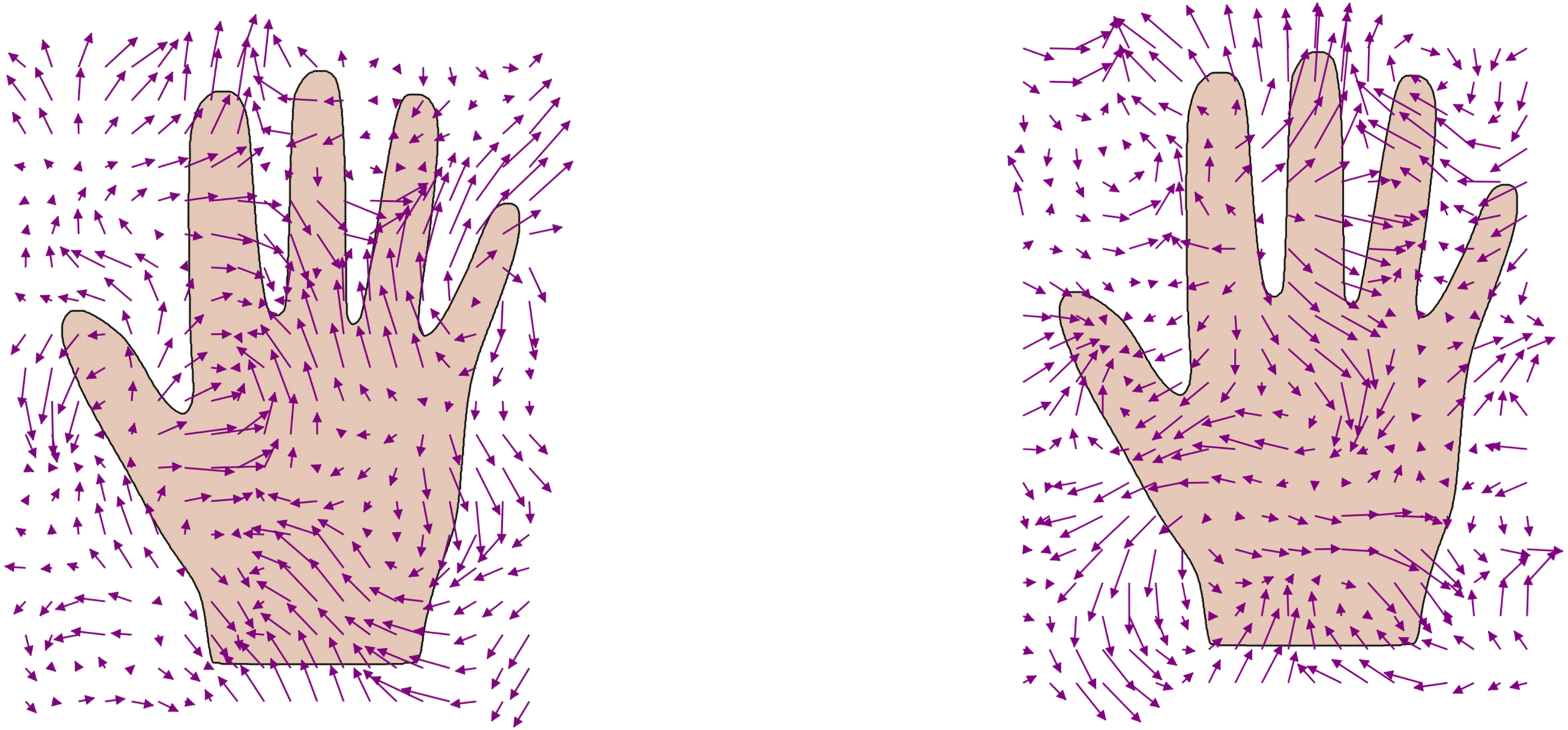
University of Basel





s_1 small, s_2 large, $\sigma_1 = \sigma_2$ large

A model for smooth deformations



$s_1 = s_2$ large, $\sigma_1 = \sigma_2$ small

Sum of two kernels

University of Basel

The sum of two positive semi-definite kernels is a positive semi-definite kernel

$$k(x, x') = g(x, x') + h(x, x')$$

- $k(x, x')$ is correlated if g(x,x') or h(x,x') are correlated
- Ideal for defining deformations on multiple scales

Product of two kernels

University of Basel

The (element-wise) product of two positive semi-definite kernels is a positive semi-definite kernel

$$k(x, x') = g(x, x') \odot h(x, x')$$

- $k(x, x')$ is correlated if g(x,x') and h(x,x') are correlated
- Ideal for localizing correlations

Other combinations of kernels

University of Basel

Rules for constructing kernels:

1. $k(x, x') = k_1(x, x') + k_2(x, x')$
2. $k(x, x') = \alpha k_1(x, x'), \alpha \in \mathbb{R}_+$
3. $k(x, x') = k_1(x, x') \odot k_2(x, x')$
4. $k(x, x') = f(x)\, f(x')^T, f: X \to \mathbb{R}^d$
5. $k(x, x') = B^T k(x, x') B, \quad \mathrm{B} \in \mathbb{R}^{r \times d}$

- These rules will be explored in the following article.