Chapter 2

Foundations of Pattern Recognition

What is Pattern Recognition?

Pattern Recognition is the theory of the best mapping of an unknown pattern or observation/monitoring $\mathbf{z}_i$ to a ***category*** or ***equivalence class*** $\mathcal{E}_j$ (classification).

An equivalence class $\mathcal{E}$ consists of a set of patterns $\{\mathbf{x}_i\}$ and a two-valued assignment/mapping (equivalence relation) with the three following properties:

a) $\mathbf{x}_i \sim \mathbf{x}_i$ reflexive (each element is equivalent to itself)

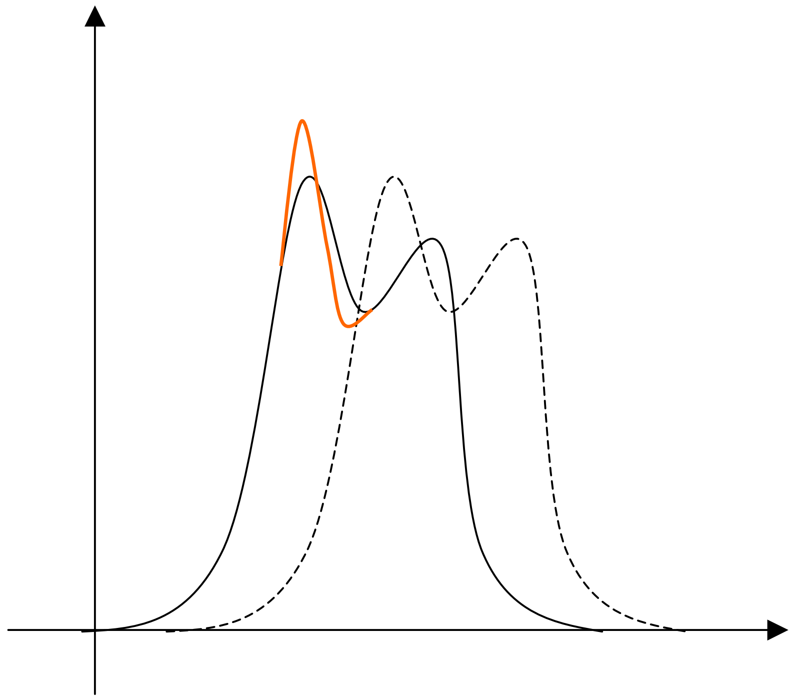
b) $\mathbf{x}_i \sim \mathbf{x}_j \Rightarrow \mathbf{x}_j \sim \mathbf{x}_i$ symmetric

c) $(x \sim y) \& (y \sim z) \Rightarrow x \sim z$ transitive

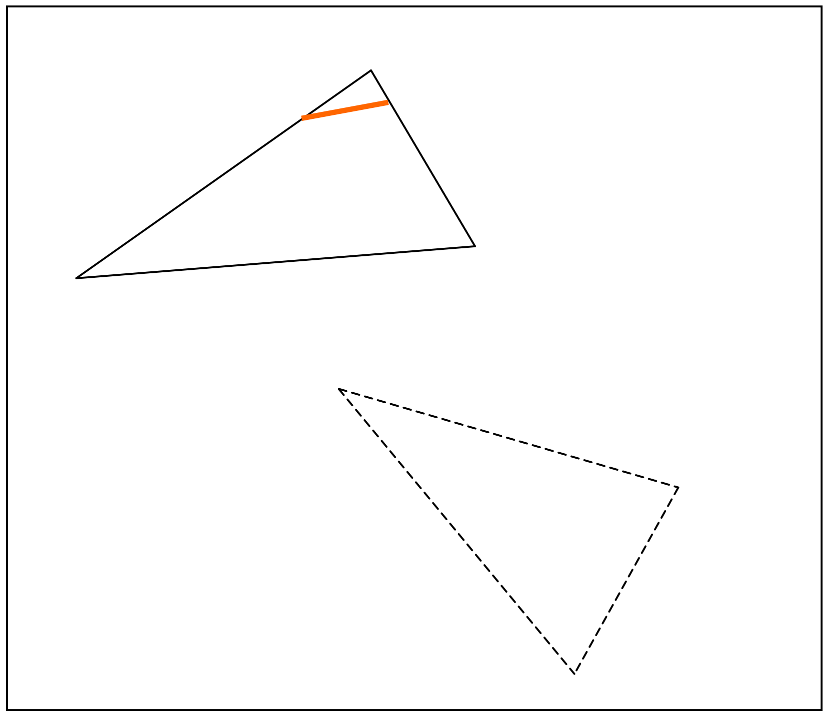
$\mathbf{x}_i \sim \mathbf{x}_j$: i.e. $\mathbf{x}_i$ is equivalent to $\mathbf{x}_j$ with regard to the relation ~

Relevant and irrelevant variations in signals and pictures

1D



2D



Categories or equivalence classes

Equivalence classes $\mathcal{E}$ can be defined in two ways, namely

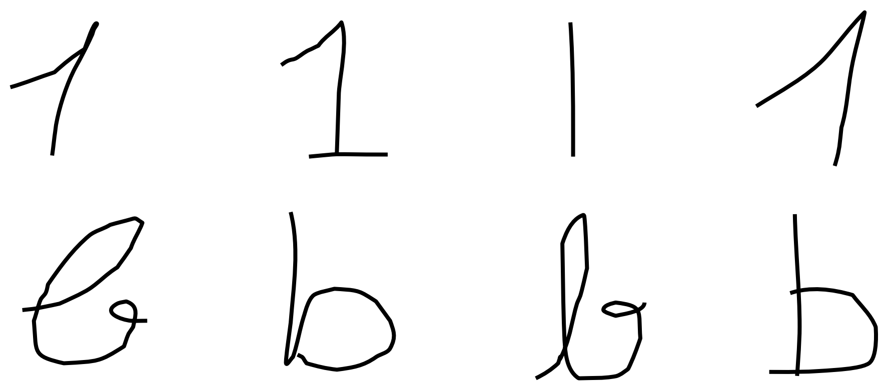
1) Declaration of all representatives of $\mathcal{E}$, since the variations cannot be expressed systematically, or:

2a) A generating element $\mathbf{x}_0$ and a mathematical group $\mathcal{G}$ (closure of the data)

2b) Closed mapping with subsequent mapping to a subspace (projection, occlusion), e.g. a moving 3D-object, that is subsequently projected to the cameraplane

Example for 1):

The set of all handwritten letters or digits:

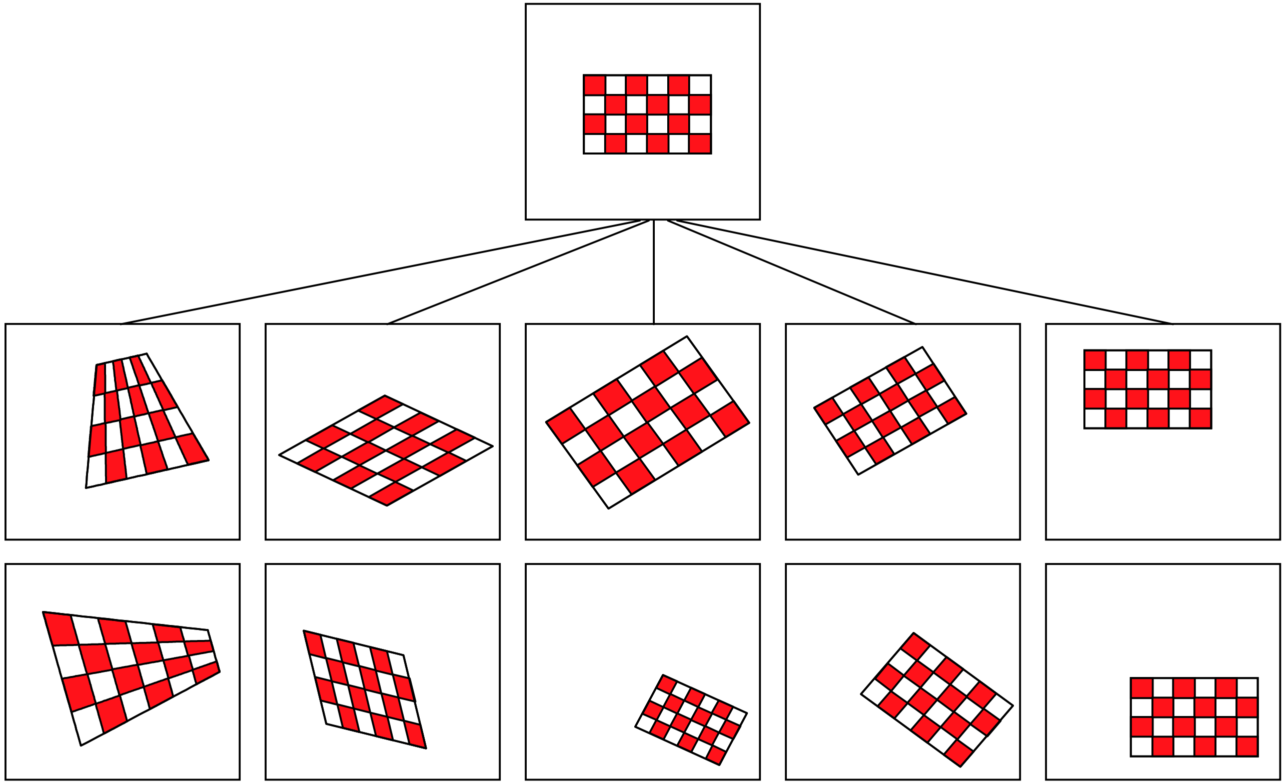


A parametric description of the equivalence class is nearly impossible.

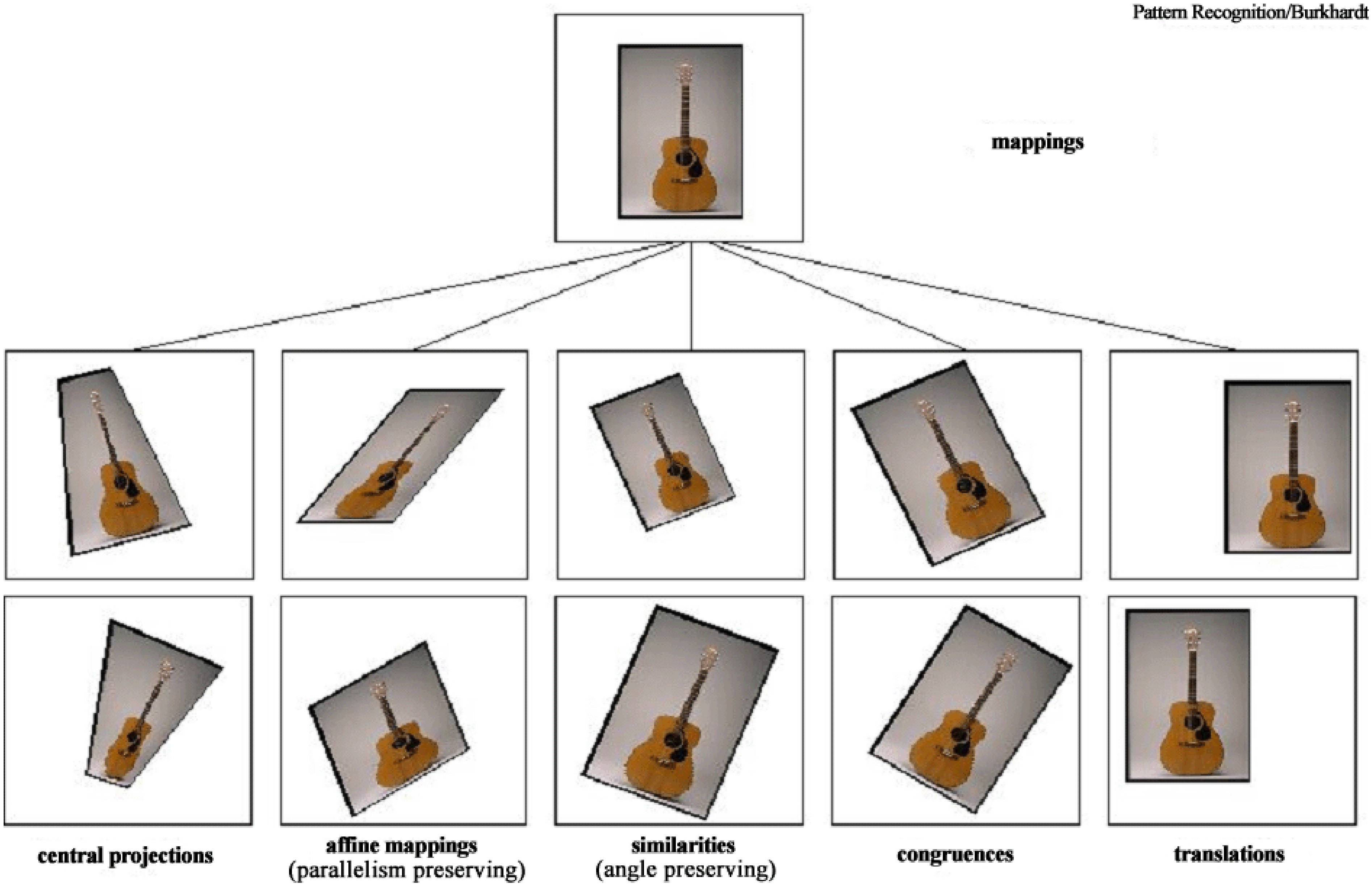
Equivalence class for the printed letter A

Figure 1.1. Variety of different images all representing the same character *A* (from Hofstadter's *Metamagical Themes*: *Questing for the Essence of Mind and Pattern* [HOF1985]).

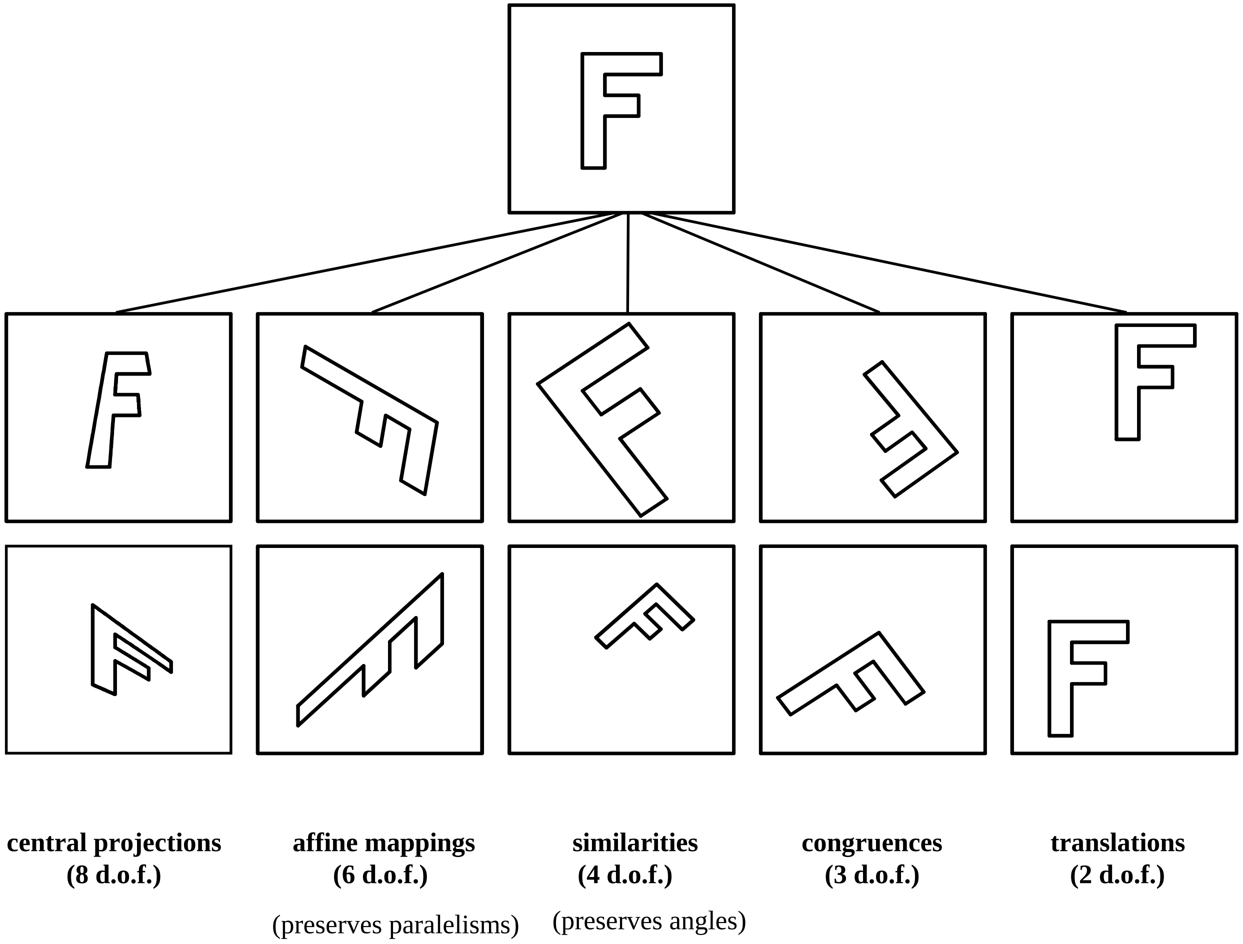
Example for 2a: geometric transformations



central projections (8 d.o.f.)	affine mappings (6 d.o.f.) (parallelism preserving)	similarities (4 d.o.f.) (angle preserving)	congruences (3 d.o.f.)	translations (2 d.o.f.)


Pattern Recognition/Burkhardt
mappings
central projections
affine mappings
(parallelism preserving)
similarities
(angle preserving)
congruences
translations

Geometric transformations



Example for 2b: Moving in Space (Translation and Rotation) followed by a mapping to a subspace; incomplete observations

Group

Def.: An algebraic structure $\mathcal{G}$ with a binary operation $\bullet$ is called a group, if the following axioms apply for any elements $a,b,c \in \mathcal{G}$:

1) $a \bullet (b \bullet c) = (a \bullet b) \bullet c$ associative
2) There is an ***identity element*** $e \in \mathcal{G}$, with
   $a \bullet e = e \bullet a = a \quad \forall a \in \mathcal{G}$
3) For all a there is an ***inverse element*** $a^{-1} \in \mathcal{G}$, with:
   $a \bullet a^{-1} = a^{-1} \bullet a = e$

From 1)-3) follows, that there is <u>exactly</u> one identity element and for all $a \in \mathcal{G}$ there is exactly one inverse element $a^{-1} \in \mathcal{G}$.

For abelian groups also applies commutativity:

$$a \bullet b = b \bullet a \qquad \forall a,b \in \mathcal{G}$$

Example for 2a):

The group $\mathcal{G}(\mathbf{p})$ of geometric mappings, that characterize the motions of objects, including projective mappings. An equivalence class can be obtained by giving a generating element $\mathbf{x}_0$ and a motion group, which can also be described parametrically with vector **p.** E.g. the group of planar (Euclidean) motions (translation and rotation) can achieve this:

$$\mathcal{E}_{\mathcal{G}}(\mathbf{x}_0) := \{ g_i(\mathbf{x}_0) \mid \forall g_i \in \mathcal{G} \}$$

Therefore, for two objects **x**,**y** of the same equivalence class applies:

$$x \overset{\mathcal{G}}{\sim} y \Leftrightarrow \exists g_i \in \mathcal{G}: x = g_i(y)$$

Affine transformations wrt. two-dimensional coordinates $\mathbf{t}=(t_0,t_1)$:

$$\mathbf{t}' = \mathbf{A}\mathbf{t} + \mathbf{a}$$

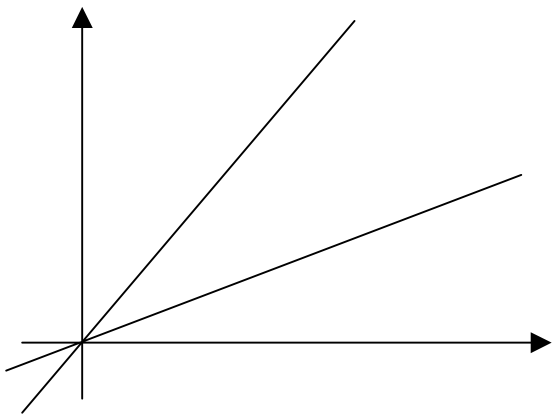
with:

$\mathbf{A}=\mathbf{I}$	the group of translations	$\mathcal{T}$	(dim($\mathbf{p}$)=2)
$\mathbf{A}^T\mathbf{A}=\mathbf{I}$	the group of congruences	$\mathcal{C}$	(dim($\mathbf{p}$)=3)
$\mathbf{A}^T\mathbf{A}=k\mathbf{I}$	the group of similarities	$\mathcal{S}$	(dim($\mathbf{p}$)=4)
$\det(\mathbf{A})\neq 0$	the group of affine mappings	$\mathcal{A}$	(dim($\mathbf{p}$)=6)

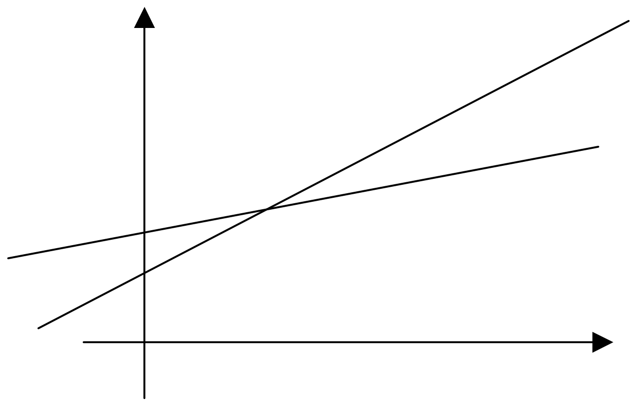
The Affine Transformation

$$\mathbf{t}' = \mathbf{A}\mathbf{t} + \mathbf{a}$$

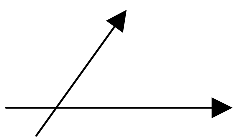
6 degrees of freedom



linear (homogeneous) transformations



affine transformations

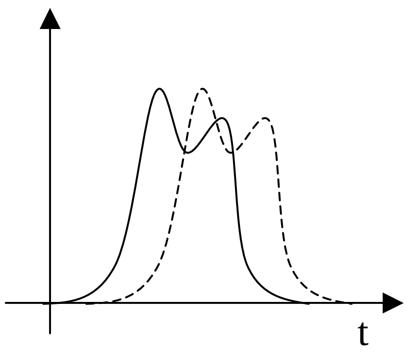


affine (non-rectangular) coordinates, as a generalization of the cartesian coordinates

The group of translations $\mathcal{T}$ for continuously defined signals and images

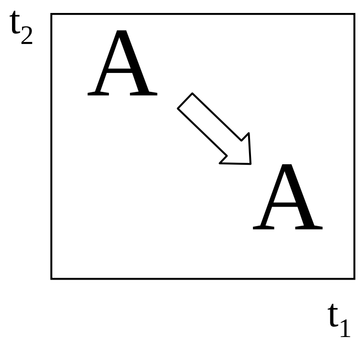
The set of translations $\{\tau\}$ form regarding the combination through composition $\tau_1(\tau_2(\ldots))=(\tau_1\bullet\tau_2)(\ldots)$ an abelian group.

1D



$$x(t') = x(t-a)$$

2D



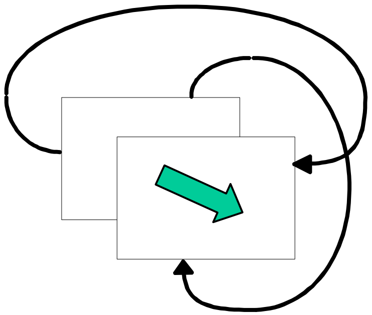
$$X(t_1', t_2') = X(\mathbf{t})$$
$$= X(t_1' - a_1, t_2' - a_2) = X(\mathbf{t} - \mathbf{a})$$

The group of translations $\mathcal{T}$ as a closed operation

On ***infinite*** coordinates the translation can be defined as a closed operation. Closure is neccessary to ensure that during operation elements do not disappear and other elements appear. It is required that:

$$\mathbf{x} \in \mathcal{X} \Rightarrow \tau(\mathbf{x}) \in \mathcal{X} \qquad \forall \tau \in \mathcal{G}$$

When translating, e.g. signals or images, that are defined in a ***finite*** domain (signal or image window), closure of data is achieved by ***cyclic shifting***. For infinte coordintes closure can be achieved by periodic continuation of a finite domain or window.



Compact group: the parameters that describe the group are limited to a finite domain!!

A mathematic group ensures closure, since an inverse element exists!

The group of translations $\mathcal{T}$ for sampled finite signals and images

Sampled finite patterns:

$$\mathbf{x} := \{x_i\} \qquad i = 0, \cdots, N-1$$
$$\mathbf{X} := \{X_{i,j}\} \qquad i, j = 0, \cdots, N-1$$

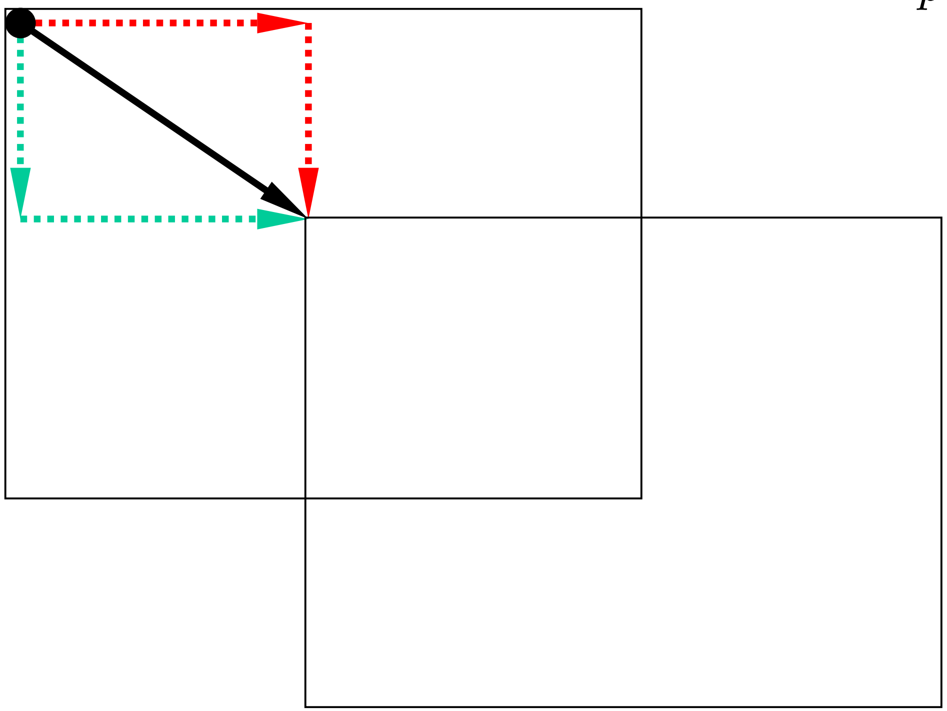
Translation as cyclic permutation:

$$\tau_k(\mathbf{x}) := \{x_{(i+k)\mathrm{mod}N}\} = \{x_{<i+k>_N}\}$$
$$\tau_{k,l}(\mathbf{X}) := \{X_{<i+k>_M, <j+l>_N}\}$$

Example:

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix} \Rightarrow \tau_{2,1}(\mathbf{A}) = \begin{bmatrix} 10 & 11 & 12 & 9 \\ 14 & 15 & 16 & 13 \\ 2 & 3 & 4 & 1 \\ 6 & 7 & 8 & 5 \end{bmatrix}$$

The 2D-translation can be factorized in two 1D-translations (row/column permutation):



$$\tau_{p,r} = \tau_p \bullet \tau_r = \tau_r \bullet \tau_p$$

The cyclic permutation of rows and columns of an image matrix using permutation matrices:

$$\tau_{1,1}(\mathbf{A}) = \overbrace{\mathbf{P}^T \underbrace{\mathbf{A}\mathbf{P}}_{\text{row permutation}}}^{\text{column permutation}} \qquad \text{resp:} \quad \tau_{p,r}(\mathbf{A}) = (\mathbf{P}^p)^T \mathbf{A} \mathbf{P}^r$$

The permutation matrix is orthogonal and thus:

$$\mathbf{P}^{-1} = \mathbf{P}^T$$

For example for N=4:

$$\mathbf{P} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \quad \textbf{with:} \; \mathbf{P}^0 = \mathbf{P}^4 = \mathbf{I}$$

The group of congruences $\mathcal{C}$ for continuously defined images

The group of *congruences* results from *translation* and *rotation*, which is also called *Euclidian* motion.

$$\mathbf{t'} = \mathbf{A}\mathbf{t} + \mathbf{a}$$

$$\mathbf{A} = \begin{pmatrix} \cos(\varphi) & \sin(\varphi) \\ -\sin(\varphi) & \cos(\varphi) \end{pmatrix}; \quad \mathbf{a} \neq \mathbf{0}$$

The rotation matrix **A** is orthogonal and thus:

$$\mathbf{A}^{-\mathbf{1}} = \mathbf{A}^T \text{ and: } det(\mathbf{A}) = \mathbf{1}$$

The group of similarities $\mathcal{S}$ for continuously defined images

The group of *similarities* results from *translation*, *rotation* and *compression*:

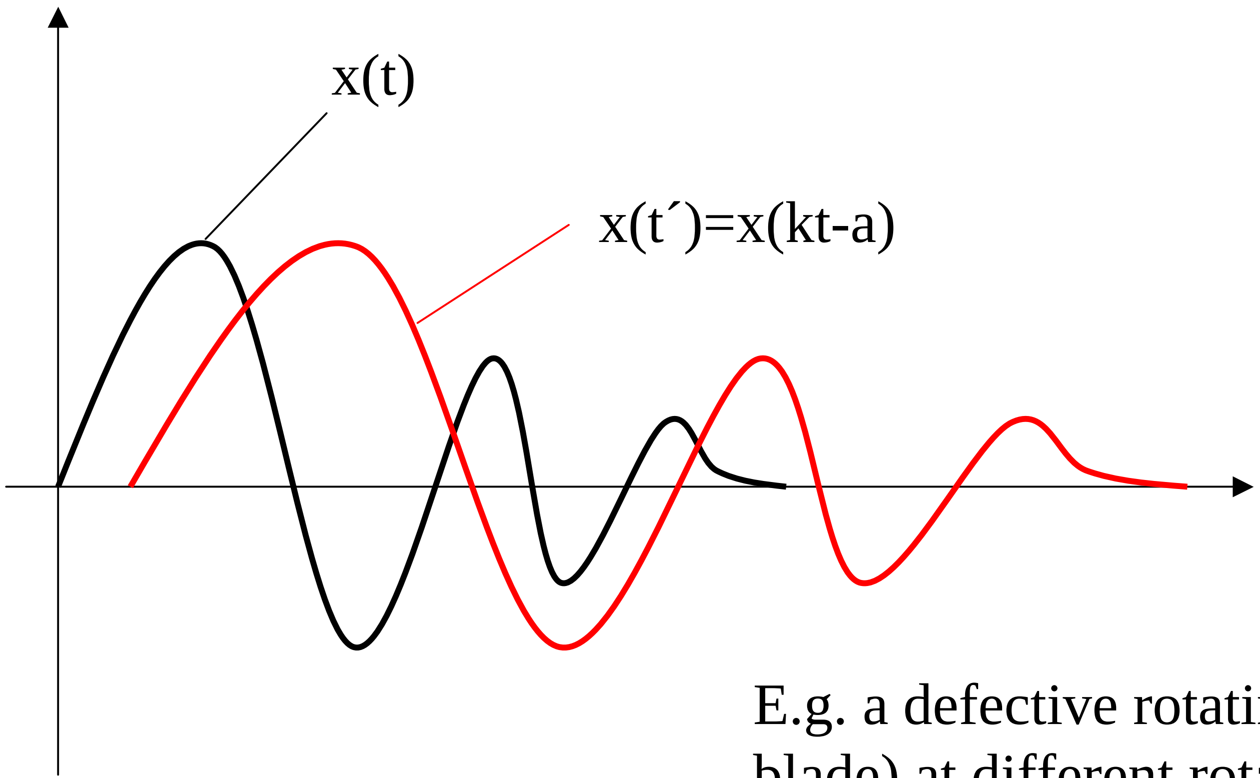
$$\mathbf{t'} = \mathbf{A}\mathbf{t} + \mathbf{a}$$

$$\mathbf{A} = k \begin{pmatrix} \cos(\varphi) & \sin(\varphi) \\ -\sin(\varphi) & \cos(\varphi) \end{pmatrix}; \quad \mathbf{a} \neq \mathbf{0}, k \neq 0$$

Hence: $\mathbf{A}\mathbf{A}^T = k^2 \mathbf{I}$ as well as: $det(\mathbf{A}) = k^2$

And: $\mathbf{A}^{-\mathbf{1}} = \frac{1}{k^2}\mathbf{A}^T$

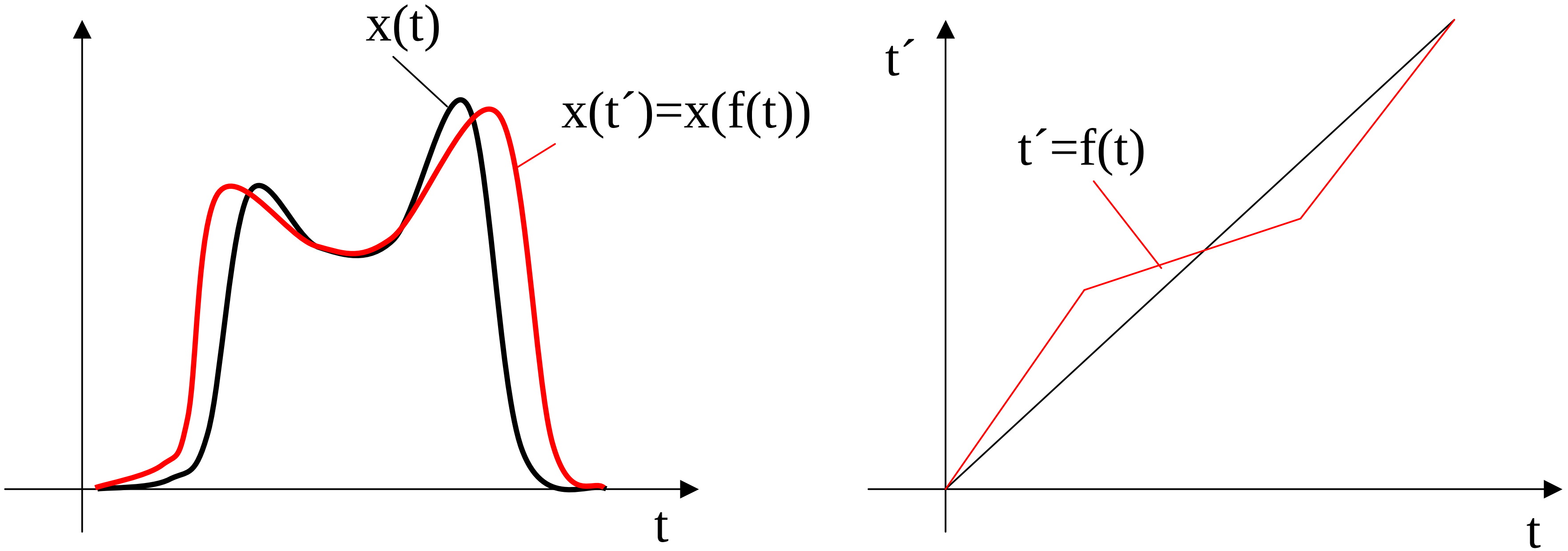
E.g.: compression and shifting of a one-dimensional signal



E.g. a defective rotating piece (e.g. turbine blade) at different rotational speed (or even timevariant rotational speed (see below).

More general equivalence classes: arbitrary time modulations

$x(t')=x(f(t))$ f(t) monotonic (causality, time does not run backwards), but other arbitrary



similarity in the sense of a generalized metric!

More general equivalence classes: arbitrary time modulation

- E.g.: different speakers speed during speech recognition or varying tempi during music recognition
- This case is more difficult than the turbine blade example, because a tachometer could easily be attached and the time jitter could be corrected