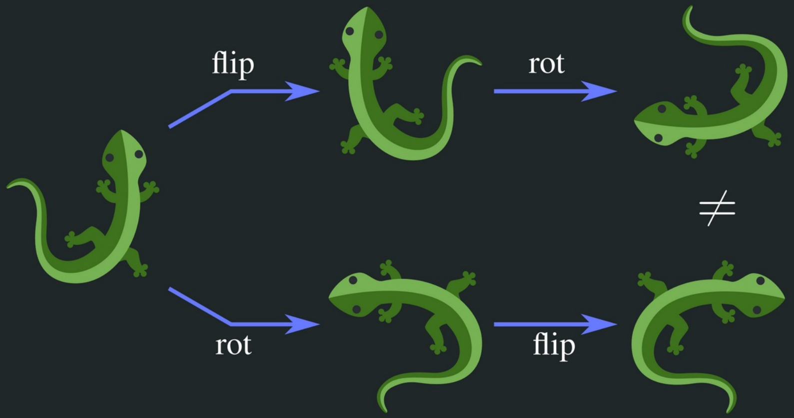
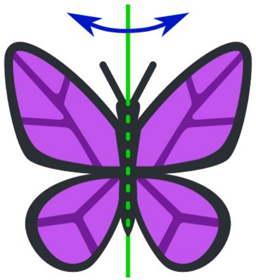
Groups, Representations & Equivariant maps

First Italian School in Geometric Deep Learning

Maurice Weiler
AMLab, QUVA Lab
University of Amsterdam
@maurice_weiler

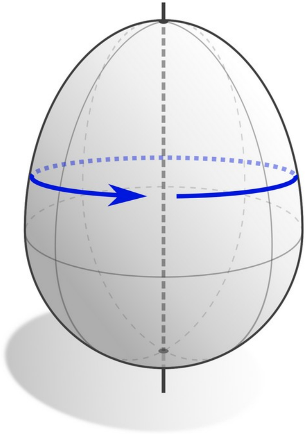


symmetry **groups** = “sets” of transformtions leaving an object invariant



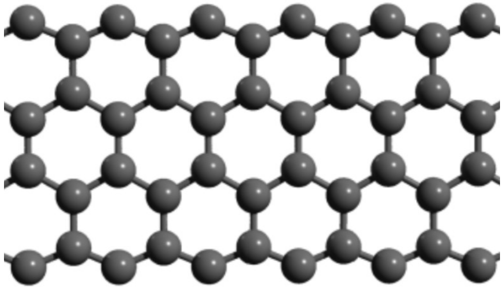
reflection symmetry

trivial symmetry

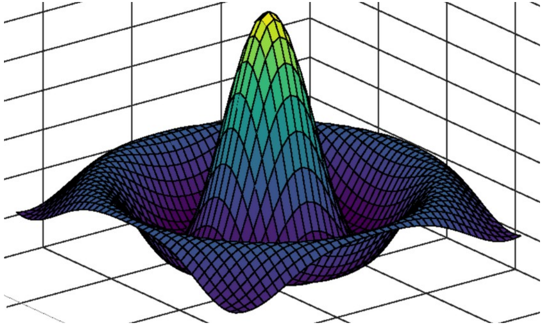


isometries of manifolds

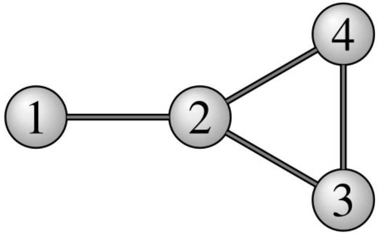
discrete translations of crystal lattice



rotationally symmetric potential

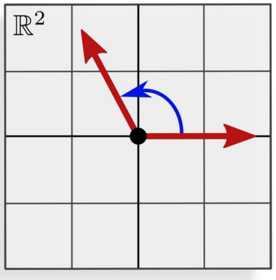


graph authomorphisms



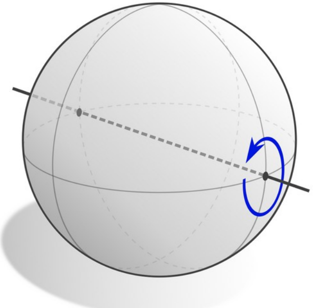
group **actions** - on different objects / spaces

example: SO(2) actions

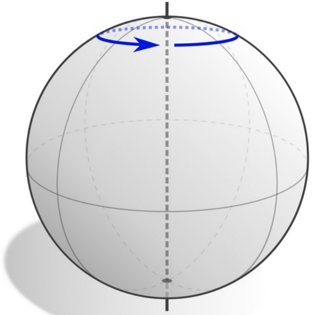


action on vectors in $\mathbb{R}^2$

actions on the 2-sphere S^2:



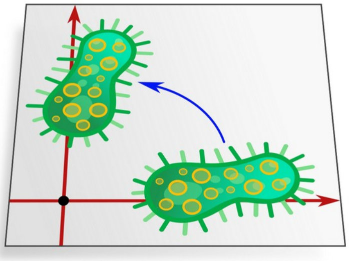
x-axis rotations



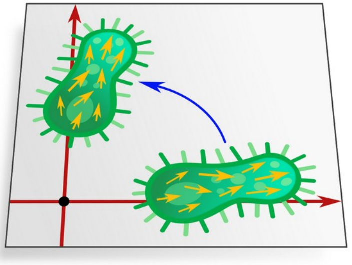
z-axis rotations

actions on signals / fields:

scalar field



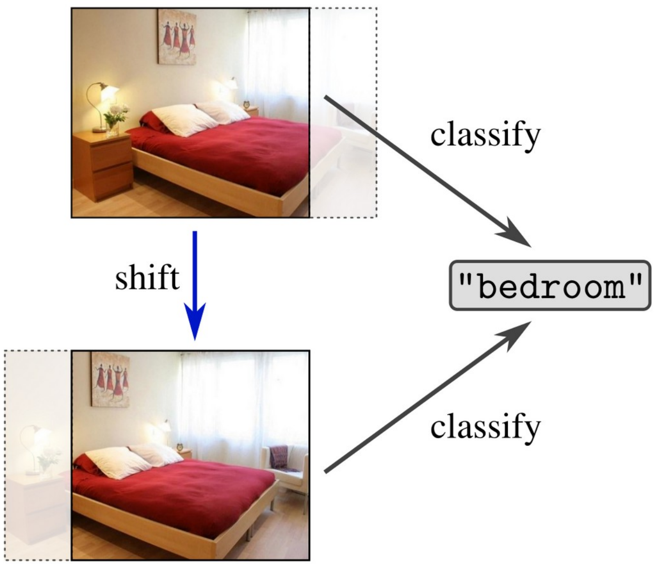
vector field



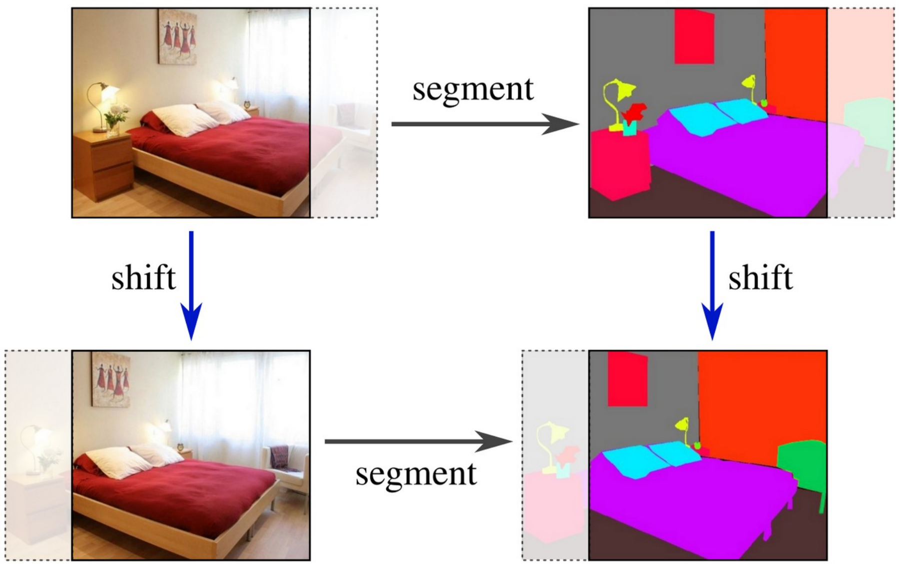
group **representations** - *linear* actions on *vector spaces* (e.g. feature vector spaces)

group **invariant** & **equivariant** functions

invariant image classification



equivariant image segmentation



Outline

Symmetry groups - basic definitions

Group actions

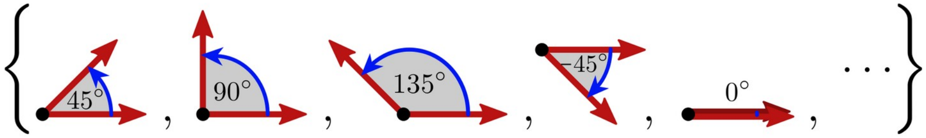
Invariant & equivariant maps

Group representations

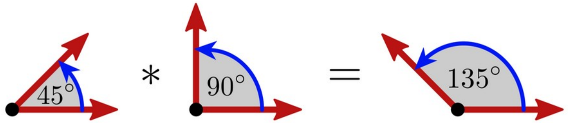
Symmetry groups

example: 2d rotation group $\mathrm{SO}(2)$

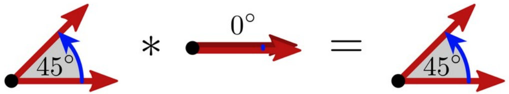
there is a **set** of rotations



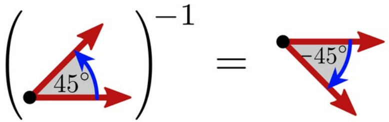
rotations can be **composed**



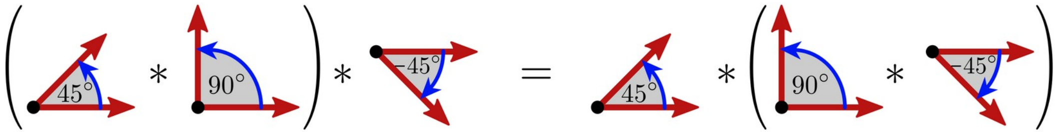
there is an **identity** rotation



each rotation has an **inverse**



the composition is **associative**



Symmetry groups

Definition B.1 (Group). *A* group *is a tuple* $(G, \cdot)$, *consisting of a set* G *and a binary operation*

$$\cdot : \ G \times G \to G, \quad (g,h) \mapsto g \cdot h$$

satisfying the following three group axioms:

associativity: *for all* $g, h, k \in G$ *one has* $(g \cdot h) \cdot k = g \cdot (h \cdot k)$

identity element: $\exists e \in G$ *such that* $\forall g \in G$ *one has* $e \cdot g = g = g \cdot e$

inverse element: $\forall g \in G \ \ \exists g^{-1} \in G$ *such that* $g \cdot g^{-1} = e = g^{-1} \cdot g$

examples:

translation group : $(\mathbb{R}^d, +)$

unitary group : $\mathrm{U}(1) := \{ e^{i\phi} \,|\, \phi \in [0, 2\pi) \}$

general linear group : $\mathrm{GL}(d) := \{ g \in \mathbb{R}^{d \times d} \,|\, \det(g) \neq 0 \}$

trivial group : $\{e\}$

check closure
+ group axioms !

Symmetry groups

Definition B.1 (Group). *A* group *is a tuple* $(G, \cdot)$, *consisting of a set* G *and a binary operation*

$$\cdot : \; G \times G \to G, \quad (g,h) \mapsto g \cdot h$$

satisfying the following three group axioms:

associativity: *for all* $g, h, k \in G$ *one has* $(g \cdot h) \cdot k = g \cdot (h \cdot k)$

identity element: $\exists e \in G$ *such that* $\forall g \in G$ *one has* $e \cdot g = g = g \cdot e$

inverse element: $\forall g \in G \;\; \exists g^{-1} \in G$ *such that* $g \cdot g^{-1} = e = g^{-1} \cdot g$

counter examples:

couter example 1: $\left\{ e^{i\phi} \,\middle|\, \phi \in [0, \pi) \right\}$ (closure)

couter example 2: $\left\{ g \in \mathbb{R}^{d \times d} \,\middle|\, \det(g) = 2 \right\}$ (closure)

couter example 3: $\left\{ g \in \mathbb{R}^{d \times d} \right\}$ (inverse)

Symmetry groups

Definition B.1 (Group). *A* group *is a tuple* $(G, \cdot)$*, consisting of a set* G *and a binary operation*

$$\cdot : \; G \times G \to G, \quad (g,h) \mapsto g \cdot h$$

satisfying the following three group axioms:

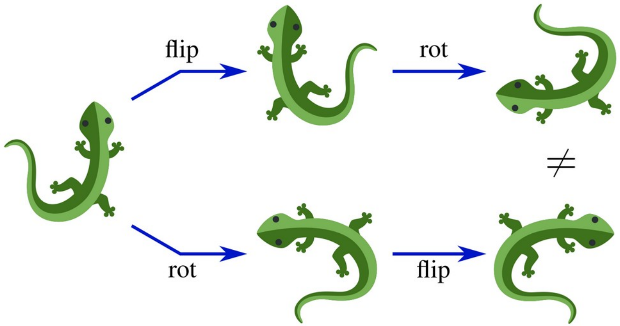
$$\begin{aligned} \textit{associativity:} &\quad \textit{for all } \; g,h,k \in G \; \textit{ one has } \; (g \cdot h) \cdot k = g \cdot (h \cdot k) \\ \textit{identity element:} &\quad \exists e \in G \; \textit{ such that } \; \forall g \in G \; \textit{ one has } \; e \cdot g = g = g \cdot e \\ \textit{inverse element:} &\quad \forall g \in G \;\; \exists g^{-1} \in G \; \textit{ such that } \; g \cdot g^{-1} = e = g^{-1} \cdot g \end{aligned}$$

groups may come with *additional structure*:

group category	structure on *G*	binary operation
topological group	topology	continuous map
Lie group	smooth manifold	smooth map
finite group	finite set	any function (between finite sets)

Abelian groups

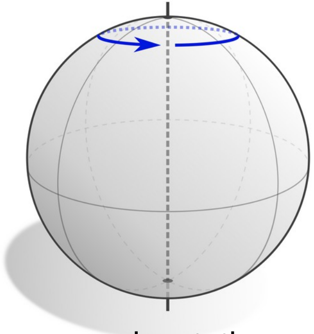
groups elements do in general *not* **commute**: $gh \neq hg$



groups with commutative elements are called **Abelian**:

Definition B.2 (Abelian group). *A group is called* abelian *if all of its elements commute, i.e. if:*

$$gh = hg \quad \forall g, h \in G$$



z-axis rotations

Subgroups

subsets of group elements may form a **subgroup**:

Definition A.1 (Subgroup). *A subset $H \subseteq G$ of a group G forms a* subgroup *if it is closed under composition and taking inverses:*

$$\text{composition:} \quad \textit{for all} \;\; g, h \in H \;\; \textit{one has} \;\; gh \in H$$

$$\text{inversion:} \quad \textit{for all} \;\; g \in H \;\; \textit{one has} \;\; g^{-1} \in H$$

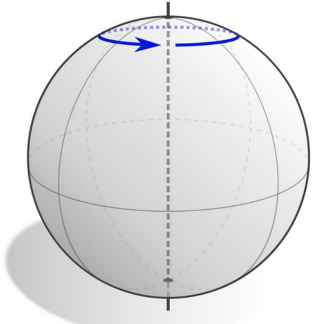
As the name suggests, subgroups are themselves groups, that is, they satisfy the three group axioms. One writes $H \leq G$.

examples:

discrete translations: $(\mathbb{Z}^d, +) \;\leq\; (\mathbb{R}^d, +)$

rotations around fixed axis: $\mathrm{SO}(2) \;\leq\; \mathrm{SO}(3)$

trivial examples: $\{e\} \leq G, \quad G \leq G$



z-axis rotations

Subgroups

subsets of group elements may form a **subgroup**:

Definition A.1 (Subgroup). *A subset $H \subseteq G$ of a group G forms a* subgroup *if it is closed under composition and taking inverses:*

$$\textit{composition:} \quad \textit{for all} \;\; g, h \in H \;\; \textit{one has} \;\; gh \in H$$

$$\textit{inversion:} \quad \textit{for all} \;\; g \in H \;\; \textit{one has} \;\; g^{-1} \in H$$

As the name suggests, subgroups are themselves groups, that is, they satisfy the three group axioms. One writes $H \leq G$.

counter examples:

$$\left\{ e^{i\phi} \,\middle|\, \phi \in [0, \pi) \right\} \not\leq \mathrm{U}(1)$$ (closure + inverse violated)

$$(\mathbb{R}_{\geq 0}, +) \not\leq (\mathbb{R}, +)$$ (inverse violated)

Products of groups

there are different **product** operations to combine groups into a supergroup

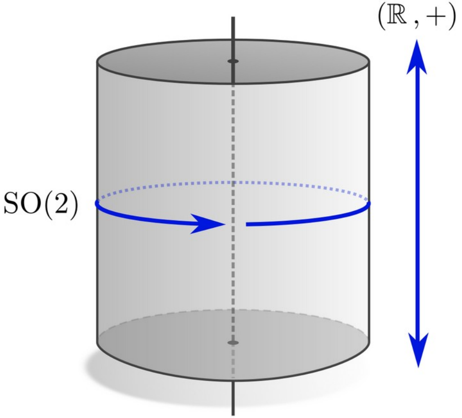
Definition A.5 (Direct product of groups). *Let $(H,\cdot)$ and $(K,\star)$ be arbitrary groups. Their (outer)* direct product *$(H,\cdot)\times(K,\star)$ is defined on the Cartesian product $H\times K$ of the underlying sets, equipped with the binary operation*

$$H\times K\to H\times K,\quad \big((\tilde h,\tilde k),\,(h,k)\big)\mapsto\big(\tilde h\cdot h,\ \tilde k\star k\big)$$

which composes the elements of the factors H and K independently from each other.

example:

$\mathrm{SO}(2)\times(\mathbb{R},+)$



Products of groups

there are different **product** operations to combine groups into a supergroup

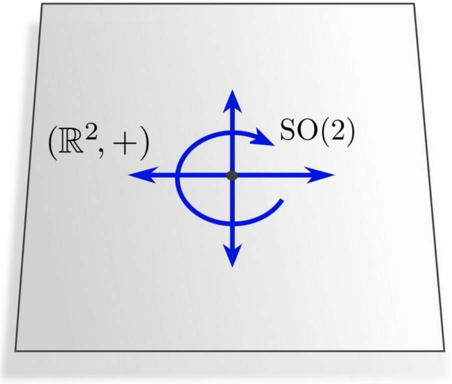
Definition A.5 (Direct product of groups). *Let* $(H, \cdot)$ *and* $(K, \star)$ *be arbitrary groups. Their (outer)* direct product $(H, \cdot) \times (K, \star)$ *is defined on the Cartesian product* $H \times K$ *of the underlying sets, equipped with the binary operation*

$$H \times K \to H \times K, \quad \big((\tilde{h}, \tilde{k}),\ (h, k)\big) \mapsto \big(\tilde{h} \cdot h,\ \tilde{k} \star k\big)$$

which composes the elements of the factors H *and* K *independently from each other.*

counter example: $\mathrm{SE}(2)$

$\neq$ ~~$\mathrm{SO}(2) \times (\mathbb{R}^2, +)$~~



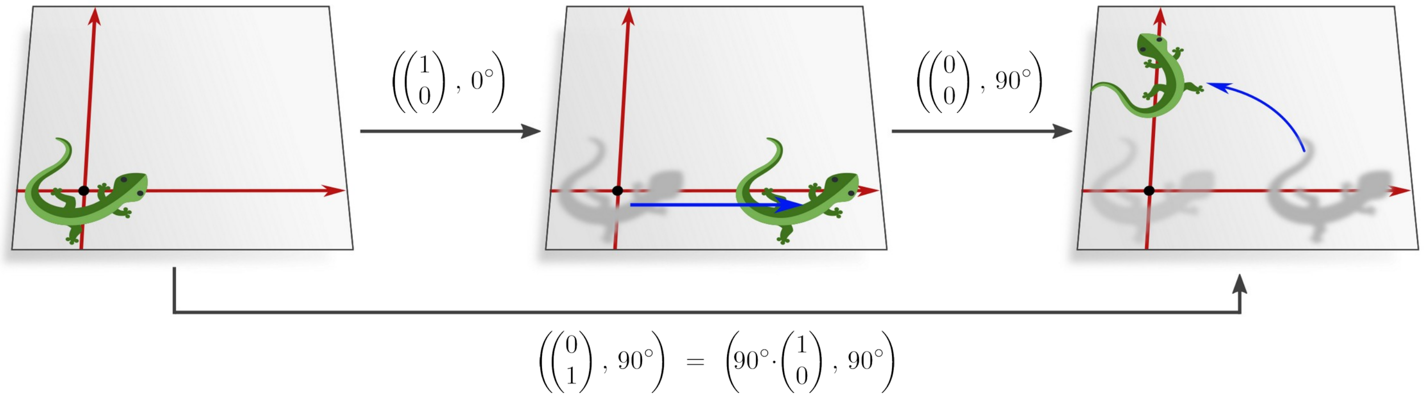
Products of groups

there are different **product** operations to combine groups into a supergroup

Idea: In a *semidirect product* $H \rtimes K$, the group K acts on H.

example: special Euclidean group $\mathrm{SE}(2) := (\mathbb{R}^2, +) \rtimes \mathrm{SO}(2)$

$$\neq (\mathbb{R}^2, +) \times \mathrm{SO}(2)$$



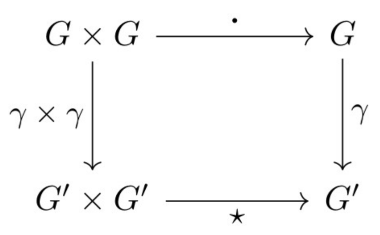
Group homomorphisms & isomorphisms

group **homomorphisms** are structure preserving maps between groups

Definition A.2 (Group homomorphism). *A group homomorphism between groups* $(G, \cdot)$ *and* $(G', \star)$ *is a map* $\gamma : G \to G'$ *such that*

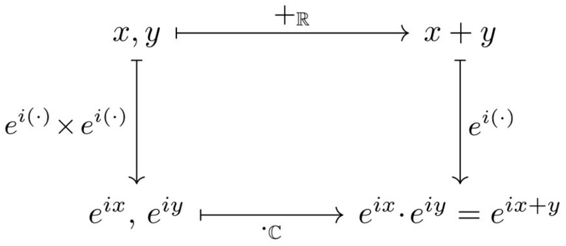
$$\gamma(g \cdot h) = \gamma(g) \star \gamma(h) \qquad \forall\, g, h \in G\,.$$

This implies generally that $\gamma(g^{-1}) = \gamma(g)^{-1}$ and that $g(e_G) = e_{G'}$.



example: complex exponentiation

$$\exp(i(\cdot))\colon\ (\mathbb{R}, +) \mapsto \mathrm{U}(1), \quad x \mapsto e^{ix}$$



$\Longrightarrow$ homomorphisms may lose information about the group $\rightarrow$ *isomorphisms*

Group homomorphisms & isomorphisms

group **isomorphisms** identify equivalent groups

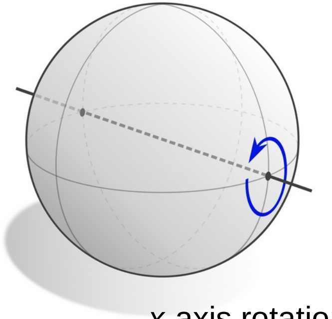
Definition A.4 (Group isomorphism). Group isomorphisms *are invertible group homomorphisms. One writes $G \cong G'$ to state that G and G' are isomorphic.*

examples:

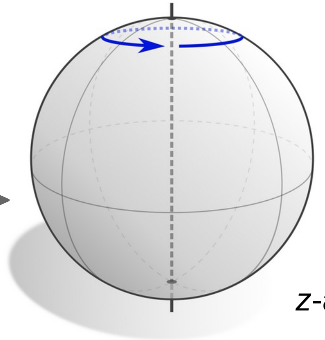
$$\begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{pmatrix}$$

$$\mathrm{U}(1) \cong \mathrm{SO}(2) \cong \mathrm{SO}(3)_z \cong \mathrm{SO}(3)_x$$

$$e^{i\phi} \qquad \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

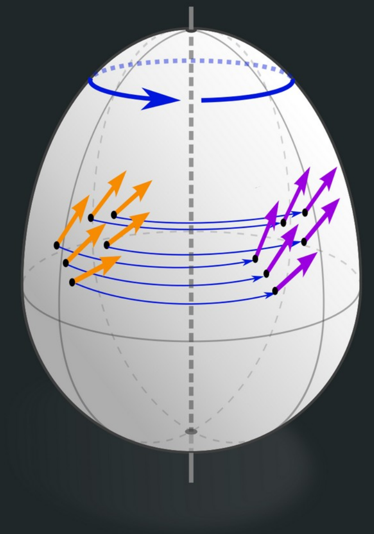


x-axis rotations



z-axis rotations

Group actions



Group actions

groups can **act** on other objects:

Definition A.2 (Group action). *Let G be a group and X be a set. A (left)* group action *is a map*

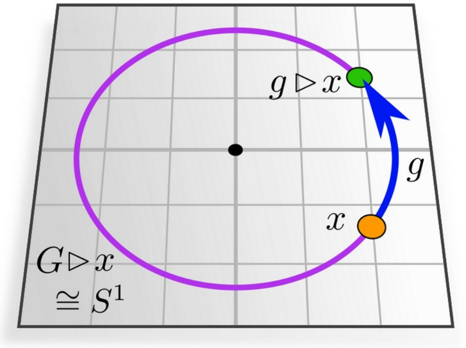
$$\triangleright : \; G \times X \to X, \quad (g, x) \mapsto g \triangleright x$$

that is compatible with the group composition and identity element:

associativity: $(gh) \triangleright x = g \triangleright (h \triangleright x)$ *for any* $g, h \in G,\; x \in X$

identity: $e \triangleright x = x$ *for any* $x \in X$

example: $\mathrm{SO}(2)$-action on $\mathbb{R}^2$



Group actions

groups can **act** on other objects:

Definition A.2 (Group action). *Let G be a group and X be a set. A (left)* group action *is a map*

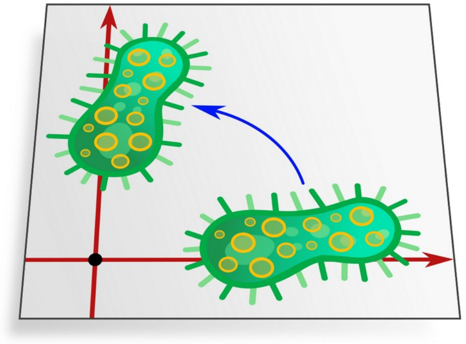
$$\triangleright : \; G \times X \to X, \quad (g, x) \mapsto g \triangleright x$$

that is compatible with the group composition and identity element:

$$\textit{associativity:} \quad (gh) \triangleright x = g \triangleright (h \triangleright x) \;\; \textit{for any} \;\; g, h \in G, \; x \in X$$

$$\textit{identity:} \quad e \triangleright x = x \;\; \textit{for any} \;\; x \in X$$

example: $\mathrm{SO}(2)$-action on scalar field



Group actions – orbits

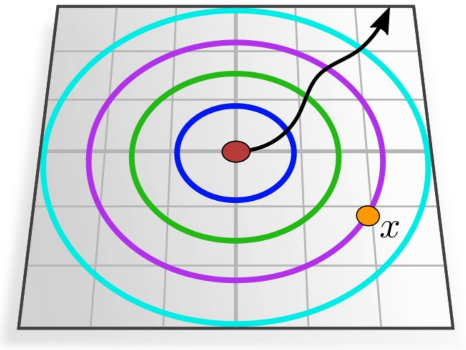
the action traces out an **orbit** in *X:*

Definition A.6 (Group orbit). *Let* $\triangleright$ *be an action of* G *on* X *and consider any element* $x \in X$*. The subset*

$$G \triangleright x \;:=\; \left\{ g \triangleright x \,\middle|\, g \in G \right\}$$

of X *is then denoted as* orbit *of* x.

example: $\mathrm{SO}(2)$-action on $\mathbb{R}^2$



Group actions – orbits

"being in the same orbit" defines an **equivalence relation**

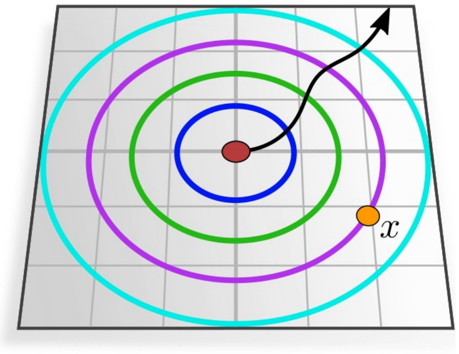
reflexivity: $x \sim_\triangleright x$, that is, x is contained in its own orbit $G \triangleright x$

symmetry: $x \sim_\triangleright y \Leftrightarrow y \sim_\triangleright x$, that is, if x is contained in y's orbit, then y is contained in x's orbit

transitivity: $x \sim_\triangleright y \wedge y \sim_\triangleright z \Rightarrow x \sim_\triangleright z$, that is, if x is contained in y's orbit and if y is contained in z's orbit, then x is contained in z's orbit

$\Longrightarrow$ we can take a *quotient* w.r.t. this equivalence relation

example: $\mathrm{SO}(2)$-action on $\mathbb{R}^2$



Group actions – quotient sets

the **quotient set** is the set of all orbits

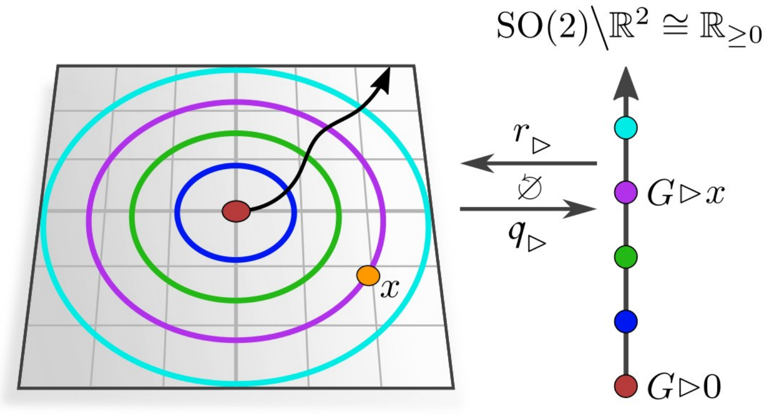
Definition A.7 (Quotient set and quotient map). *The* quotient set *induced by a G-action* $\triangleright$ *on* X *is the set of all orbits:*

$$G\backslash X \;:=\; \{G \triangleright x \mid x \in X\}$$

The corresponding quotient map *collapses elements of* X *to their orbit:*

$$q_\triangleright : X \to G\backslash X, \quad x \mapsto G \triangleright x$$

example: $\mathrm{SO}(2)$-action on $\mathbb{R}^2$



Group actions – orbit representatives

quotient maps are generally non-invertible, but one may choose **orbit representatives**

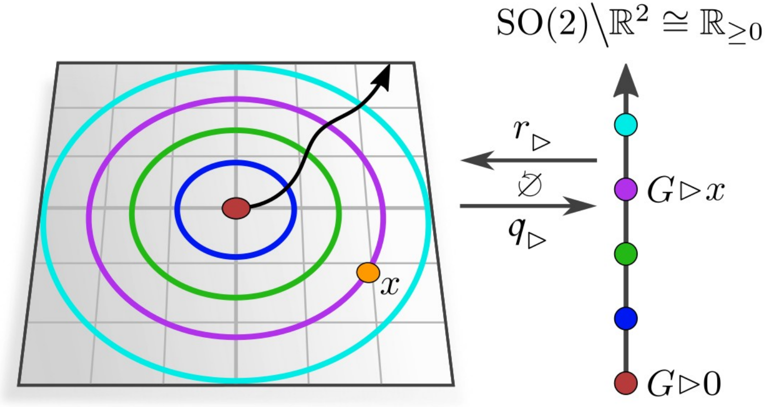
Definition A.8 (Orbit representative). Orbit representatives *are specified by a map*

$$r_\triangleright : G\backslash X \to X \qquad \text{such that} \qquad q_\triangleright \circ r_\triangleright(G \triangleright x) = G \triangleright x \quad \forall\, G \triangleright x \in G\backslash X\,,$$

i.e. such that the following diagram commutes:

$$G\backslash X \xrightarrow{\;r_\triangleright\;} X \xrightarrow{\;q_\triangleright\;} G\backslash X, \qquad q_\triangleright \circ r_\triangleright = \mathrm{id}_{G\backslash X}$$

example: $\mathrm{SO}(2)$-action on $\mathbb{R}^2$



Transitivity, homogeneity & stabilizers

Definition A.12 (Transitive action / homogeneous space).
A G-action $\triangleright$ *on a space* X *is called* transitive *iff it satisfies*

$$\forall\, x, y \in X \quad \exists\, g \in G \quad \text{such that} \quad y = g \triangleright x\,.$$

X *is then called a* homogeneous space.

Definition A.13 (Stabilizer subgroup). *Let* $\triangleright$ *be a G-action on a set* X*.*
The stabilizer subgroup *of some element* $x \in X$ *is defined as:*

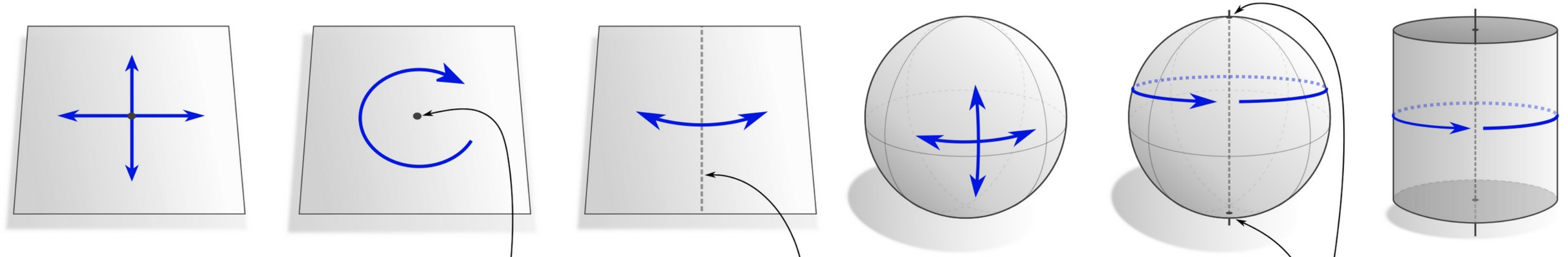
$$\mathrm{Stab}_x \;:=\; \{g \in G \,|\, g \triangleright x = x\} \;\leq\; G$$

A General Theory of Equivariant CNNs on Homogeneous Spaces

Taco S. Cohen
Qualcomm AI Research*
Qualcomm Technologies Netherlands B.V.
tacos@qti.qualcomm.com

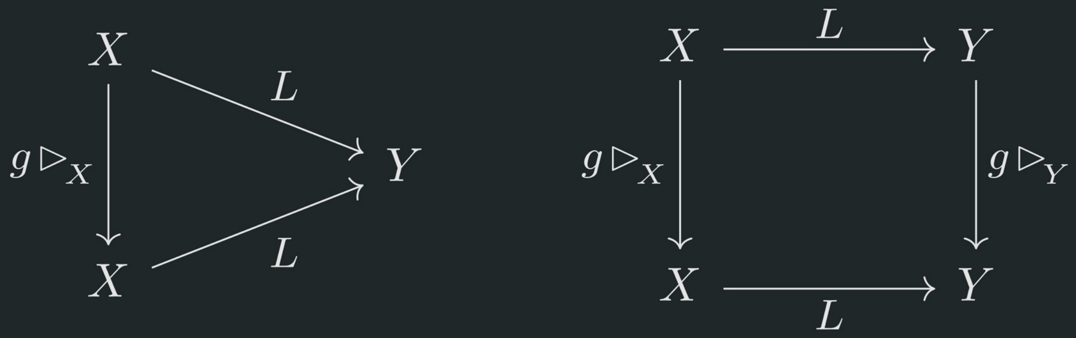
Mario Geiger
PCSL Research Group
EPFL
mario.geiger@epfl.ch

Maurice Weiler
QUVA Lab
U. of Amsterdam
m.weiler@uva.nl



group	$(\mathbb{R}^2, +)$	SO(2)	reflections	SO(3)	SO(2)	SO(2)
transitive	✓	×	×	✓	×	×
stabilizers	$\{e\}$	$\{e\}$, SO(2)	$\{e\}$, reflections	SO(2)	$\{e\}$, SO(2)	$\{e\}$

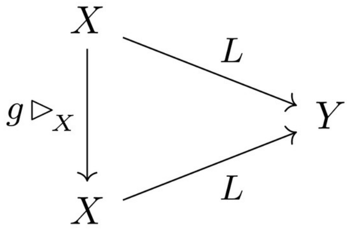
Invariant & Equivariant maps



Invariant maps

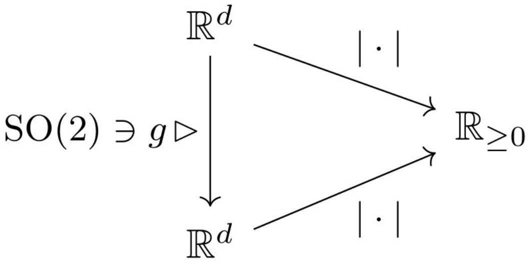
Definition A.9 (Invariant map). *Let $\triangleright_X$ be a group action on a set X. A function $L : X \to Y$ is called G-*invariant, *iff it satisfies*

$$L(g \triangleright_X x) \;=\; L(x) \qquad \forall\, g \in G,\;\; x \in X\,.$$



example:

rotation invariant vector norm

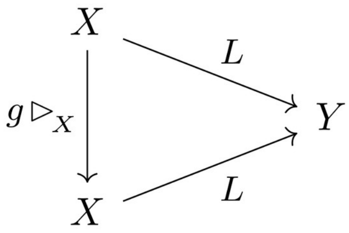


since $|g \triangleright v| = |v|$ for $g \in \mathrm{SO}(2)$

Invariant maps

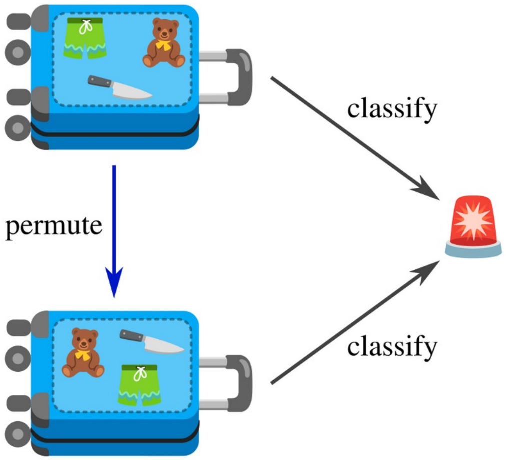
Definition A.9 (Invariant map). *Let* $\triangleright_X$ *be a group action on a set* X. *A function* $L : X \to Y$ *is called* G-invariant, *iff it satisfies*

$$L(g \triangleright_X x) \; = \; L(x) \qquad \forall \, g \in G, \;\; x \in X \, .$$



example:

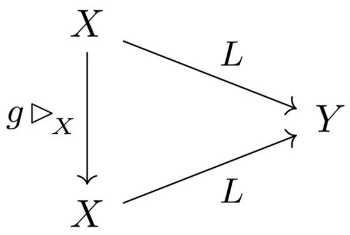
permutation invariant luggage classification



Invariant maps

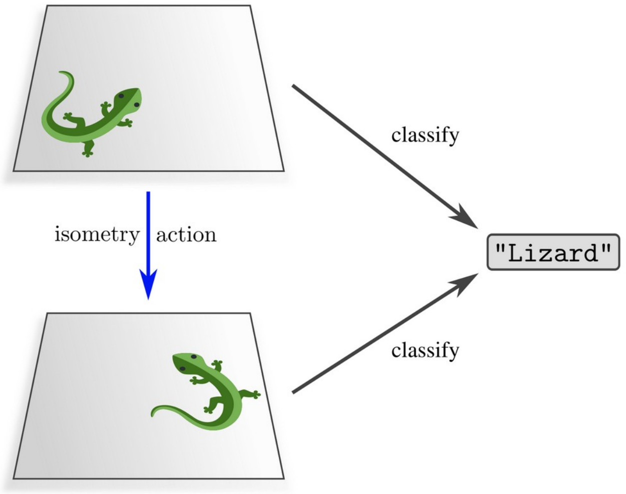
Definition A.9 (Invariant map). *Let* $\triangleright_X$ *be a group action on a set* X*. A function* $L: X \to Y$ *is called* G-invariant, *iff it satisfies*

$$L(g \triangleright_X x) \;=\; L(x) \qquad \forall\, g \in G,\;\; x \in X\,.$$



example:

isometry invariant image classification



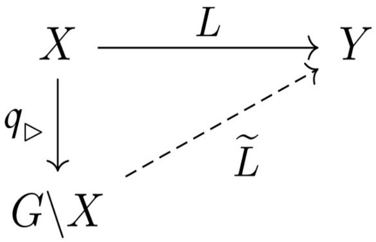
Invariant maps - universal property

universal property: invariant maps *“descent to the quotient”*

i.e. for any *G*-invariant map $L : X \to Y$

there exists a unique map $\widetilde{L} : G\backslash X \to Y$

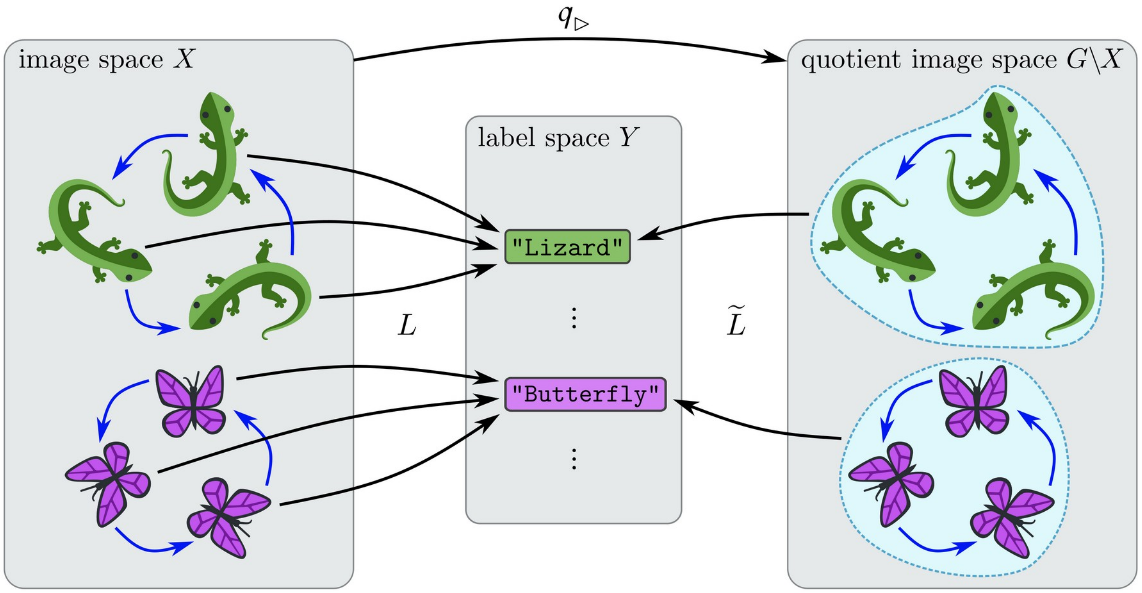
such that $L = \widetilde{L} \circ q_{\triangleright}$



example:

invariant image classification

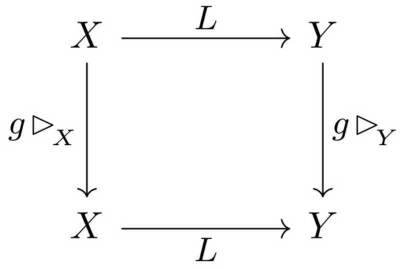
$\Rightarrow$ reduced hypothesis space of invariant models !

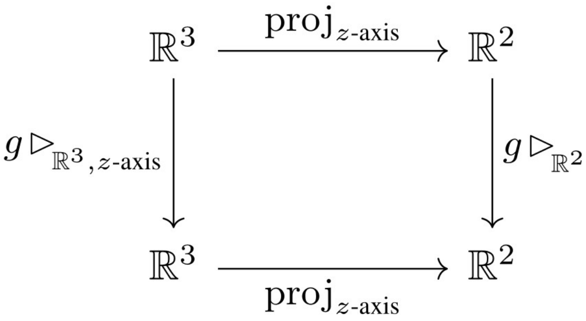


Equivariant maps

Definition A.10 (Equivariant map). *Let $\triangleright_X$ and $\triangleright_Y$ be group actions on sets X and Y. A function $L : X \to Y$ is said to be G-*equivariant *iff it commutes with these actions:*

$$L(g \triangleright_X x) \;=\; g \triangleright_Y L(x) \qquad \forall\, g \in G,\;\; x \in X$$





example:

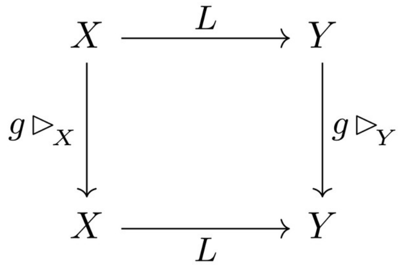
SO(2)-rotations around a projection axis commute with the projection

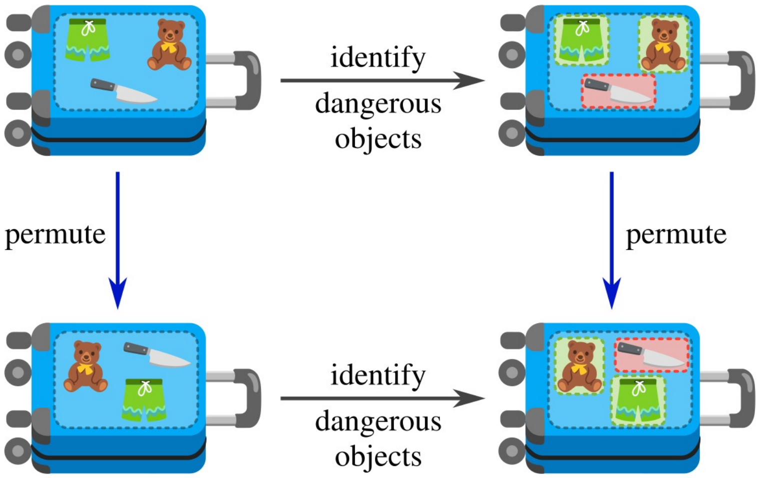
$$\underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}}_{\mathrm{proj}_{z\text{-axis}}} \underbrace{\begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{g \triangleright_{\mathbb{R}^3,\, z\text{-axis}}} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \underbrace{\begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}}_{g \triangleright_{\mathbb{R}^2}} \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}}_{\mathrm{proj}_{z\text{-axis}}} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Equivariant maps

Definition A.10 (Equivariant map). *Let $\rhd_X$ and $\rhd_Y$ be group actions on sets X and Y. A function $L : X \to Y$ is said to be G-*equivariant *iff it commutes with these actions:*

$$L(g \rhd_X x) \;=\; g \rhd_Y L(x) \qquad \forall\, g \in G,\;\; x \in X$$





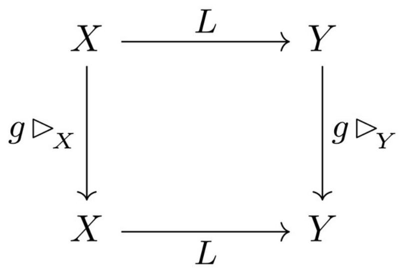
example:

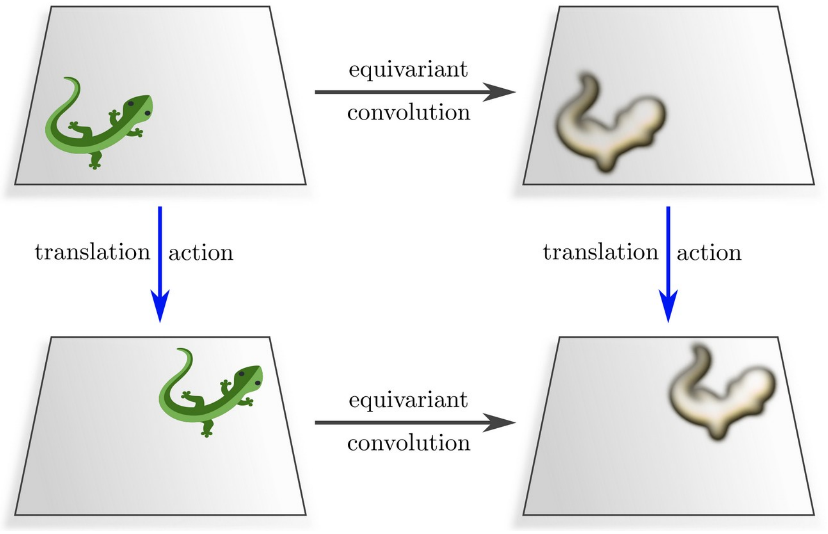
permutation equivariant labeling
of dangerous objects

Equivariant maps

Definition A.10 (Equivariant map). *Let $\triangleright_X$ and $\triangleright_Y$ be group actions on sets X and Y. A function $L : X \to Y$ is said to be G-*equivariant *iff it commutes with these actions:*

$$L(g \triangleright_X x) \;=\; g \triangleright_Y L(x) \qquad \forall\, g \in G,\;\; x \in X$$





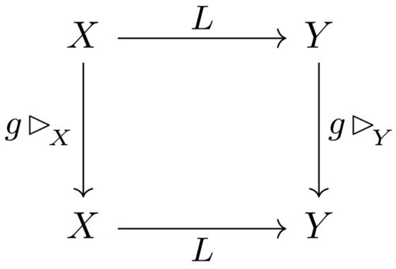
example:

translation equivariant convolution

Equivariant maps

Definition A.10 (Equivariant map). *Let* $\triangleright_X$ *and* $\triangleright_Y$ *be group actions on sets* X *and* Y. *A function* $L : X \to Y$ *is said to be* G-equivariant *iff it commutes with these actions:*

$$L(g \triangleright_X x) \;=\; g \triangleright_Y L(x) \qquad \forall\, g \in G,\;\; x \in X$$



$$K*:\; L^2(\mathbb{R}) \to L^2(\mathbb{R}), \quad f \;\mapsto\; K*f \;:=\; \int_{\mathbb{R}} dy\, K(x-y)\, f(y)$$

$$\begin{aligned}
\big(K*(g \triangleright f)\big)(x) \;&=\; \int_{\mathbb{R}} dy\, K(x-y)\,(g \triangleright f)(y) \\
&= \int_{\mathbb{R}} dy\, K(x-y)\, f(y-g) \\
&= \int_{\mathbb{R}} dz\, K((x-g)-z)\, f(z) \\
&= \big(K*f\big)(x-g) \\
&= \big(g \triangleright (K*f)\big)(x)
\end{aligned}$$

substitute

$z := y - g$

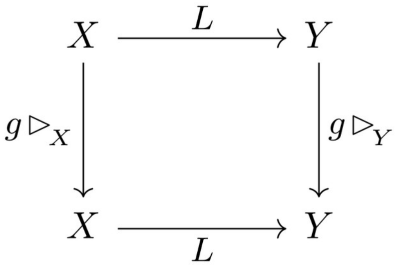
example:

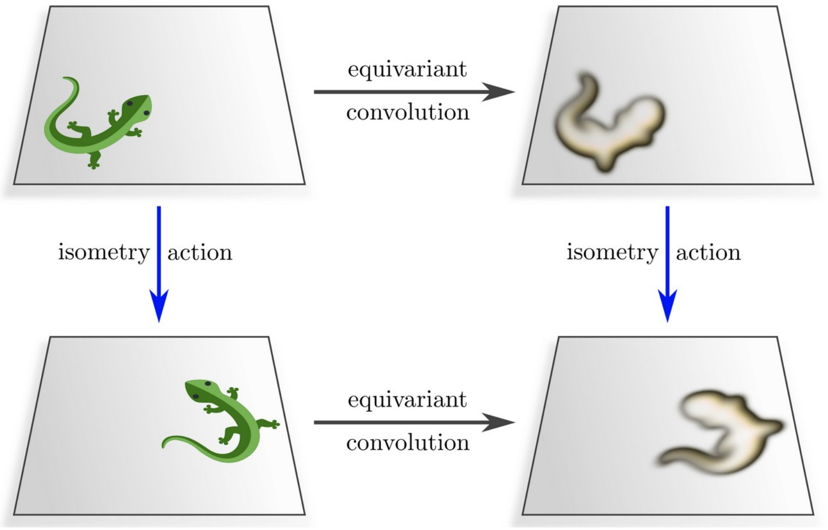
translation equivariant convolution

Equivariant maps

Definition A.10 (Equivariant map). *Let $\triangleright_X$ and $\triangleright_Y$ be group actions on sets X and Y. A function $L : X \to Y$ is said to be G-*equivariant *iff it commutes with these actions:*

$$L(g \triangleright_X x) \;=\; g \triangleright_Y L(x) \qquad \forall\, g \in G,\;\; x \in X$$





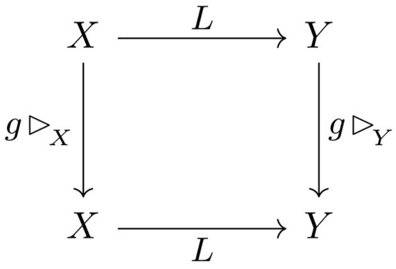
example:

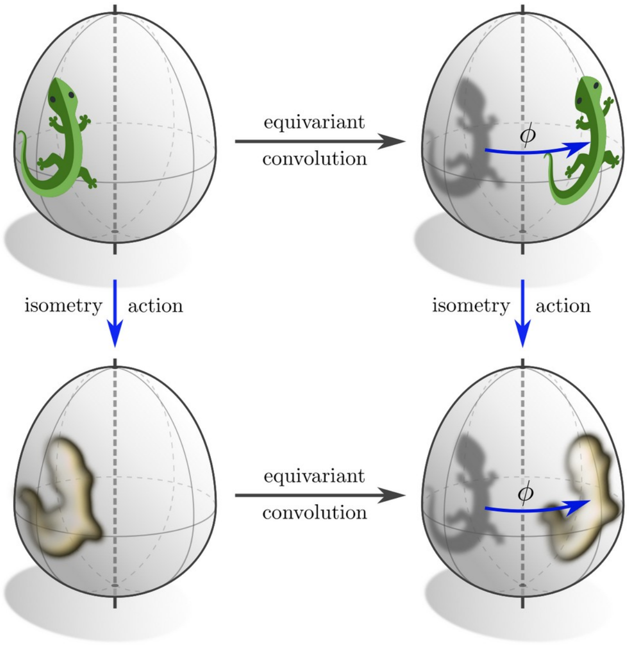
isometry equivariant convolution

Equivariant maps

Definition A.10 (Equivariant map). *Let* $\triangleright_X$ *and* $\triangleright_Y$ *be group actions on sets* X *and* Y*.* *A function* $L : X \to Y$ *is said to be* G-equivariant *iff it commutes with these actions:*

$$L(g \triangleright_X x) \;=\; g \triangleright_Y L(x) \qquad \forall\, g \in G,\;\; x \in X$$



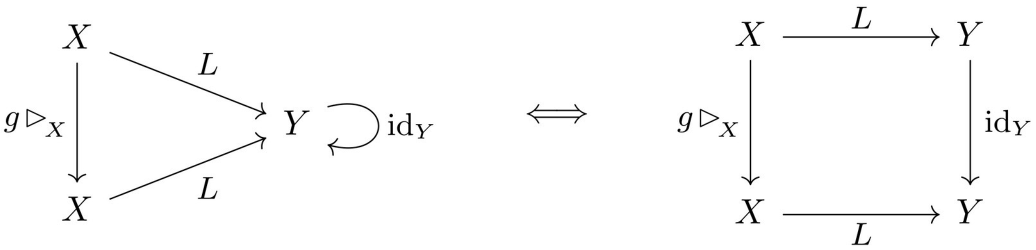


example:

isometry equivariant convolution on Riemannian manifold

Invariance $\Leftrightarrow$ Equivariance

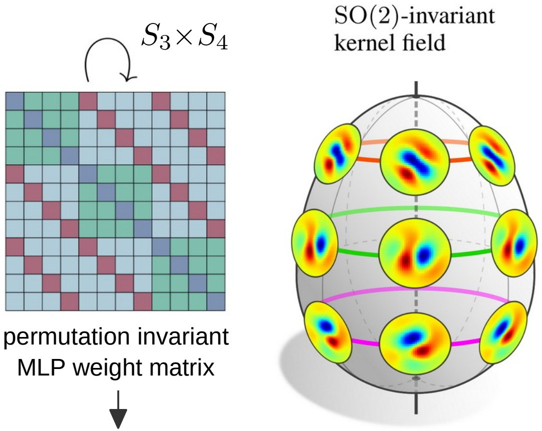
invariant maps are a special case of *equivariant* maps with a *trivial action* id_Y *on their codomain:*



equivariant maps are themselves *invariants* under the group action:

$$L \circ (g \triangleright_X (\cdot)) \;=\; (g \triangleright_Y (\cdot)) \circ L$$

$$\Longleftrightarrow \quad (g^{-1} \triangleright_Y (\cdot)) \circ L \circ (g \triangleright_X (\cdot)) \;=\; L$$



***G*-equivariant NNs $\Longleftrightarrow$ *G*-invariant neural connectivity**

Hartford et al., *Deep Models of Interactions Across Sets*, ICML 2018

Equivariant Neural Networks

(feed forward) neural networks are sequences of layers:

$$\mathcal{F}_0 \xrightarrow{L_1} \mathcal{F}_1 \xrightarrow{L_2} \mathcal{F}_2 \xrightarrow{L_3} \cdots \xrightarrow{L_{N-1}} \mathcal{F}_{N-1} \xrightarrow{L_N} \mathcal{F}_N$$

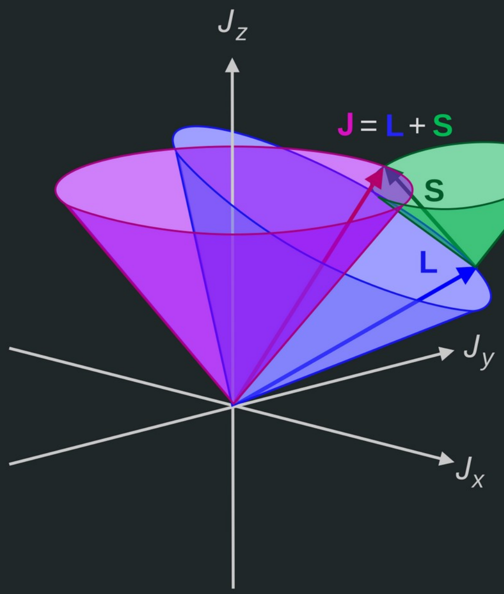
equivariant NNs are sequences of equivariant layers:

$$\begin{array}{ccccccccccc}
\mathcal{F}_0 & \xrightarrow{L_1} & \mathcal{F}_1 & \xrightarrow{L_2} & \mathcal{F}_2 & \xrightarrow{L_3} & \cdots & \xrightarrow{L_{N-1}} & \mathcal{F}_{N-1} & \xrightarrow{L_N} & \mathcal{F}_N \\
{\scriptstyle g\,\triangleright_0}\big\downarrow & & {\scriptstyle g\,\triangleright_1}\big\downarrow & & {\scriptstyle g\,\triangleright_2}\big\downarrow & & & & {\scriptstyle g\,\triangleright_{N-1}}\big\downarrow & & \big\downarrow{\scriptstyle g\,\triangleright_N} \\
\mathcal{F}_0 & \xrightarrow[L_1]{} & \mathcal{F}_1 & \xrightarrow[L_2]{} & \mathcal{F}_2 & \xrightarrow[L_3]{} & \cdots & \xrightarrow[L_{N-1}]{} & \mathcal{F}_{N-1} & \xrightarrow[L_N]{} & \mathcal{F}_N
\end{array}$$

invariant NNs are usually built from 1) equivariant layers, 2) an invariant map and 3) a final MLP:

$$\begin{array}{ccccccccccc}
\mathcal{F}_0 & \xrightarrow{L_1} & \cdots & \xrightarrow{L_N} & \mathcal{F}_N & \searrow^{L_{\text{inv}}} & & & & & \\
{\scriptstyle g\,\triangleright_0}\big\downarrow & & & & {\scriptstyle g\,\triangleright_N}\big\downarrow & & \widetilde{\mathcal{F}}_0 & \xrightarrow{\widetilde{L}_1} & \cdots & \xrightarrow{\widetilde{L}_M} & \widetilde{\mathcal{F}}_M \\
\mathcal{F}_0 & \xrightarrow[L_1]{} & \cdots & \xrightarrow[L_N]{} & \mathcal{F}_N & \nearrow_{L_{\text{inv}}} & & & & &
\end{array}$$

Group representation theory



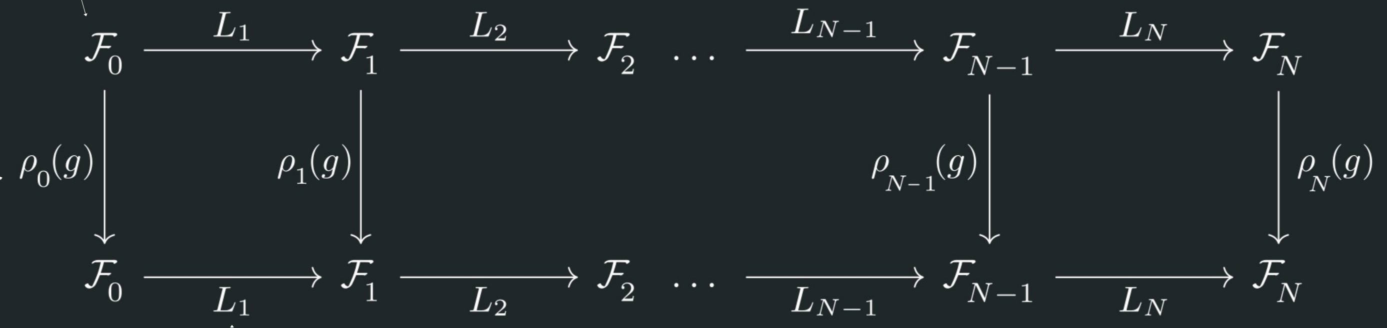
$$\rho_L \otimes \rho_S \;\cong\; \bigoplus_{J=|L-S|}^{L+S} \rho_J$$

Group representation theory

motivation: systematic investigation of equivariant NNs in terms of representation theory

feature vector spaces = representation spaces

linear group actions
=
group representations



equivariant maps between representations
(e.g. intertwiners)

Linear group representations

group **representations** ... ... model group elements as matrices (or linear operators)

... act on vector spaces

Definition A.1 (Linear group representation). *A* linear group representation *of a group G on a real vector space $\mathbb{R}^N$ is a group homomorphism*

$$\rho : G \to \mathrm{GL}(\mathbb{R}^N)$$

group of invertible $n{\times}n$-matrices

group composition

matrix multiplication

Recall that *homomorphisms* satisfy:

$$\begin{aligned} \textit{composition:} \quad & \rho(gh) = \rho(g)\rho(h) \quad \forall\, g, h \in G \\ \textit{inverse:} \quad & \rho(g^{-1}) = \rho(g)^{-1} \quad \forall\, g \in G \\ \textit{identity:} \quad & \rho(e) = \mathrm{id}_V \ , \end{aligned}$$

Linear group representations

group **representations** ... ... model group elements as matrices (or linear operators)

... act on vector spaces

Definition A.1 (Linear group representation). *A* linear group representation *of a group G on a real vector space $\mathbb{R}^N$ is a group homomorphism*

$$\rho : G \to \mathrm{GL}(\mathbb{R}^N)$$

group of invertible $n \times n$-matrices

Recall that *homomorphisms* satisfy:

group composition

matrix multiplication

$$\begin{aligned} \textit{composition:} \quad & \rho(gh) = \rho(g)\rho(h) \quad \forall\, g,h \in G \\ \textit{inverse:} \quad & \rho(g^{-1}) = \rho(g)^{-1} \quad \forall\, g \in G \\ \textit{identity:} \quad & \rho(e) = \mathrm{id}_V \ , \end{aligned}$$

Linear group representations

group **representations** ... ... model group elements as matrices (or linear operators)

... act on vector spaces

Definition A.1 (Linear group representation). *A* linear group representation *of a group G on a vector space V is a group homomorphism*

$$\rho : G \to \mathrm{GL}(V)$$

from G to the general linear group $\mathrm{GL}(V)$ (invertible linear maps) of the vector space. V is referred to as representation space.

group composition

composition of linear maps

Recall that *homomorphisms* satisfy:

$$\begin{aligned} \textit{composition:} &\quad \rho(gh) = \rho(g)\rho(h) \quad \forall\, g,h \in G \\ \textit{inverse:} &\quad \rho(g^{-1}) = \rho(g)^{-1} \quad \forall\, g \in G \\ \textit{identity:} &\quad \rho(e) = \mathrm{id}_V \;, \end{aligned}$$

Linear group representations

group **representations** ... ... model group elements as matrices (or linear operators)

... act on vector spaces

Definition A.1 (Linear group representation). *A* linear group representation *of a group G on a vector space V is a group homomorphism*

$$\rho : G \to \mathrm{GL}(V)$$

examples for $G = \mathrm{SO}(2)$:

trivial rep: $\rho_{\text{triv}} : \mathrm{SO}(2) \to \mathrm{GL}(1), \quad g \mapsto (1) = \mathrm{id}_{\mathbb{R}^1}$

defining rep: $\rho_{\text{def}} : \mathrm{SO}(2) \to \mathrm{GL}(2), \quad g \mapsto g = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}$

(real) irreducible reps: $\rho_{\text{irrep},k} : \mathrm{SO}(2) \to \mathrm{GL}(2), \quad g \mapsto g^k = \begin{pmatrix} \cos k\phi & -\sin k\phi \\ \sin k\phi & \cos k\phi \end{pmatrix}$ for any $k \in \mathbb{N}$

(2nd order) tensor rep: $\rho_{\text{tensor},2} : \mathrm{SO}(2) \to \mathrm{GL}(4), \quad g \mapsto g \otimes g = \begin{pmatrix} \cos\phi \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} & -\sin\phi \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \\ \sin\phi \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} & \cos\phi \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \end{pmatrix}$

Restricted representation

representations can be **restricted** to subgroups

Definition A.2 (Restricted representation). *Let (ρ, V) be a G-representation and let $H \leq G$ be a subgroup. The restricted representation of ρ is the H-representation*

$$\mathrm{Res}_H^G \rho : \; H \to \mathrm{GL}(V), \quad h \mapsto \rho(h)$$

example: restriction from continuous rotations in SO(2) to 90° rotations in C_4 :

ϕ	0	$\frac{\pi}{2}$	π	$\frac{3\pi}{2}$
$\mathrm{Res}_{\mathrm{C}_4}^{\mathrm{SO}(2)} \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$	$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$

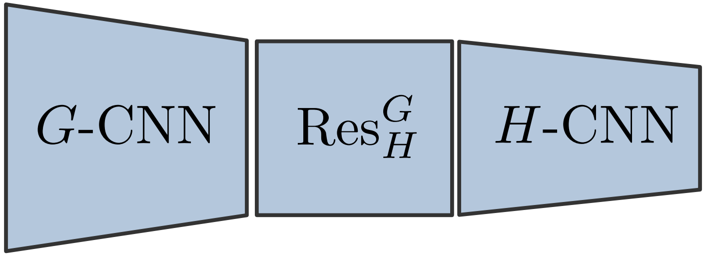
Restricted representation

representations can be **restricted** to subgroups

Definition A.2 (Restricted representation). *Let* (ρ, V) *be a* G*-representation and let* $H \leq G$ *be a subgroup. The restricted representation of* ρ *is the* H*-representation*

$$\mathrm{Res}_H^G\,\rho : \; H \to \mathrm{GL}(V), \quad h \mapsto \rho(h)$$

example: for models with *varying level of equivariance* with depth, turn *G*-representation features into *H*-representation features



Weiler & Cesa, *General E(2)-equivariant Steerable CNNs*, NeurIPS 2019

Direct sum of representations

the direct sum $V \oplus W$ of vector spaces V and W contains "stacked" vectors:

$$\begin{pmatrix} | \\ v \\ | \end{pmatrix} \oplus \begin{pmatrix} | \\ w \\ | \end{pmatrix} = \begin{pmatrix} | \\ v \\ | \\ w \\ | \end{pmatrix}$$

there is a corresponding **direct sum of representations**:

Definition A.3 (Direct sum representation). *Let (ρ_1, V_1) and (ρ_2, V_2) be G-representations. Their* direct sum $(\rho_1 \oplus \rho_2,\ V_1 \oplus V_2)$ *is defined by:*

$$(\rho_1 \oplus \rho_2)(g)\,(v_1 \oplus v_2) \;:=\; \rho_1(g)\,v_1 \oplus \rho_2(g)\,v_2$$

The two subspaces V_1 and V_2 of $V_1 \oplus V_2$ are transforming independently under this representation.

for matrix representations:

$$(\rho_1 \oplus \rho_2)(g) \;=\; \begin{pmatrix} \rho_1(g) & 0 \\ 0 & \rho_2(g) \end{pmatrix} \qquad \text{e.g.} \qquad \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \oplus \begin{pmatrix} 1 \end{pmatrix} = \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Direct sum of representations

the direct sum $V \oplus W$ of vector spaces V and W contains "stacked" vectors:

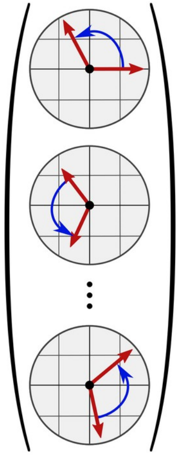
$$\begin{pmatrix} | \\ v \\ | \end{pmatrix} \oplus \begin{pmatrix} | \\ w \\ | \end{pmatrix} = \begin{pmatrix} | \\ v \\ | \\ w \\ | \end{pmatrix}$$

there is a corresponding **direct sum of representations**:

Definition A.3 (Direct sum representation). *Let* (ρ_1, V_1) *and* (ρ_2, V_2) *be* G*-representations. Their* direct sum $(\rho_1 \oplus \rho_2,\, V_1 \oplus V_2)$ *is defined by:*

$$(\rho_1 \oplus \rho_2)(g)\,(v_1 \oplus v_2) \;:=\; \rho_1(g)\, v_1 \oplus \rho_2(g)\, v_2$$

example: $\mathrm{SO}(2)$-equivariant MLP, processing a batch (= *direct sum*) of vectors in $\mathbb{R}^2$



Tensor product of representations

representations can also be combined by taking their **tensor product**:

Definition A.4 (Tensor product representation). *Let* (ρ_1, V_1) *and* (ρ_2, V_2) *be two* G*-representations. The* tensor product representation $(\rho_1 \otimes \rho_2,\, V_1 \otimes V_2)$ *acts on the tensor product of vector spaces as follows:*

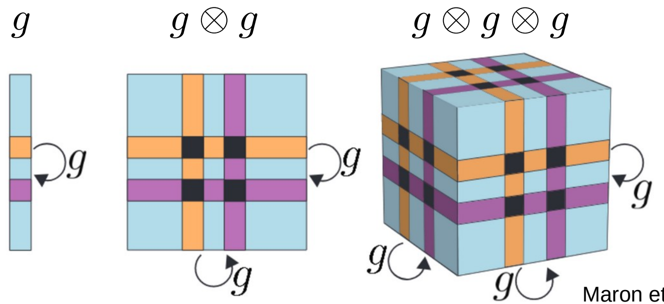
$$(\rho_1 \otimes \rho_2)(g)\,(v_1 \otimes v_2) \;:=\; \rho_1(g)\,v_1 \otimes \rho_2(g)\,v_2$$

for matrix representations:

$$(\rho_1 \otimes \rho_2)(g) \;=\; \begin{pmatrix} \rho_1(g)_{11} \cdot \rho_2(g) & \cdots & \rho_1(g)_{1\mathcal{V}} \cdot \rho_2(g) \\ \vdots & \ddots & \vdots \\ \rho_1(g)_{\mathcal{V}1} \cdot \rho_2(g) & \cdots & \rho_1(g)_{\mathcal{V}} \cdot \rho_2(g) \end{pmatrix} \qquad \text{e.g.} \qquad \begin{pmatrix} \cos\phi \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} & -\sin\phi \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \\ \sin\phi \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} & \cos\phi \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \end{pmatrix}$$

example:

higher order data tensors
(e.g. adjacency matrix):



Maron et al., *On the Universality of Invariant Networks*, ICML 2019

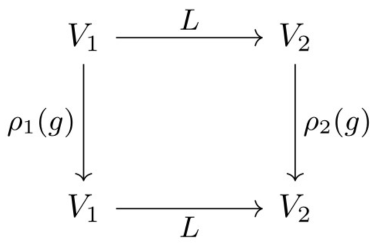
Intertwiners

linear equivariant maps between representations are called **intertwiners**

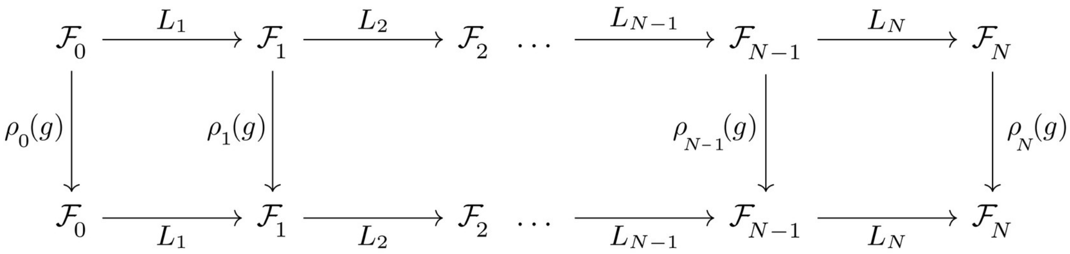
Definition A.7 (Intertwiner). *Let* (ρ_1, V_1) *and* (ρ_2, V_2) *be two* G*-representations. An* intertwiner *between them is an equivariant linear map* $L : V_1 \to V_2$*. It satisfies*

$$L \circ \rho_1(g) \;=\; \rho_2(g) \circ L \qquad \forall\, g \in G\,,$$

The vector space of intertwiners is usually denoted as $\mathrm{Hom}_G(V_1, V_2)$.



intertwiners are the main building blocks of equivariant NNs (interleaved with equivariant nonlinearities)



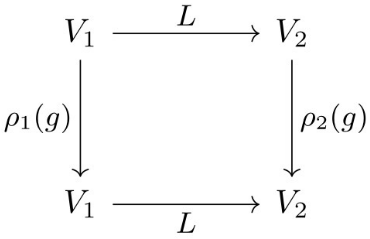
Intertwiners

linear equivariant maps between representations are called **intertwiners**

Definition A.7 (Intertwiner). *Let* (ρ_1, V_1) *and* (ρ_2, V_2) *be two* G*-representations. An* intertwiner *between them is an equivariant linear map* $L : V_1 \to V_2$*. It satisfies*

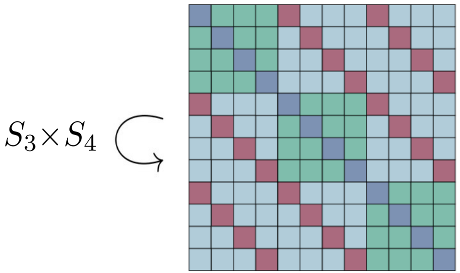
$$L \circ \rho_1(g) \;=\; \rho_2(g) \circ L \qquad \forall\, g \in G\,,$$

The vector space of intertwiners is usually denoted as $\mathrm{Hom}_G(V_1, V_2)$.



example: equivariant MLPs with… … representation spaces as feature vector spaces

… intertwiners as weight matrices



Hartford et al., *Deep Models of Interactions Across Sets*, ICML 2018

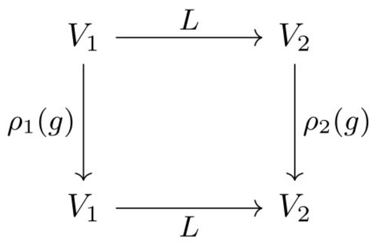
Intertwiners

linear equivariant maps between representations are called **intertwiners**

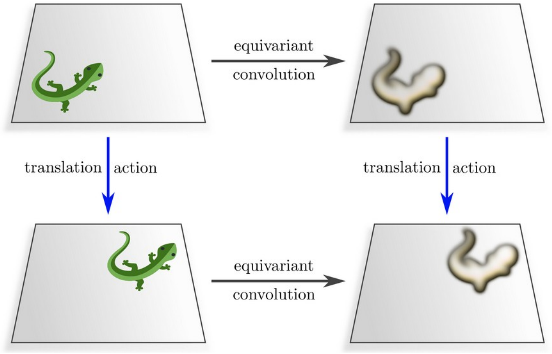
Definition A.7 (Intertwiner). *Let* (ρ_1, V_1) *and* (ρ_2, V_2) *be two* G*-representations. An* intertwiner *between them is an equivariant linear map* $L : V_1 \to V_2$*. It satisfies*

$$L \circ \rho_1(g) \;=\; \rho_2(g) \circ L \qquad \forall\, g \in G\,,$$

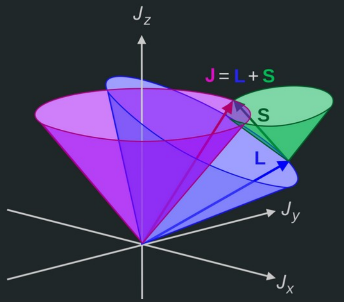
The vector space of intertwiners is usually denoted as $\mathrm{Hom}_G(V_1, V_2)$.



example: convolutions are linear + equivariant $\Longrightarrow$ they are translation intertwiners



Group representation theory



further topics:

- irreducible representations
- isomorphic representations
- Schur’s lemma
- Complete reducibility of unitary representations
- Clebsch-Gordan decomposition
- Peter-Weyl theorem and Fourier transforms

} useful for solving for intertwiner spaces

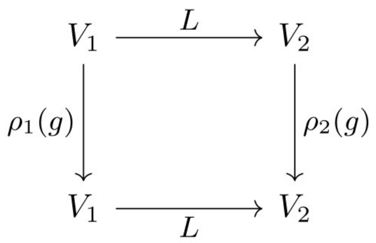
Intertwiners

linear equivariant maps between representations are called **intertwiners**

Definition A.7 (Intertwiner). *Let* (ρ_1, V_1) *and* (ρ_2, V_2) *be two* G*-representations. An* intertwiner *between them is an equivariant linear map* $L : V_1 \to V_2$*. It satisfies*

$$L \circ \rho_1(g) \;=\; \rho_2(g) \circ L \qquad \forall\, g \in G\,,$$

The vector space of intertwiners is usually denoted as $\mathrm{Hom}_G(V_1, V_2)$.



we need to solve for intertwiner spaces to build equivariant networks

Schur's lemma characterizes intertwiner spaces for *irreducible representations*

on the next slides:
- irreducible representations
- isomorphic representations
- Schur's lemma
- complete reducibility of unitary representations

Invariant subspaces, subrepresentations & irreps

representations may contain invariant **subrepresentations**

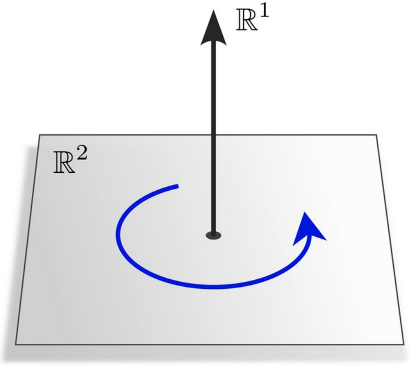
Definition A.5 (Invariant subspace, subrepresentation).

- *let* (ρ, V) *be a* G*-representation and consider a vector subspace* $W \subseteq V$
- W *is called* invariant *if it is closed under the action of* ρ*, i.e., if* $\rho(g)w \in W \quad \forall\, w \in W,\ g \in G.$
- *the restriction* $\rho|_W : G \to \mathrm{GL}(W)$ *of* ρ *to* W *is denoted as* subrepresentation

examples:

z-axis rotations: $\begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}$ has the z-axis and the xy-plane as invariant subspaces

trivial examples: the full subspace $W = V$ and the empty subspace $W = 0$ are always invariant



Invariant subspaces, subrepresentations & irreps

representations may contain invariant **subrepresentations**

Definition A.5 (Invariant subspace, subrepresentation).

- *let* (ρ, V) *be a* G*-representation and consider a vector subspace* $W \subseteq V$
- W *is called* invariant *if it is closed under the action of* ρ, *i.e., if* $\rho(g)w \in W \quad \forall\, w \in W,\ g \in G.$
- *the restriction* $\rho|_W : G \to \mathrm{GL}(W)$ *of* ρ *to* W *is denoted as* subrepresentation

Definition A.6 (Irreducible representation (irrep)). *A representation* (ρ, V) *is called* irreducible representation *(irrep) if it has only the two trivial subrepresentations* $W = V$ *and* $W = 0$.

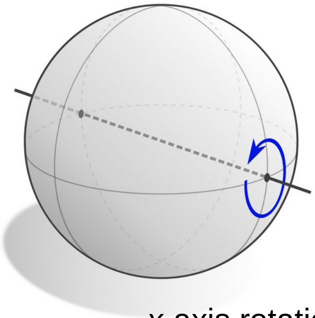
example: the *real* irreps of $\mathrm{SO}(2)$ are frequency-k rotation matrices $\begin{pmatrix} \cos k\phi & -\sin k\phi \\ \sin k\phi & \cos k\phi \end{pmatrix}, \quad k \in \mathbb{N}$

Isomorphic representations

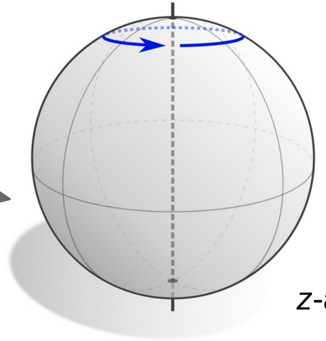
Definition A.7 (Equivalent (isomorphic) representations). *Two G-representations (ρ_1, V_1) and (ρ_2, V_2) are said to be* equivalent *or* isomorphic *if there exists an* invertible intertwiner, *i.e. a vector space isomorphism $L : V_1 \xrightarrow{\sim} V_2$ satisfying $L \circ \rho_1(g) = \rho_2(g) \circ L \;\; \forall\, g \in G$, between them.*

$$\begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \not\cong_{\mathrm{SO(2)\text{-}rep}} \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \cong_{\mathrm{SO(2)\text{-}rep}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{pmatrix}$$

not a vector
space isomorphism



x-axis rotations



z-axis rotations

Schur's lemma

intertwiners between irreducible representations are characterized by *Schur's lemma:*

Lemma A.18 (Schur's lemma). *Let* (ρ_1, V_1) *and* (ρ_2, V_2) *be irreducible* G*-representations, then:*

1) *non-trivial intertwiners* $V_1 \to V_2$ *exist only if the irreps are isomorphic, i.e. if* $\rho_1 \cong \rho_2$.

2) *if the representations agree,* $(\rho_1, V_1) = (\rho_2, V_2)$*, then:*

- for $\mathbb{C}$-reps: *the irrep intertwiner is a scalar multiple* $\lambda \cdot \mathrm{id}$ *of the identity*
- for $\mathbb{R}$-reps: *the irrep intertwiner is an endomorphism (easy to find)*

application example:

intertwiner constraint
reduction to irrep constraints,

then solving via Schur's lemma

$$W \cdot \rho_{\text{in}}(g) \;=\; \rho_{\text{out}}(g) \cdot W$$

$$\Leftrightarrow \qquad W \;=\; \rho_{\text{out}}(g) \cdot W \cdot \rho_{\text{in}}(g)^{-1}$$

irrep decomposition

$$\underbrace{\begin{pmatrix} W_{\text{irrep}}^{J_1 l_1} & W_{\text{irrep}}^{J_1 l_2} & \cdots \\ W_{\text{irrep}}^{J_2 l_1} & W_{\text{irrep}}^{J_2 l_2} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}}_{W_{\text{irrep}}} = \underbrace{\begin{pmatrix} \rho_{J_1}(g) & & \\ & \rho_{J_2}(g) & \\ & & \ddots \end{pmatrix}}_{\bigoplus_{J \in I_{\text{out}}} \rho_J(g)} \cdot \underbrace{\begin{pmatrix} W_{\text{irrep}}^{J_1 l_1} & W_{\text{irrep}}^{J_1 l_2} & \cdots \\ W_{\text{irrep}}^{J_2 l_1} & W_{\text{irrep}}^{J_2 l_2} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}}_{W_{\text{irrep}}} \cdot \underbrace{\begin{pmatrix} \rho_{l_1}(g)^{-1} & & \\ & \rho_{l_2}(g)^{-1} & \\ & & \ddots \end{pmatrix}}_{\bigoplus_{l \in I_{\text{in}}} \rho_l(g)^{-1}}$$

Compact groups & unitary representations

for *compact groups*, one may w.l.o.g. consider **unitary representations**

→ e.g. any subgroups of $\mathrm{O}(d)$ or $\mathrm{U}(d)$

Definition A.7 (Unitary group). *Let V be an inner product space. The* unitary group $\mathrm{U}(V)$ *is the group formed by all unitary transformations from V to itself:*

$$\mathrm{U}(V) \;=\; \big\{ g \in \mathrm{GL}(V) \,\big|\, \langle gv, gw \rangle_V = \langle v, w \rangle_V \quad \forall\, v, w \in V \big\} \;\leq\; \mathrm{GL}(V)$$

Definition A.8 (Unitary representation). *A* unitary representation *on an inner product space V is a (continuous) homomorphism*

$$\rho : G \to \mathrm{U}(V)\,.$$

Theorem A.9. *Every linear representation of a compact group on an inner product space is equivalent to a unitary representation.*

Complete reducibility

Theorem A.10 (Complete reducibility). *Let (ρ, V) be a finite dimensional unitary G-representation. It decomposes into a direct sum $\rho \cong \bigoplus_i \rho_i$ of unitary irreps ρ_i.*

Complete reducibility

Theorem A.10 (Complete reducibility). *Let (ρ, V) be a finite dimensional unitary G-representation. It decomposes into a direct sum $\rho \cong \bigoplus_i \rho_i$ of unitary irreps ρ_i.*

example: **Clebsch-Gordan decomposition** of *irrep tensor products*

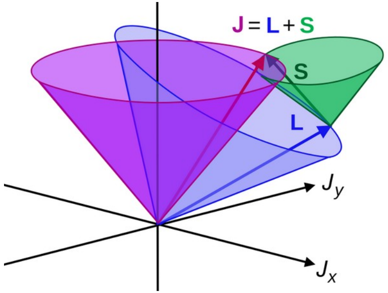
$G = \mathrm{SO}(2)$:

$$\rho_j \otimes \rho_l \;\cong\; \rho_{|j-l|} \oplus \rho_{j+l}$$

$$\begin{pmatrix} \cos j\phi \begin{pmatrix} \cos l\phi & -\sin l\phi \\ \sin l\phi & \cos l\phi \end{pmatrix} & -\sin j\phi \begin{pmatrix} \cos l\phi & -\sin l\phi \\ \sin l\phi & \cos l\phi \end{pmatrix} \\ \sin j\phi \begin{pmatrix} \cos l\phi & -\sin l\phi \\ \sin l\phi & \cos l\phi \end{pmatrix} & \cos j\phi \begin{pmatrix} \cos l\phi & -\sin l\phi \\ \sin l\phi & \cos l\phi \end{pmatrix} \end{pmatrix} \cong \begin{pmatrix} \begin{pmatrix} \cos(|j-l|\phi) & -\sin(|j-l|\phi) \\ \sin(|j-l|\phi) & \cos(|j-l|\phi) \end{pmatrix} & \\ & \begin{pmatrix} \cos((j+l)\phi) & -\sin((j+l)\phi) \\ \sin((j+l)\phi) & \cos((j+l)\phi) \end{pmatrix} \end{pmatrix}$$

$G = \mathrm{SO}(3)$:

$$\rho_L \otimes \rho_S \cong \bigoplus_{J=|L-S|}^{L+S} \rho_J$$

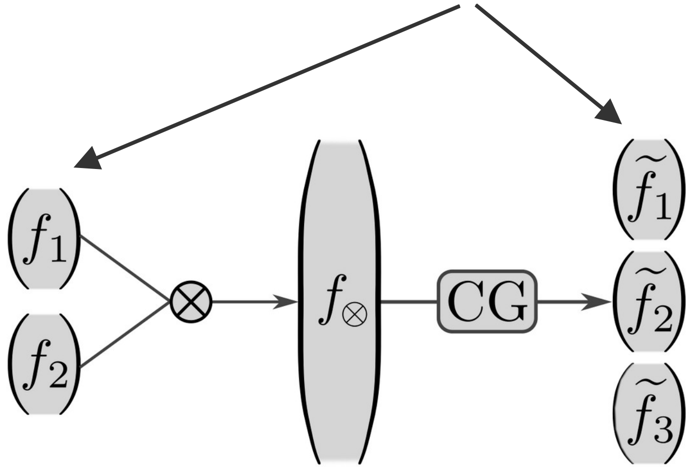


Complete reducibility

Theorem A.10 (Complete reducibility). *Let (ρ, V) be a finite dimensional unitary G-representation. It decomposes into a direct sum $\rho \cong \bigoplus_i \rho_i$ of unitary irreps ρ_i.*

example: **Clebsch-Gordan decomposition** of *irrep tensor products*

application: tensor product nonlinearities, acting on irrep-features



Weiler et al., *3D Steerable CNNs*, NeurIPS 2018

Anderson et al., *Cormorant: Covariant Molecular Neural Networks*, NeurIPS 2018

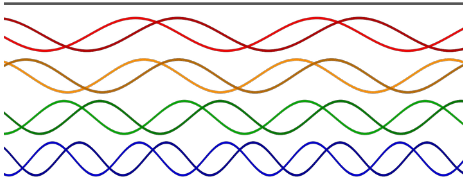
Peter-Weyl theorem

Peter-Weyl theorem / Fourier transforms on *homogeneous spaces*

fct on homogeneous space
regular/quotient representation
(in general non-finite dimensional)

harmonics
irrep spaces

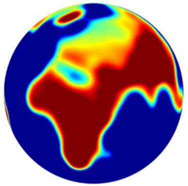
$G = (\mathbb{R}, +):$

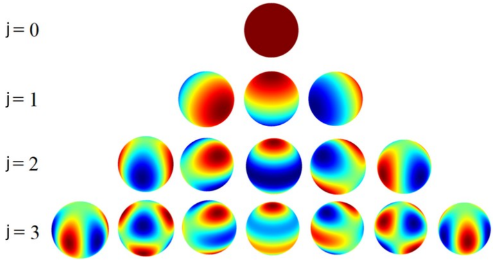


Peter-Weyl
irrep decomposition

(aka Fourier transform)

$G = \mathrm{SO}(3):$





Peter-Weyl for homogeneous spaces: Lang & Weiler, *A Wigner-Eckart theorem for group equivariant convolution kernels, ICLR 2021*