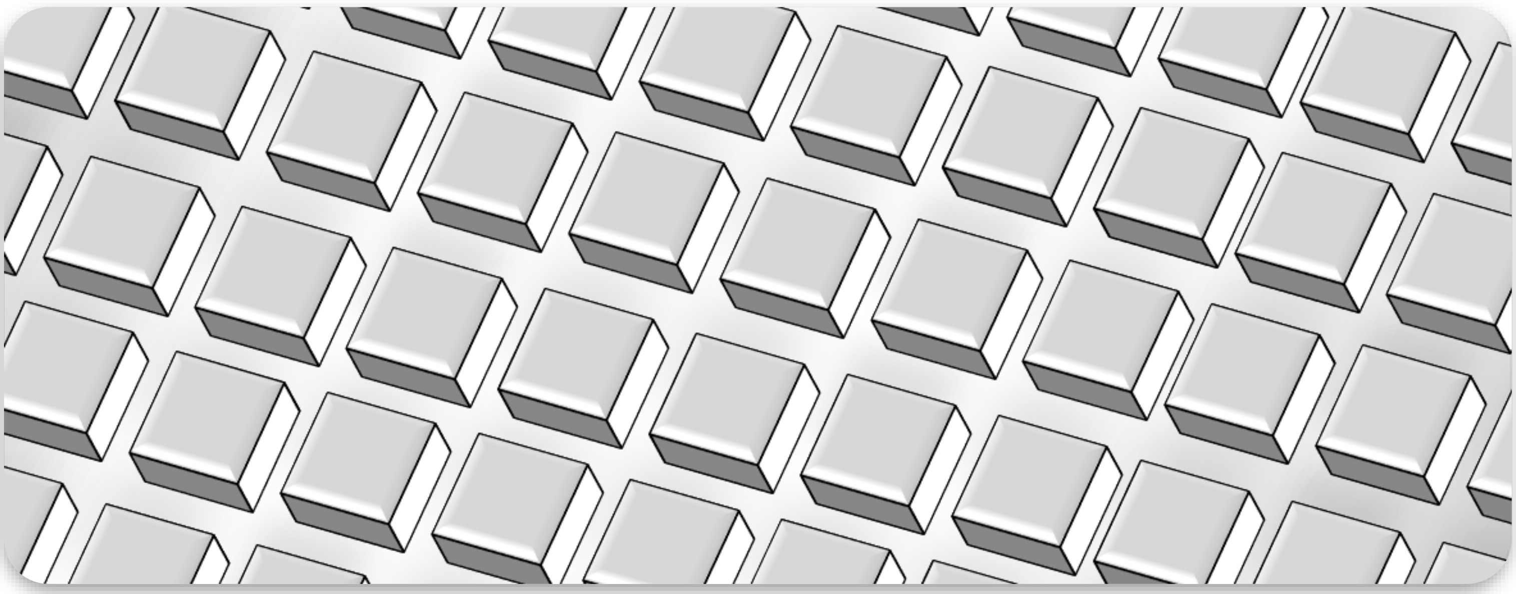
Modelling 2

STATISTICAL DATA MODELLING



Chapter 11
Symmetry

Michael Wand · Institut für Informatik, JGU Mainz · michael.wand@uni-mainz.de

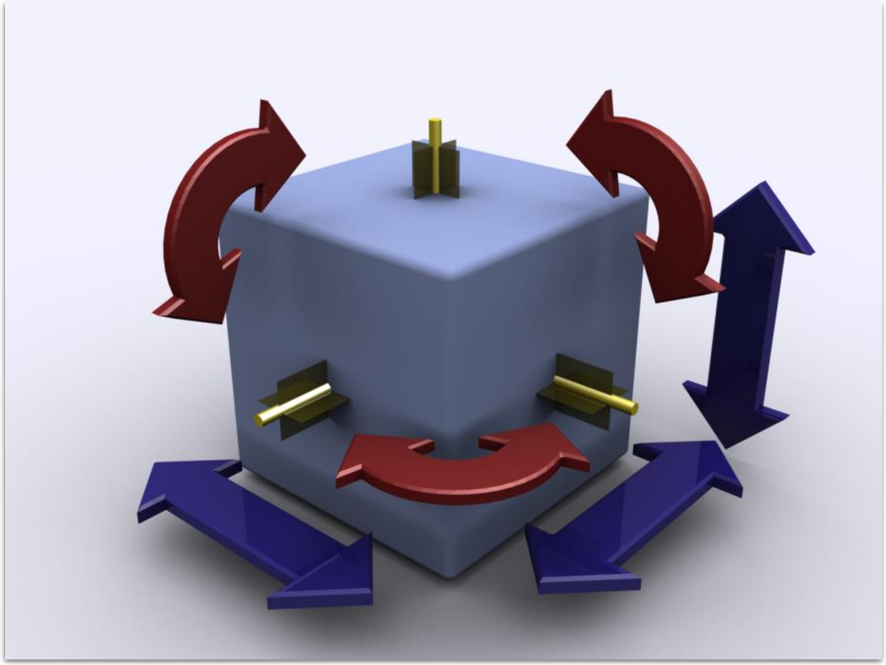
Video #10

Symmetry

- **Symmetry is the absence of information**
- **Group Theory**
- **Equivariance & Networks**

Symmetry is the absence of Information

Symmetry

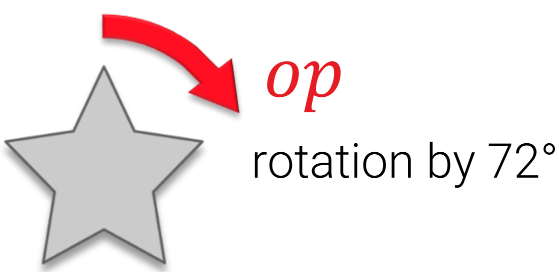


Symmetry

Intuition: Geometry

- Object $\mathcal{X} \subset \mathbb{R}^d$ (Geometry)

- Operations that do not change the geometry:



Symmetry

All operations

- Set of operations that do not change the object

- Rotation by $\{0°, 72°, 144°, 216°, 288°\}$

Symmetry

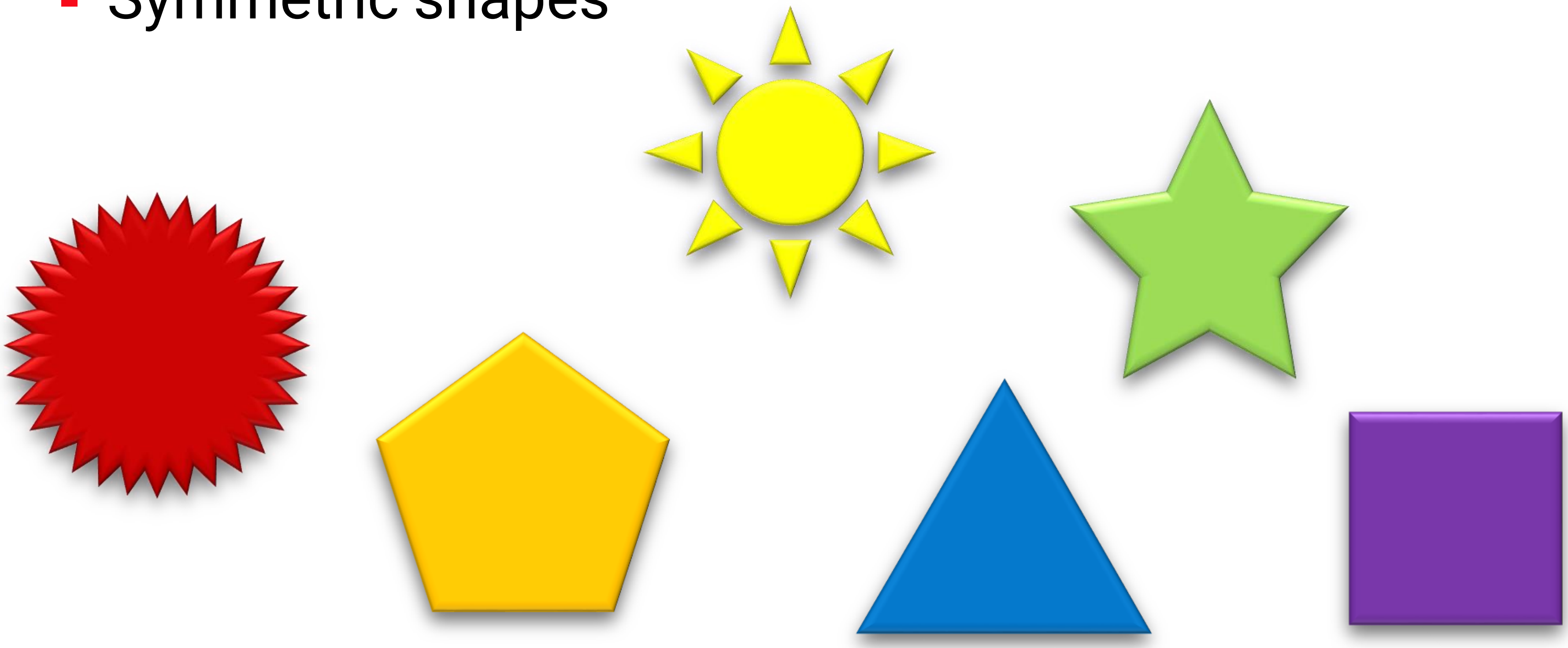
Geometric point of view

- *Symmetry* is the *absence of information*
- Rotation has no effect (on the subset of $\mathbb{R}^2$)
- Information „*does not exist*"

Examples

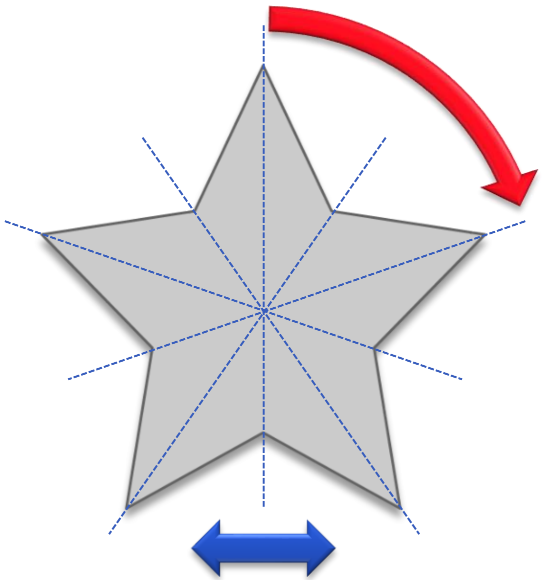
Geometry

- Symmetric shapes



More Examples

Star (2D)

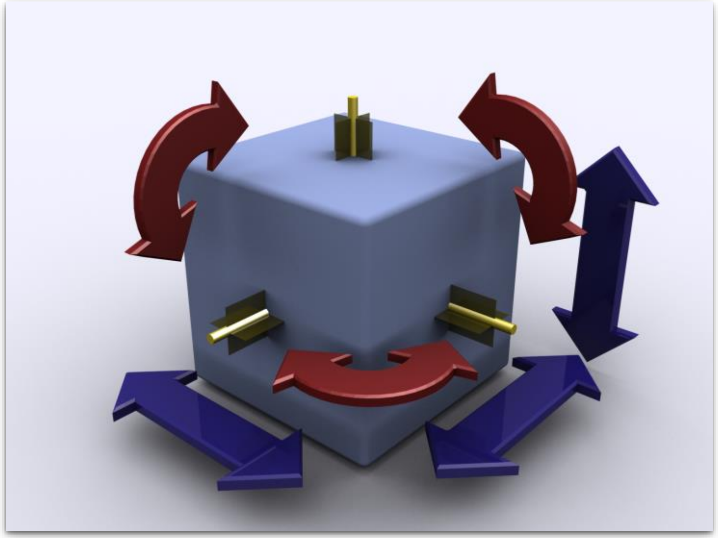


5 rotations R

10 rotations and reflections

$$R = \{0°, 72°, 144°, 216°, 288°\}$$

Cube (3D)

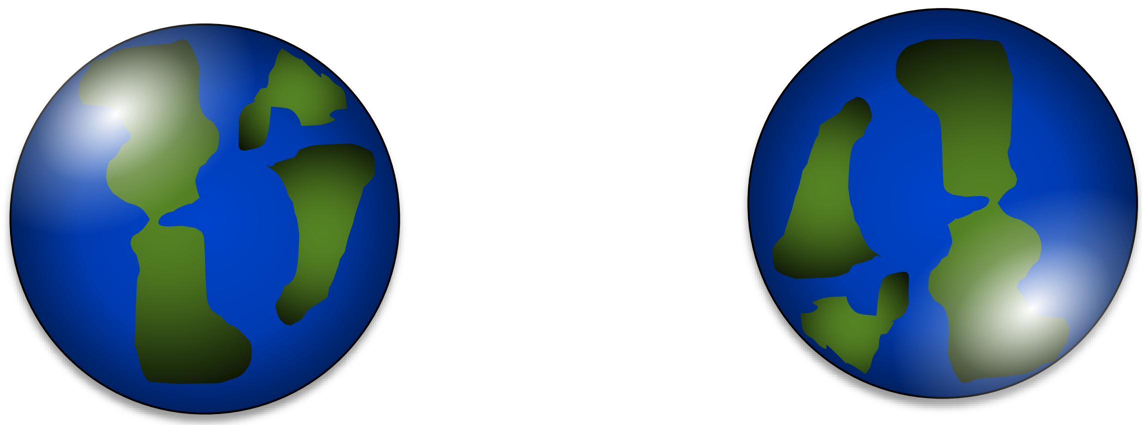


24 rotations

48 rotation + reflections

(Physical) Modelling

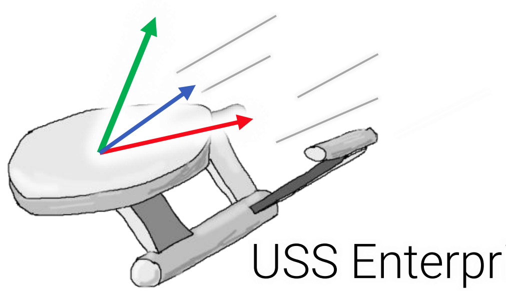
Symmetry in Nature



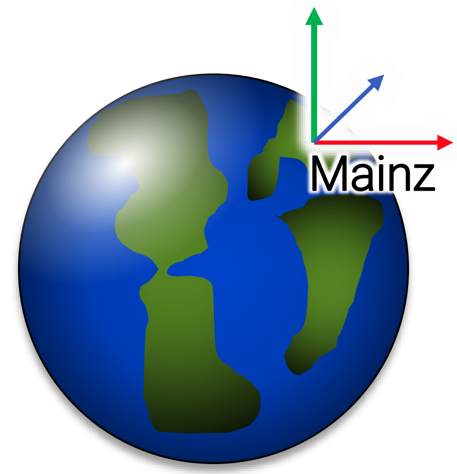
Invariance of Physical Laws

- Physical laws are symmetric
  - Rotations, translations(, reflections) do not matter
- “Galilean” invariance:
  - Choice of coordinate frame irrelevant

Symmetry in Physics



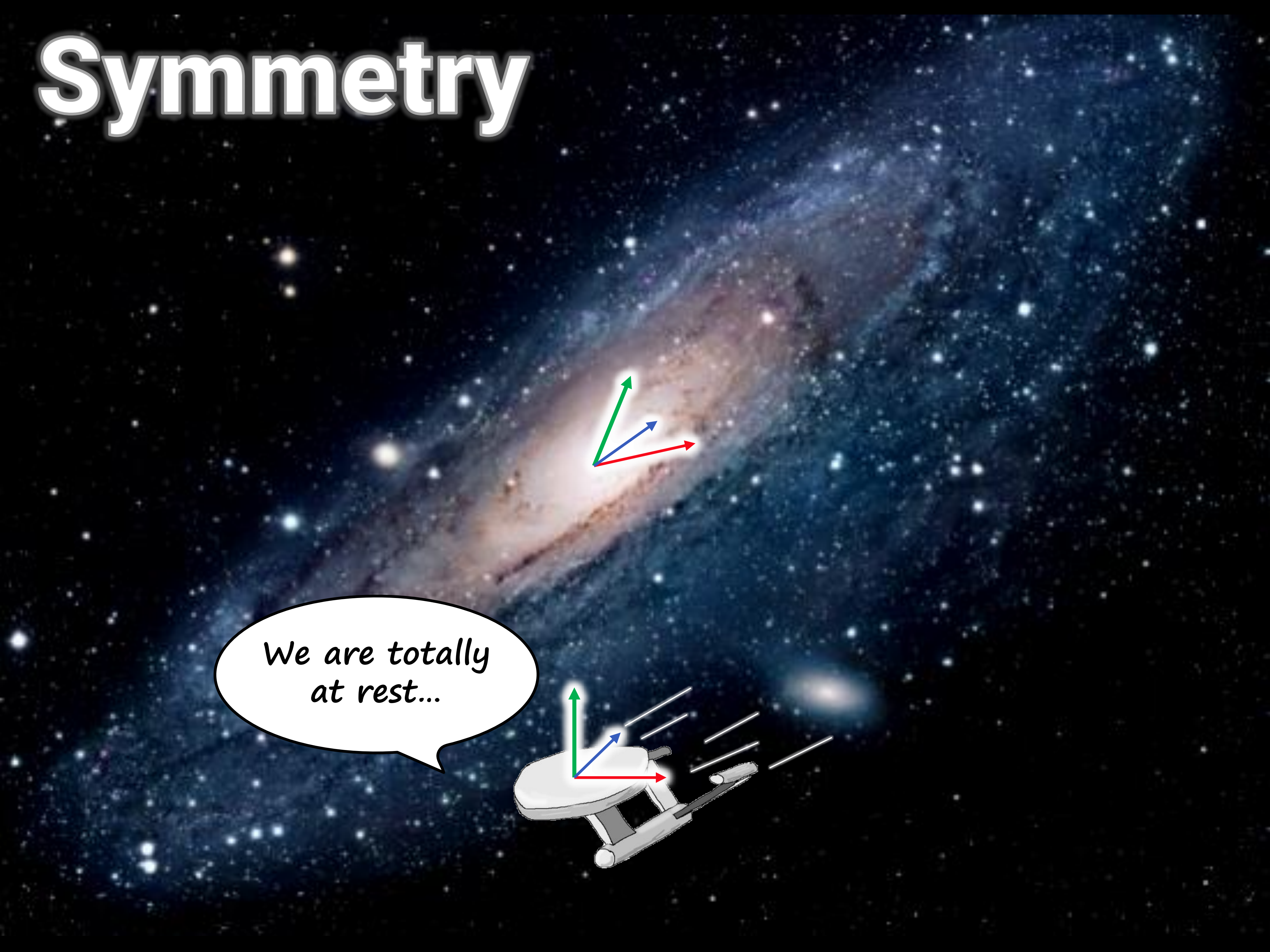
USS Enterprise,
space, the final frontier



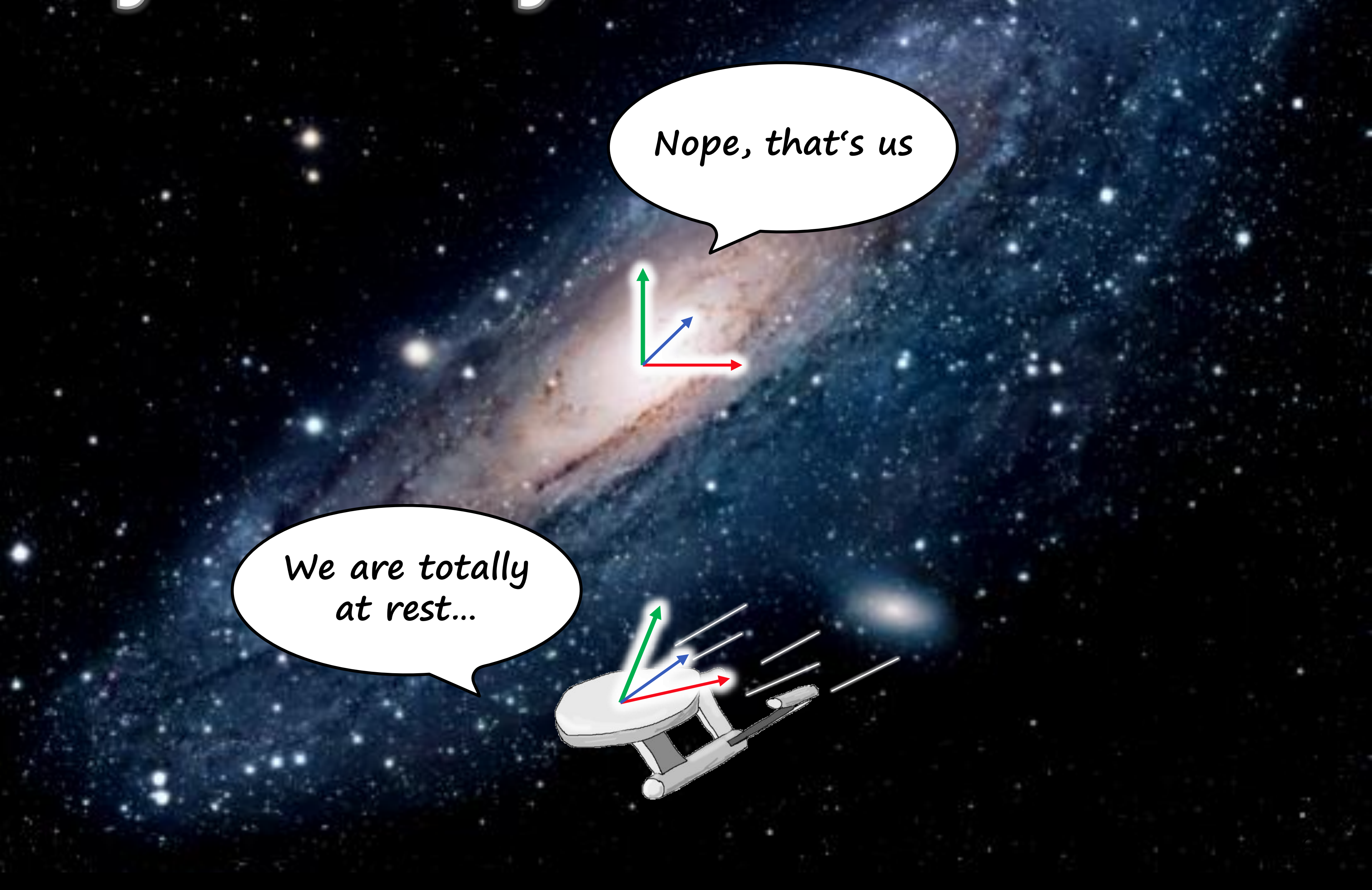
Relativity

- 4D Space-time symmetry: “Poincare group”
  - Rotations, translations, “boosts”
  - Time is different from space (Minkowski-space)
- Change of velocity leaves physical laws unchanged
  - Including *speed of light* (information propagation)
  - There is no absolute velocity

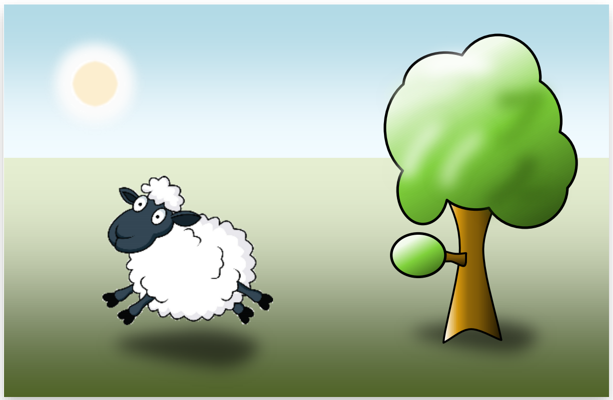
Symmetry

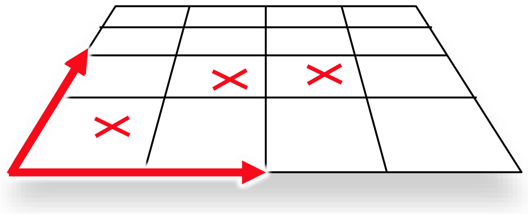


Symmetry

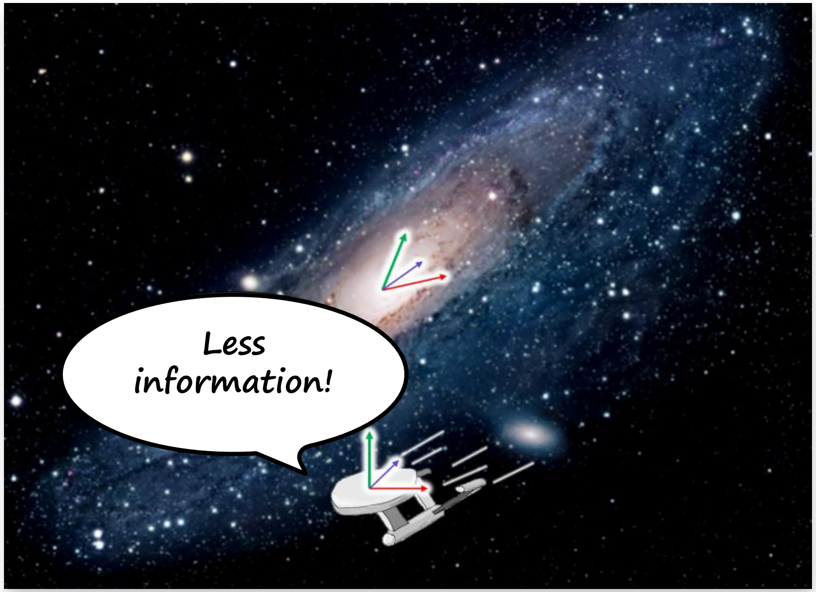


Redundant Model



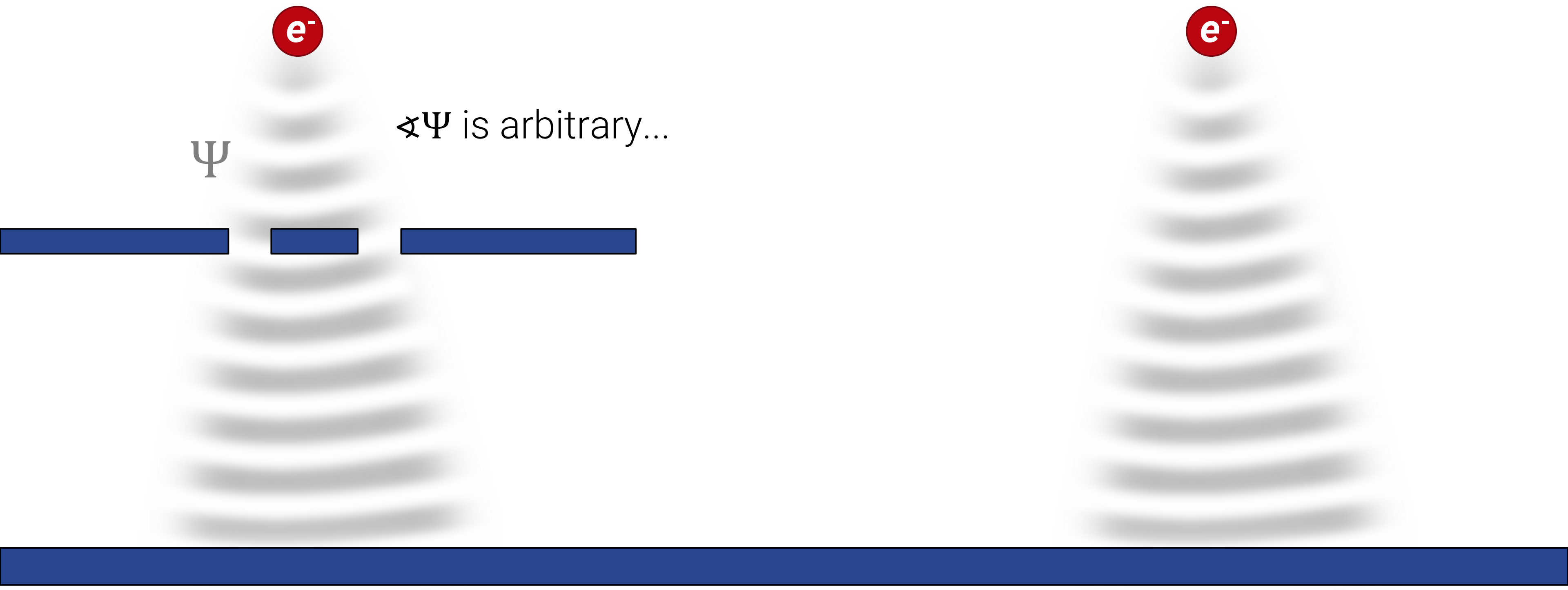


Evolutionary Concept
Absolute reference points

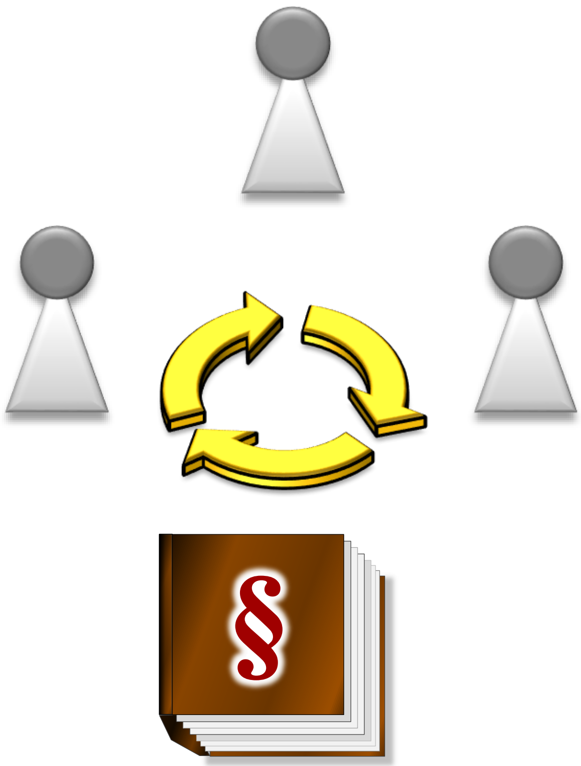


New Model
No absolute reference

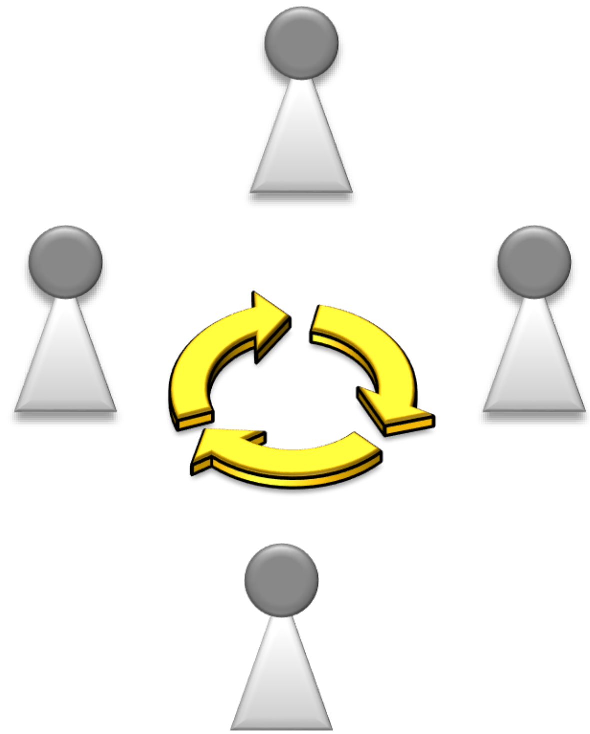
Symmetry in Quantum Physics



Social Sciences*)



Democracy



Communism

*) *do not take this too seriously…*

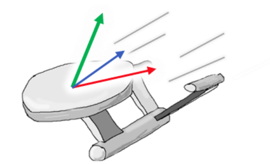
Summary

Symmetry

Symmetry

- Object remains invariant under transformations
- Information changed by transformation is irrelevant
  - Concretely (5-fold symmetric "Star")

  - Semantically (choice of coordinate system)

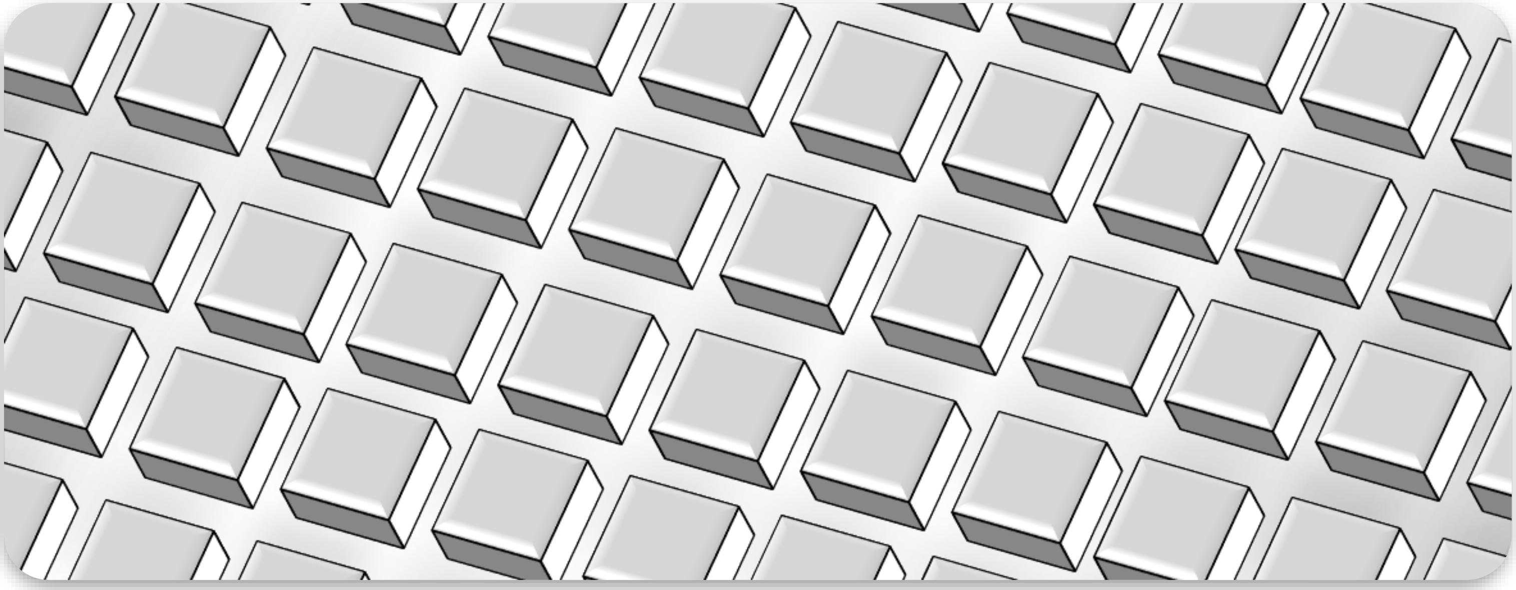


Symmetry in Empirical Science

- The same law applies in different scenarios
- Symmetry = invariance
  - Or equivariance, as we will see later

Modelling 2

STATISTICAL DATA MODELLING



Chapter 10
Symmetry

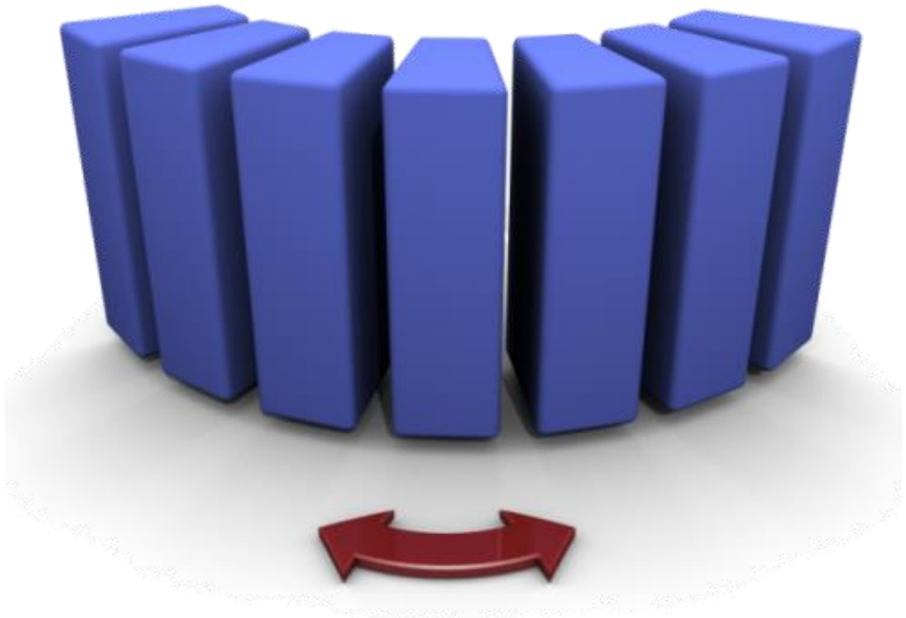
Michael Wand · Institut für Informatik, JGU Mainz · michael.wand@uni-mainz.de

Video #10

Symmetry

- Symmetry is the absence of information
- **Group Theory**
- Equivariance & Networks

Mathematical Model



Definition

Group Axioms

- Closed:
  Set G, closed mapping

$$\text{“}\circ\text{”}: G, G \rightarrow G$$

- Associative:

$$(g_1 \circ g_2) \circ g_3 = g_1 \circ (g_2 \circ g_3)$$

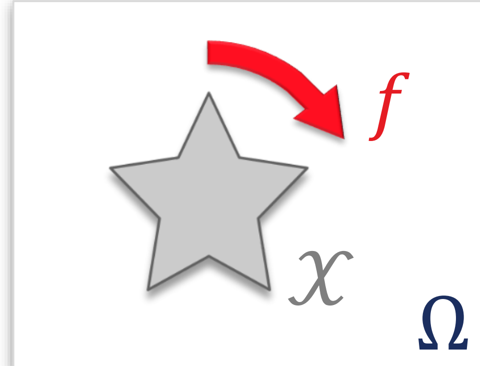
- Neutral element:

$$id \in G : id \circ g = g \circ id = g$$

- Inverse: For each $g \in G$ exists an $g^{-1} \in G$:

$$g \circ g^{-1} = g^{-1} \circ g = id$$

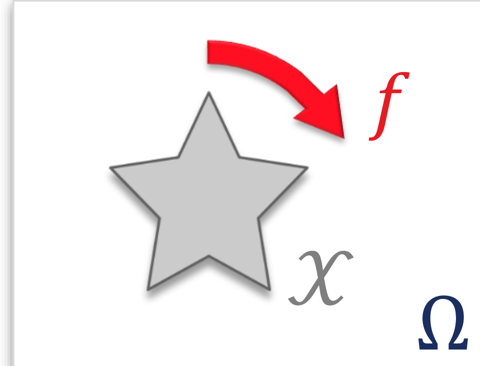
Symmetry



Formalization of symmetry

- Transformations of Domain Ω
  - $f: \Omega \to \Omega$, f bijective
- Object $\mathcal{X} \subseteq \Omega$
- Set G of operation f that leave $\mathcal{X}$ intact
  - $f(\mathcal{X}) = \mathcal{X} \quad \Rightarrow \quad f \in G$
- Consider all concatenations of such operations
  - Such as $f \circ g \circ h \circ f \circ h$ with $f, g, h \in G$
- The set G of operations forms a group wrt. "$\circ$"
  - $G \coloneqq \{f: \Omega \to \Omega | f(\mathcal{X}) = \mathcal{X}\}$
  - Operation "$\circ$" = function concatenation

Transformation Groups



Transformation Group

- $G = \{f : \Omega \to \Omega \mid f(\mathcal{X}) = \mathcal{X}\}$
- G is a group
  - Algebraically closed: $[f, g \in G] \Rightarrow [f \circ g \in G]$
  - Associative (trivial!): $(f \circ g) \circ h = f \circ (g \circ h)$
  - Neutral element: $id \in G \;\; (id(\mathcal{X}) = \mathcal{X})$
  - Inverse element: $[f \in G] \Rightarrow [f^{-1} \in G]$

General Groups? Group Actions!

- Group not made of transformations
  - Associate $g \in G$ with transformation of something else
- Example: $\mathbb{Z} \bmod 5$ and 72° rotations

Transformation Groups and General Groups

Symmetry & Group Theory

Discrete Groups

Example

- Addition in $\mathbb{N}$ modulo 5
- Multiplication in $\mathbb{N}^+$ modulo 5

“Cayley tables” (multiplication tables)

+	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

*	1	2	3	4
1	1	2	3	4
2	2	4	1	3
3	3	1	4	2
4	4	3	2	1

Permutation Groups

Permutations (Bijections)

- Swapping element of a set S
- Set $\Pi(S)$ of all permutations of a set S forms a group

Theorem

- All groups are isomorphic to a permutation group
  - Bijection group in the ∞-case
- Proof: multiplication table as representation
  - (finite case)

Structural Insight

Groups $\equiv$ **Symmetry** (Transformation groups)

„$\Leftarrow$“

- Symmetry transformations
  - Fully Information preserving
  - Always applicable
- Always a group structure

„$\Rightarrow$“

- Groups isomorphic to permutation of sets
- Most abstract notion of transformations/symmetry

Important Transformation Groups

Names for transformation groups

- $GL(d)$ – invertible linear maps in $\mathbb{R}^d$
  - invertible matrices:
  - $\mathbf{x} \mapsto \mathbf{A}\mathbf{x},\ \det \mathbf{A} \neq 0$
- $O(d)$ – orthogonal transforms in $\mathbb{R}^d$
  - reflections, rotations:
  - $\mathbf{x} \mapsto \mathbf{A}\mathbf{x},\ \mathbf{A} = \mathbf{A}^{\mathrm{T}}$
- $E(d)$ – isometries of the Euclidean space $\mathbb{R}^d$
  - translations, reflections, rotations
  - $\mathbf{x} \mapsto \mathbf{A}\mathbf{x} + \mathbf{t},\ \mathbf{A} = \mathbf{A}^{\mathrm{T}}$
- Prefix "S" removes reflections: $SO(d)$, $SE(d)$
  - $\det \mathbf{A} > 0$

Equivalence Classes

– ignoring stuff –

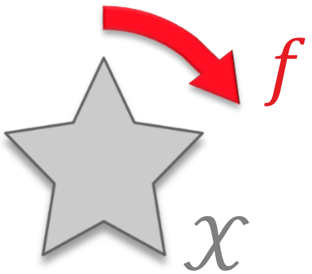
Equivalence Classes

Modeling "irrelevant" information

- Transformation group

$$G := \{f \in \mathcal{T} \,|\, f: \Omega \to \Omega\}$$

- Transformations might change $\mathcal{X}$
  - But in aspects we do not care about



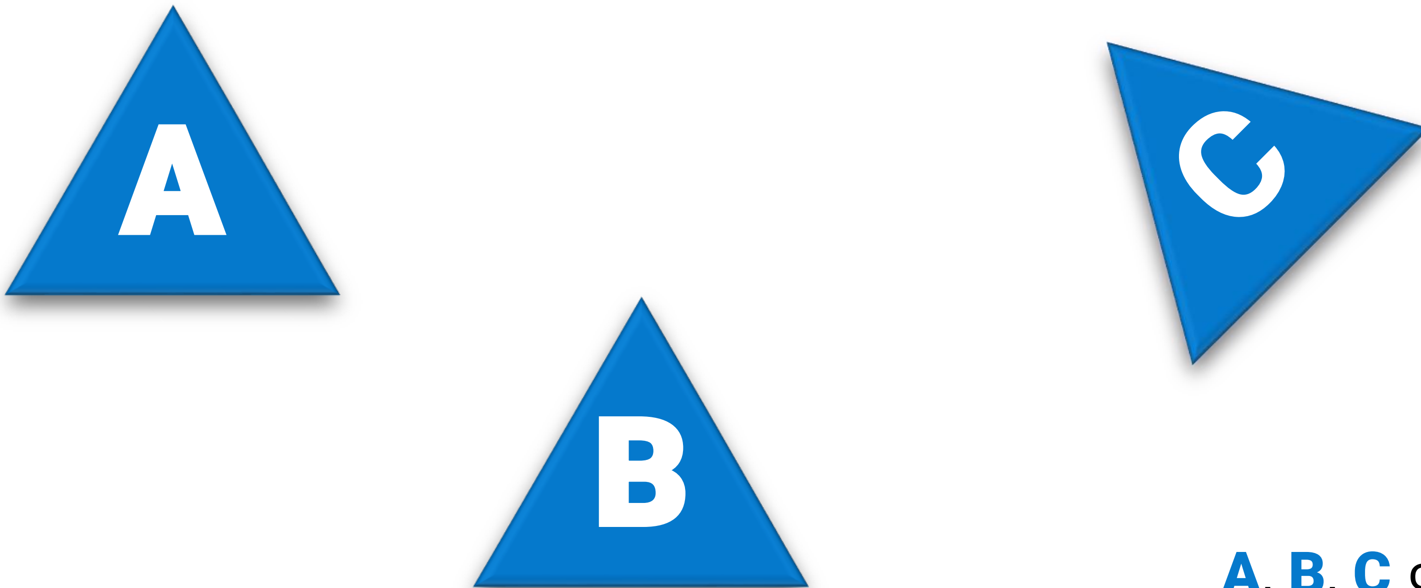
- Ignore similar states

$$\mathcal{Y} \equiv \mathcal{X} \text{ if and only if } \exists f \in G: \mathcal{Y} = f(\mathcal{X})$$

- Written as

$$\mathcal{Y} = \mathcal{X} \bmod G$$

Example



$\mathbf{A}, \mathbf{B}, \mathbf{C} \subset \mathbb{R}^2$

$$\mathbf{A} \equiv \mathbf{B} \equiv \mathbf{C} \bmod SE(2)$$

(rigid copies)

Generators

– building groups –

Generators for Groups

Subgroups

- Subsets of groups that are groups
  - I.e., subset is algebraically closed

(Discretely) Generated Groups

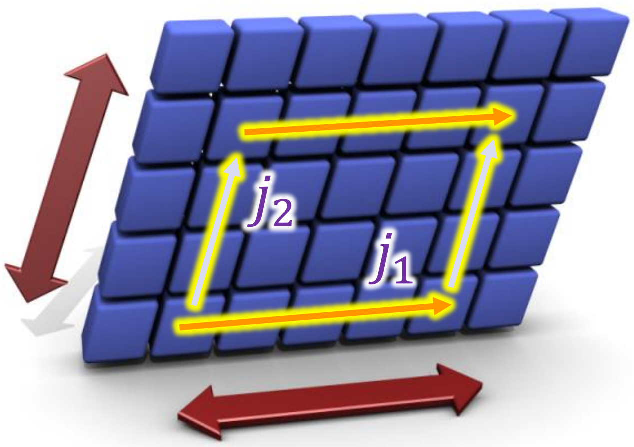
- Let G be a group, and $g_1, \dots, g_n \in G$
- The set of all objects

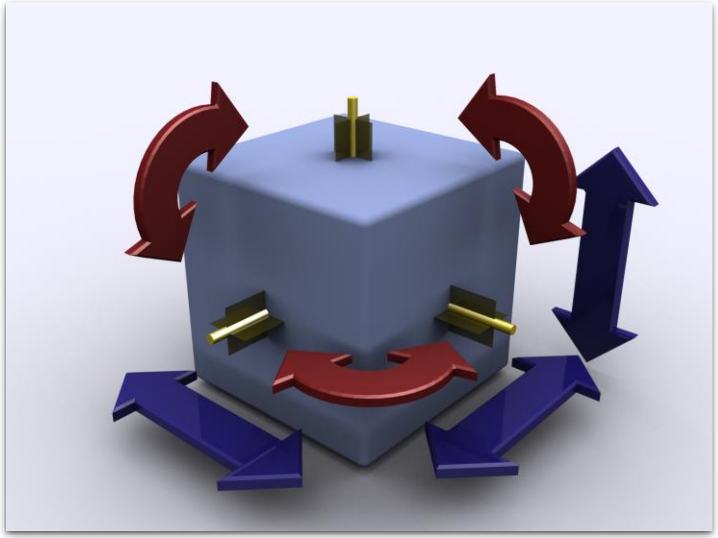
$$\langle g_1, \dots, g_n \rangle := g_{i_1}^{j_1} \circ \cdots \circ g_{i_k}^{j_k},$$

$$i_1, \dots, i_k \in \{1, \dots, n\}, j_1, \dots, j_k \in \mathbb{Z}$$

is called the sub group of G *generated* by $g_1, \dots, g_n$

To commute or not to commute…





$$g_1 \circ g_2 = g_2 \circ g_1$$

not commutative

„flat“ grid

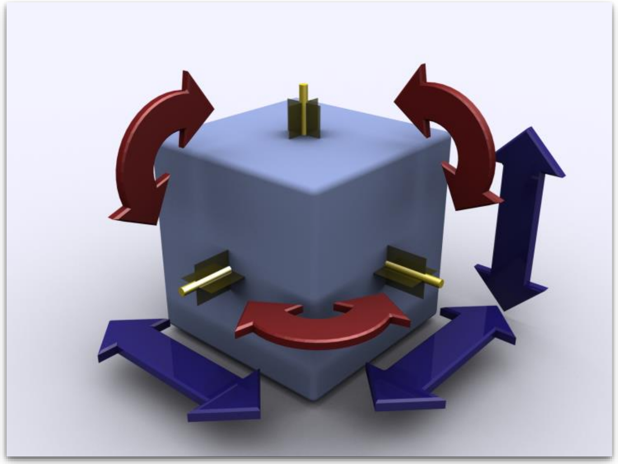
$$\underbrace{g_1^{j_1} \circ \cdots \circ g_k^{j_k}}$$

can be sort: element coordinates

Generators

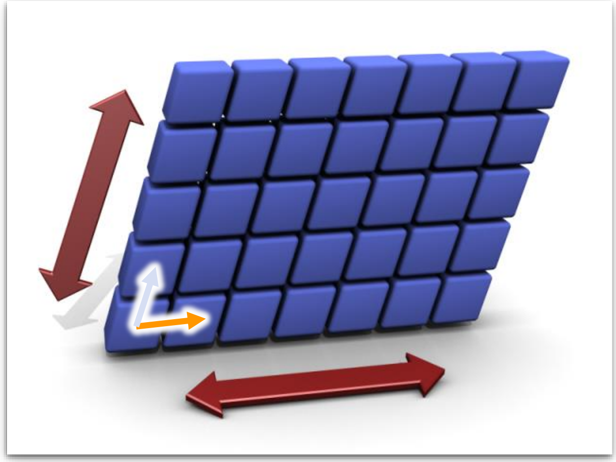
Example

- Group $SO(3)$
- Subgroup: $O_h \subset SO(3)$
- Example generators:
  x-rotation 90°, y-rotation 90°,
  z-rotation 90°, x-reflection

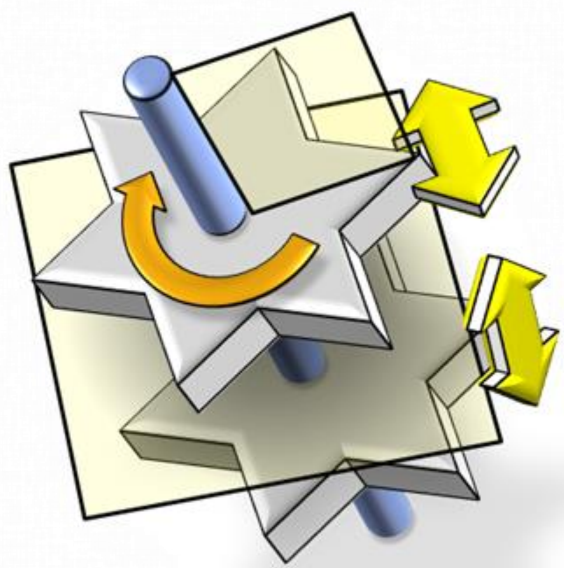


Commutative example

- Two translations
- Group represented as a grid

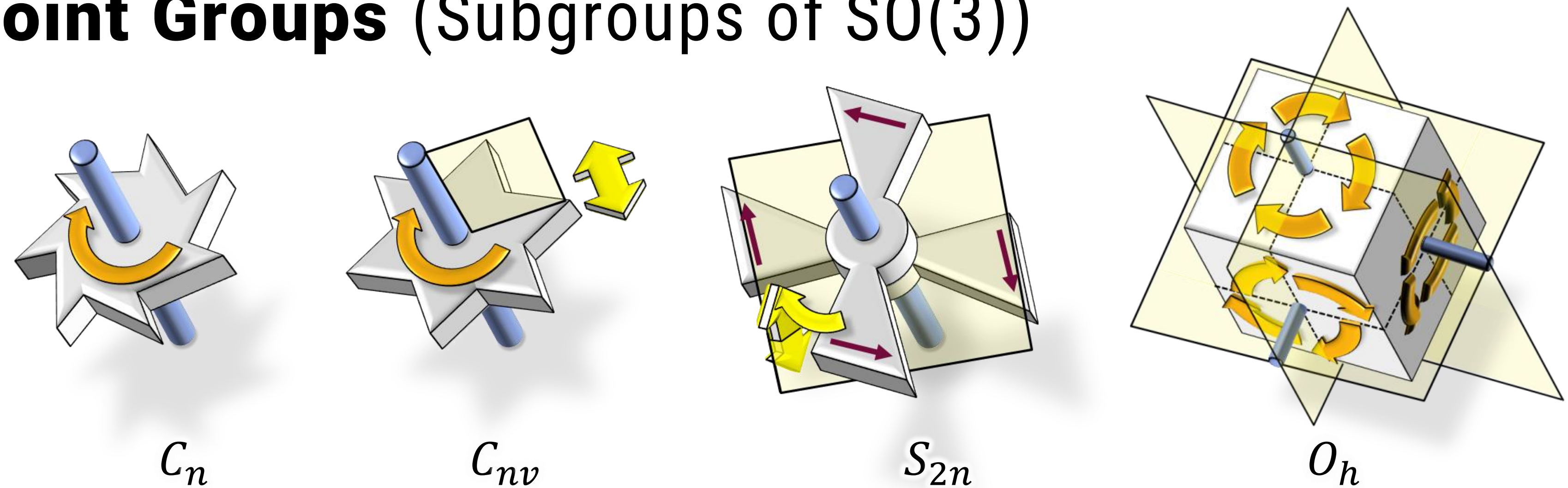


Example: Christallographic Groups



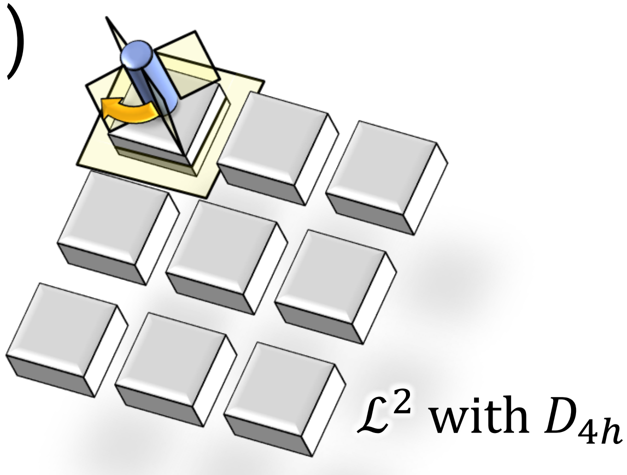
Euclidean Symmetry Groups

Point Groups (Subgroups of SO(3))



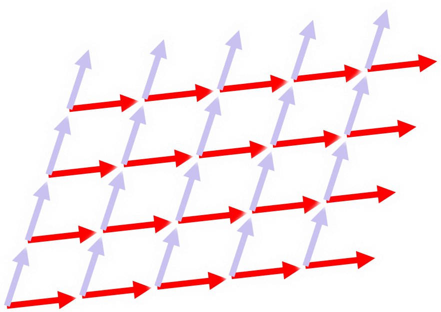
(Christallographic) Lattices (E(3))

- Translations
  - 1 Translation in 1D
  - Up to 2 in 2D, up to 3 in 3D
- Combination w/point group

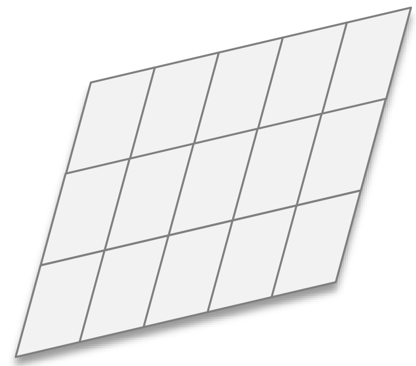


Wallpaper Groups (2D Chrystals)

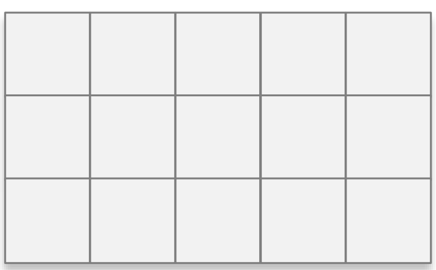
Building Blocks: 2D Grids (Transl. Lattice)



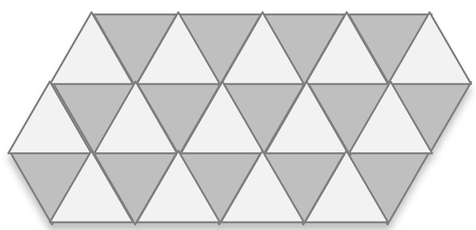
$t_1^i \circ t_2^j$



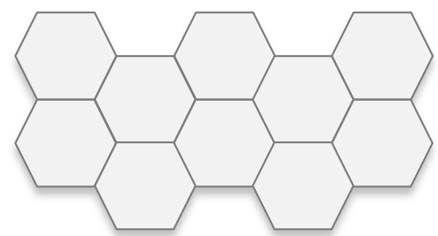
Combination with Rotations / Reflections



2-fold / 4-fold



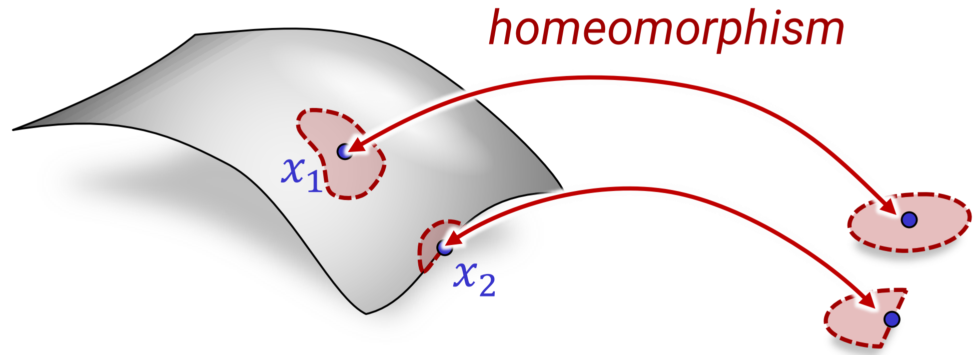
3-fold



6-fold

Lie-Groups

Manifold



Now considering

- Groups that are also manifolds

Generators for Groups

(Continuously) **Generated Groups**

- Let G be a group, and $g_1, \dots, g_n \in G$
- The set of all objects

$$\langle g_1, \dots, g_d \rangle := g_1^{x_1} \circ \cdots \circ g_d^{x_d},$$

$$x_1, \dots, x_d \in \mathbb{R}$$

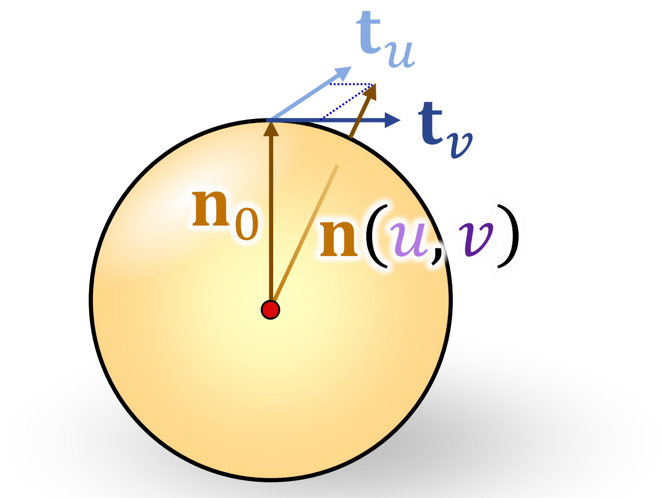
  is called the sub group of G *generated* by $g_1, \dots, g_n$
- Vector space $\mathbb{R}^d$ of vectors $(x_1, \dots, x_k)$

(Think of $g_1, \dots, g_n$ as matrices)

Manifolds & Tangent Spaces

Local Parameterization of a manifold

- Point on a sphere
- Tangent parameterization
- New variables *u*, *v*



Walk on manifold?

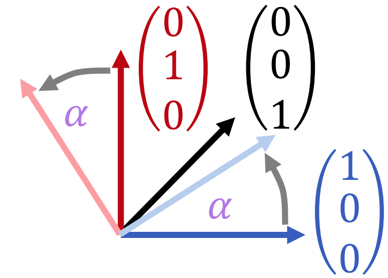
- Walk in tangent space
- Normalize result
  - Project on manifold
- Example application:
  - Non-linear optimization

$$\mathbf{n}'(u, v) = \mathbf{n}_0 + u \cdot \mathbf{t}_u + v \cdot \mathbf{t}_v$$

$$\mathbf{n}(u, v) = \frac{\mathbf{n}'(u, v)}{\|\mathbf{n}'(u, v)\|}$$

[Hoffer et al. 04]

Example: Rotations in SO(3)



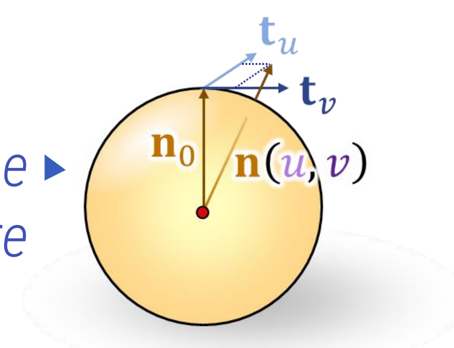
Rotation group SO(3)

$$\mathbf{R}_{\alpha,\beta,\gamma} = \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ \sin\beta & 0 & \beta \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\gamma & -\sin\gamma \\ 0 & \sin\gamma & \cos\gamma \end{pmatrix}$$

First order Taylor (Tangent Vectors)

$$\nabla_{\alpha,\beta,\gamma}\mathbf{R}_{\alpha,\beta,\gamma} = \begin{pmatrix} 0 & \alpha & \beta \\ -\alpha & 0 & \gamma \\ -\beta & -\gamma & 0 \end{pmatrix}$$

c.f: tangent space of the sphere ▶



$$\mathbf{n}(u,v) = \mathbf{n}_0 + u \cdot \mathbf{t}_u + v \cdot \mathbf{t}_v$$

Example: Rotations in SO(3)

Taylor Expansion

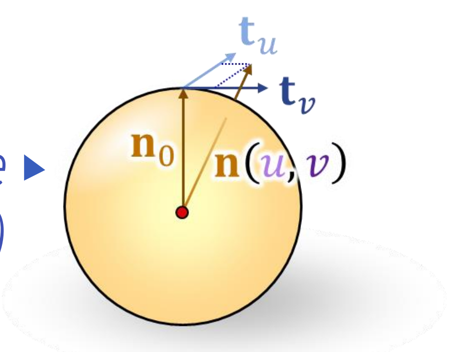
$$\mathbf{R}_{\lambda\alpha,\lambda\beta,\lambda\gamma} \doteq \mathbf{I} + \lambda \nabla_{\alpha,\beta,\gamma} \mathbf{R}_{\alpha,\beta,\gamma}$$

Exponential Map (geodesic in tangent direction)

$$\mathbf{R}_{\lambda\alpha,\lambda\beta,\lambda\gamma} = \exp\left(\mathbf{I} + \lambda \nabla_{\alpha,\beta,\gamma} \mathbf{R}_{\alpha,\beta,\gamma}\right) = \begin{pmatrix} 1 & \alpha & \beta \\ -\alpha & 1 & \gamma \\ -\beta & -\gamma & 1 \end{pmatrix}^{\lambda}$$

- Geodesic in tangent direction
  - Equivalent: Matrix exponential by Taylor-series

c.f: sphere (tangent space) ▸



$$\mathbf{n}(u,v) = \mathbf{n}_0 + u \cdot \mathbf{t}_u + v \cdot \mathbf{t}_v$$

Lie-Groups

Lie Algebra

- Vector space of “tangent directions”
  - Same structure at every $\mathbf{T} \in \mathbf{G}$ (symmetry)
  - Linear combinations of “directions”

Lie Group

- Continuous manifold contains group elements
- Exponential map
  - Follow the path on the manifold
- Theorem
  - Compact, connected Lie groups: exponential map surjective
  - Maps of vector space reach ”the whole group”

Summary

Brief Excerpt of Group Theory

Understanding symmetry

- Symmetry Theory “=“ Group Theory

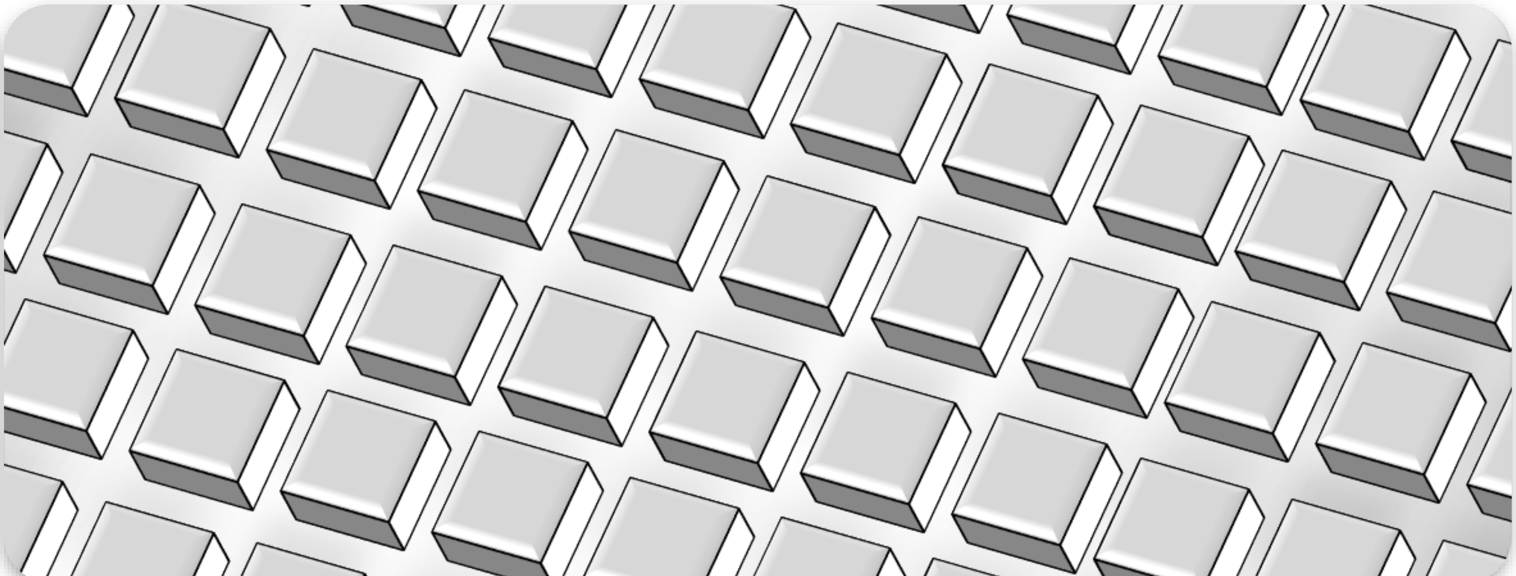
Mathematical structures

- Definition of groups
  - Group actions if group elements are not explicitly transformation
- Taken quotients & equivalence classes
- Generators, structure of commutative groups
- Lie Groups: Continuously generated groups

Next: Equivariance

Modelling 2

STATISTICAL DATA MODELLING



Chapter 10
Symmetry

Michael Wand · Institut für Informatik, JGU Mainz · michael.wand@uni-mainz.de

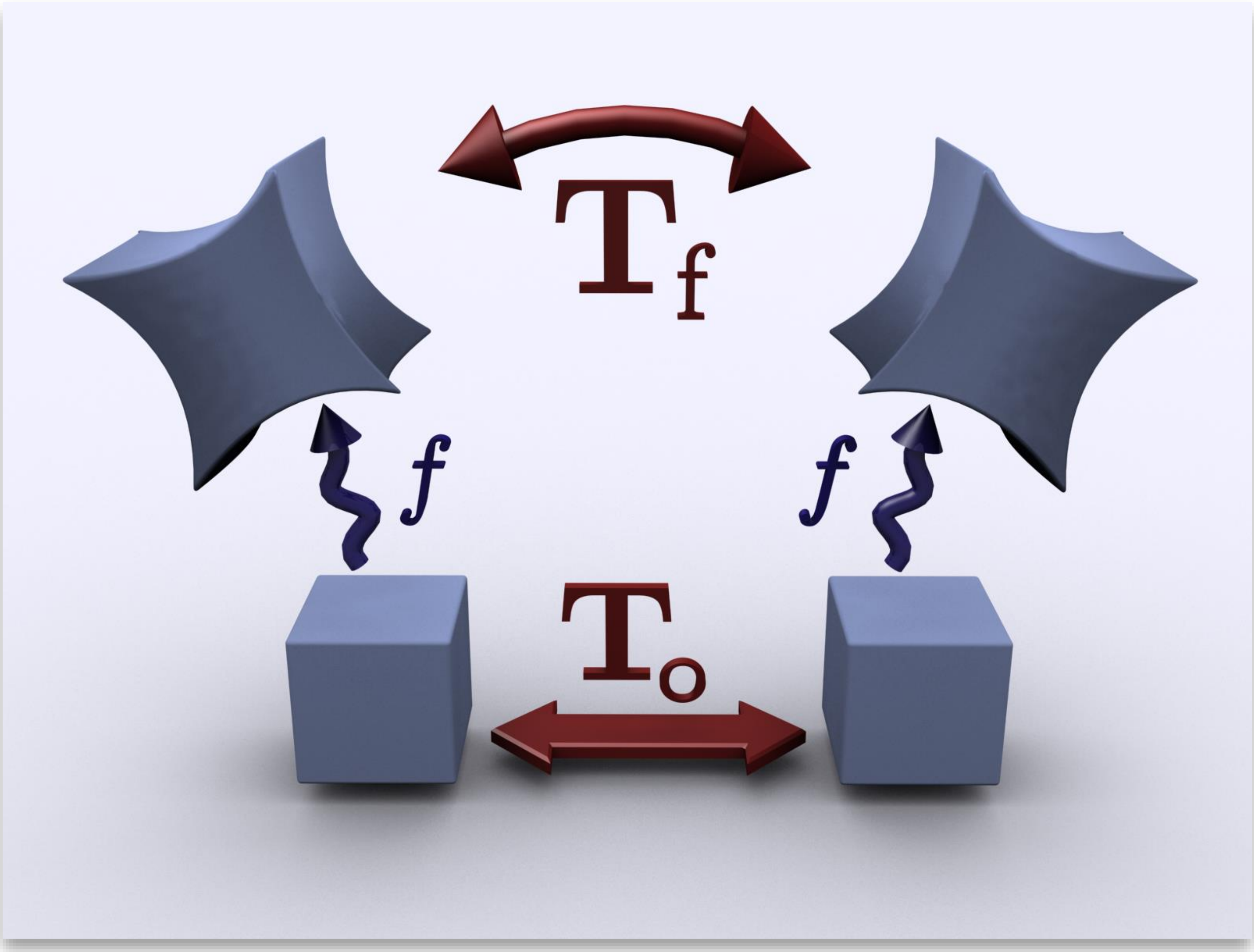
Video #10

Symmetry

- Symmetry is the absence of information
- Group Theory
- **Equivariance & Networks**

Symmetry Preservation

Preservation of Symmetry



Maps f that Preserve Symmetry

Symmetry preservation

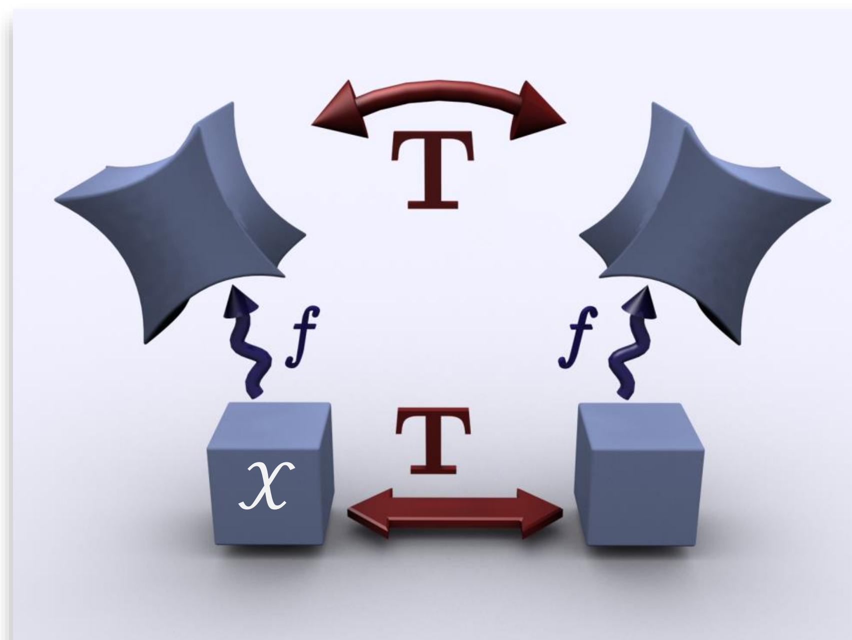
- Mapping f that preserve symmetry from group G
- For example: Geometric deformation

Symmetric f

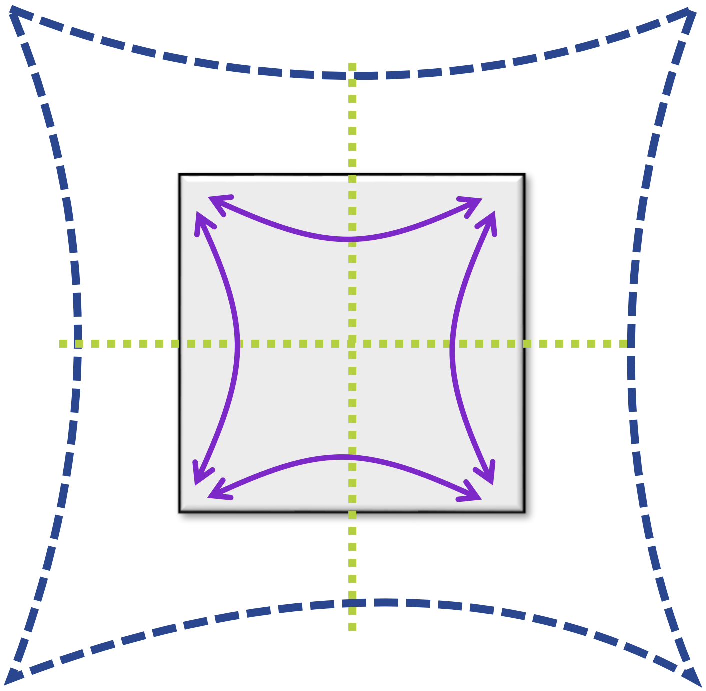
For all $\mathbf{x} \in \mathcal{X}$:

For all $\mathbf{T} \in G$:

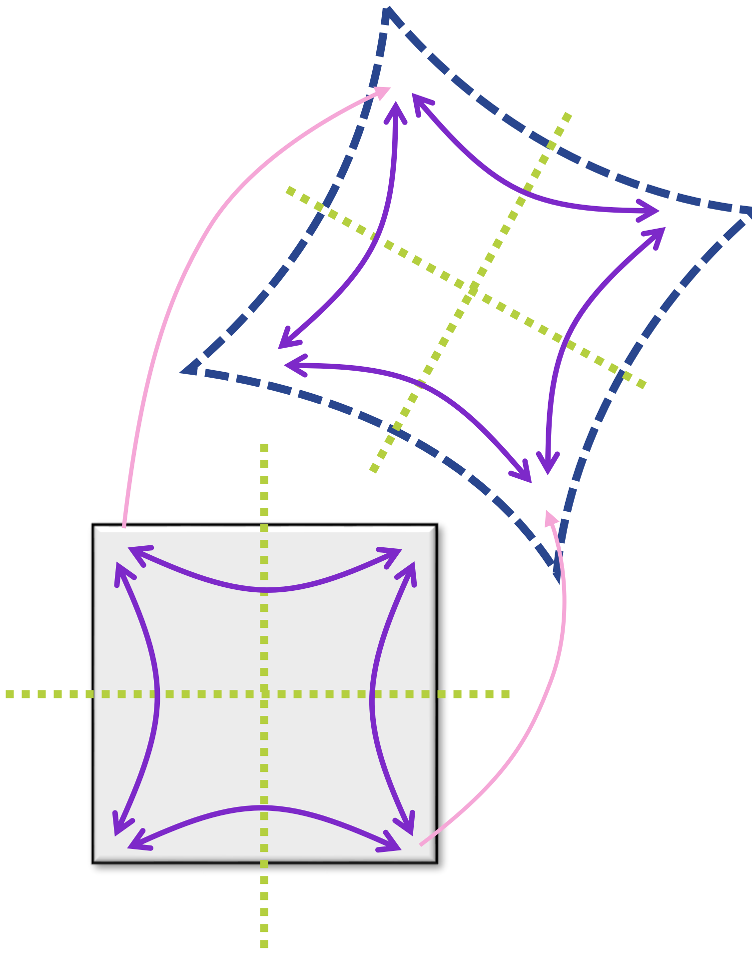
$\mathbf{T}f(\mathbf{x}) = f(\mathbf{T}\mathbf{x})$



Keep Invariants

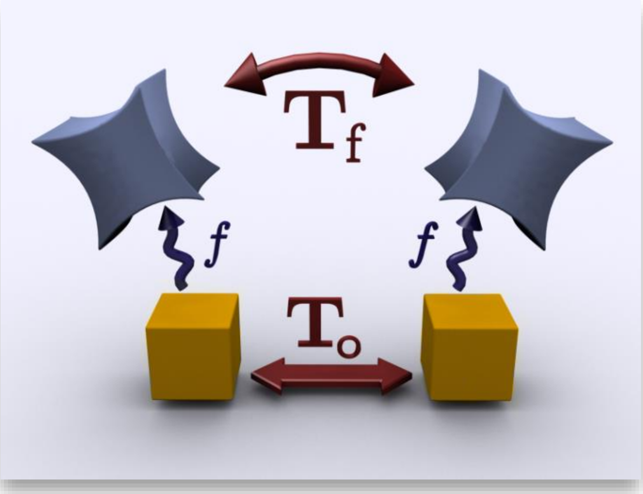


Symmetry Preservation



Symmetry Structure Preservation

Symmetry-Preserving 3D Deformation



[joint work with
Martin Bokeloh, 2011]

Scan Matching

[joint work with Christian Kurz, Xiaokun Wu, Thorsten Thormälen, Pushmeet Kohli, 2014]

Mathematical Formalization

Setting

- "Input" group G_{in}
- "Output" Group G_{out}

Group Homomorphism

- $h: G_{in} \rightarrow G_{out}$
- $h(a \circ b) = h(a) \circ h(b), \qquad a, b \in G_{in}$
- Algebraic structure preserved
  - Special case: group isomorphism – exact same structure
  - Isomorphism = bijective homomorphism

Useful Theorems

Definitions

- $\ker(h) = \{g \in G_{in} | h(g) = id\}$
- $\mathrm{im}(h) = \{f \in G_{out} | \exists g \in G_{in}: f = h(g)\}$

Theorems

- $\ker(h)$ is a (normal) subgroup of G_{in}
- $\mathrm{im}(h)$ is a subgroup of G_{out}
- $\mathrm{im}(h)$ is isomorphic to $G_{in} \setminus \ker(h)$

Normal Subgroup

- Normal subgroup $N \subseteq G: \forall g \in G: g \circ N = N \circ g$

Useful Theorems

Definitions

- $A \backslash G$ = "Quotient group of A mod G"
  - $a_1 \equiv a_2 \bmod G$ as new equality operator
  - G is a normal subgroup of A
- $a \equiv b \bmod G :\Leftrightarrow \exists g_l, g_r \in G: a = b \circ g_r = g_l \circ b$

normal subgroup $N \subseteq G$: $\forall g \in G: g \circ N = N \circ g$

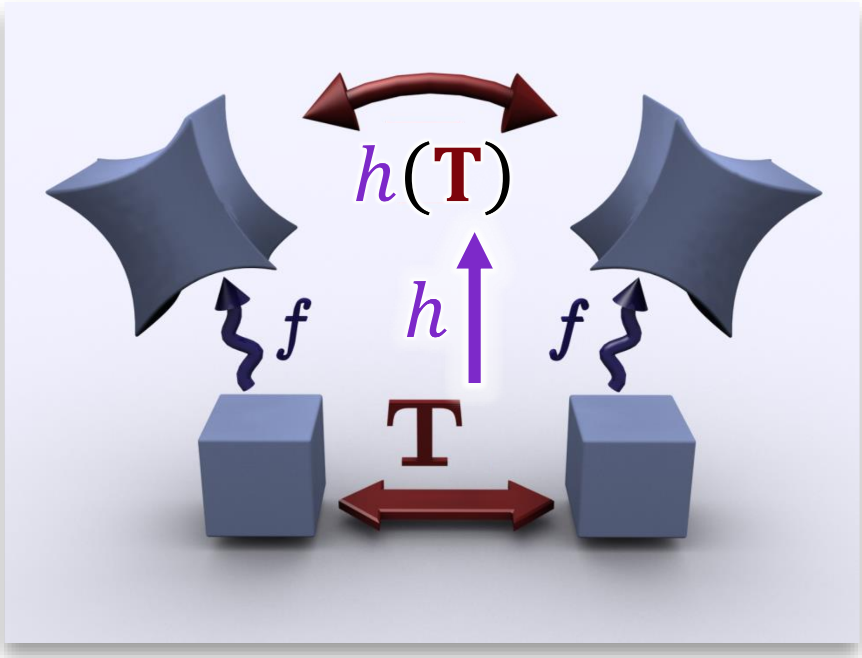
Deformation Example / CNNs

Symmetry preservation

- f is group-equivariant if

$$T_f = h(T_o)$$

  for a group homomorphism h



Important special cases

- f is co-variant if $T_f = T_o$
- f is invariant if $T_f = id$

Deformation Example / CNNs

Neural networks

- CNNs are group equivariant
  - "Pixel-wise" translation group $(\mathbb{Z}^d, +)$
- Co-variant output of convolutional layers
  - Translations of input images translates output
- Invariant output of CNN classifiers
  - Translations do not change class

Deformation Example / CNNs

General concept

- Group-convolutional neural network layer
  - Input function $x: \mathbb{R}^d \to \mathbb{R}^k$
  - Set of filters $w: \mathbb{R}^d \to \mathbb{R}^k$
  - Transformation group: $G \subseteq \{g: \mathbb{R}^d \to \mathbb{R}^d \,|\, g \text{ bijective}\}$
- Compute cross-correlations for all $g \in G$:

$$f_g(x) = \langle x, w \circ g \rangle$$

- Deeper Layers
  - Input function $x: G \to \mathbb{R}^k$
  - Set of filters $w: G \to \mathbb{R}^k$
- CNNs do exactly this for $G = (\mathbb{Z}^d, +)$

Other Applications

Symmetry & Networks

Representational Symmetries

- ReLu networks are symmetric under
  - Permutations of hidden neurons
  - Rescaling of adjacent layers
- And nothing else
  - If you want exact symmetry
  - Approximate symmetry: unknown

Learning Symmetries

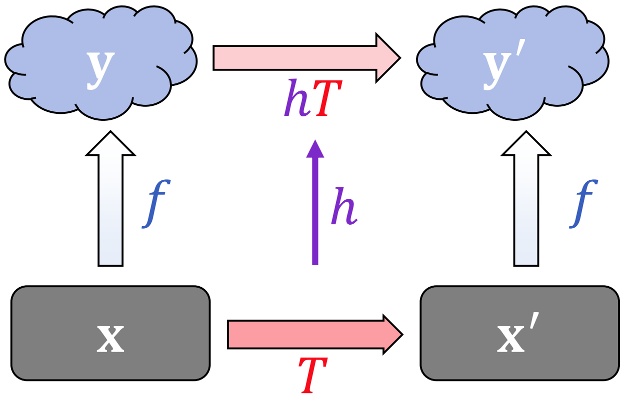
- Finding general symmetry groups in data
- Instead of “just” using CNNs / G-CNNS

Summary

Equivariance

Two operations

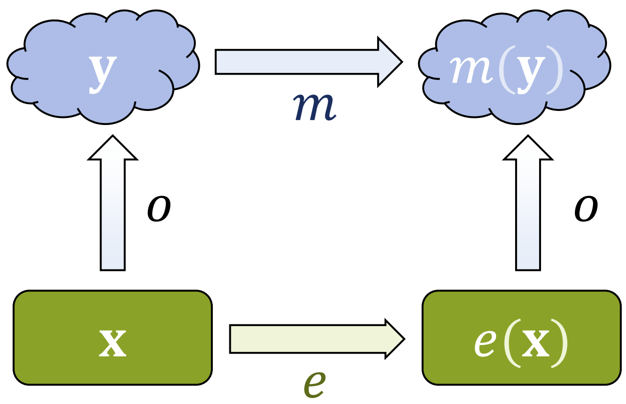
- Transformations / changes T
- Property / structure extractor f



Symmetry transfers

- Objects $\mathbf{x}$ have symmetries / invariants
- Extracted property $f(\mathbf{x})$ has the “same symmetries”
  - Same algebraic structure
- f is equivariant
- Abstract basis for data driven methods

Formalization of Modelling



Modeling

- Model m, state-of-the world $\mathbf{x}$
- “Symmetric” setup of experiment e
- Observations o must preserve symmetry

$$m \circ o \approx o \circ e$$

  - Choose m accordingly (not o, obviously)