On Heads and Tails (2)

Foundations seminar

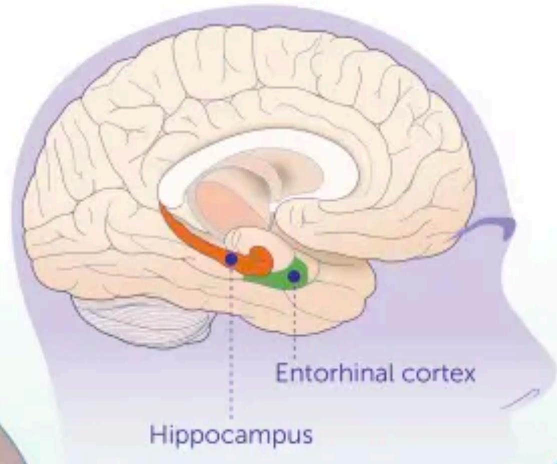
Raul Vicente, 18th March 2025

frontiers

How our body's internal navigation system works

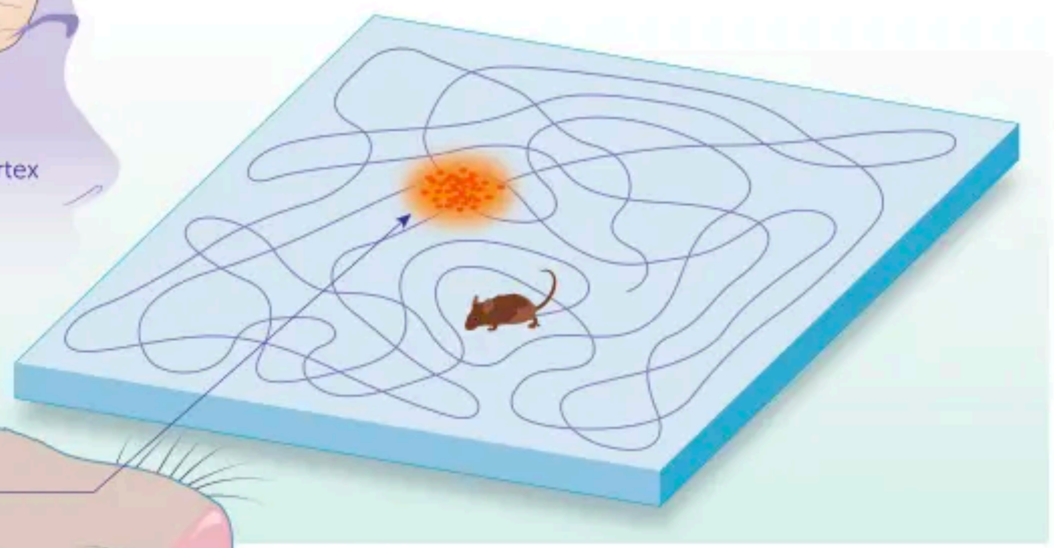
Different types of cells in the brain help us create an internal spatial representation of our environment or "cognitive map" and play a unique role in navigation.

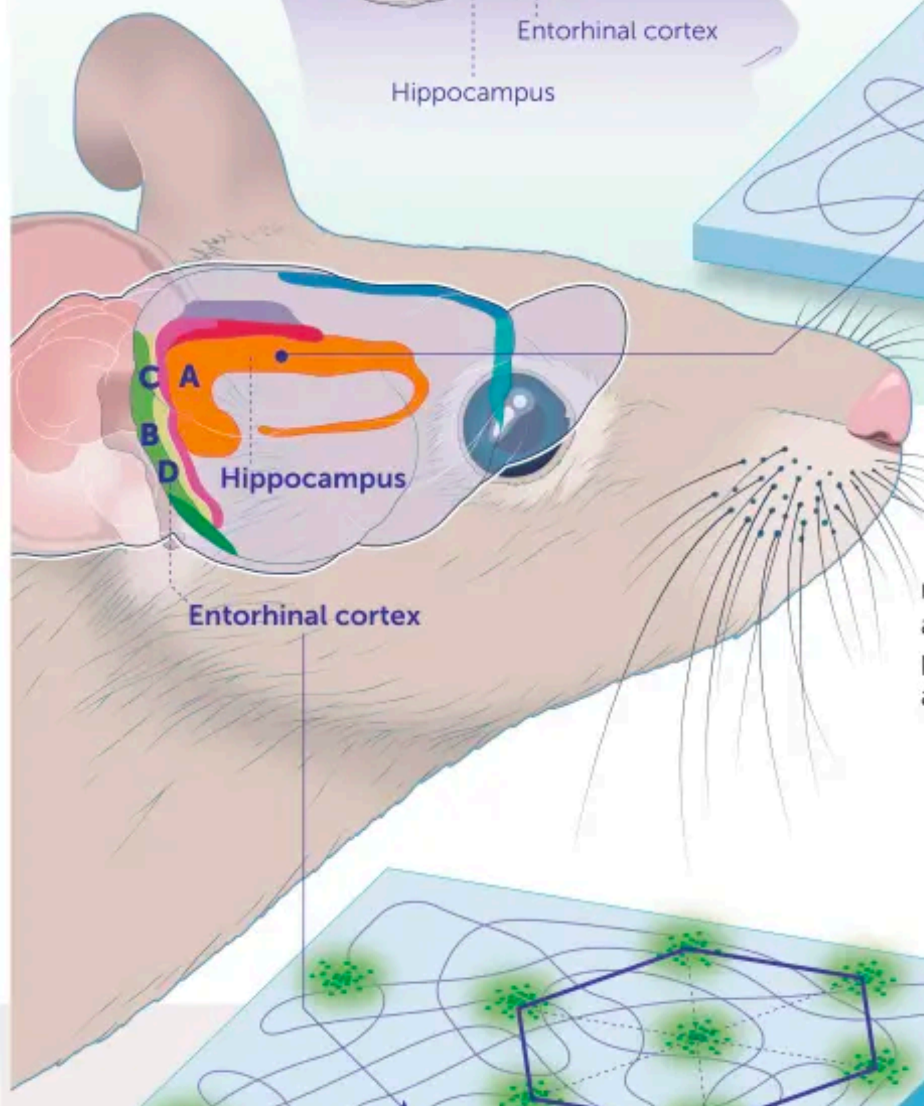
Location-related brain cells studied in rats have also been found in the human brain.



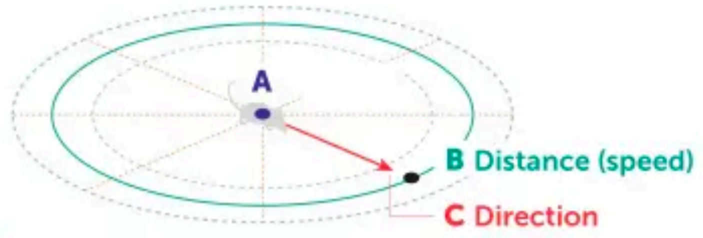
Place cells

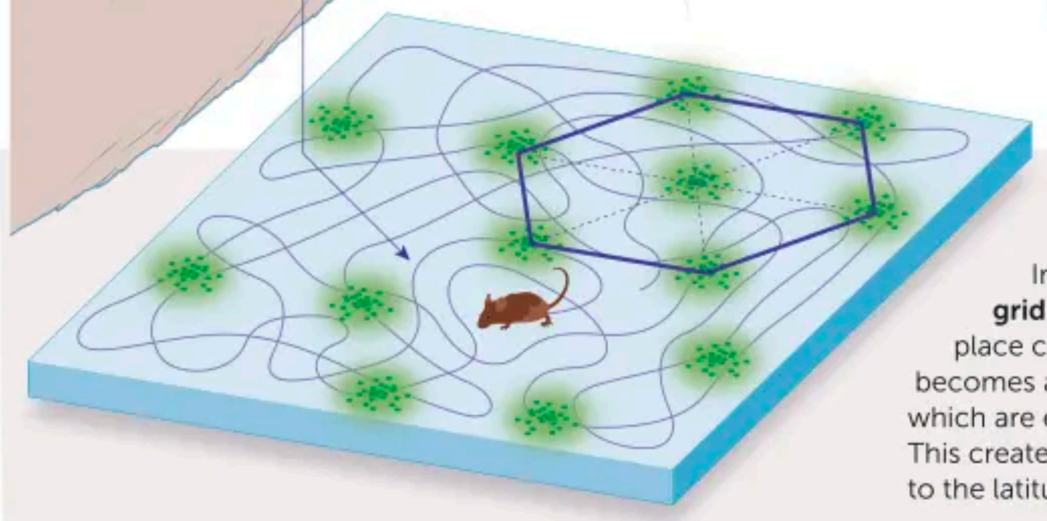
In 1971, researchers John O'Keefe and John Dostrovsky worked with rats to discover place cells. Located in the hippocampus region of the brain, different place cells are acitvated only at specific locations. They tell us where we are.





To navigate successfully in the environment, you need to know your starting position, walking direction, and speed. In addition to **place cells (A)**, other cell types have a unique role in navigation. **Speed cells (B)** increase their activity when an animal moves faster, helping it calculate the distance it passes when it moves. **Head direction cells (C)** inform the animal in which direction it is heading.



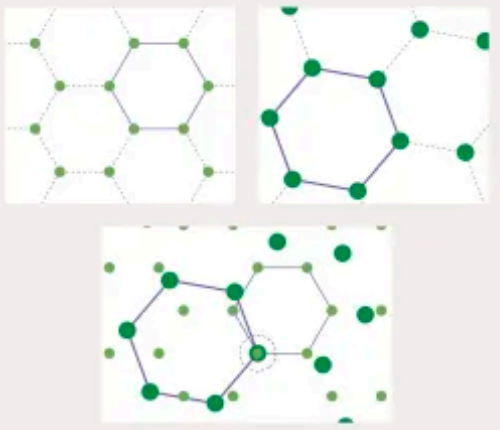


Grid cells

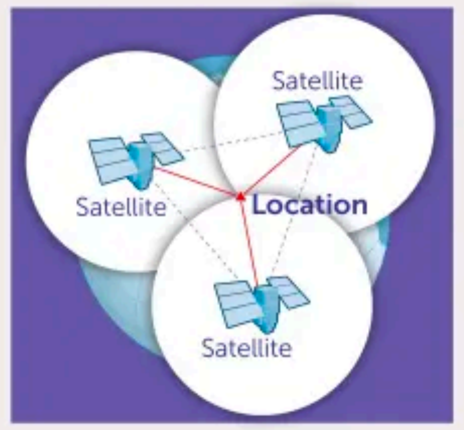
In 2005 May-Britt and Edvard Moser discovered **grid cells**, located in the entorhinal cortex **(D)**. Unlike place cells, active at only one location, each grid cell becomes active in multiple locations in an environment, which are evenly spaced in a repeating hexagonal pattern. This creates a system of coordinates with a metric, similar to the latitude-longitude grid we use in maps.

Grid cells help determine location ...

Different grid cells produce different hexagonal patterns at different scales and shifted with respect to the grids of other cells. Using the overlapping grids of several cells, location can be uniquely identified by coincident firing of multiple grid cells.



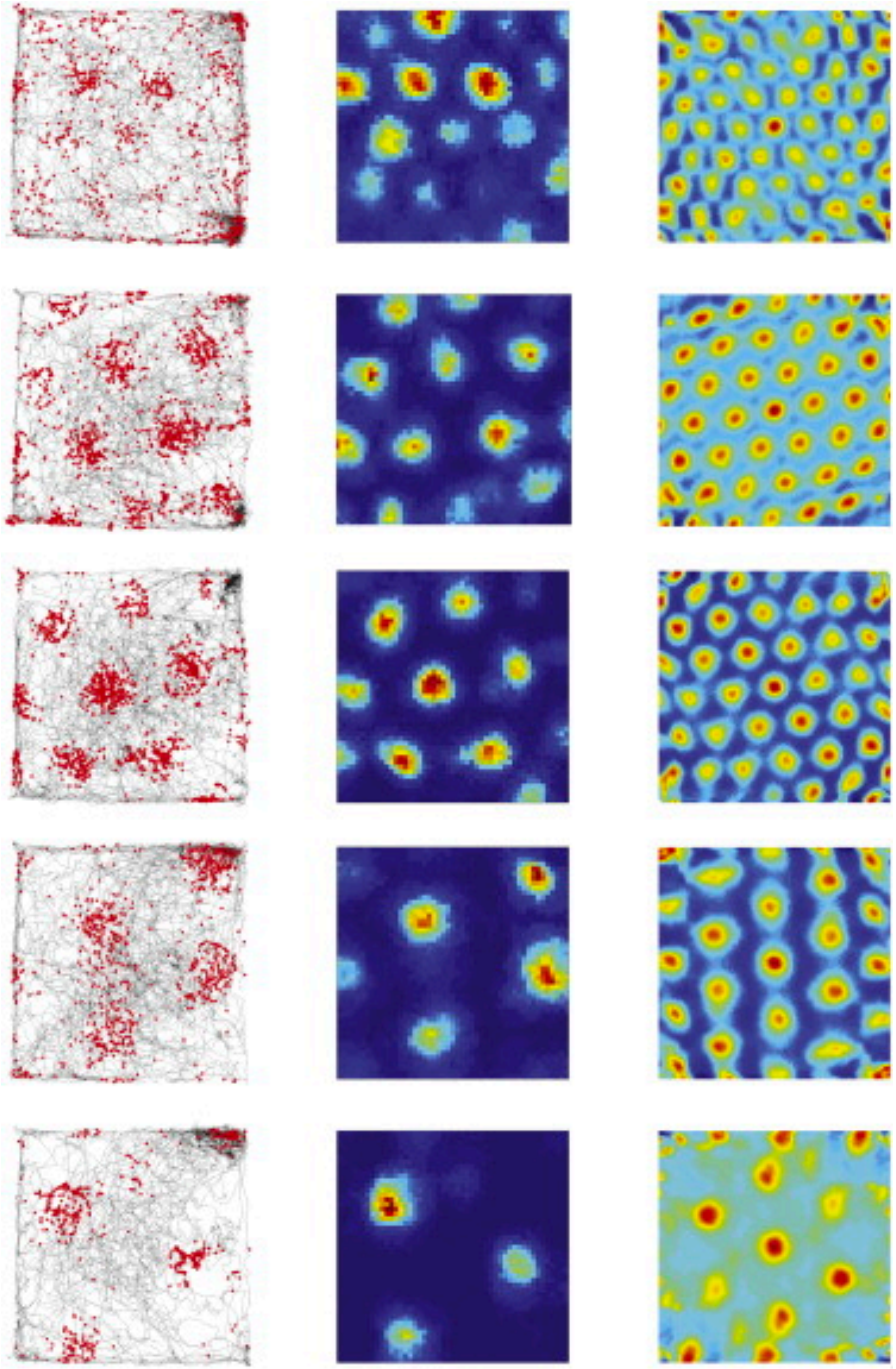
... similarly to GPS systems

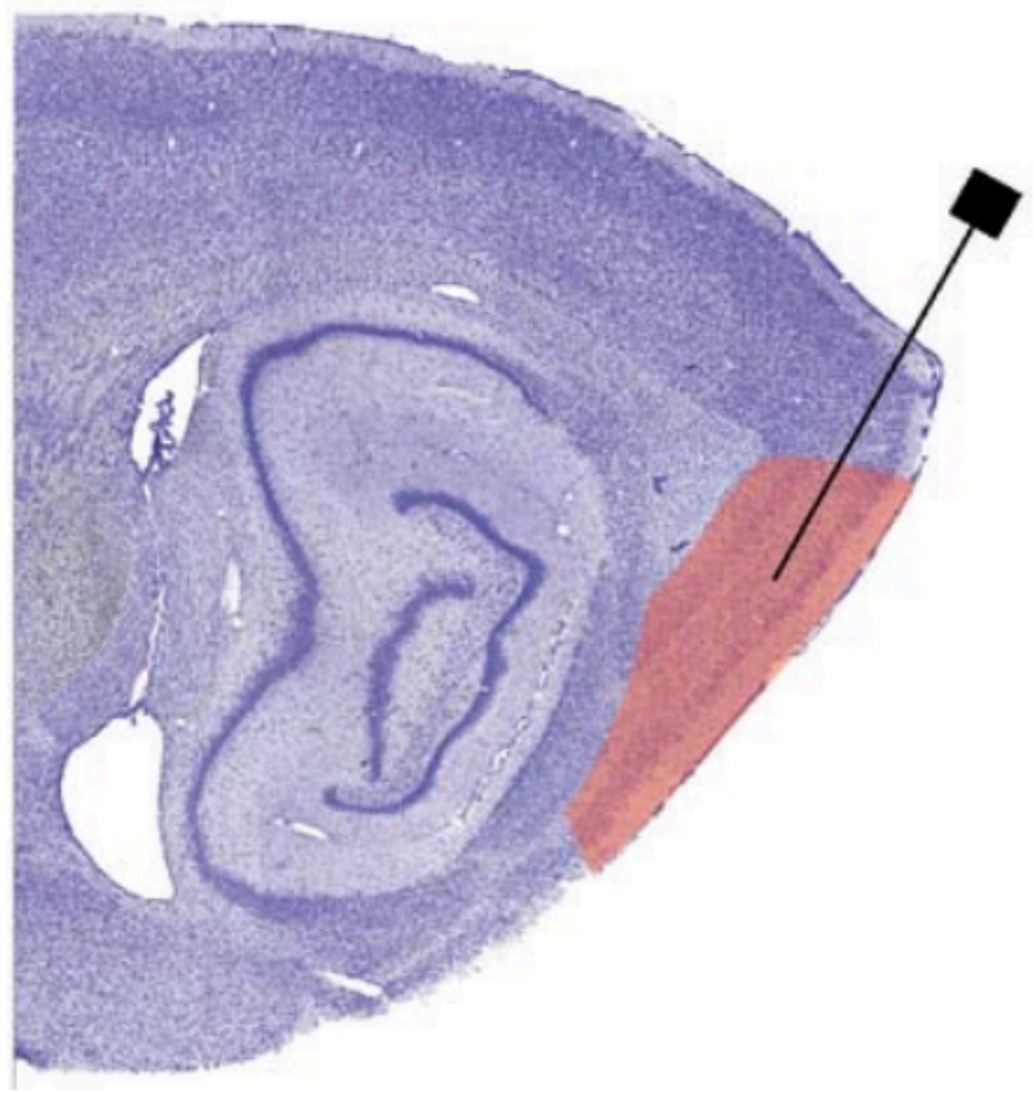


In a similar way, in GPS systems the distance measuring of one satelite alone can't pinpoint a location, but the overlap of at last three satelites is used to pinpoint the exact location of a receiver.

Source: *How do we Find our Way? Grid Cells in the Brain*, by May-Britt Moser. Published by Frontiers for Young Minds, Sep. 2021

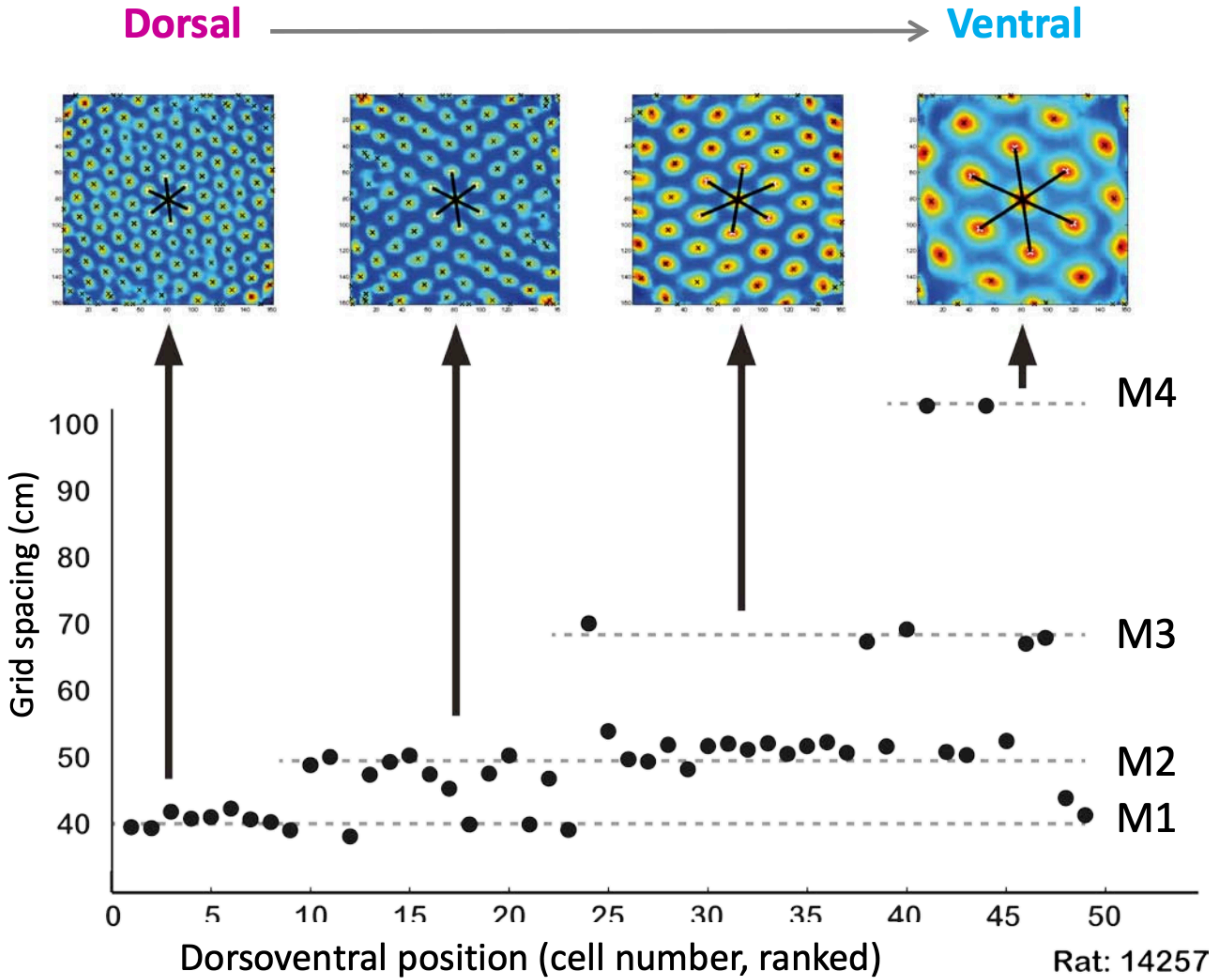
Infographic by 5W Infographic





Tor & Hanne Stensola

Trygve Solstad
Kristian Frøland



Stensola et al. *Nature*, 492, 72-78 (2012)

Grid **pattern** -> Superposition of Fourier modes (**Eigenfunctions** of Laplacian **on the domain**)

Grid **scale** -> **Eigenvalue** associated to contributing eigenfunctions

1 How do these neurons activate on a hexagonal pattern?

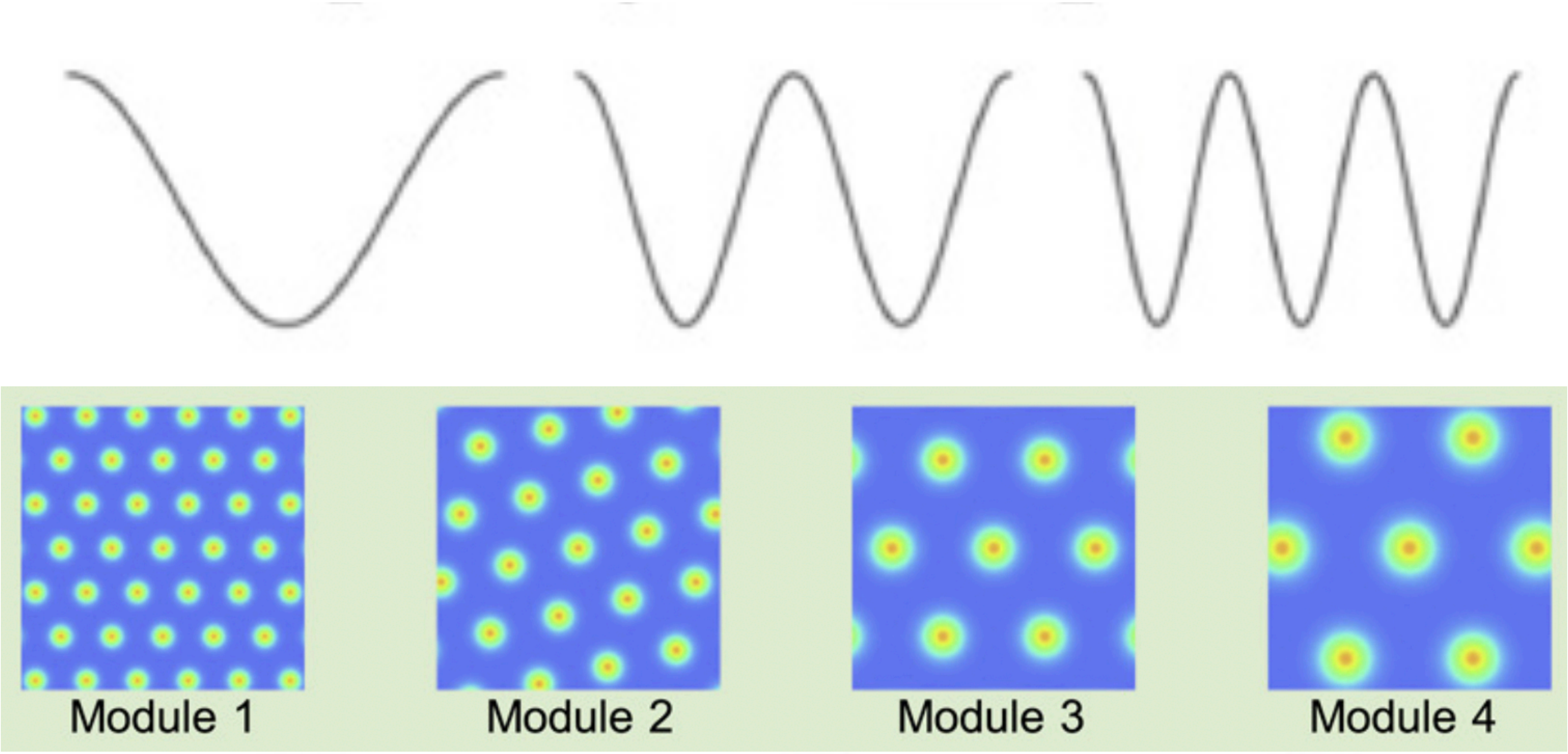
2 What is their functional role?

1 How do these neurons activate on a hexagonal pattern?

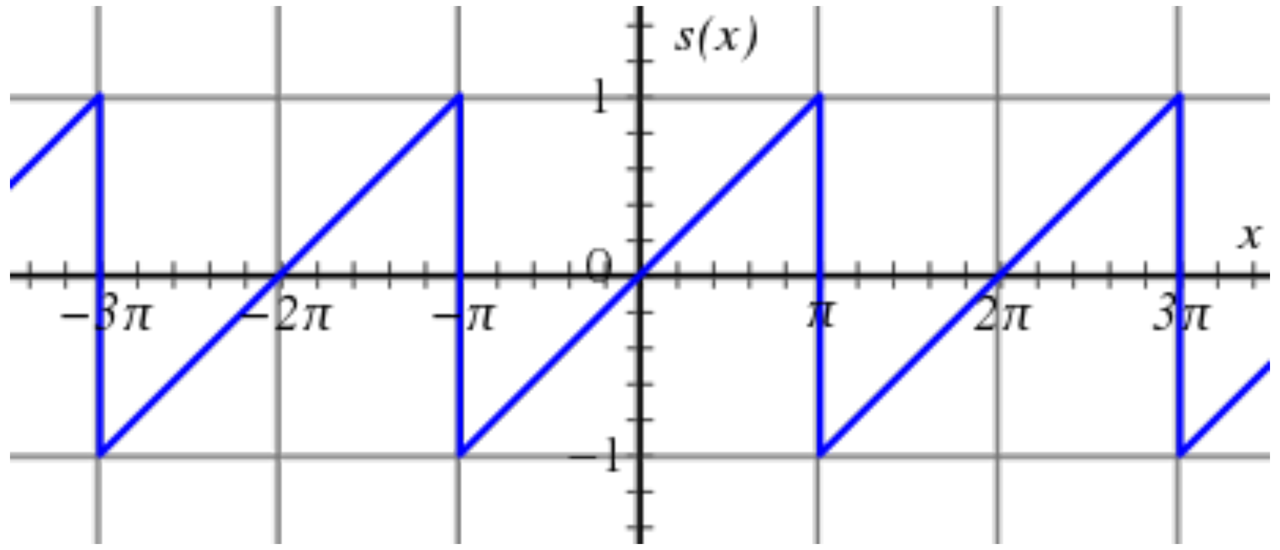
2 What is their functional role?

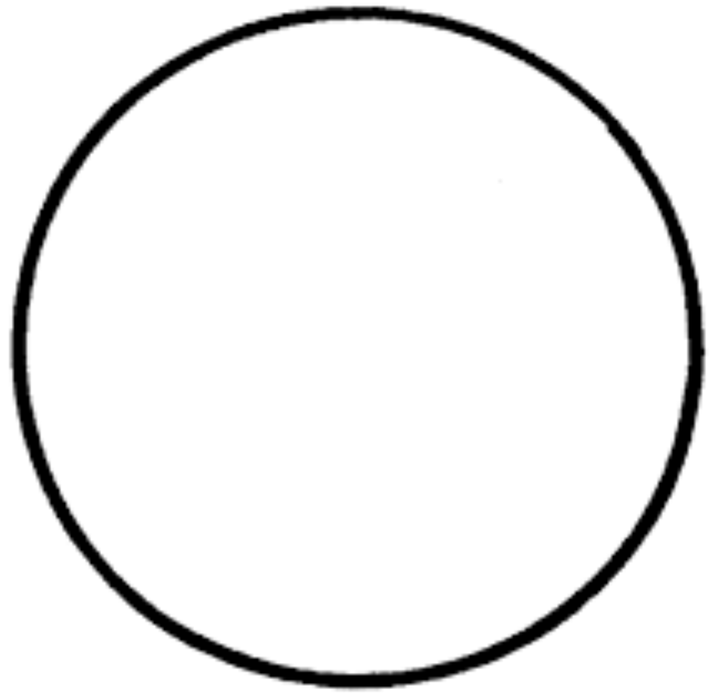
Research hypothesis

Grid cells = Fourier basis for the 2D environments



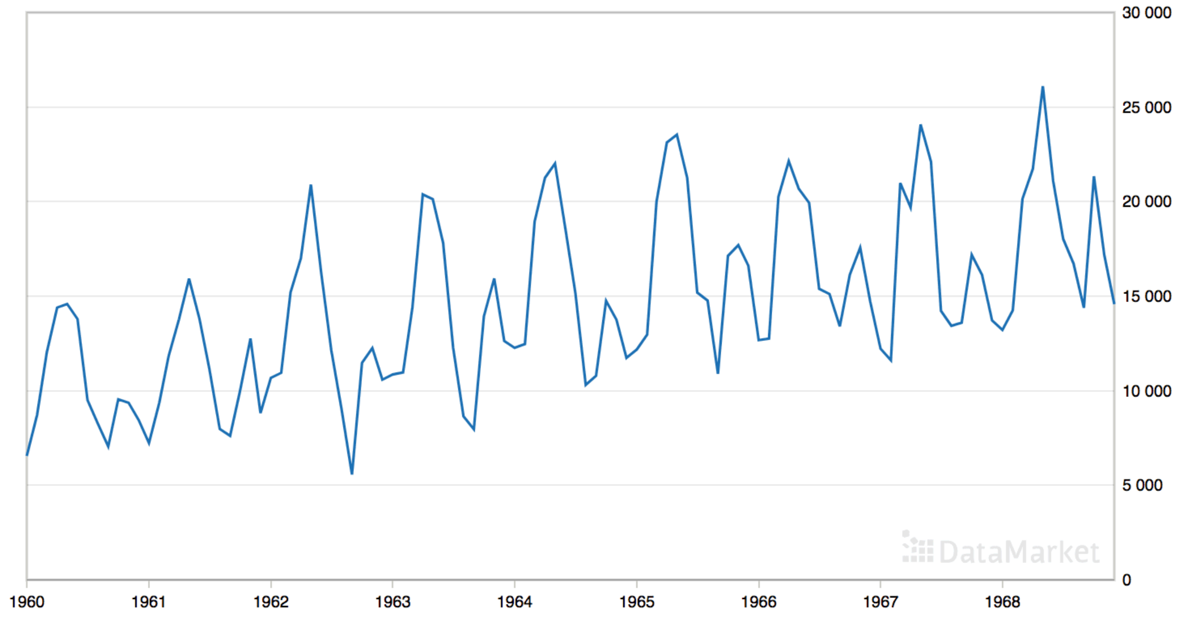
Fourier generalization to manifolds





$$\frac{d^2}{dx^2}\sin(wx) = -w^2\sin(wx)$$

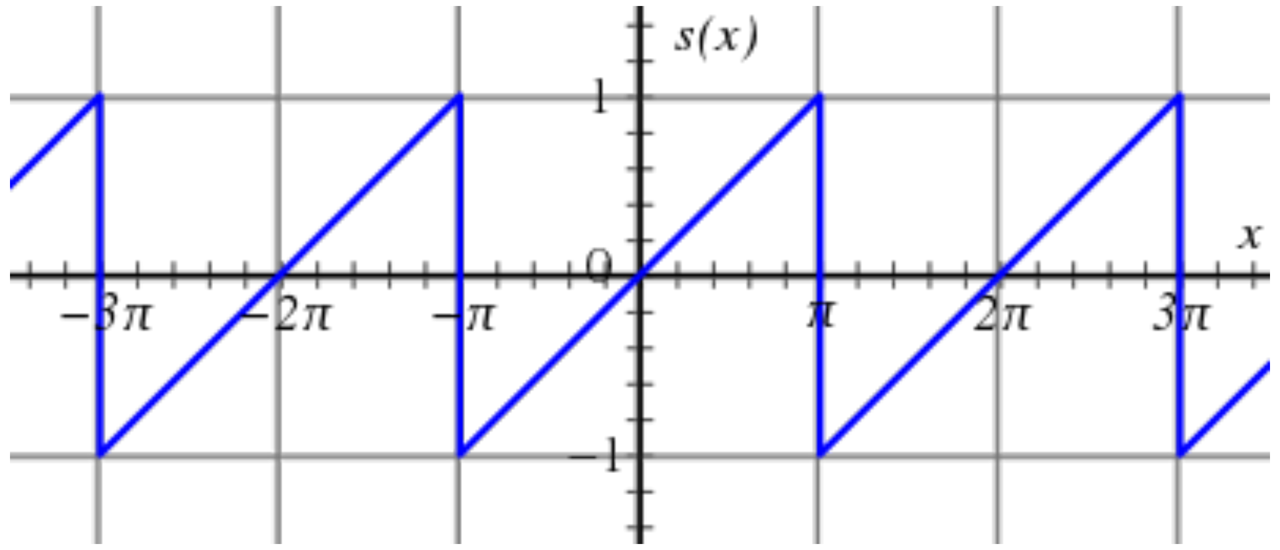
$$\frac{d^2}{dx^2}\cos(wx) = -w^2\cos(wx)$$

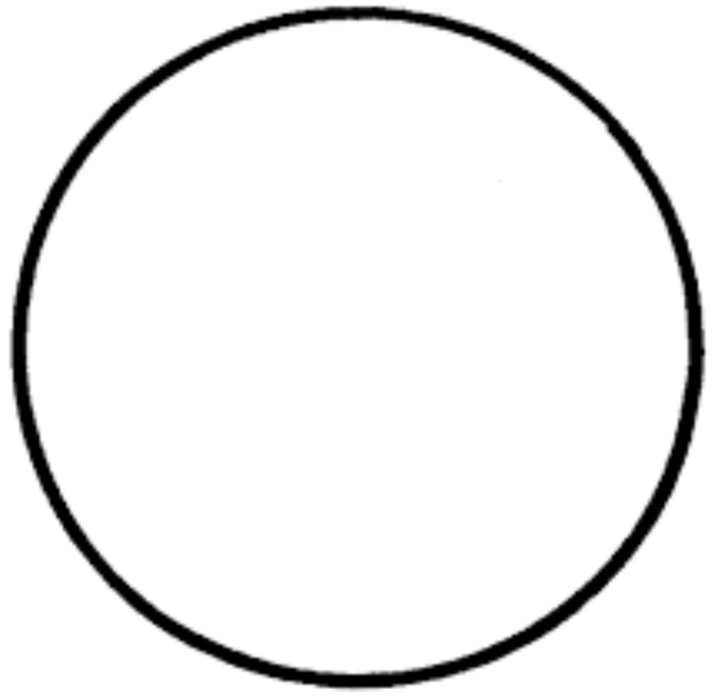


$$\frac{d^2}{dx^2}\sin(wx) = -w^2\sin(wx)$$

$$\frac{d^2}{dx^2}\cos(wx) = -w^2\cos(wx)$$

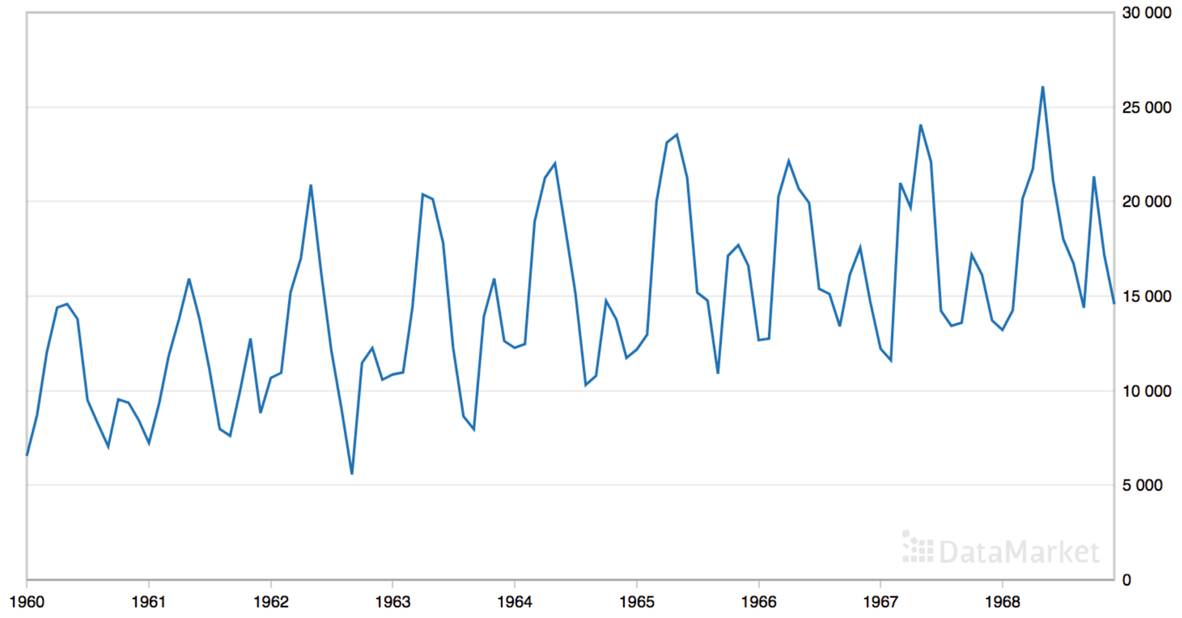
Fourier generalization to manifolds





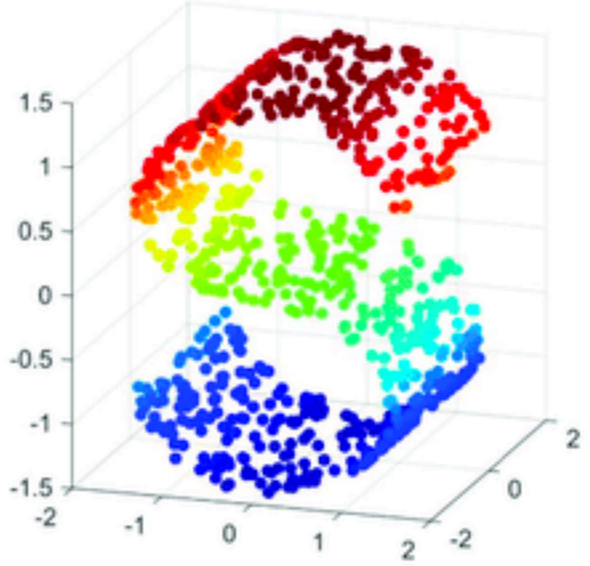
$$\frac{d^2}{dx^2}\sin(wx) = -w^2\sin(wx)$$

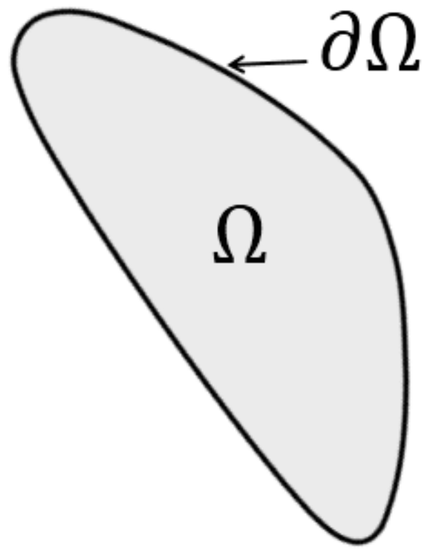
$$\frac{d^2}{dx^2}\cos(wx) = -w^2\cos(wx)$$



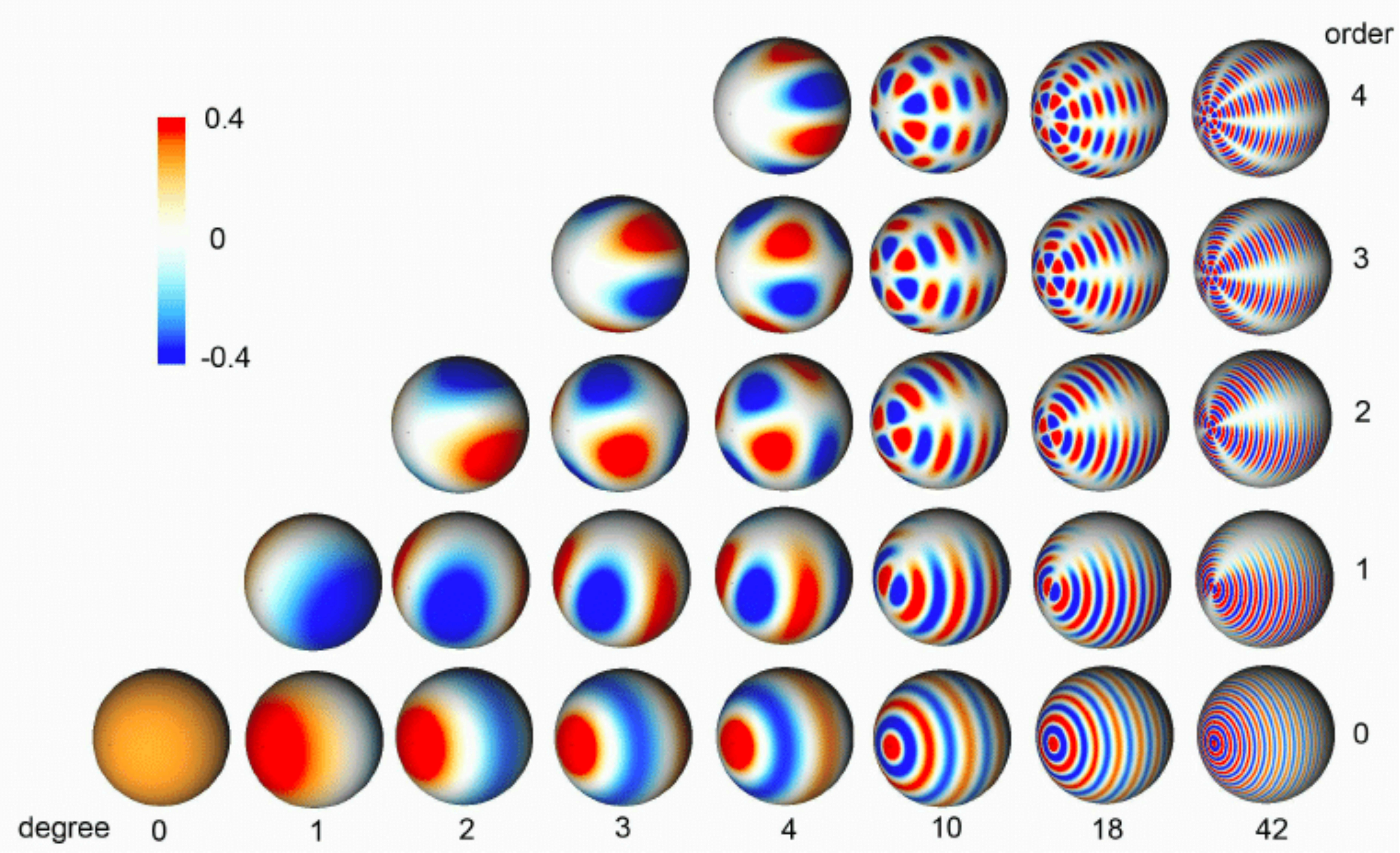
$$\frac{d^2}{dx^2}\sin(wx) = -w^2\sin(wx)$$

$$\frac{d^2}{dx^2}\cos(wx) = -w^2\cos(wx)$$



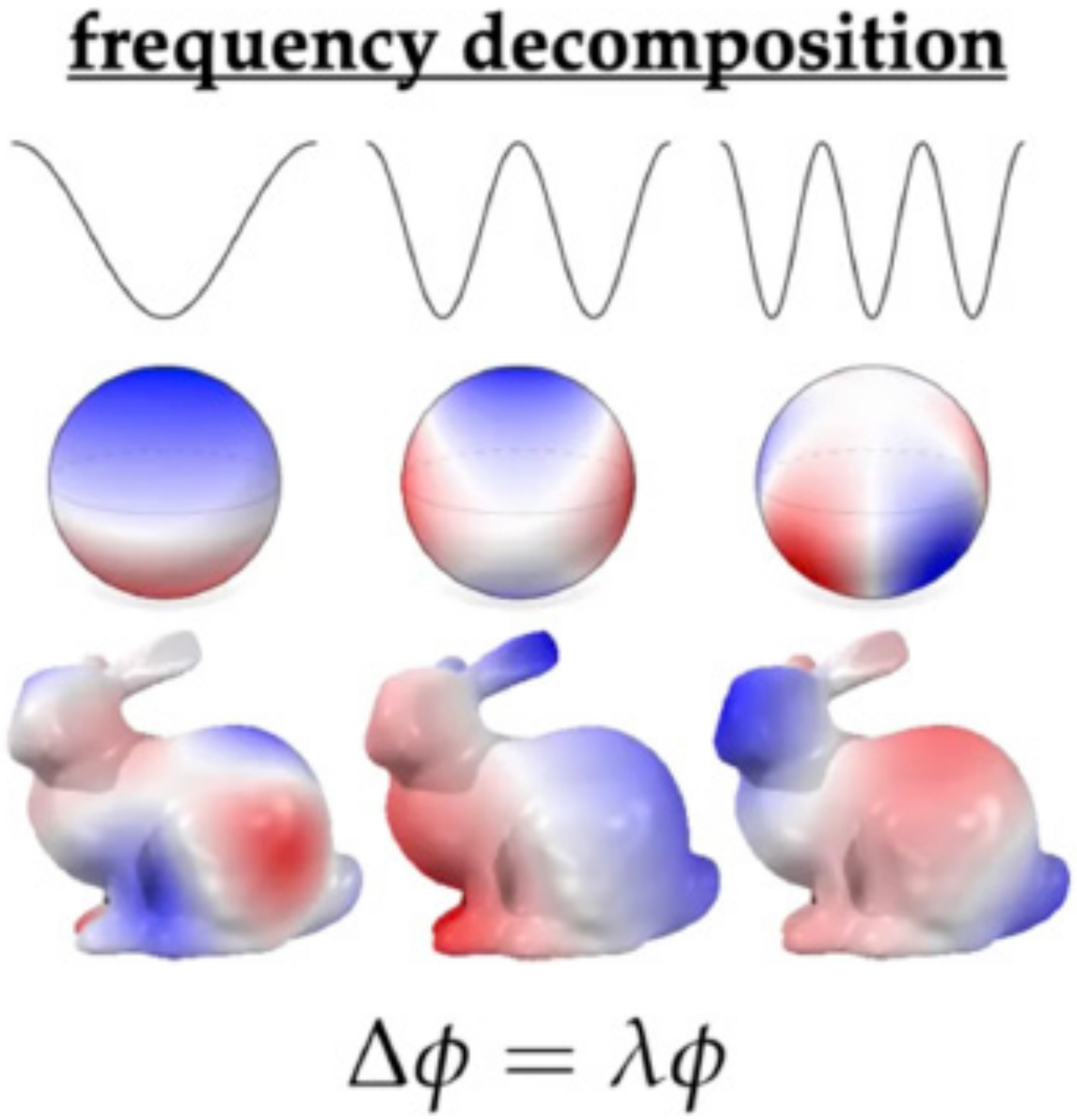


$$\nabla \cdot \nabla f = \nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$



order
4
3
2
1
0
0.4
0
-0.4
degree
0
1
2
3
4
10
18
42

Generalised Fourier



- **Basis set to represent functions**
- Filtering of scales
- Smoothing and interpolation
- Spectral coordinates
- ...

Spectral Geometry

Relation between geometry/topology and
eigenvalues and eigenfunctions of the Laplacian operator (Quantum mechanics)

Graphs

Relations between graph properties and spectra
of the graph Laplacian

Hypercube

Fourier decomposition of Boolean functions:
eigenvectors of Graph Laplacian are the
characters of $\mathbb{Z}_2^n$

$$f : \{-1,1\}^n \to \mathbb{R}$$

$$f(x) = \sum_{S \subseteq [n]} \hat{f}(S)\chi_S(x), \quad \chi_S(x) = \prod_{i \in S} x_i$$

$$\hat{f}(S) = \langle f, \chi_S \rangle$$

Manifolds

$$(M,g) \to Spec(M,g)$$

Eigenfunctions as minimizers of
Dirichlet form

Flat torus (Lattices) $\mathbb{R}^n/\Delta$

Poisson type formula relating the norms
of lattice points (lengths of closed geodesics)
to frequencies (eigenvalues)

$$\frac{1}{(4\pi t)^{d/2}} Vol(\Delta) \sum_{\lambda \in \Delta} e^{-|\lambda|^2/4t} = \sum_{\iota \in \Delta^*} e^{-4\pi|\iota|^2 t}$$

Geodesic Flows

Relation between geometry/topology a
and periodic geodesics (Classical mechanics)

Manifolds with boundary

Billiard dynamics

Manifolds without boundary

Relation between lengths of
closed geodesics and
eigenvalues of Laplacian

Representation Theory

Non-Abelian groups

Elements of the group are represented
by invertible matrices and the group operation
by matrix multiplication

Pontryagin Duality

Functions on locally compact Abelian groups $(\mathbb{Z}_2^n)$

Extension of Fourier for functions on groups
based on Pontryagin duality (group characters
generalize complex exponentials as basis functions)

$$\hat{G} := \mathrm{Hom}(G, \mathbb{R}/\mathbb{Z})$$

$$\chi : G \to U(1) \mid \chi(g_1 g_2) = \chi(g_1)\chi(g_2)$$

$$\hat{f}(\chi) = \int_G f(x)\bar{\chi}(x)d\mu(x)$$

Classical Fourier Analysis

Functions on $\mathbb{R}$ $\mathbb{R}/\mathbb{Z}$ $\mathbb{Z}$ $\mathbb{Z}_n$

Representing functions as sum of
trigonometric functions

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x)e^{-i2\pi\xi x}dx$$

Spectral geometry

Relation between geometry/topology and
eigenvalues and eigenfunctions of the Laplacian operator

9.3.3 Facts

The Laplacian on any compact Riemannian manifold provides us with all the tools of Fourier analysis on our Riemannian manifold. Let us call a function ϕ an *eigenfunction* with *eigenvalue* the number λ if

$$\Delta f = \lambda f \ .$$

The set of all eigenvalues of Δ is an infinite discrete subset of $\mathbb{R}^+$ called the *spectrum* of Δ

$$\operatorname{Spec}(M) = \{\lambda_k\} = \{0 < \lambda_1 < \lambda_2 < \dots\} \tag{9.5}$$

with λ_k tending to infinity with k.

For each eigenvalue λ_i, the vector space of eigenfunctions ϕ satisfying

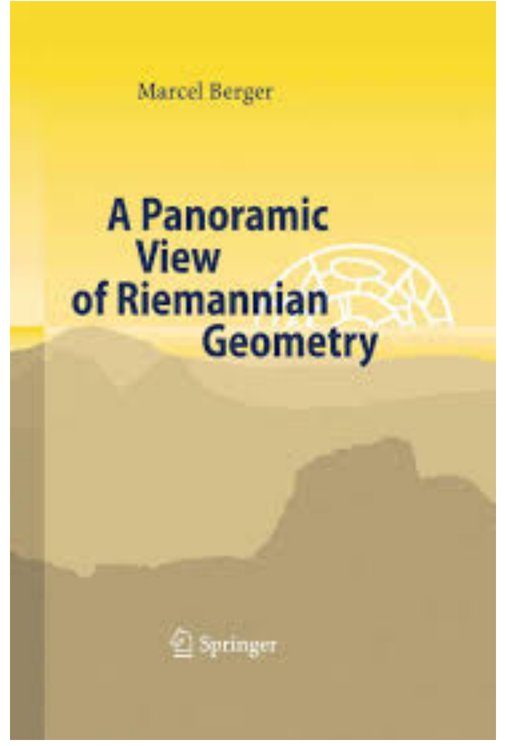
$$\Delta f = \lambda_i f$$

is always finite dimensional and its dimension is called the *multiplicity* of λ_i. Once we have a basis of the eigenfunctions with this eigenvalue written out, it is trivial to find an orthonormal basis

$$\{\phi_k\}$$

(where k runs from 1 to the multiplicity) of eigenfunctions. Here the orthonormalcy is to be understood for the global scalar product

$$\langle f, g \rangle_{L^2(M)} = \int_M f g \ .$$



As for classical Fourier series, any reasonable function

$$f : M \to \mathbb{R}$$

has Fourier coefficients

$$a_i = \int_M f\phi_i$$

and f is recovered from these coefficients by the converging series

$$f = \sum_i a_i\phi_i \ .$$

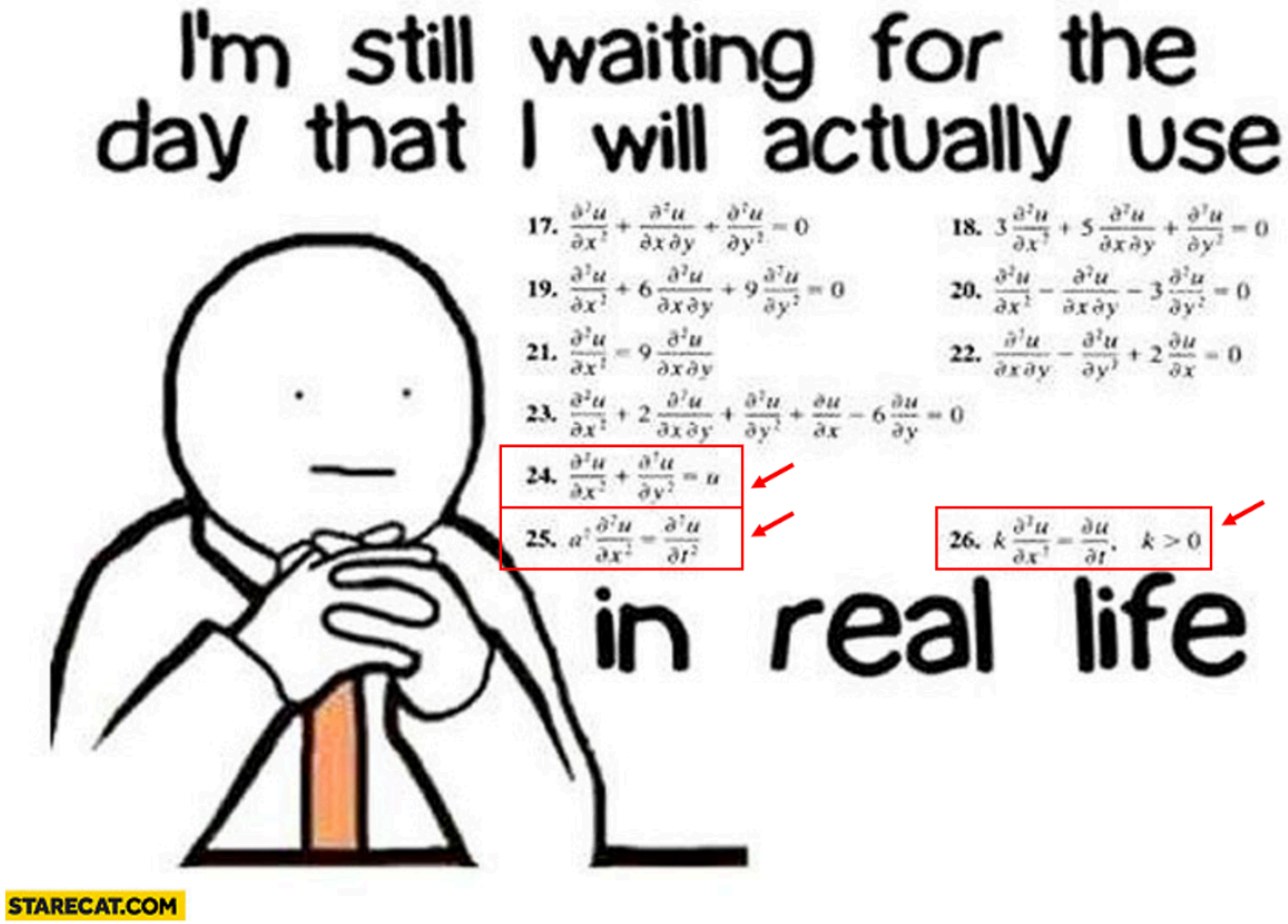
In the same spirit, the scalar product of two functions is the sum of products of their coefficients:

$$\int_M fg = \sum_i a_i b_i$$

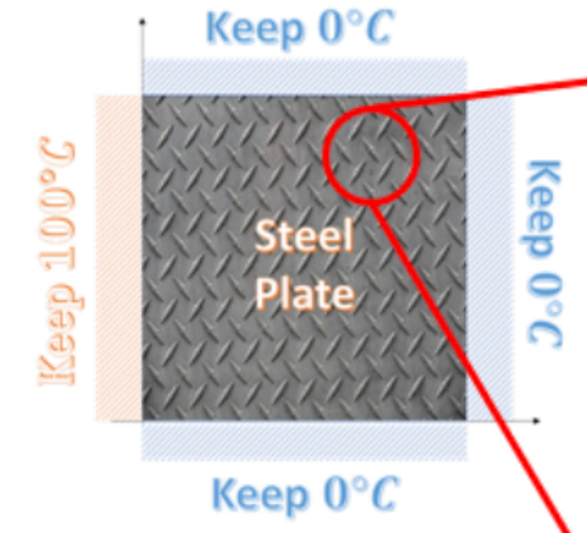
where

$$f = \sum_i a_i\phi_i$$

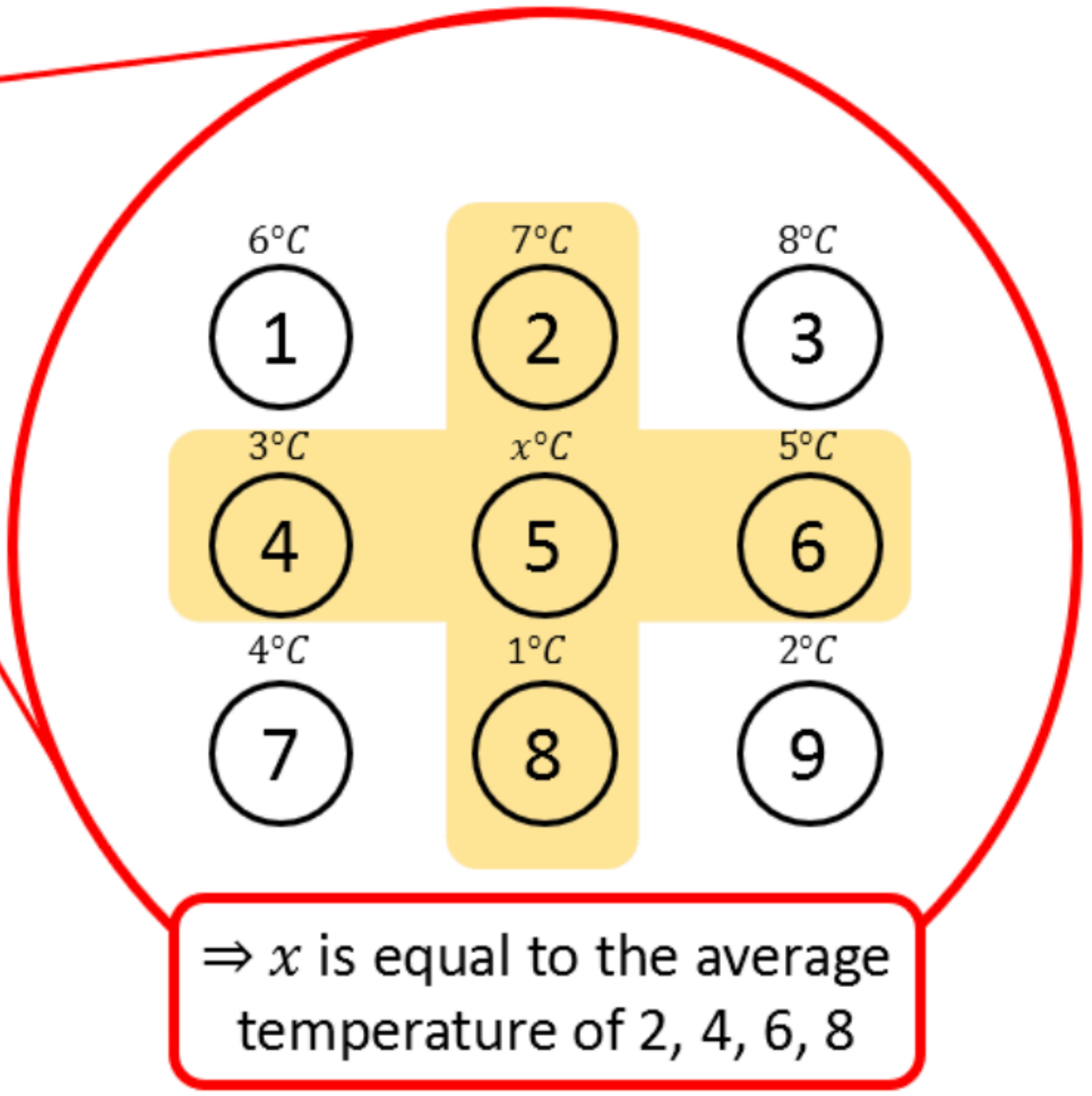
$$g = \sum_i b_i\phi_i \ .$$



$$\nabla \cdot \nabla f = \nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$



This is what Laplace's Equation is saying!



9.3.4 Heat, Wave and Schrödinger Equations

We will follow the same steps that we did in §1.8: defining heat, wave and Schrödinger equations on Riemannian manifolds. The heat equation for the heat $f(m,t)$ at time t at a point m of the Riemannian manifold M is

$$\Delta f = -\frac{\partial f}{\partial t} \,. \tag{9.6}$$

The wave equation for the height $f(m,t)$ of the "water" after time t at a point m is

$$\Delta f = -\frac{\partial^2 f}{\partial t^2} \,. \tag{9.7}$$

where if M were a surface, you would consider M covered in a thin sheet of water, or for M of three dimensions, M is a place through which sound is propagating. The wave equation can also be considered as describing the manifold M as a vibrating membrane object. Finally the Schrödinger equation uses complex valued functions and is written

$$\hbar^2 \Delta f = i\hbar \frac{\partial f}{\partial t} \tag{9.8}$$

where $i = \sqrt{-1}$ and $\hbar$ is Planck's constant.

To solve these equations, at least formally, one uses the same trick as in §§1.8.1. To solve such an equation depending both on time t and a point $m \in M$, the initial idea is to use the fact that, roughly by the Stone–Weierstraß approximation theorem, we need only to consider product functions

$$f(m,t) = g(m)h(t) .$$

One will subsequently consider series of them (as in the theory of Fourier series). Look for example at the heat equation. The function $f = gh$ satisfies the heat equation precisely when the functions g and h satisfy

$$\frac{\Delta g}{g} = -\frac{h'}{h} \tag{9.9}$$

where

$$h'(t) = \frac{dh}{dt}$$

is the usual derivative.

Since the first fraction depends only on the point $m \in M$ and the second only on the time t their common value has to be a constant, call it λ. Then the function

$$g : M \to \mathbb{R}$$

is an eigenfunction of Δ with eigenvalue λ, while h is an exponential decay at rate λ. If all eigenfunctions and eigenvalues of Δ are known, we can then solve the heat equation explicitly. Note that the time dependence $h(t)$ is

$$h(t) = \begin{cases} e^{-\lambda t} & \text{for the heat equation} \\ e^{i\lambda t} & \text{for the Schrödinger equation} \\ e^{i\sqrt{\lambda}t} & \text{for the wave equation.} \end{cases}$$

Another way to write the solution $f(m,t)$ with initial temperature $f(m,0)$ is to compute the Riemannian Fourier series

$$f(m,0) = \sum_{k=1}^{\infty} a_k \phi_k$$

and then

$$f(m,t) = \sum_{k=1}^{\infty} a_k \phi_k(m) e^{-\lambda_k t} \; .$$

For the wave equation, the fundamental solution similar to equation 9.10 requires imaginary terms, i.e.

$$e^{i\sqrt{\lambda_k}t}$$

which are linear combinations of

$$\cos\left(\sqrt{\lambda_k}t\right) \text{ and } \sin\left(\sqrt{\lambda_k}t\right) \; .$$

But the dramatic difference between the heat equation and the wave equation is that waves demand not converging series, but distributions. Heat spreads out uniformly with time, while waves bounce up and down forever. This ma-

There are very few examples where the spectrum or the eigenfunctions can be determined explicitly. Two old standards are rectangles and disks. In both cases, separation of variables disentangles the eigenfunctions. Using again the Stone–Weierstraß theorem, and because the boundary condition agrees with the separation, on a rectangle one need only look for product functions $f(x,y) = g(x)h(y)$, and there will be no other eigenfunctions. If the rectangle has side lengths a and b respectively, then the eigenfunctions are

$$\sin \frac{\pi m x}{a} \sin \frac{\pi n y}{b} \tag{1.21}$$

where m and n are any integers, yielding the set of eigenvalues

$$\lambda = \pi^2 \left(\frac{m^2}{a^2} + \frac{n^2}{b^2} \right).$$

i.e. to obtain an eigenvalue at most λ we need m and n to be integer points inside a certain ellipse. We will see below that a simple expression yields an easy first order approximation of $N(\lambda)$ when $\lambda \to \infty$, but the second order term in λ is related to deep number theory and is still not completely understood today. It is believed that the number of integers m, n with

$$m^2 + n^2 \leq \lambda^2$$

is asymptotic to $\pi^2\lambda^2 + O\left(\lambda^{\varepsilon+1/2}\right)$ as $\lambda \to \infty$, for any $\varepsilon > 0$, but there is still no proof. This is called the Gauß circle problem; see §§1.8.5. However, it is known that $\pi^2\lambda^2 + O\left(\lambda^{1/2}\right)$ is too small.

How about a second order approximation? In 1954, Pleijel got the next order approximation. In his paper Kac 1966 [775], Kac works quite hard to get the third term, guessing that the right formula should be:

$$\sum_{i=1}^{\infty} e^{-\lambda_i t} \sim \frac{\text{Area}(D)}{2\pi t} - \frac{\text{length}(\partial D)}{\sqrt{2\pi t}} + \frac{1}{6}(1-r)$$

where r is the number of holes inside D. The second term is Pleijel's. Note that can one hear the area and the perimeter of D, hence the isoperimetric inequality yields again the fact that disks are characterized by their spectrum.

Kac could only prove the third term for polygons. It was proven the next year, 1967, in the very general context of Riemannian manifolds with boundary by McKean and Singer in their fundamental paper of 1967 [910]. You will read much more about it in chapter 9.

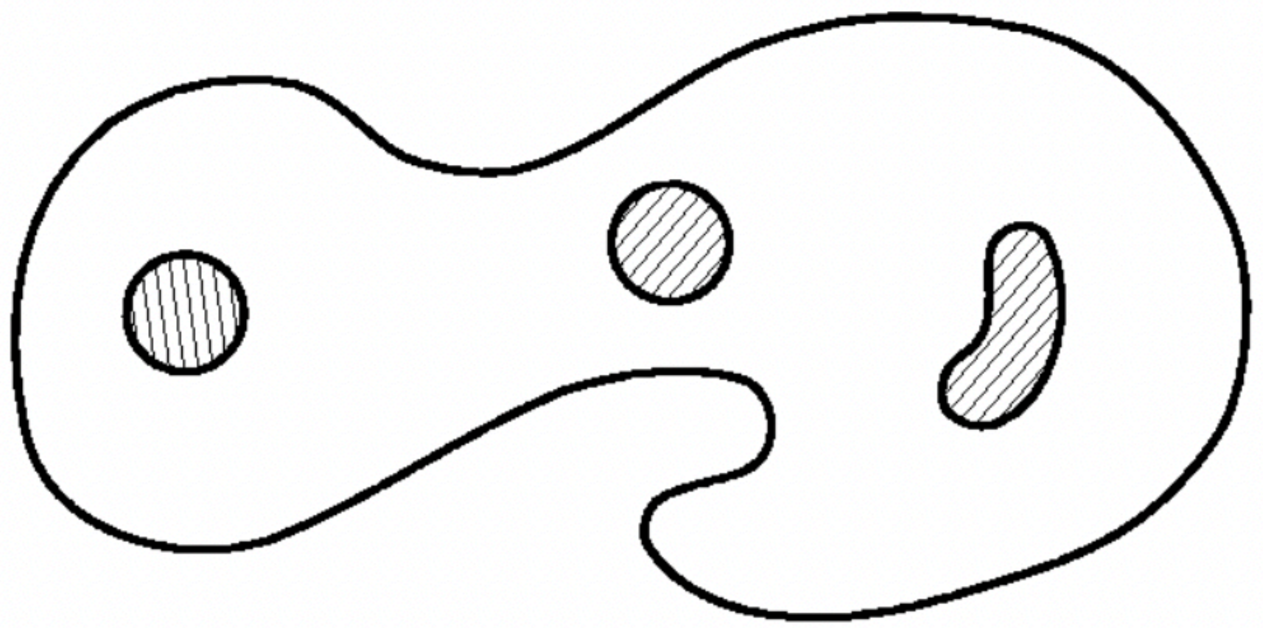


Fig. 1.97. One can hear the number of holes

The Tauberian theorem above shows that the first terms of $N(\lambda)$ and $\sum_{i=1}^{\infty} \exp(-\lambda_i t)$, can each be acquired from the other. But this does not

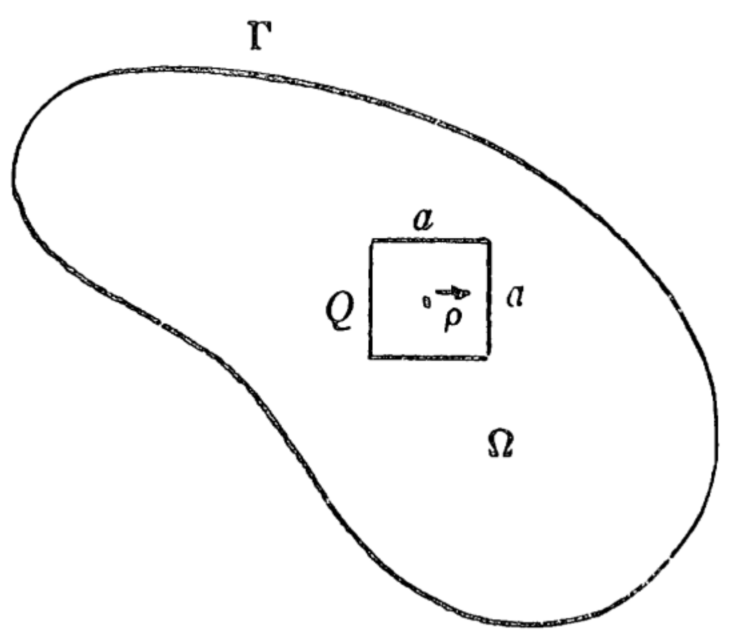
CAN ONE HEAR THE SHAPE OF A DRUM?

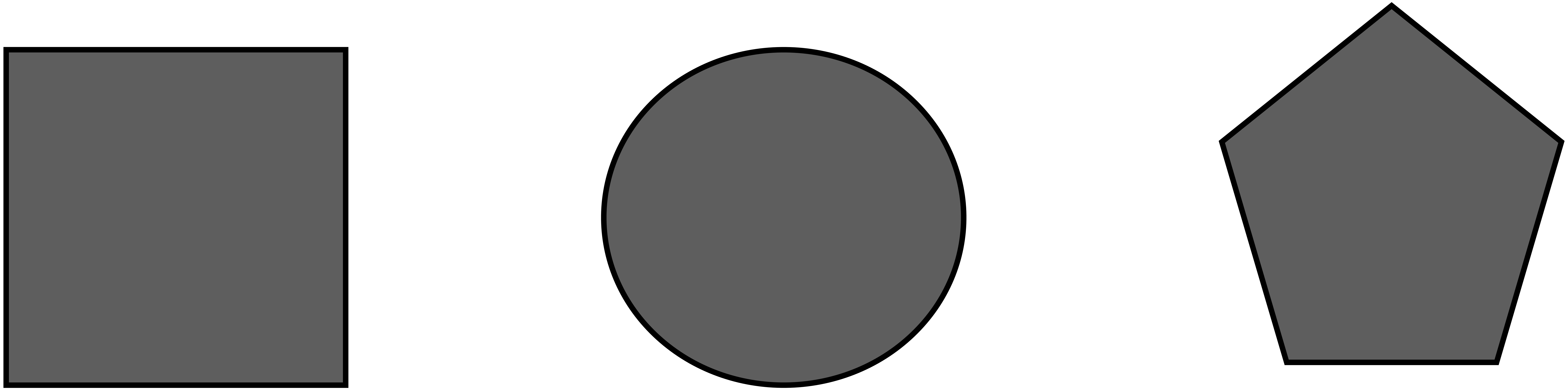
MARK KAC, The Rockefeller University, New York

To George Eugene Uhlenbeck on the occasion of his sixty-fifth birthday

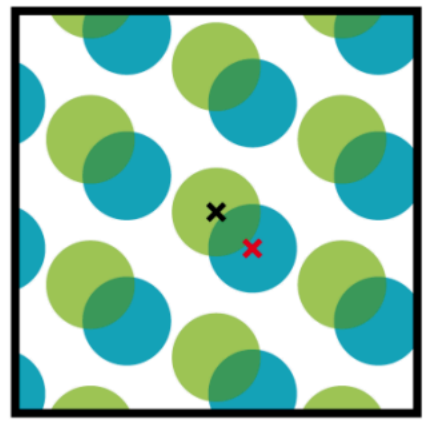
> "La Physique ne nous donne pas seulement l'occasion de résoudre des problèmes . . . , elle nous fait presentir la solution." H. POINCARÉ.

Before I explain the title and introduce the theme of the lecture I should like to state that my presentation will be more in the nature of a leisurely excursion than of an organized tour. It will not be my purpose to reach a specified destination at a scheduled time. Rather I should like to allow myself on many occasions the luxury of stopping and looking around. So much effort is being spent on streamlining mathematics and in rendering it more efficient, that a solitary transgression against the trend could perhaps be forgiven.

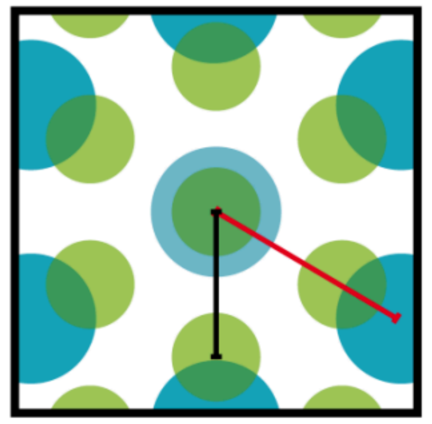




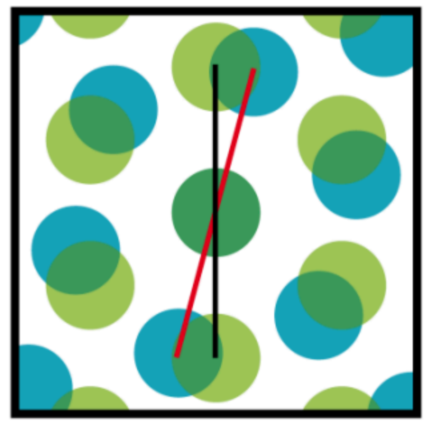
1. Phase

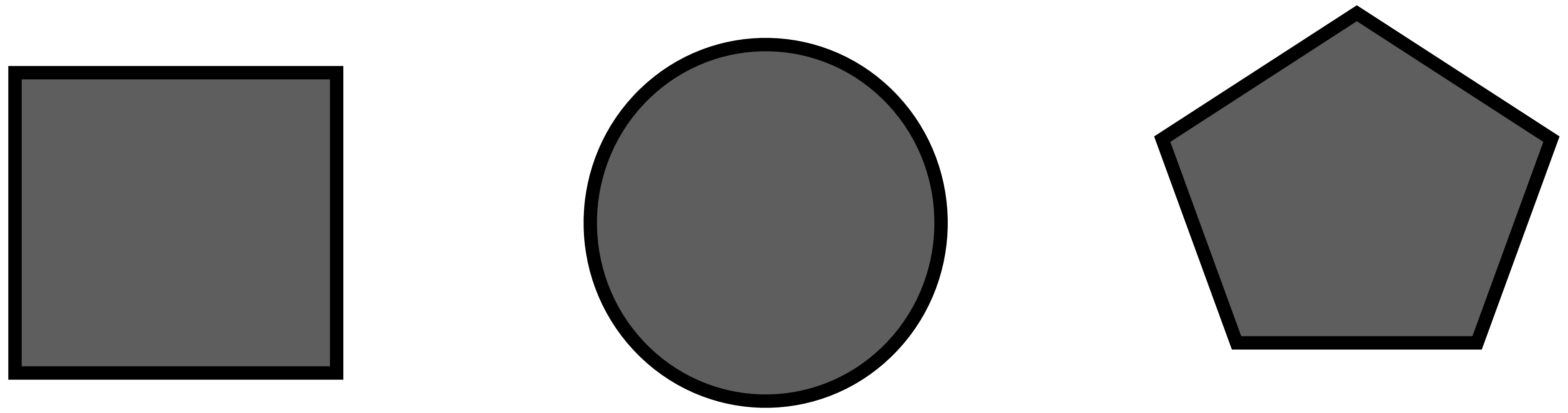


2. Scale



3. Orientation





Faber-Krahn theorem

Rayleigh–Faber–Krahn inequality

From Wikipedia, the free encyclopedia

In spectral geometry, the **Rayleigh–Faber–Krahn inequality**, named after its conjecturer, Lord Rayleigh, and two individuals who independently proved the conjecture, G. Faber and Edgar Krahn, is an inequality concerning the lowest Dirichlet eigenvalue of the Laplace operator on a bounded domain in $\mathbb{R}^n$, $n \geq 2$.[1] It states that the first Dirichlet eigenvalue is no less than the corresponding Dirichlet eigenvalue of a Euclidean ball having the same volume. Furthermore, the inequality is rigid in the sense that if the first Dirichlet eigenvalue is equal to that of the corresponding ball, then the domain must actually be a ball. In the case of $n = 2$, the inequality essentially states that among all drums of equal area, the circular drum (uniquely) has the lowest voice.

Grid **pattern** -> Superposition of Fourier modes (**Eigenfunctions** of Laplacian **on the domain**)

Grid **scale** -> **Eigenvalue** associated to contributing eigenfunctions

Rayleigh–Faber–Krahn inequality

1 language

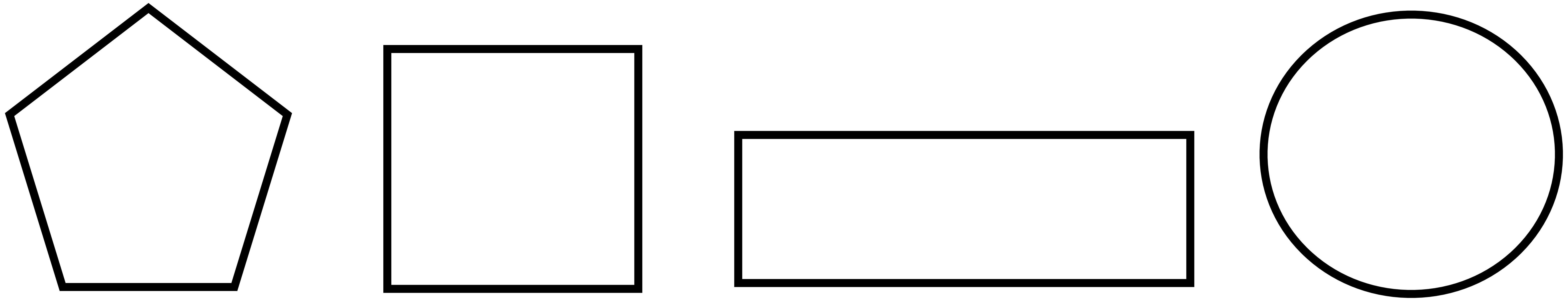
Article View the content page [ctrl-option-c] Talk

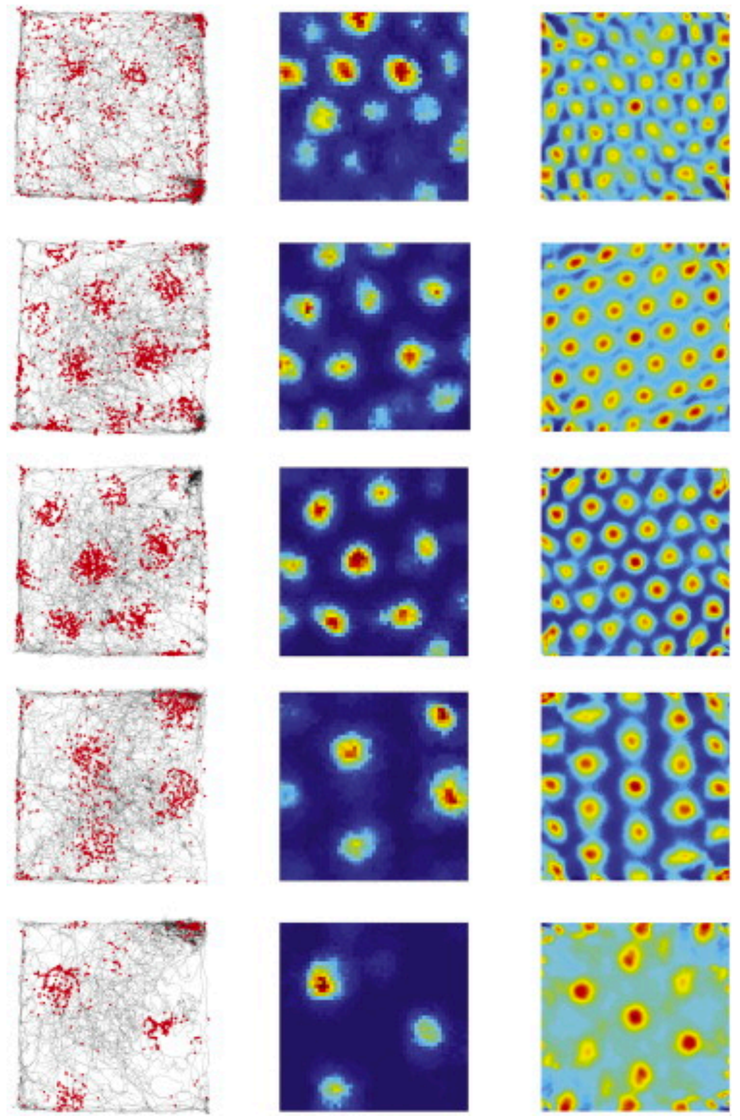
Read Edit View history Tools

From Wikipedia, the free encyclopedia

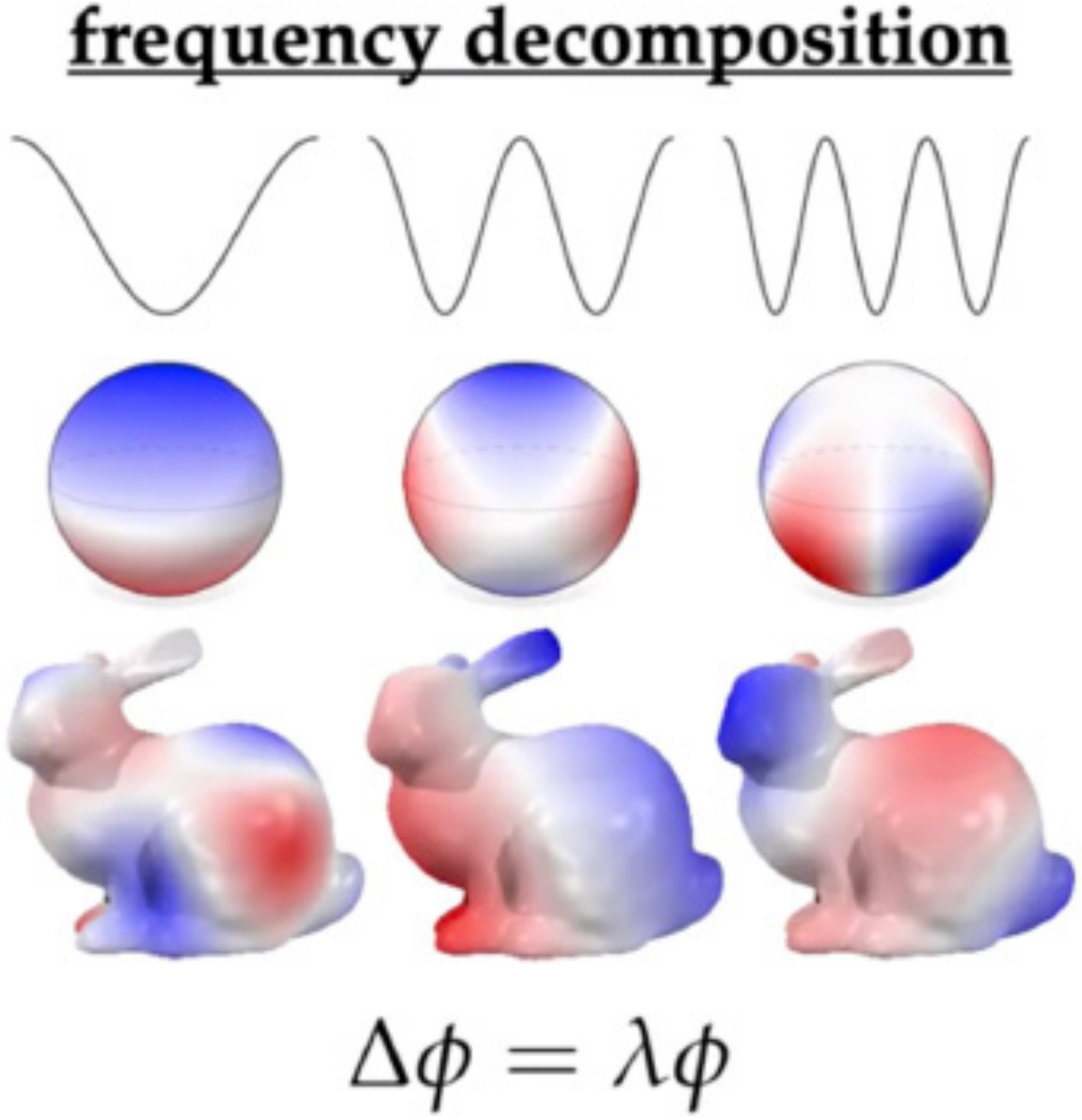
In spectral geometry, the **Rayleigh–Faber–Krahn inequality**, named after its conjecturer, Lord Rayleigh, and two individuals who independently proved the conjecture, G. Faber and Edgar Krahn, is an inequality concerning the lowest Dirichlet eigenvalue of the Laplace operator on a bounded domain in $\mathbb{R}^n$, $n \geq 2$.[1] It states that the first Dirichlet eigenvalue is no less than the corresponding Dirichlet eigenvalue of a Euclidean ball having the same volume. Furthermore, the inequality is rigid in the sense that if the first Dirichlet eigenvalue is equal to that of the corresponding ball, then the domain must actually be a ball. In the case of $n = 2$, the inequality essentially states that among all drums of equal area, the circular drum (uniquely) has the lowest voice.

Do grid cells for circular environments have lowest spatial frequency?



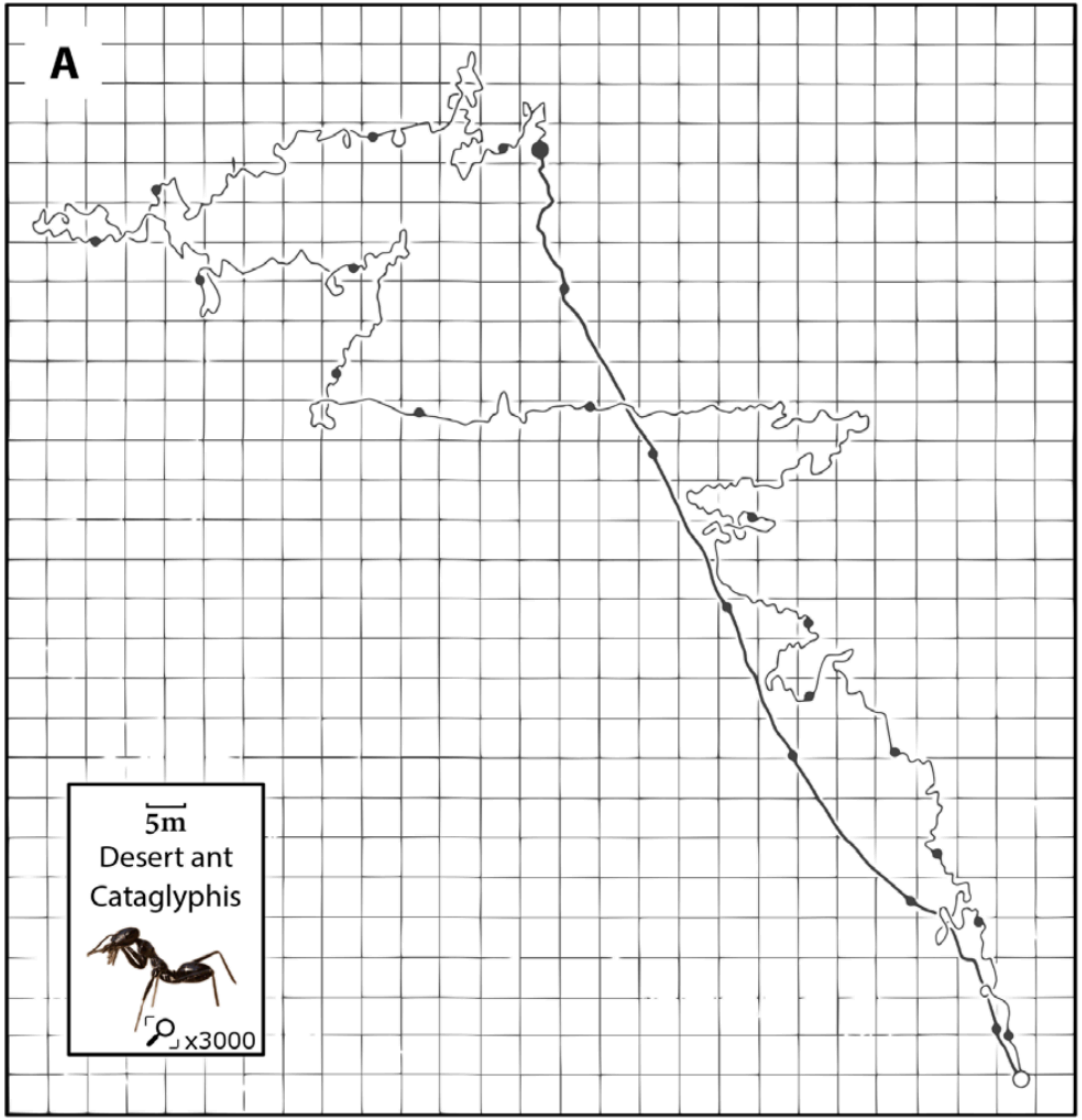


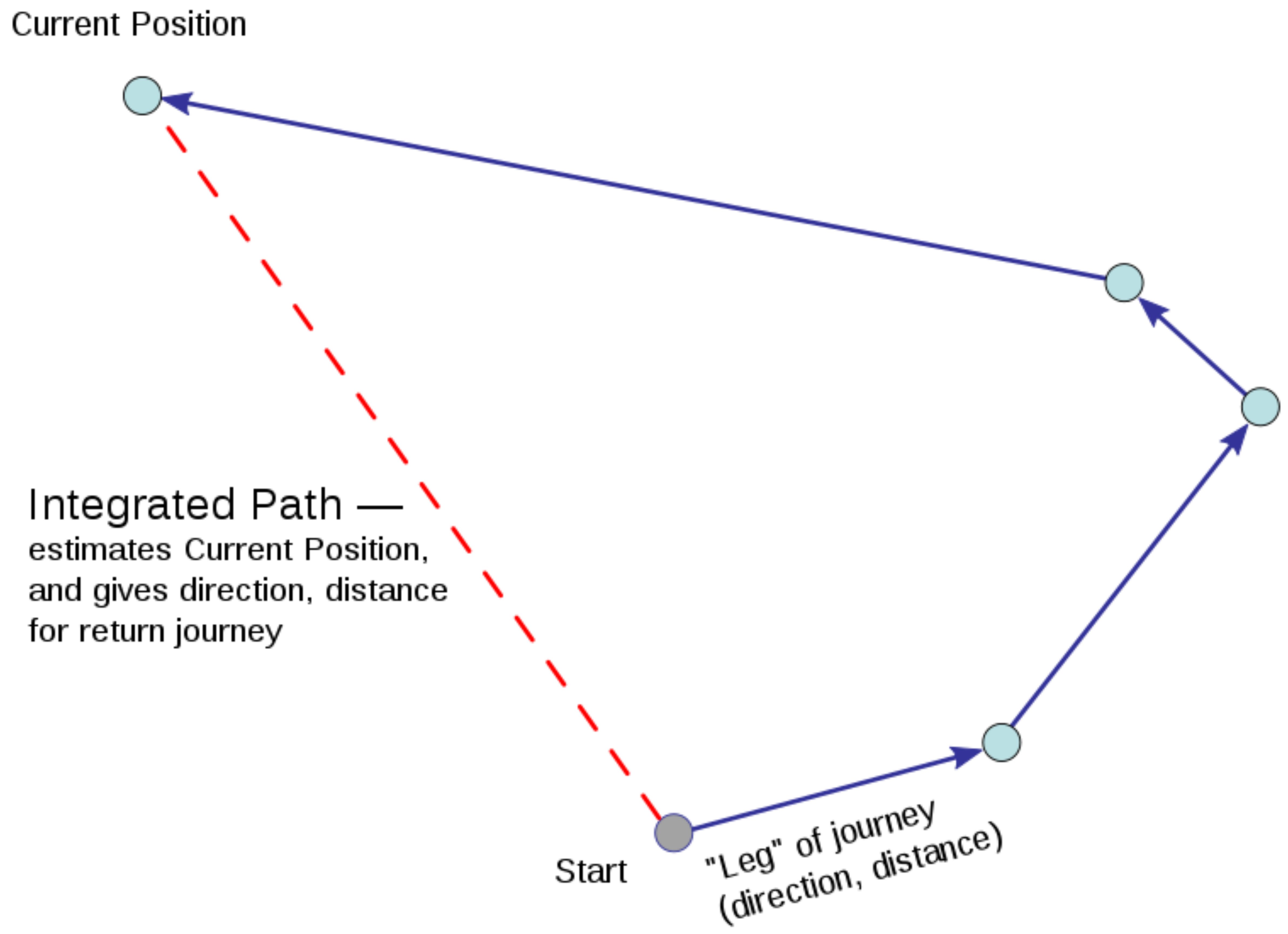
Generalised Fourier

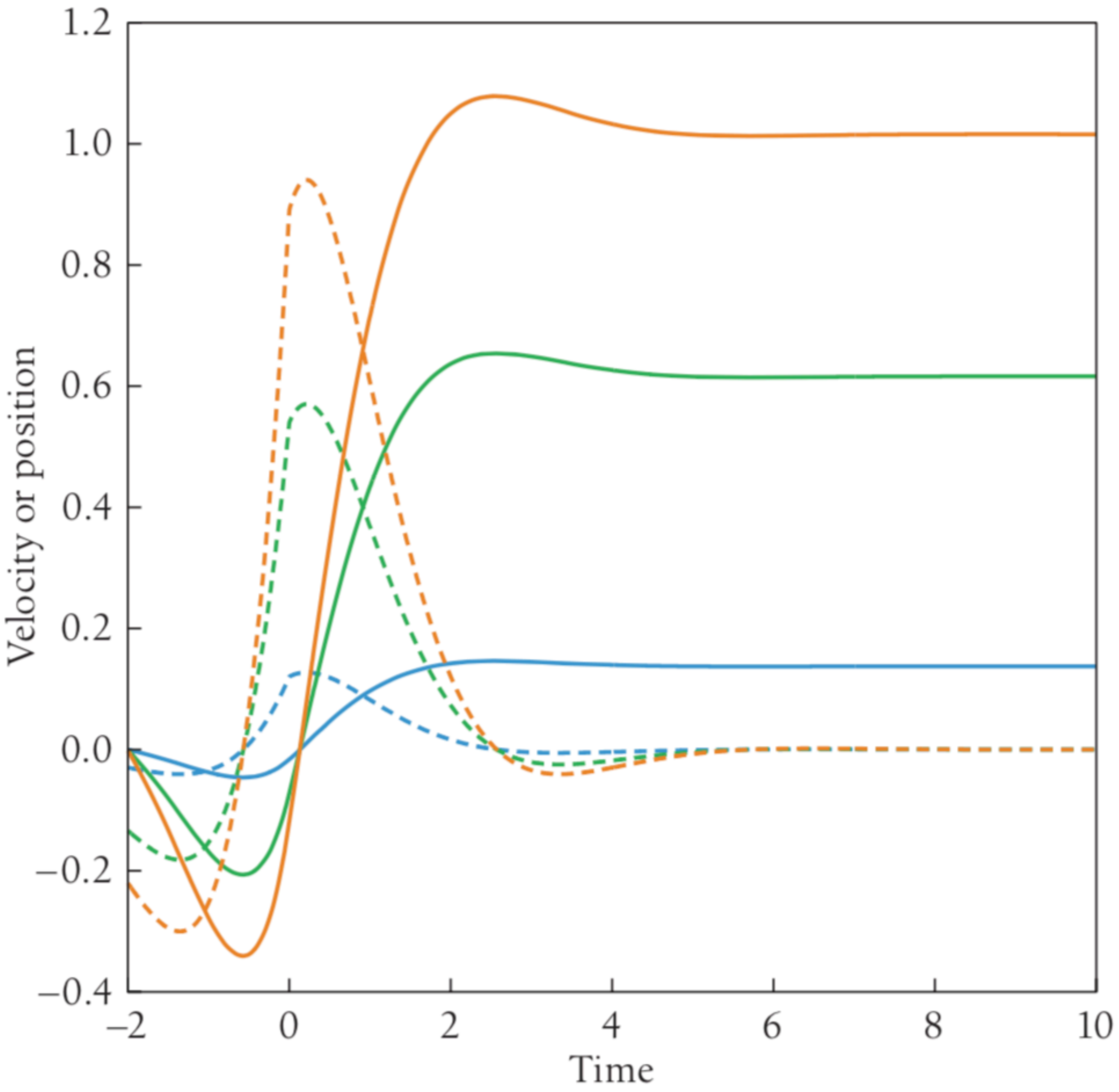


- Basis set to represent functions
- Filtering of scales
- Smoothing and interpolation
- **Spectral coordinates**
- ...

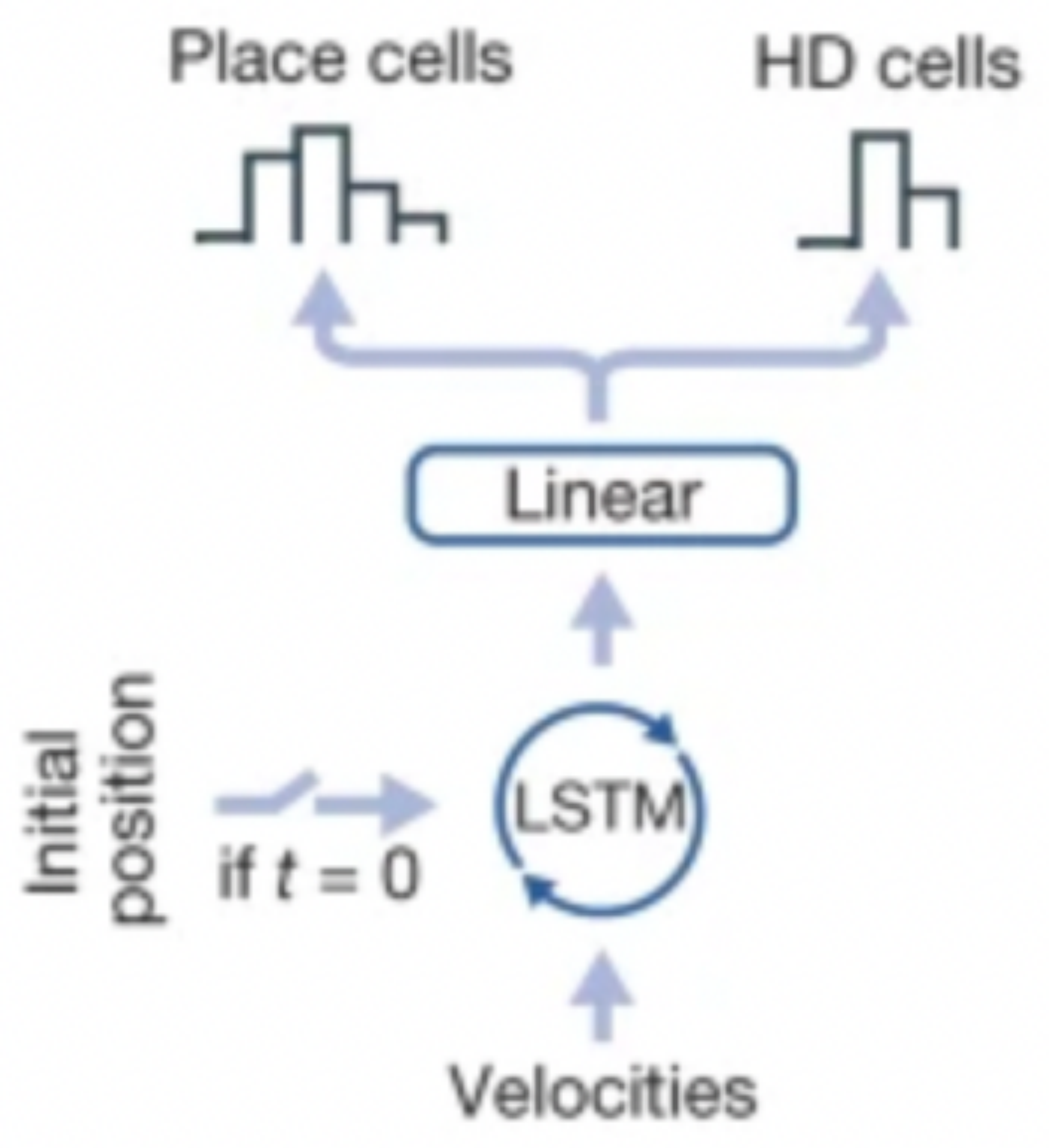
Path integration




Current Position
Integrated Path —
estimates Current Position,
and gives direction, distance
for return journey
Start
"Leg" of journey
(direction, distance)



Velocity or position
1.2
1.0
0.8
0.6
0.4
0.2
0.0
–0.2
–0.4
–2
0
2
4
6
8
10
Time

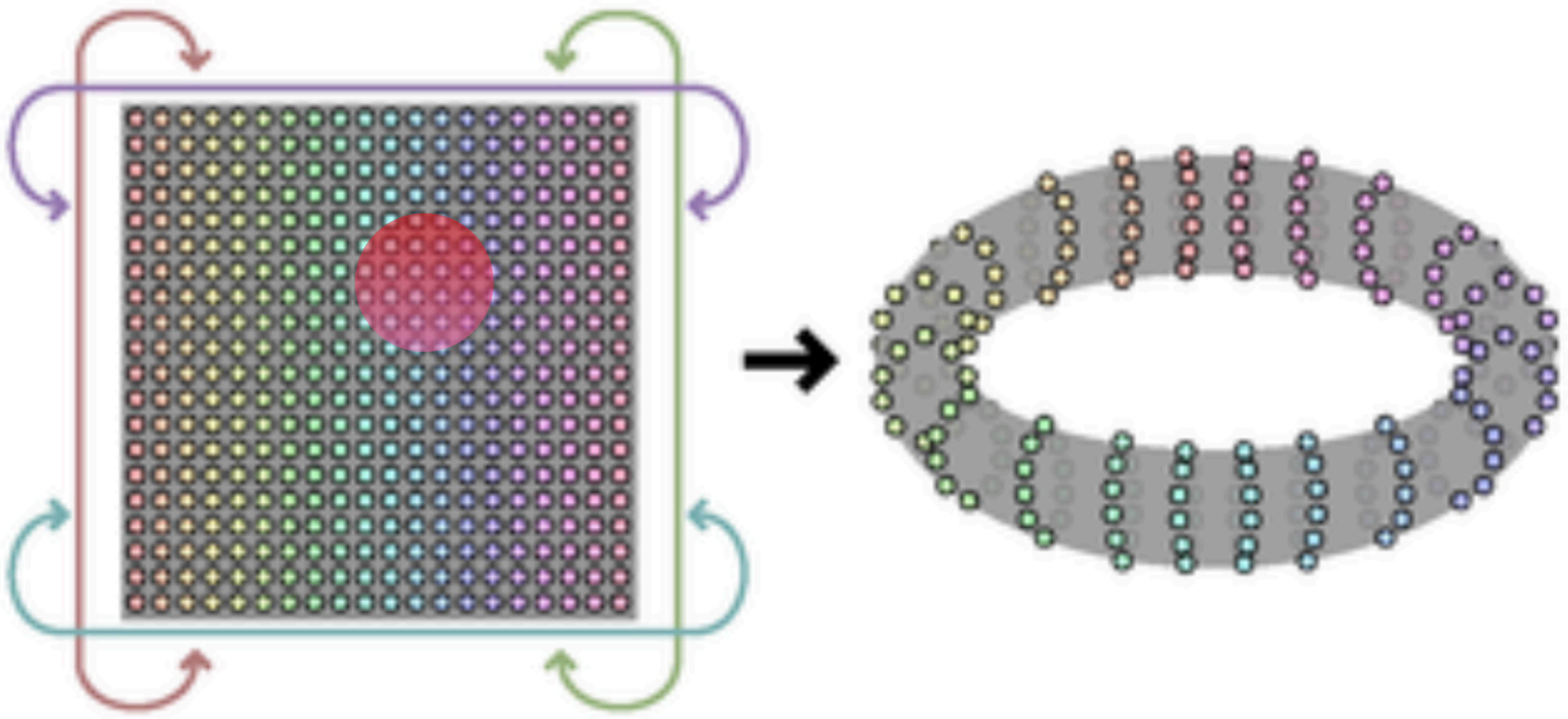


Path integration amounts to compute the integral of **velocity**

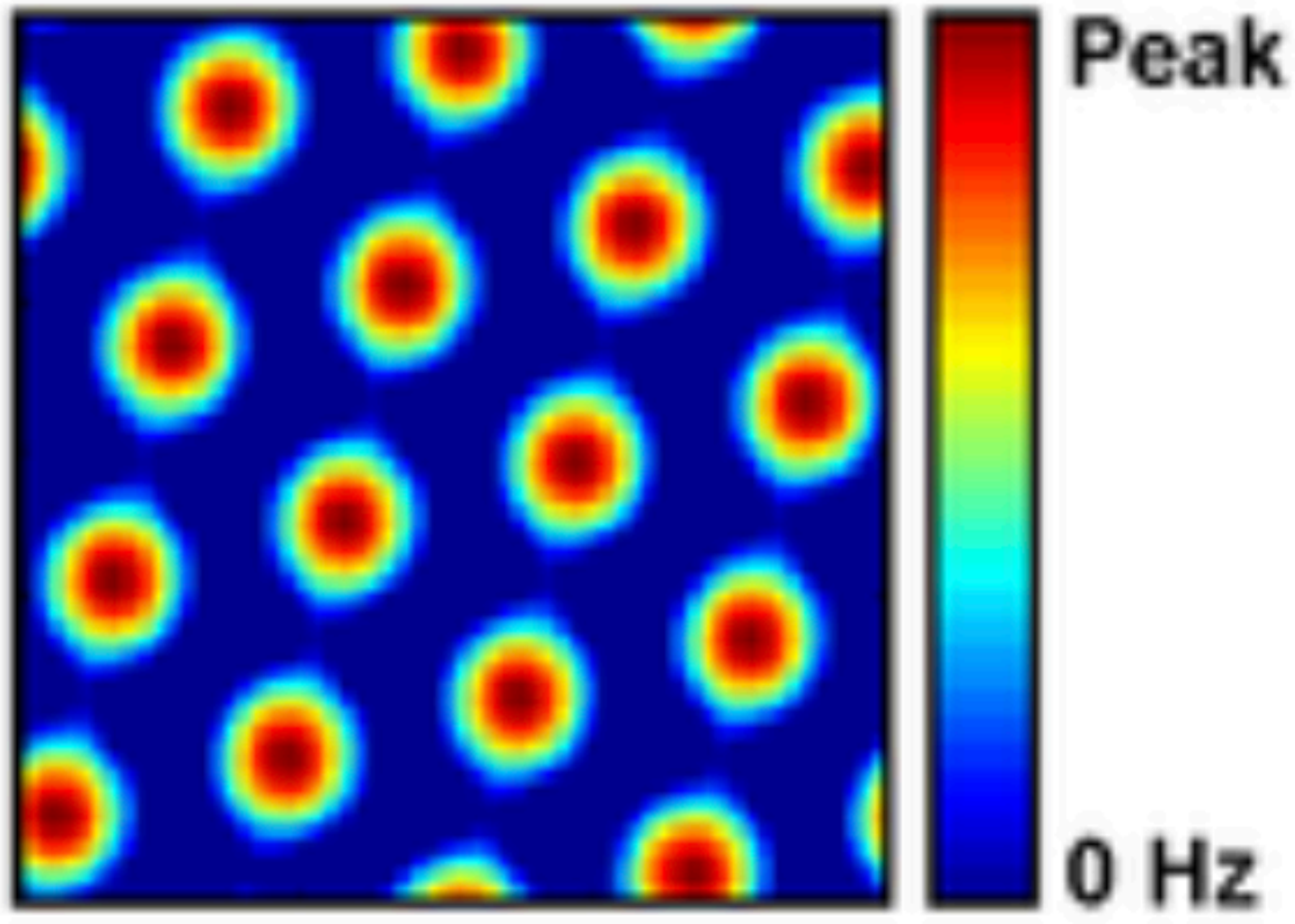
Place cells as **one-hot encoding of position**

Classical theory of grid cells

- Toroidal connectivity of neurons
- Bump of activity updated by velocity cues (to solve path integration)



Neuronal space



Physical space

Path integration just a particular case of driven dynamical system?

Path integration just a particular case of driven dynamical system?

	Derivative Form	Integral Form
Position	$r(t)$	$r(t) = r_0 + \int_0^t v\,dt'$

Position and Velocity

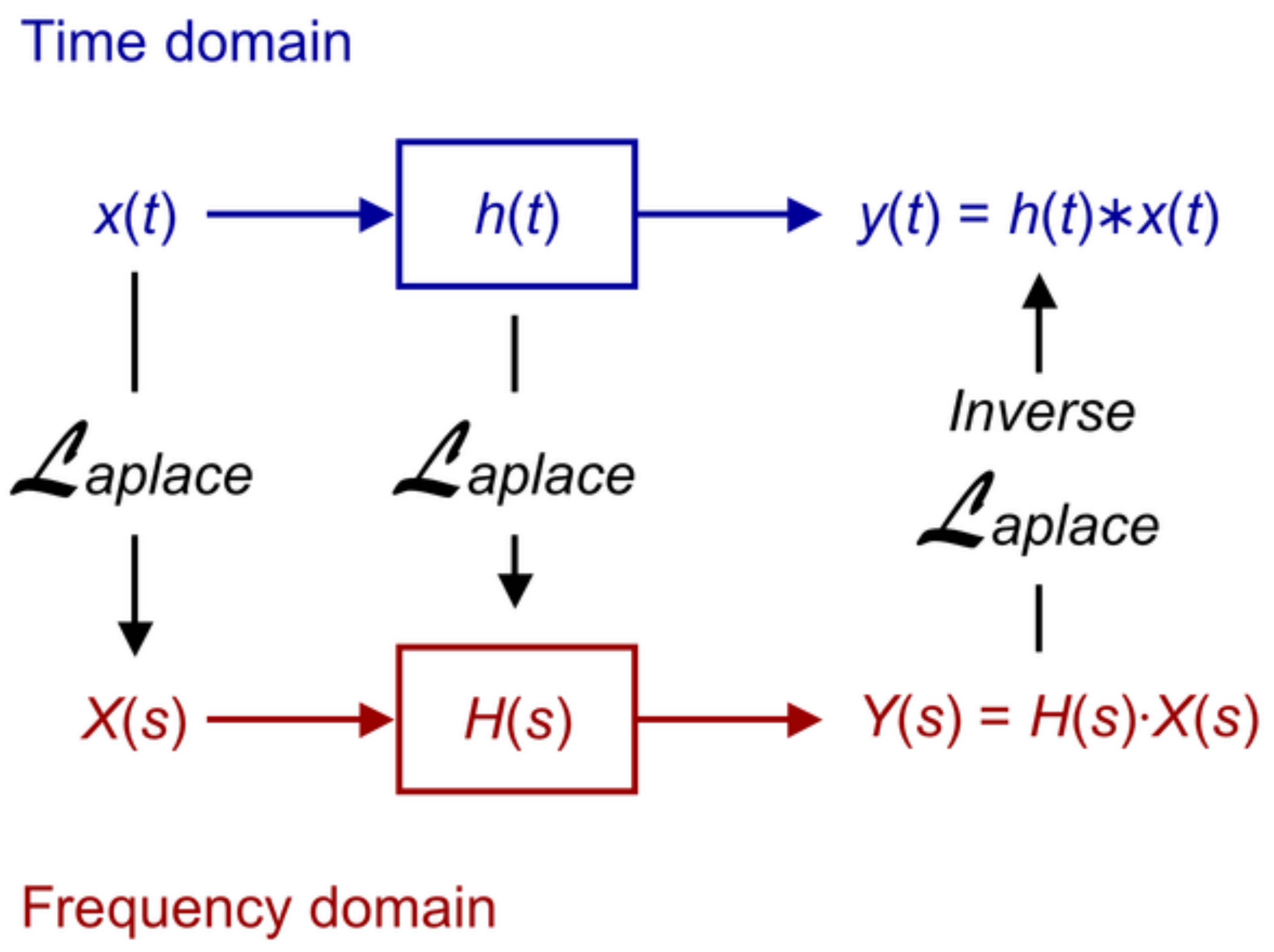
$$v(t) = \int a(t)\,dt + C$$

$$x(t) = \int v(t)\,dt + C$$

Position and Velocity

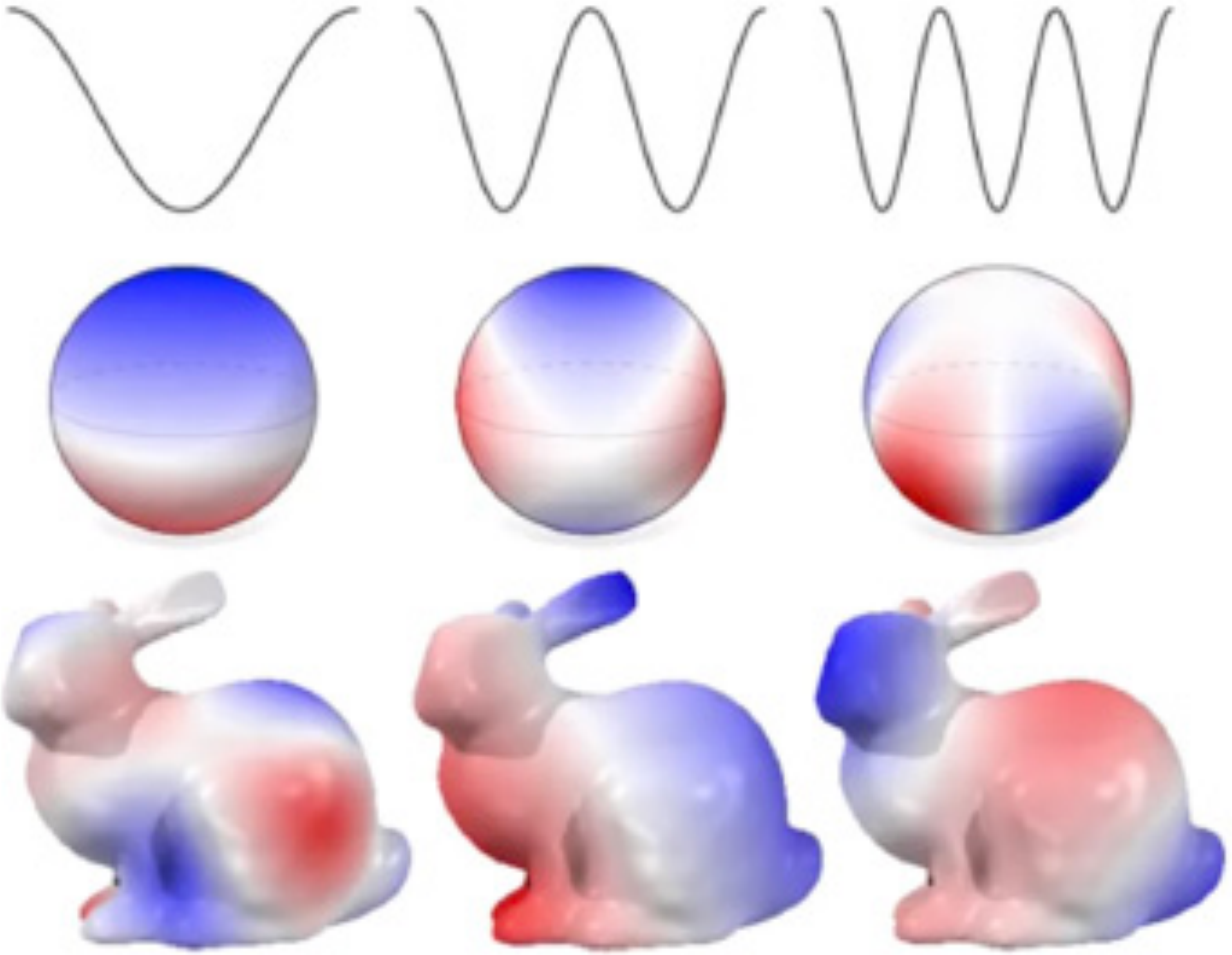
$$v(t) = \int a(t)\,dt + C$$

$$x(t) = \int v(t)\,dt + C$$



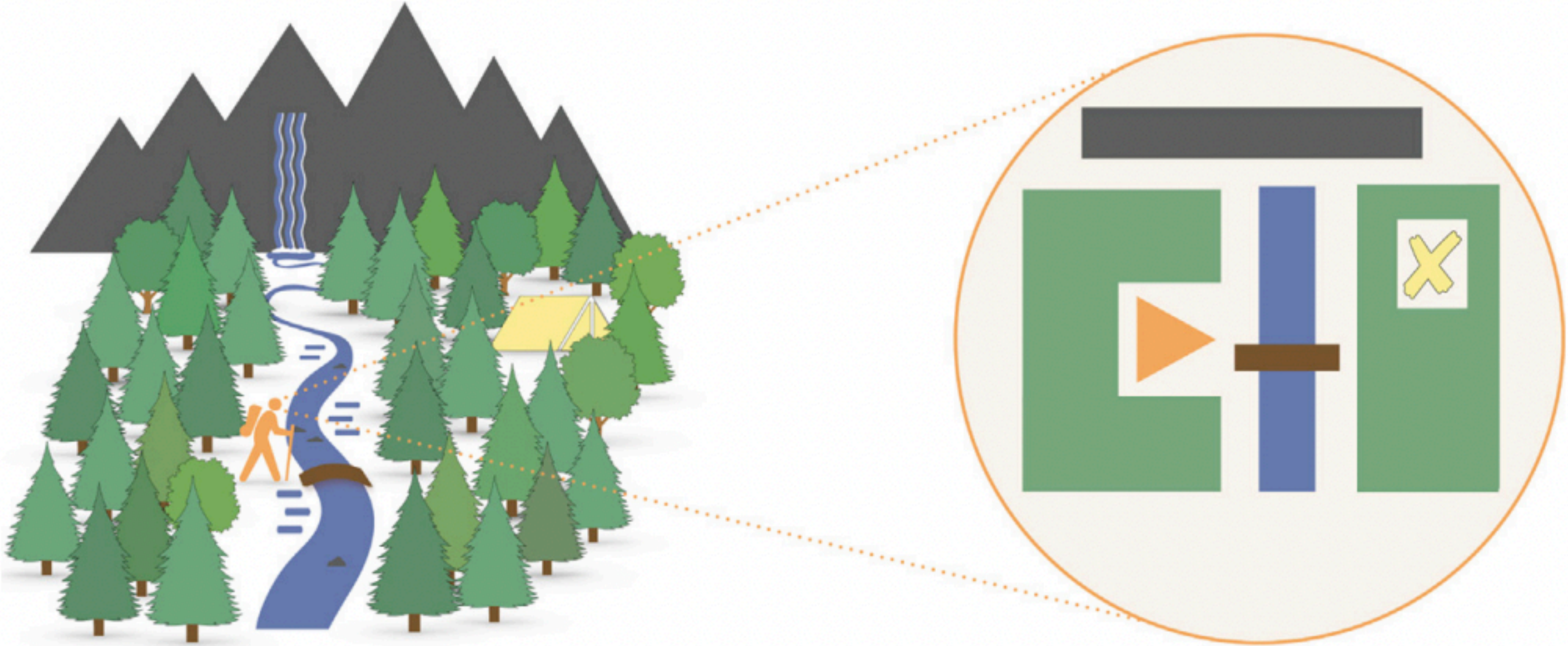
Generalised Fourier

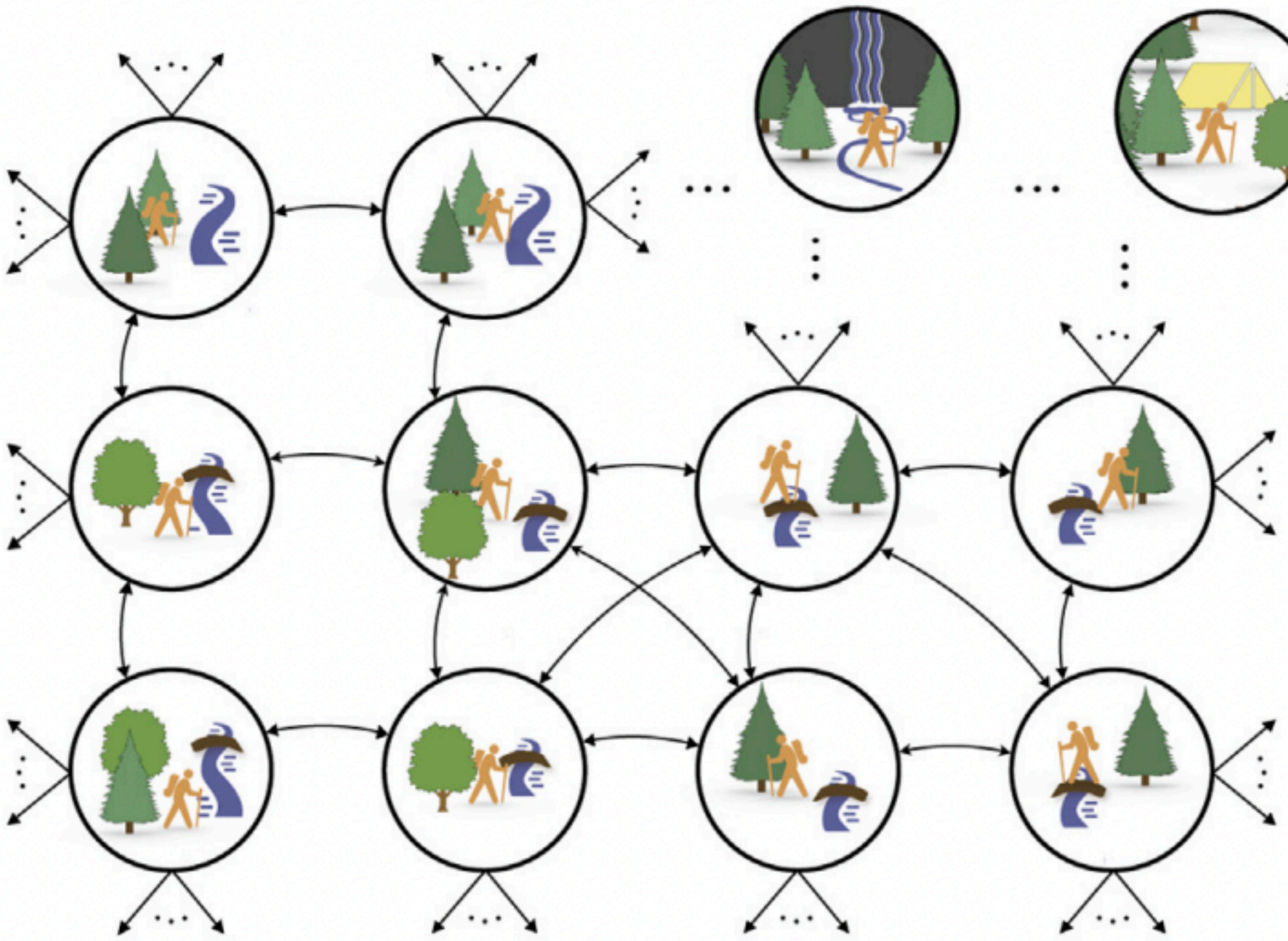
frequency decomposition

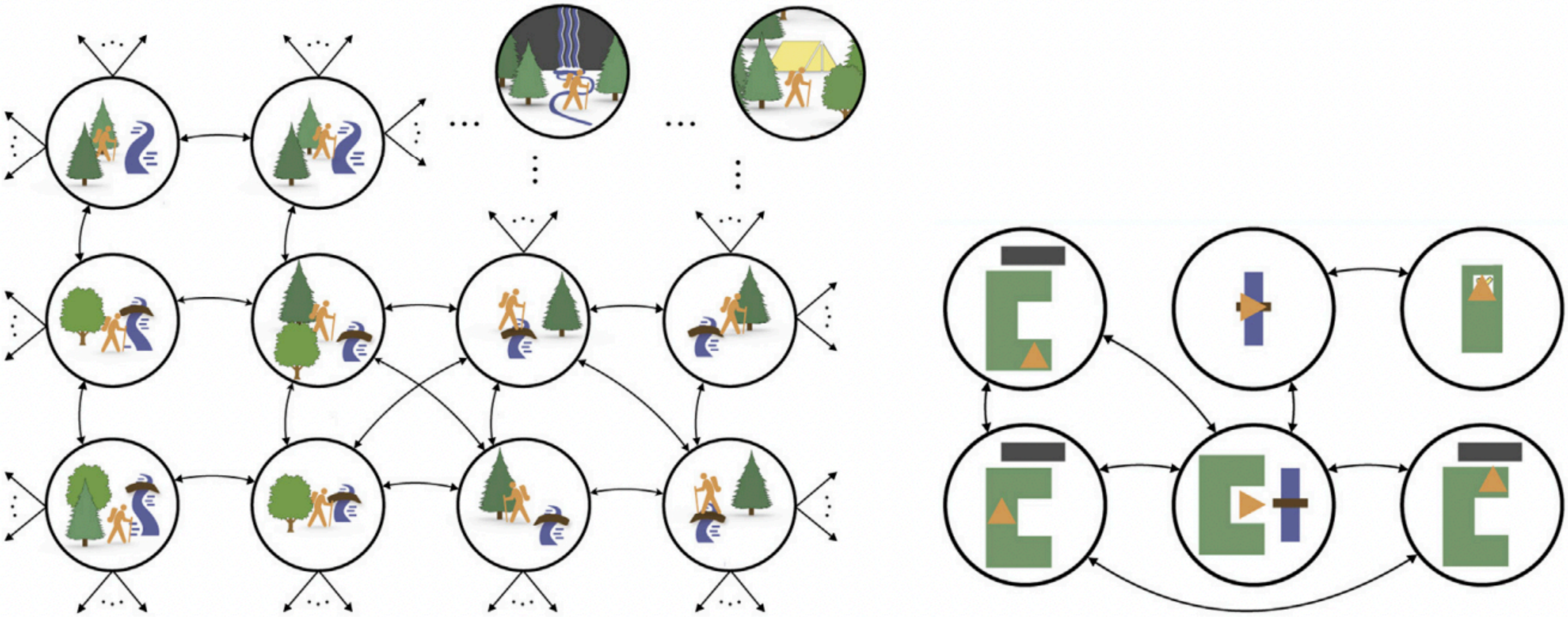


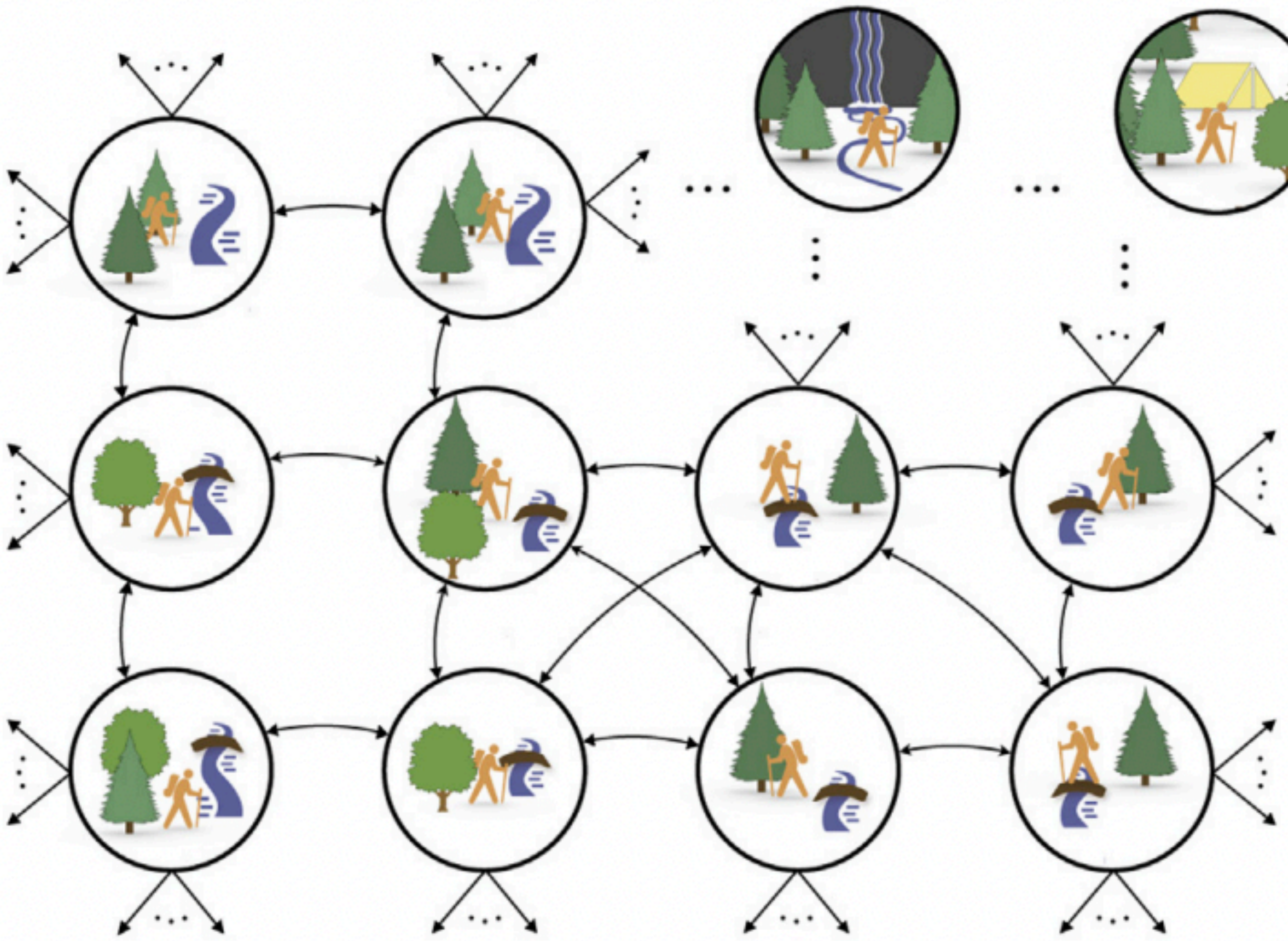
$$\Delta\phi = \lambda\phi$$

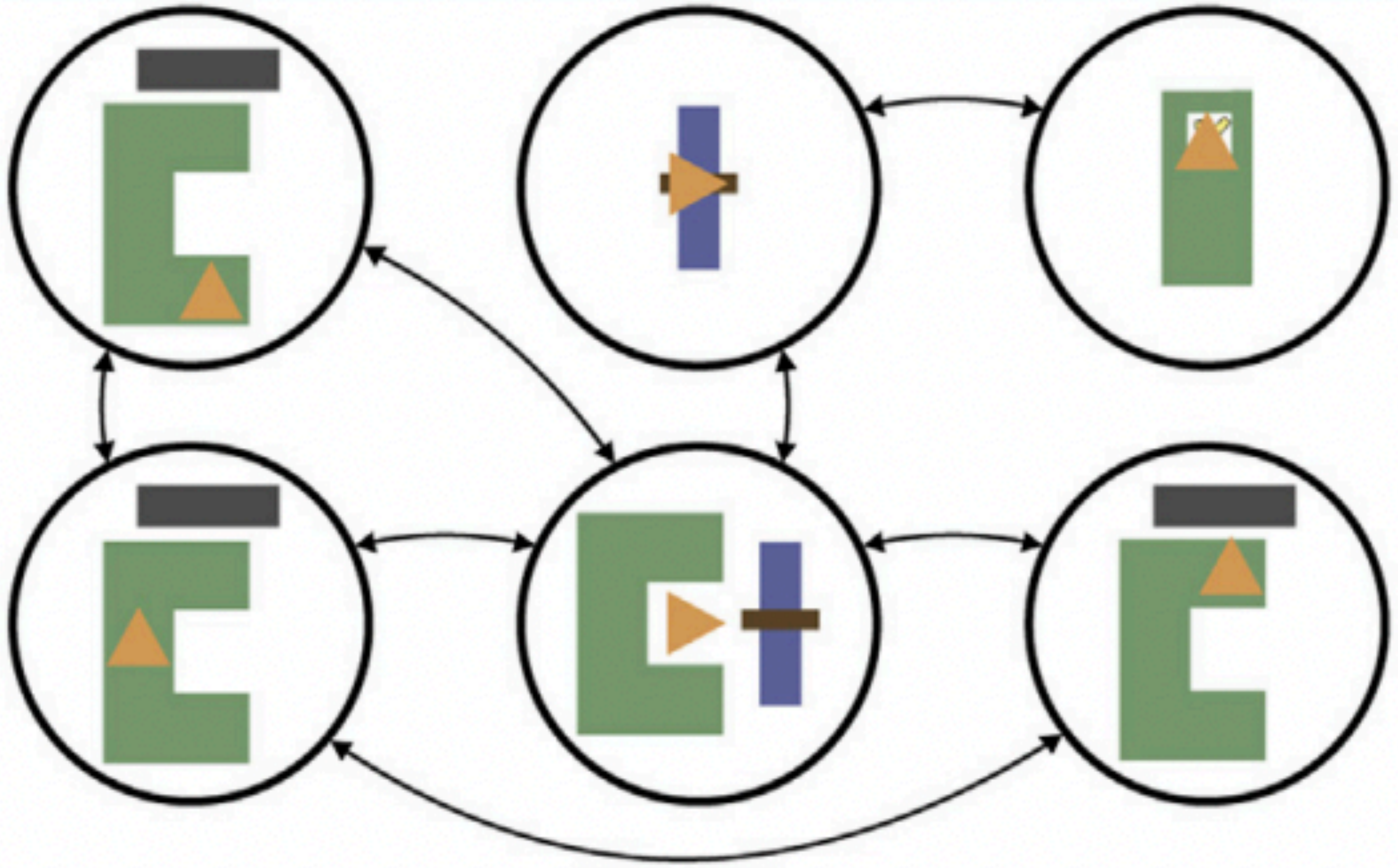
- Basis set to represent functions
- Filtering of scales
- Smoothing and interpolation
- **Spectral coordinates**
- ...



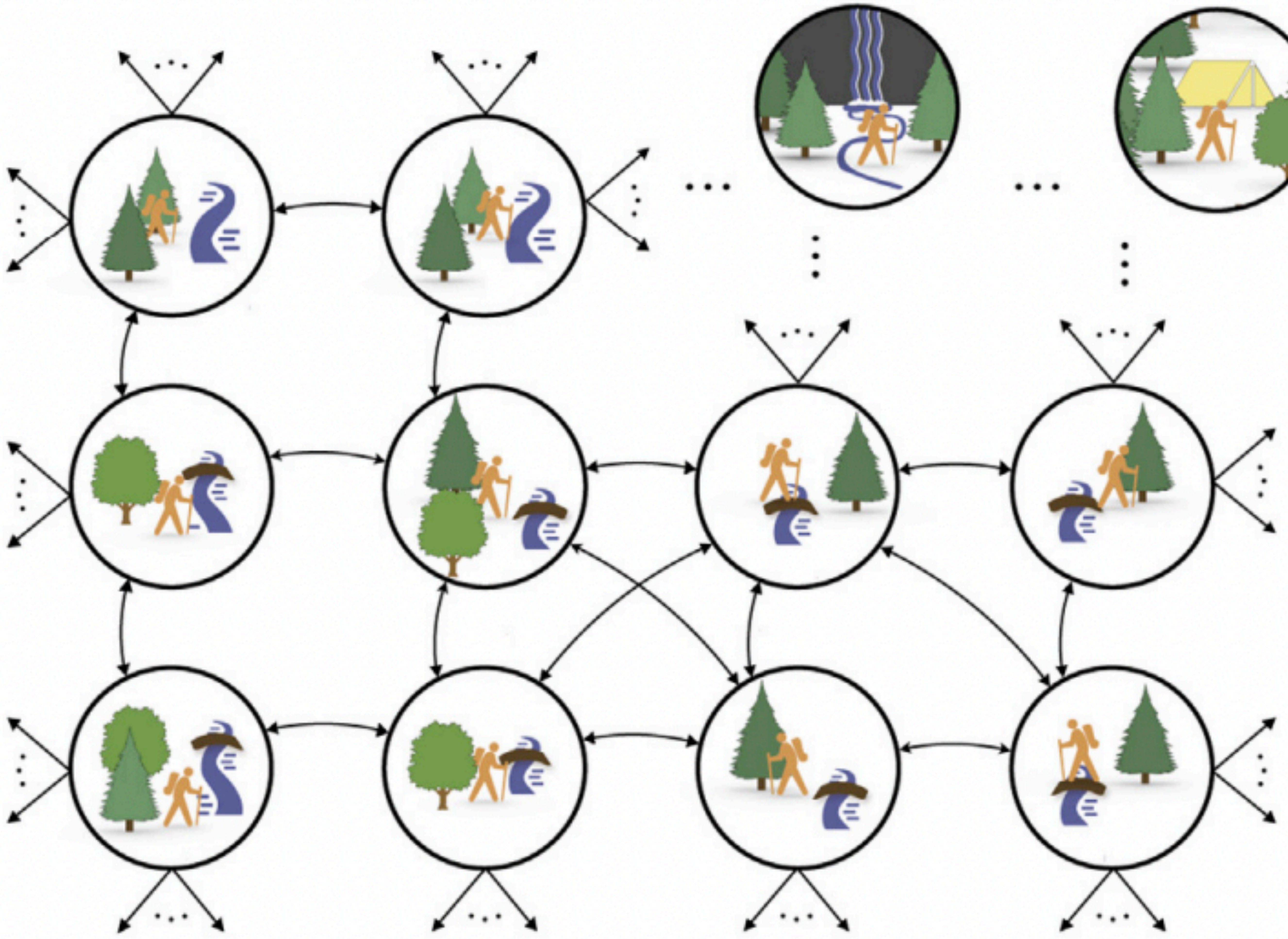




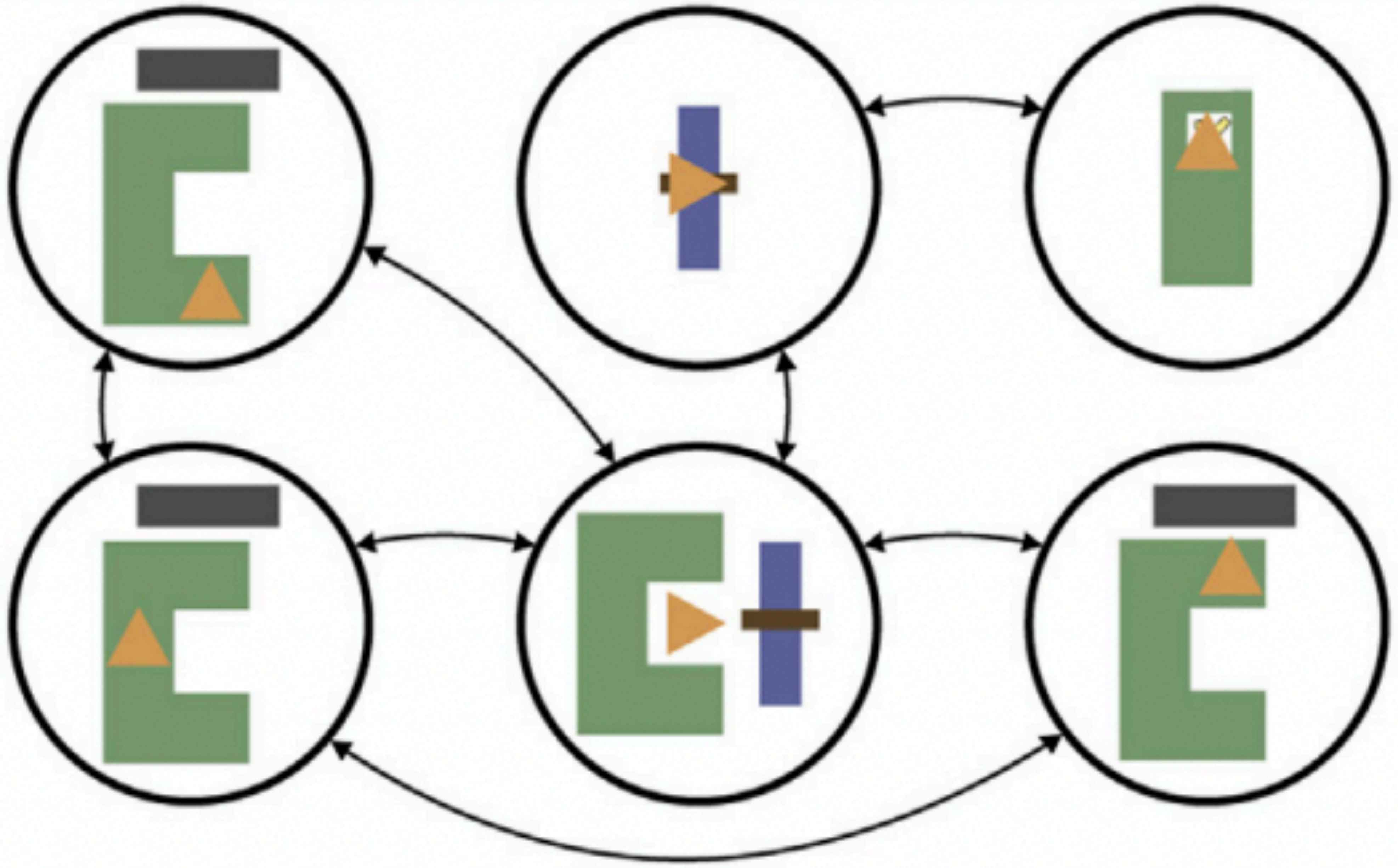




“At the current state, rotate your left leg 25 degrees, place it down beside the rock on the path, then swing your arm forward...”



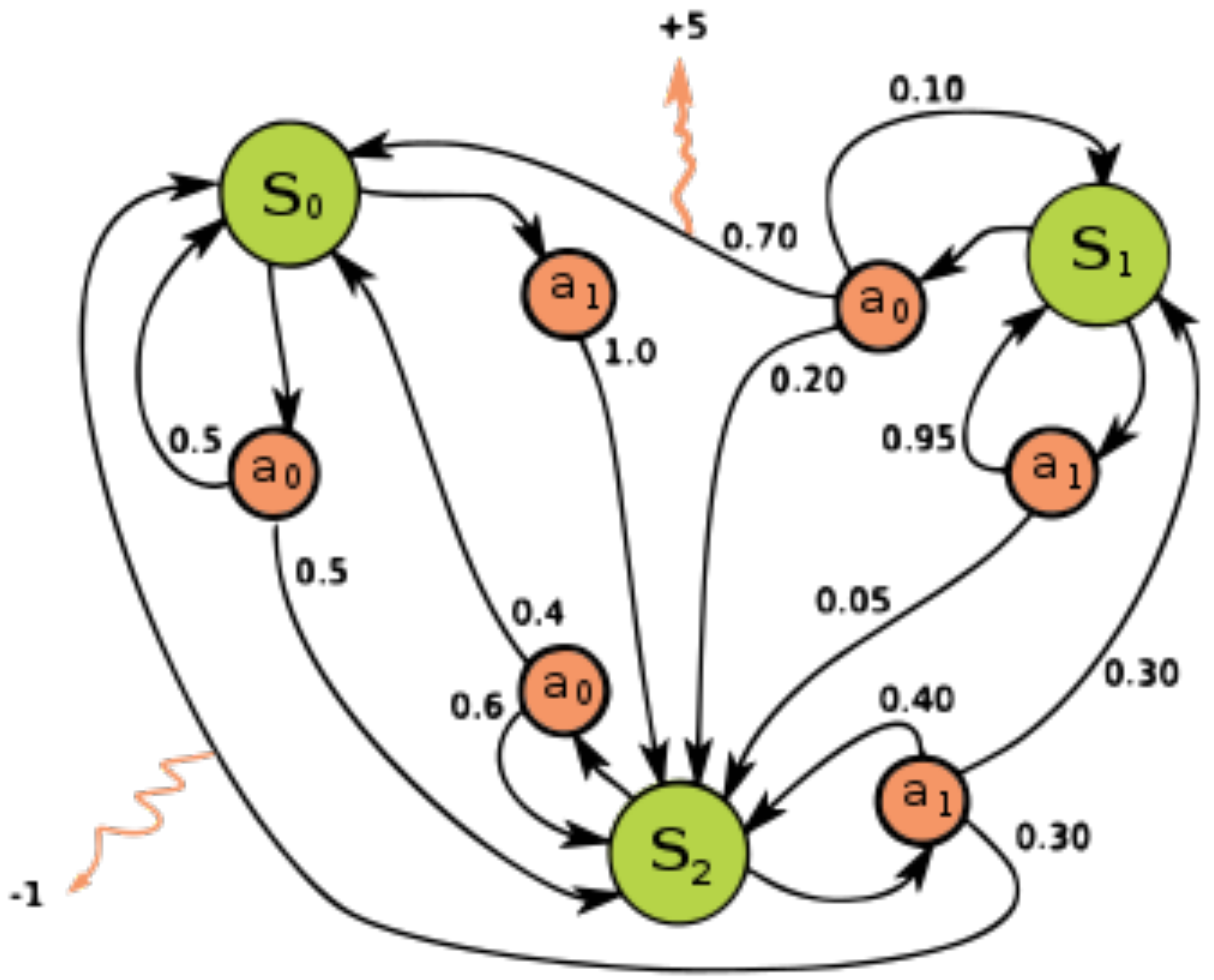
"At the current state, rotate your left leg 25 degrees, place it down beside the rock on the path, then swing your arm forward..."



"Since I am on the west side of the river, I probably want to cross the river to get to the east side of the river and then walk towards my campsite."

Reinforcement Learning

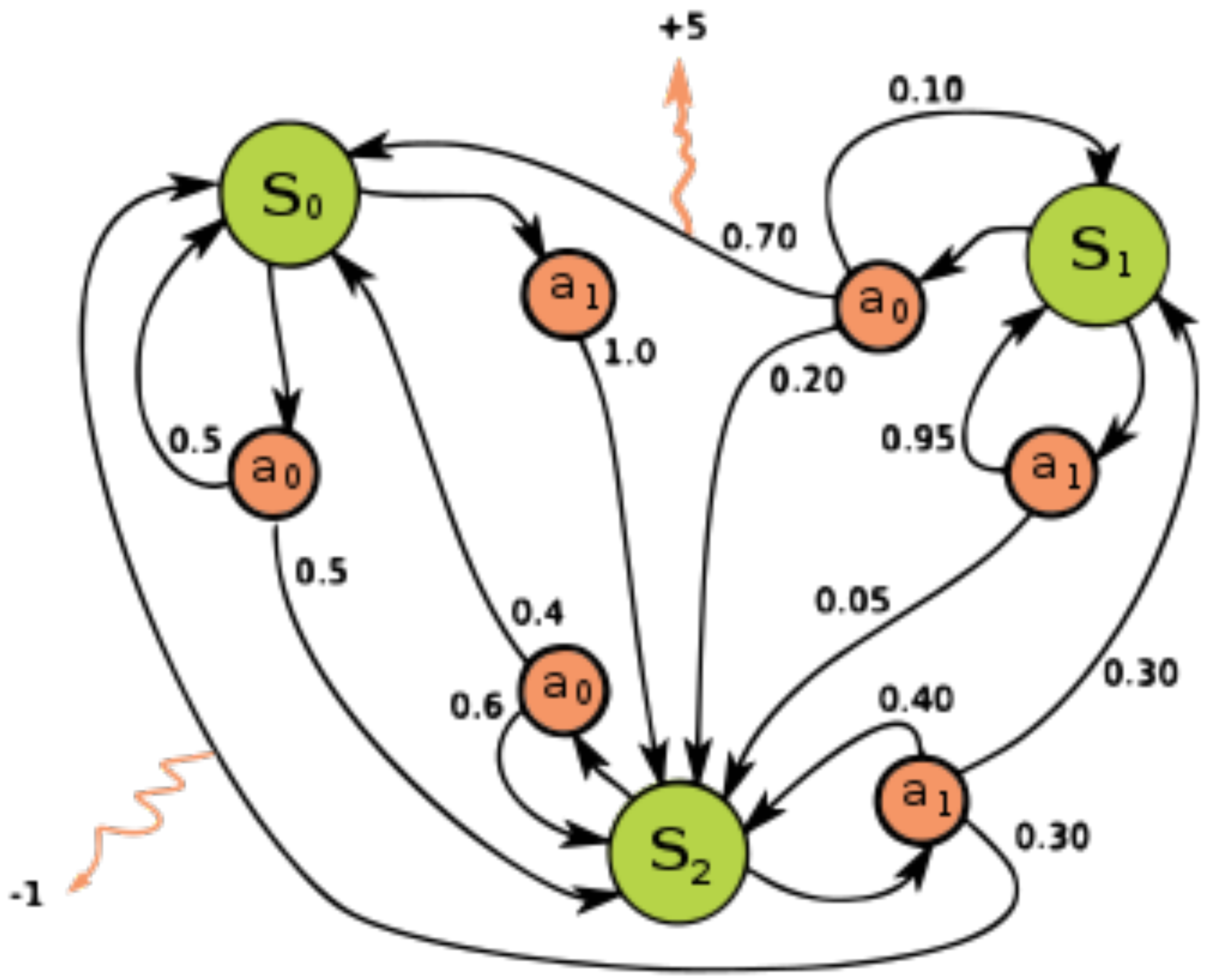
- RL: process to learn long-term future rewards
- Core problem: interdependency of optimal actions
- Tractable when representations encode structure of the environment/task that allows efficient discovery of optimal policy, search of short-paths, replanning, etc…

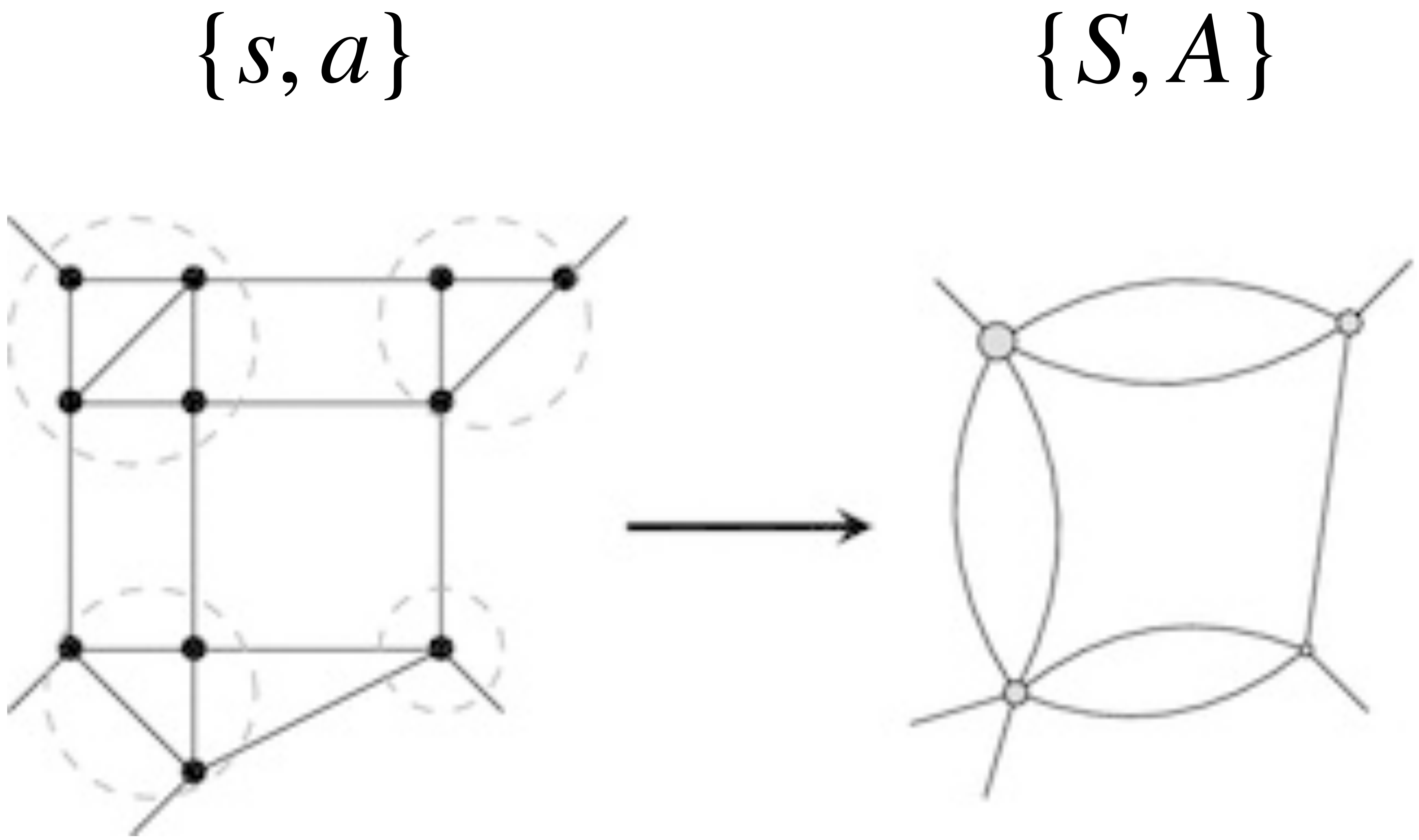


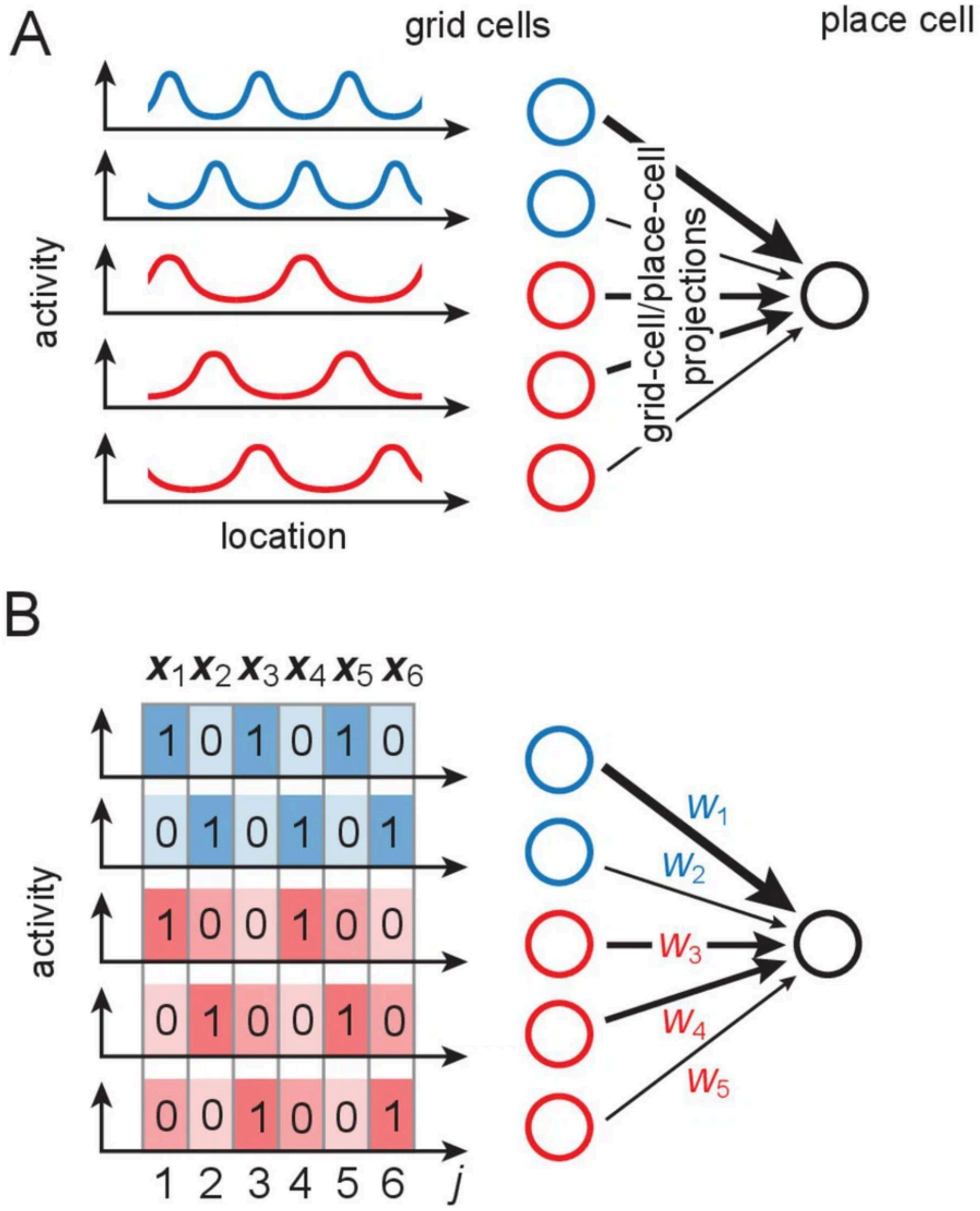
Markov Decision Process

State abstractions: treats certain configurations of the environment as similar by aggregating them. Ex: "being west of the river".

Temporal abstractions: temporally extended macro-actions that describe a general course of action. Ex: "cross the bridge".







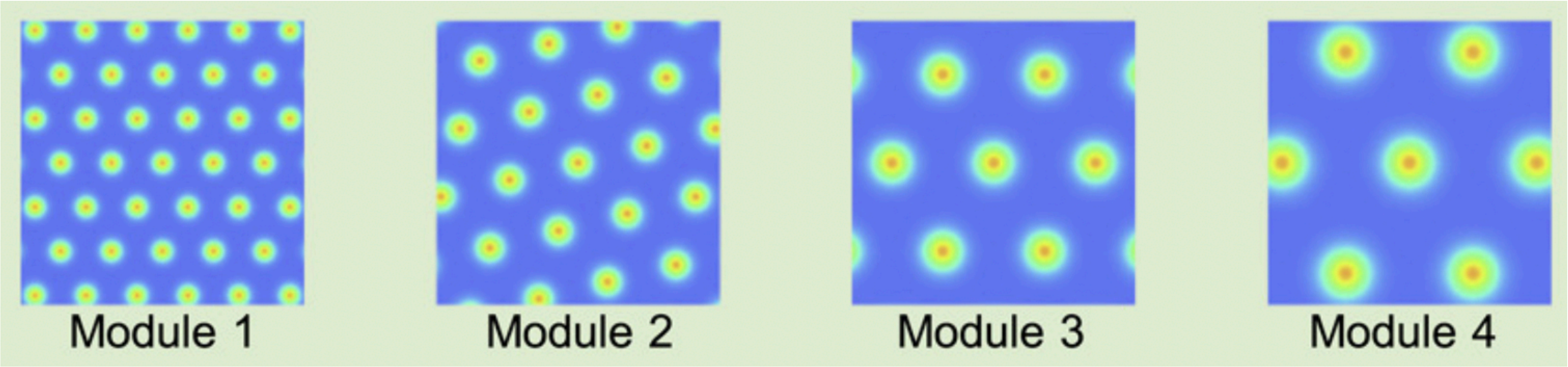
Place cells ⟶ One-hot encoding

Grid cells ⟶ Spectral encoding

$$s(x) = (\varphi_1(x), \varphi_2(x), \ldots, \varphi_N(x)) \text{ for vertex } x.$$

Grid cells

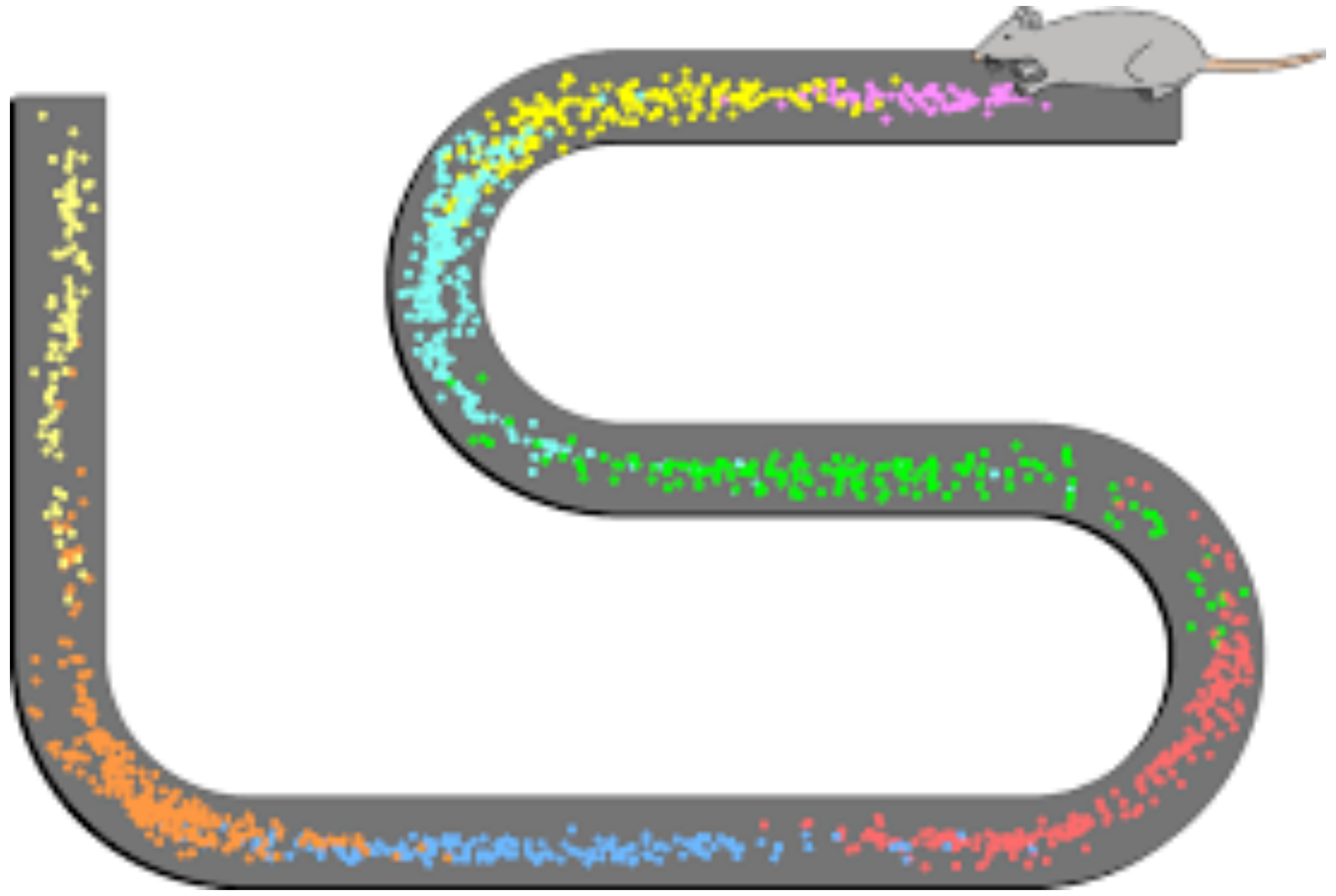
Distributed encoding

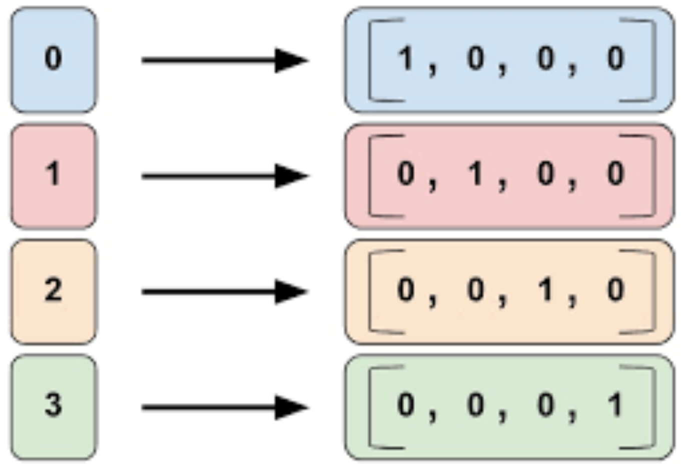


$s(x) = (\varphi_1(x), \varphi_2(x), \ldots, \varphi_N(x))$ for vertex x

Place cells

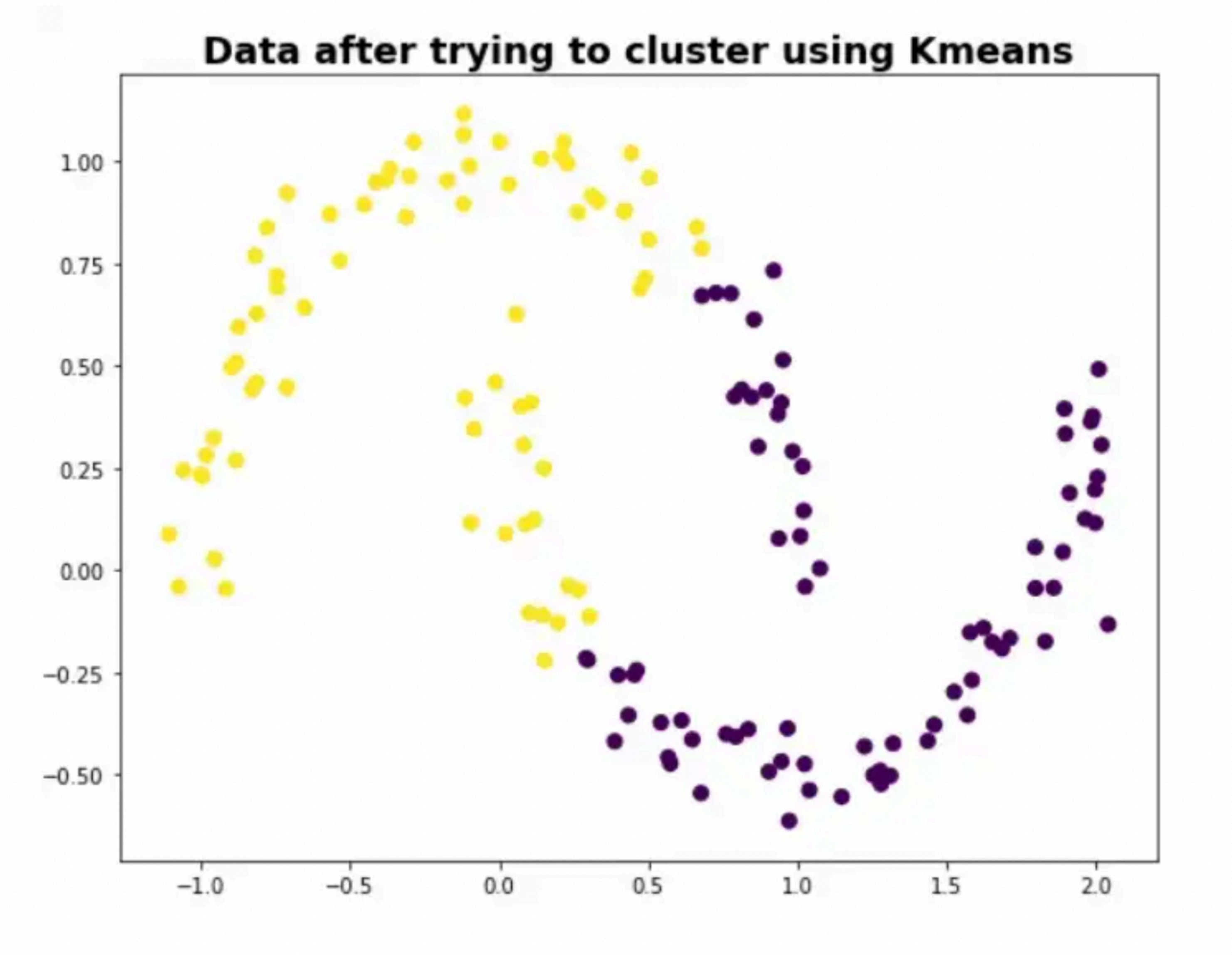
One-hot encoding

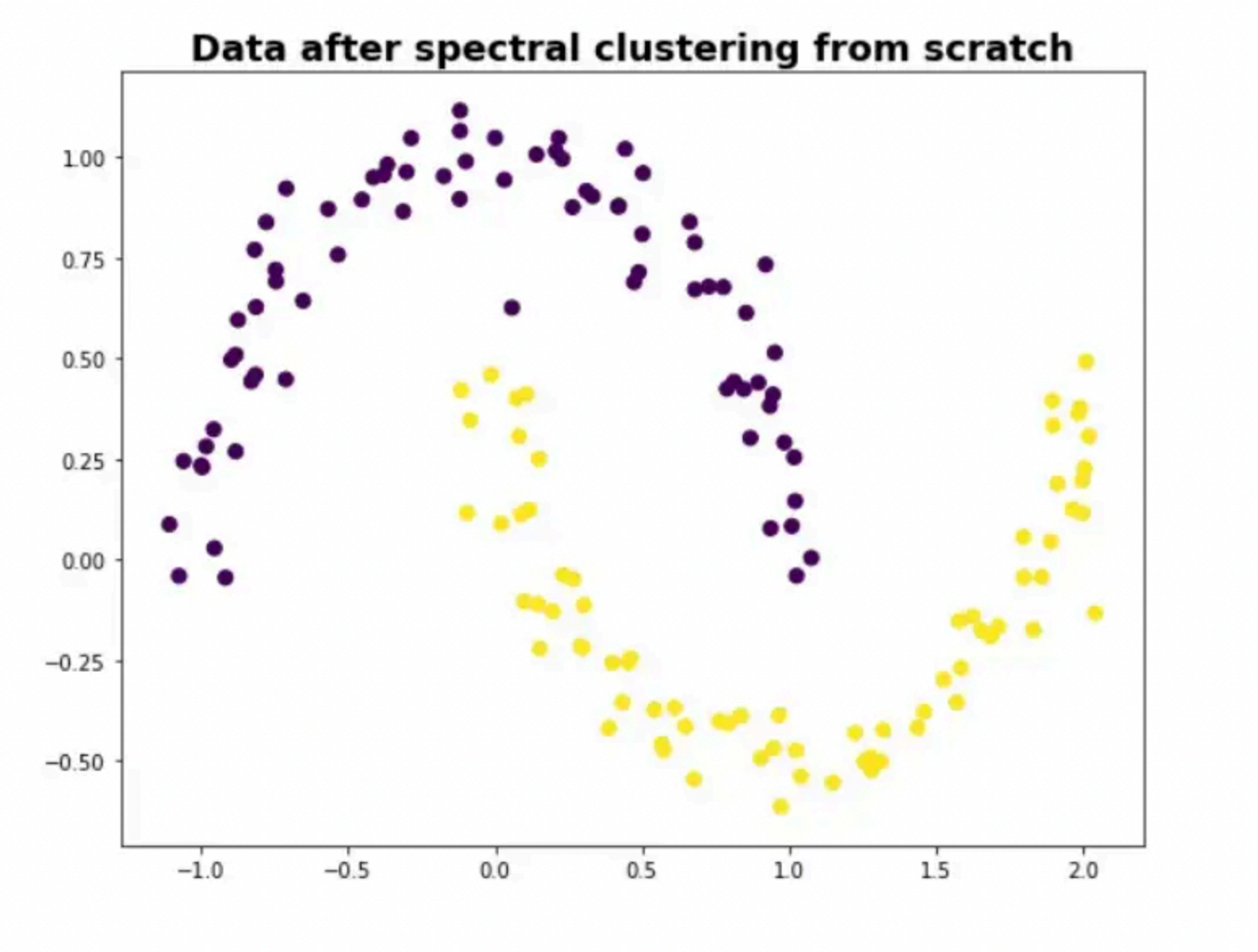




Spectral clustering

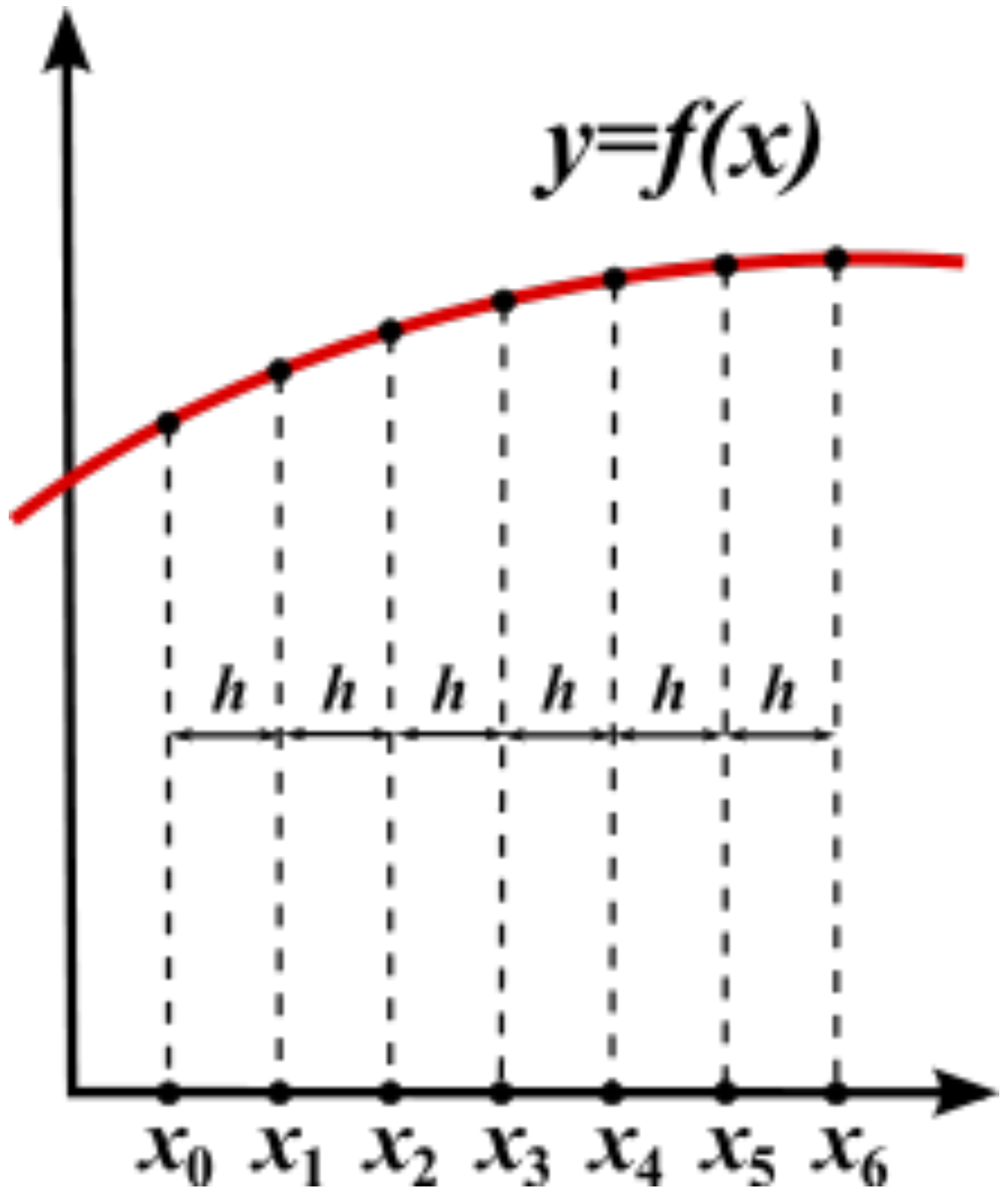
1. Calculate the Laplacian L (or the normalized Laplacian)
2. Calculate the first k eigenvectors (the eigenvectors corresponding to the k smallest eigenvalues of L)
3. Consider the matrix formed by the first k eigenvectors; the l-th row defines the features of graph node l
4. Cluster the graph nodes based on these features (e.g., using k-means clustering)





Ng, Andrew Y.; Jordan, Michael I.; Weiss, Yair (2002). "On spectral clustering: analysis and an algorithm" *Advances in Neural Information Processing Systems*.

Graph Laplacian operator



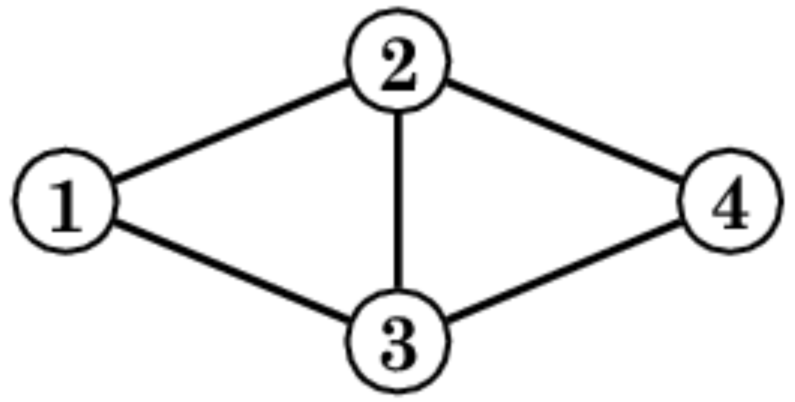
Second-order central

$$f''(x) \approx \frac{\delta_h^2[f](x)}{h^2} = \frac{\frac{f(x+h)-f(x)}{h} - \frac{f(x)-f(x-h)}{h}}{h} = \frac{f(x+h) - 2f(x) + f(x-h)}{h^2}.$$

Kernel (1, -2, 1)

On a general graph:

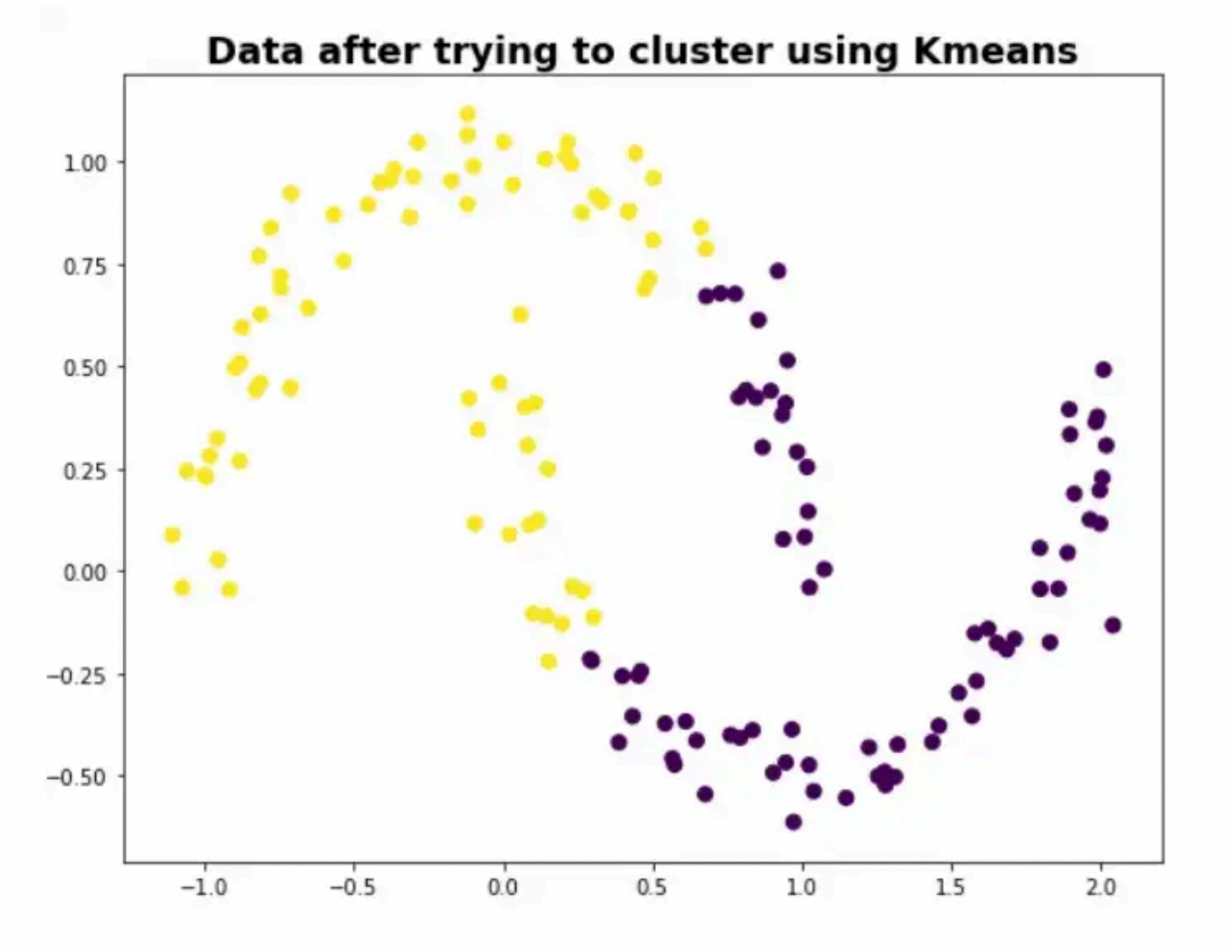
$$D = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

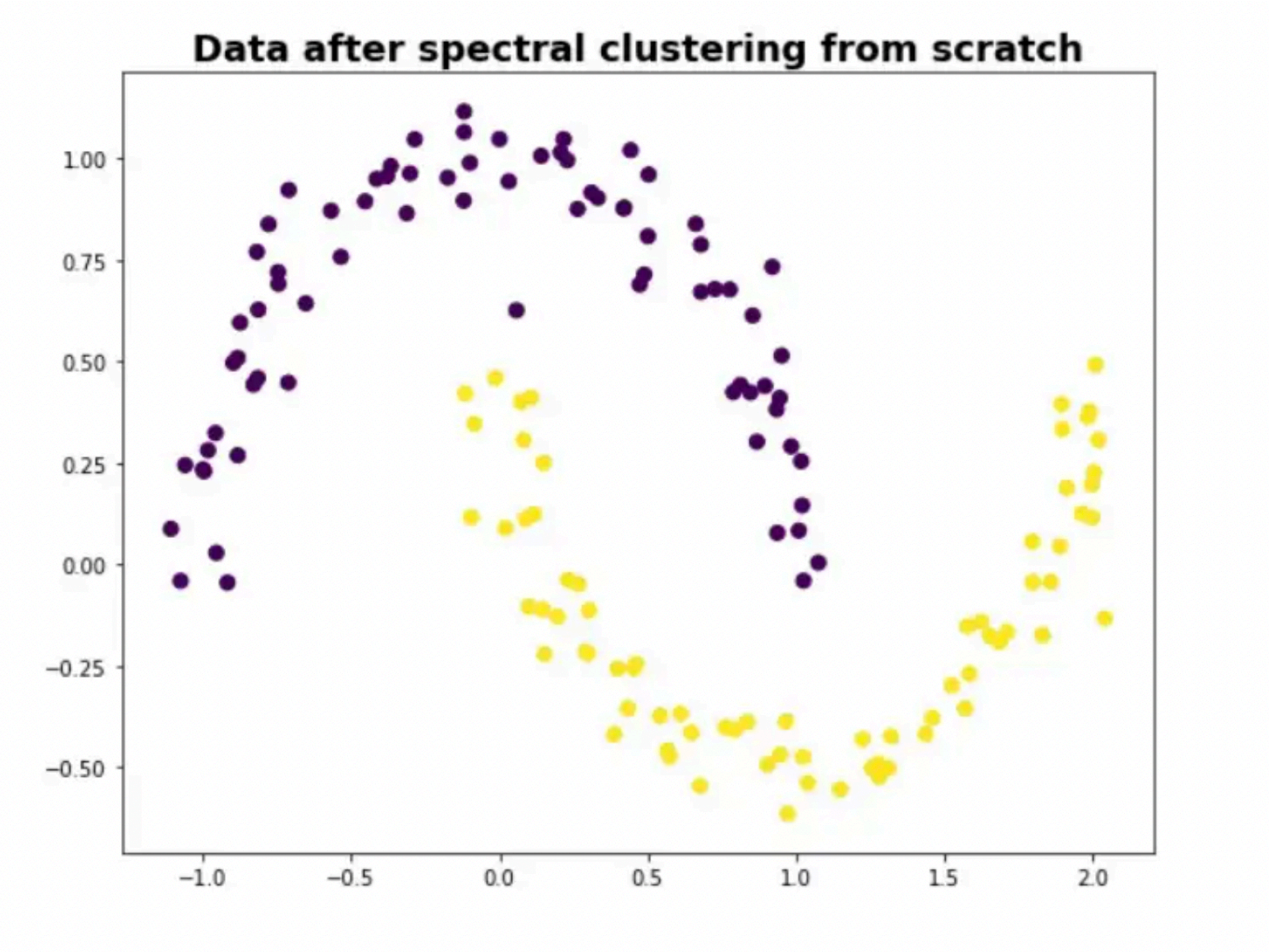


$$A = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} \qquad L = \begin{bmatrix} 2 & -1 & -1 & 0 \\ -1 & 3 & -1 & -1 \\ -1 & -1 & 3 & -1 \\ 0 & -1 & -1 & 2 \end{bmatrix}$$

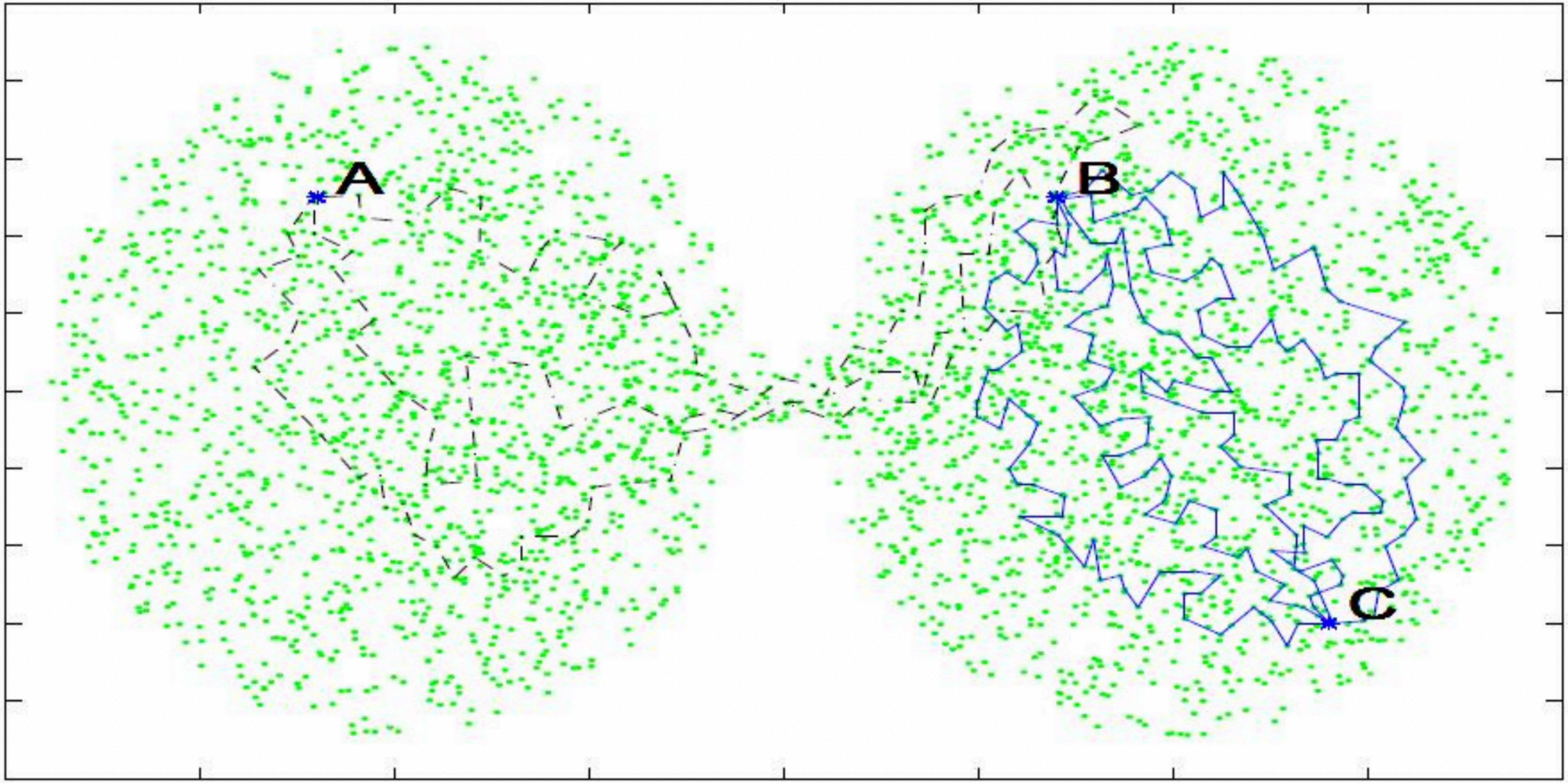
Spectral clustering

1. Calculate the Laplacian L (or the normalized Laplacian)
2. Calculate the first k eigenvectors (the eigenvectors corresponding to the k smallest eigenvalues of L)
3. Consider the matrix formed by the first k eigenvectors; the l-th row defines the features of graph node l
4. Cluster the graph nodes based on these features (e.g., using k-means clustering)

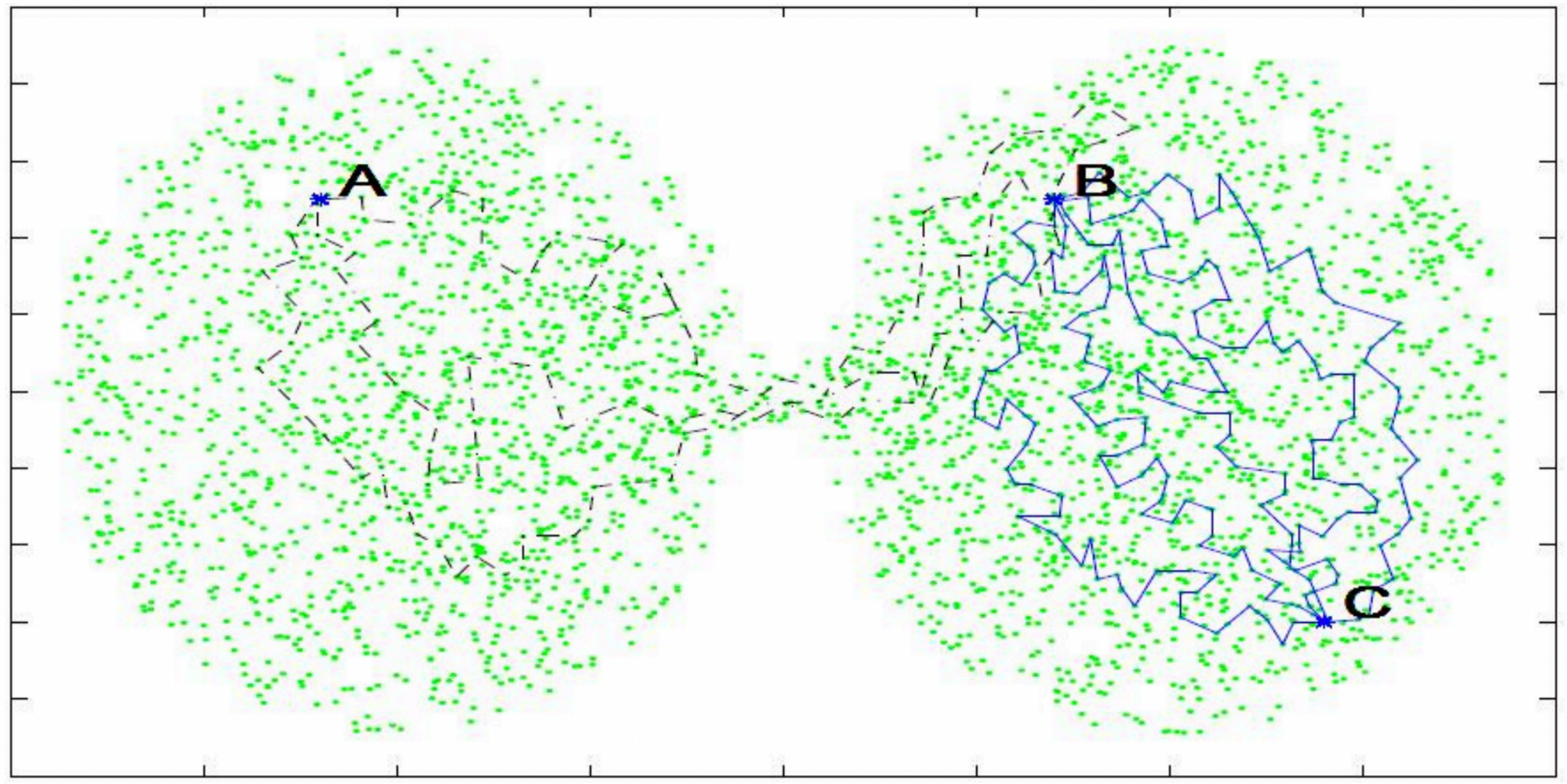




Ng, Andrew Y.; Jordan, Michael I.; Weiss, Yair (2002). "On spectral clustering: analysis and an algorithm" *Advances in Neural Information Processing Systems*.

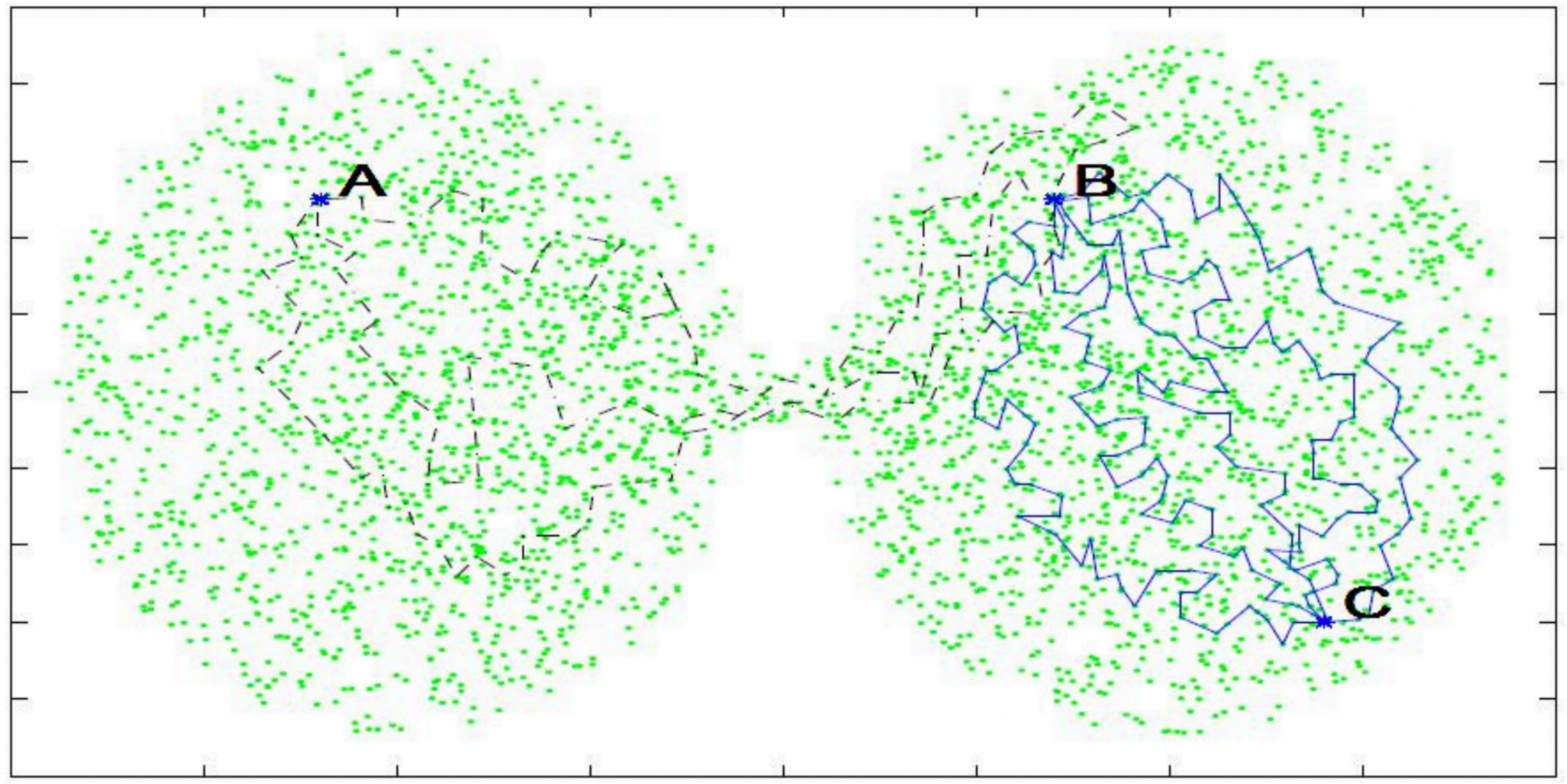

A
B
C

Diffusion maps



- Dimensionality reduction technique
- Embedding of data set into Euclidean space (low dimension)
- Coordinates computed from eigenvectors and eigenvalues of diffusion operator
- **Geometric structure** of high-dimensional data by modeling a diffusion process on a graph.

Diffusion maps



- Dimensionality reduction technique

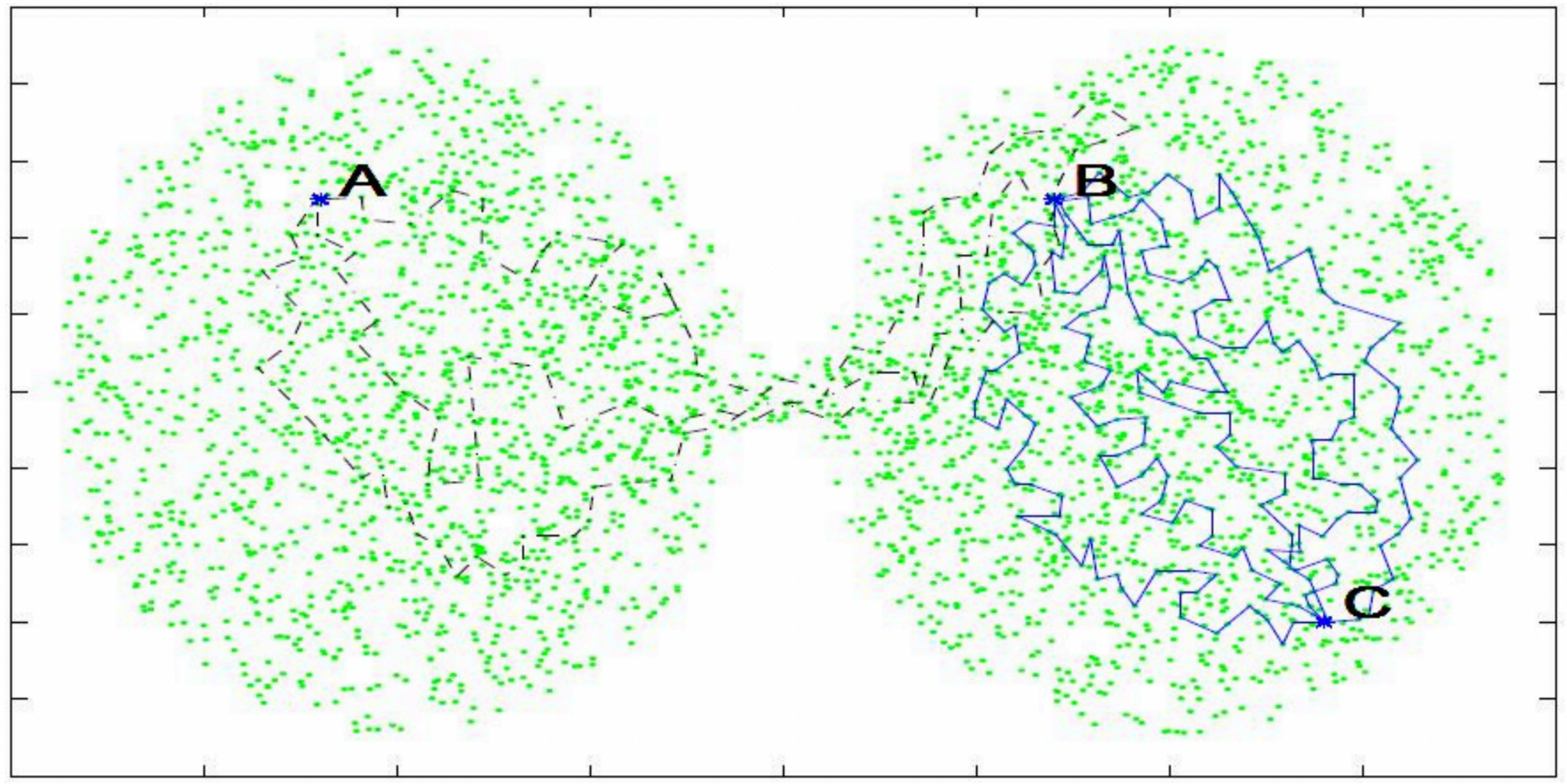
- Embedding of data set into Euclidean space (low dimension)

- Coordinates computed from eigenvectors and eigenvalues of diffusion operator

- **Geometric structure** of high-dimensional data by modeling a diffusion process on a graph.

Matrix	Definition	Purpose in Diffusion Maps
Affinity Matrix W	$W_{ij} = e^{-\frac{\|x_i - x_j\|^2}{\sigma^2}}$	Defines similarity between points
Degree Matrix D	$D_{ii} = \sum_j W_{ij}$	Measures total connectivity of each point
Laplacian $L = D - W$	Standard graph Laplacian	Captures local diffusion
Random Walk Laplacian $L_{\text{rw}} = I - D^{-1}W$	Scales diffusion process	Used to define Markov matrix
Markov Matrix $P = D^{-1}W$	Defines transition probabilities	Governs diffusion dynamics
Eigenvectors of P **or** L_{rw}	Basis for diffusion maps	Used for dimensionality reduction

Diffusion maps



- Dimensionality reduction technique

- Embedding of data set into Euclidean space (low dimension)

- Coordinates computed from eigenvectors and eigenvalues of diffusion operator

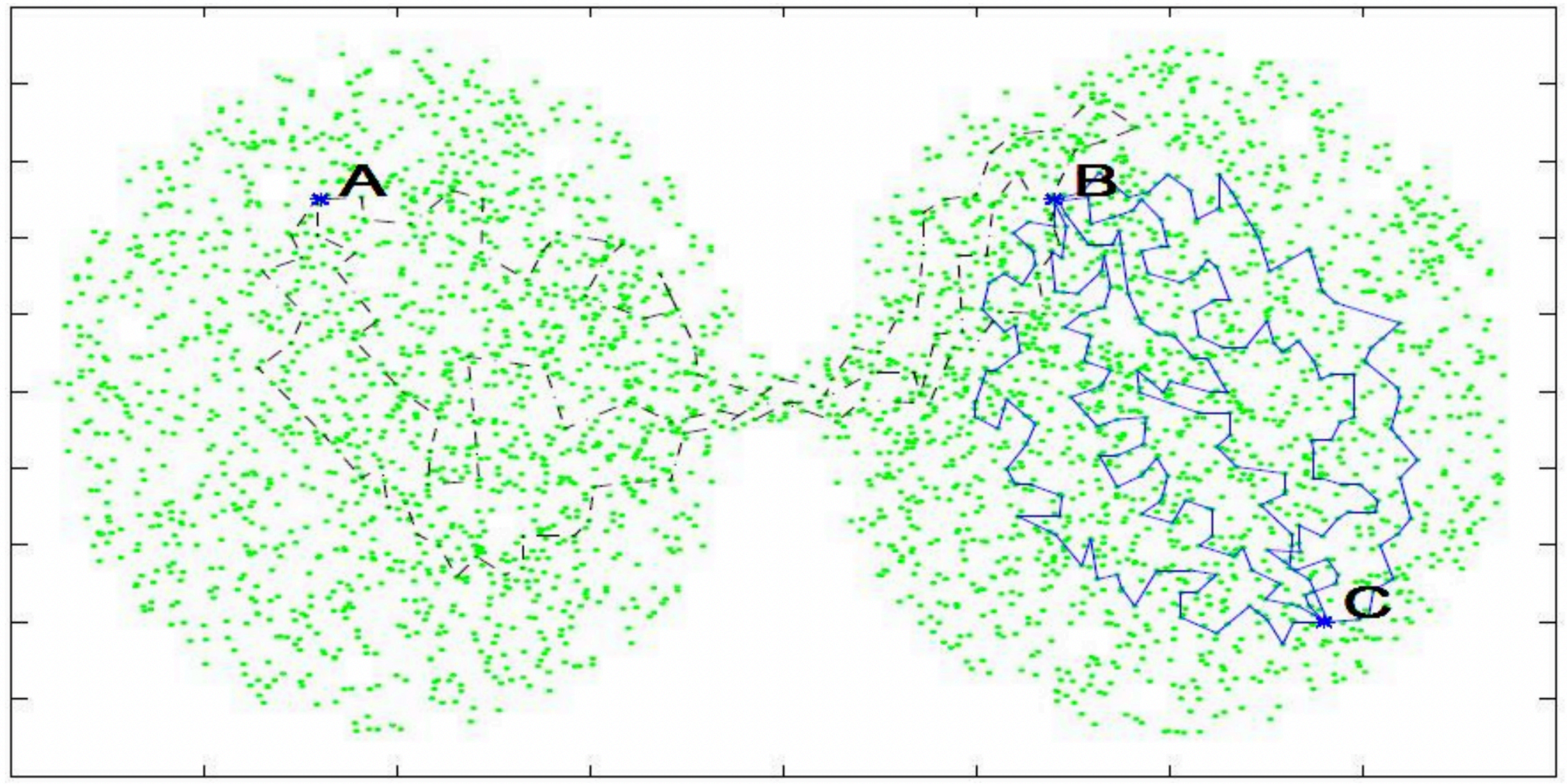
- **Geometric structure** of high-dimensional data by modeling a diffusion process on a graph.

Matrix	Definition	Purpose in Diffusion Maps
Affinity Matrix W	$W_{ij} = e^{-\frac{\|x_i - x_j\|^2}{\sigma^2}}$	Defines similarity between points
Degree Matrix D	$D_{ii} = \sum_j W_{ij}$	Measures total connectivity of each point
Laplacian $L = D - W$	Standard graph Laplacian	Captures local diffusion
Random Walk Laplacian $L_{\text{rw}} = I - D^{-1}W$	Scales diffusion process	Used to define Markov matrix
Markov Matrix $P = D^{-1}W$	Defines transition probabilities	Governs diffusion dynamics
Eigenvectors of P or L_{rw}	Basis for diffusion maps	Used for dimensionality reduction

- Diffusion distance: small if there is a large number o

$$p(x_j, t | x_i) = P^{\,t}_{i,j} \qquad D_t(x_i, x_j)^2 = \sum_y \frac{(p(y,t|x_i) - p(y,t|x_j))^2}{\phi_0(y)}$$

Diffusion maps



- Dimensionality reduction technique

- Embedding of data set into Euclidean space (low dimension)

- Coordinates computed from eigenvectors and eigenvalues of diffusion operator

- **Geometric structure** of high-dimensional data by modeling a diffusion process on a graph.

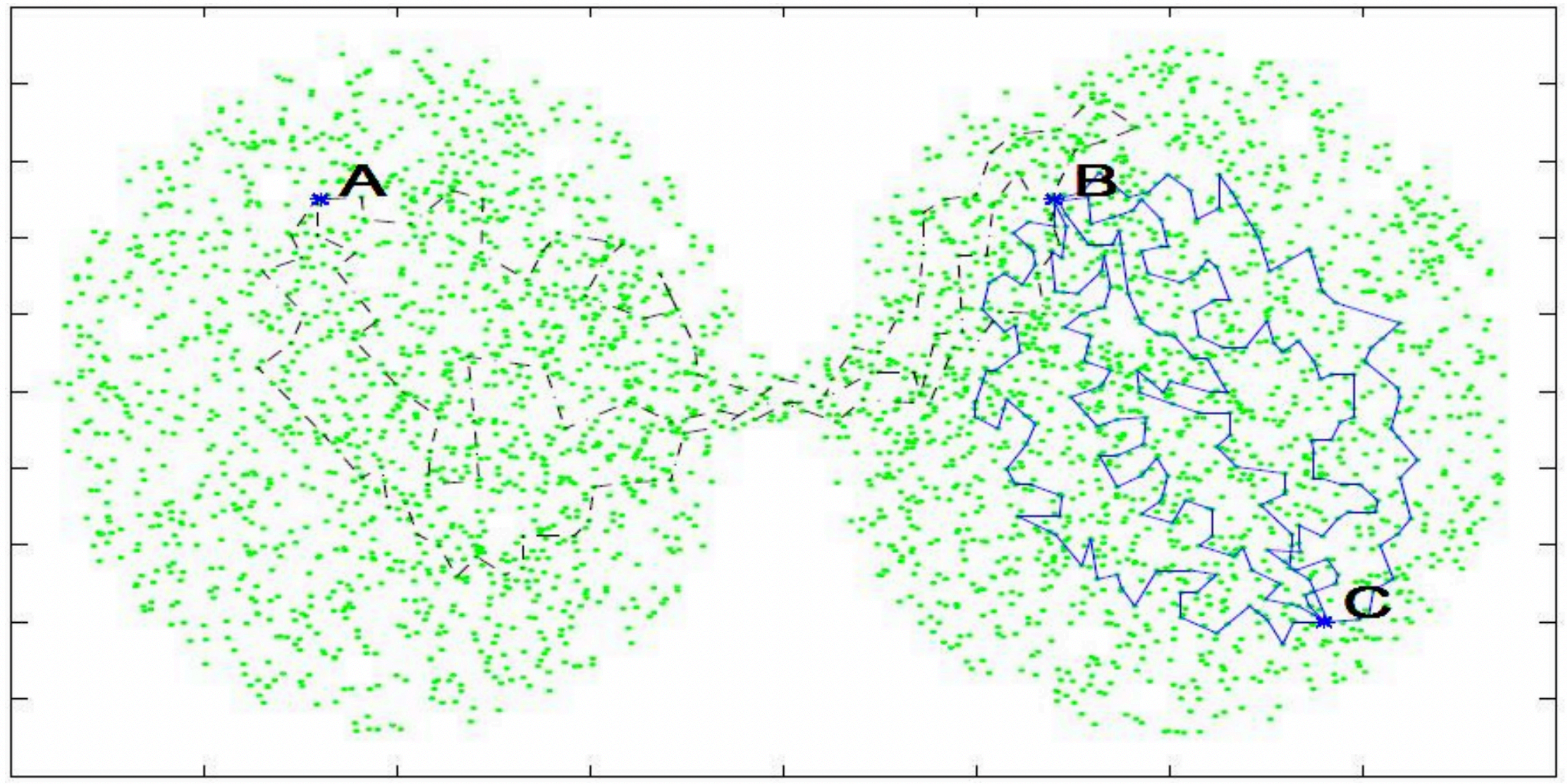
Matrix	Definition	Purpose in Diffusion Maps
Affinity Matrix W	$W_{ij} = e^{-\frac{\|x_i - x_j\|^2}{\sigma^2}}$	Defines similarity between points
Degree Matrix D	$D_{ii} = \sum_j W_{ij}$	Measures total connectivity of each point
Laplacian $L = D - W$	Standard graph Laplacian	Captures local diffusion
Random Walk Laplacian $L_{\mathrm{rw}} = I - D^{-1}W$	Scales diffusion process	Used to define Markov matrix
Markov Matrix $P = D^{-1}W$	Defines transition probabilities	Governs diffusion dynamics
Eigenvectors of P or L_{rw}	Basis for diffusion maps	Used for dimensionality reduction

- Diffusion distance: small if there is a large number o

$$p(x_j, t | x_i) = P^{\,t}_{i,j} \qquad D_t(x_i, x_j)^2 = \sum_y \frac{(p(y,t|x_i) - p(y,t|x_j))^2}{\phi_0(y)}$$

$$\Psi_t(x) = (\lambda_1^t \psi_1(x), \lambda_2^t \psi_2(x), \dots, \lambda_k^t \psi_k(x))$$

Diffusion maps



- Dimensionality reduction technique
- Embedding of data set into Euclidean space (low dimension)
- Coordinates computed from eigenvectors and eigenvalues of diffusion operator
- **Geometric structure** of high-dimensional data by modeling a diffusion process on a graph.

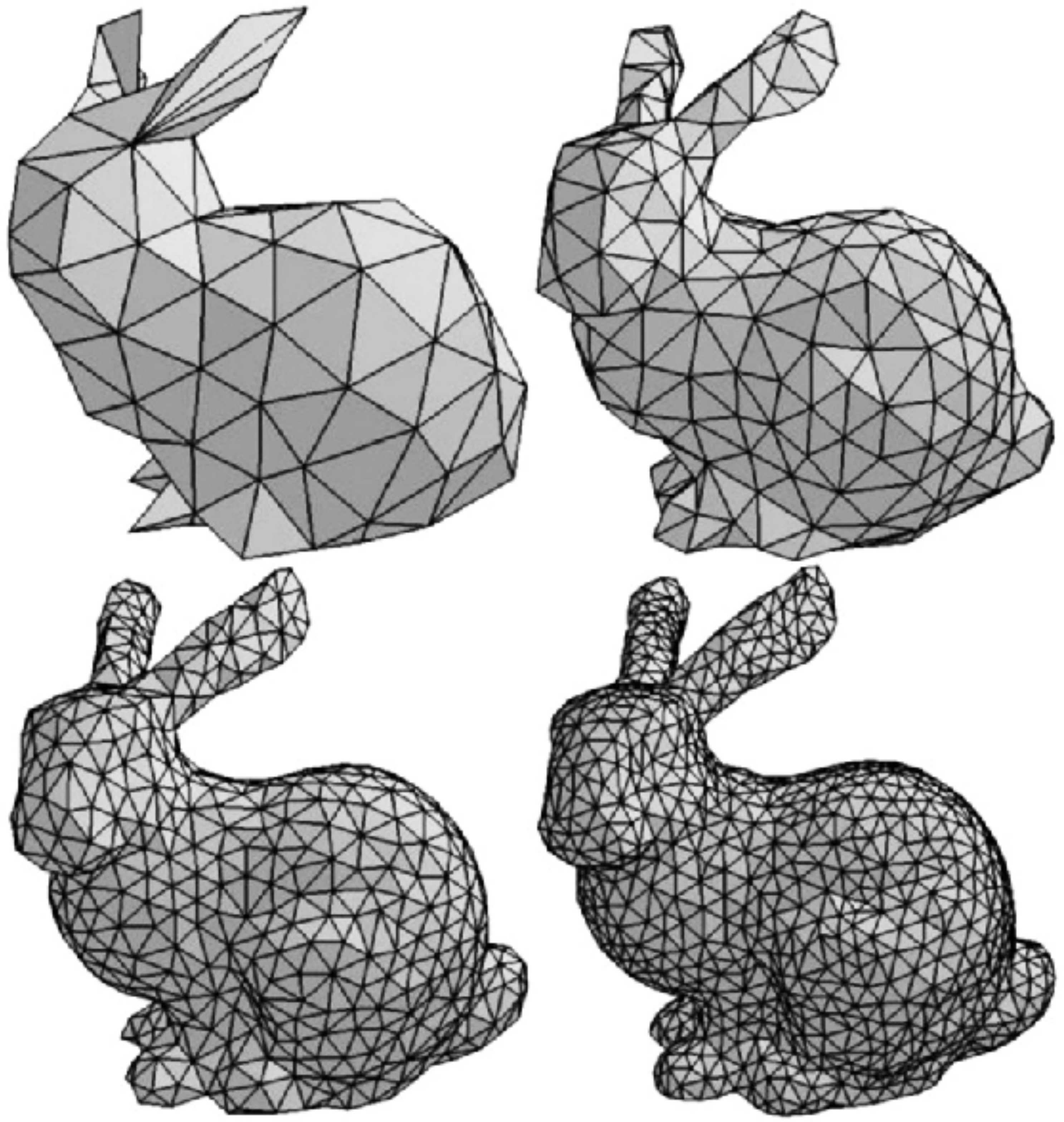
Matrix	Definition	Purpose in Diffusion Maps
Affinity Matrix W	$W_{ij} = e^{-\frac{\|x_i - x_j\|^2}{\sigma^2}}$	Defines similarity between points
Degree Matrix D	$D_{ii} = \sum_j W_{ij}$	Measures total connectivity of each point
Laplacian $L = D - W$	Standard graph Laplacian	Captures local diffusion
Random Walk Laplacian $L_{\text{rw}} = I - D^{-1}W$	Scales diffusion process	Used to define Markov matrix
Markov Matrix $P = D^{-1}W$	Defines transition probabilities	Governs diffusion dynamics
Eigenvectors of P or L_{rw}	Basis for diffusion maps	Used for dimensionality reduction

- Diffusion distance: small if there is a large number o

$$p(x_j, t | x_i) = P^{\,t}_{i,j} \qquad D_t(x_i, x_j)^2 = \sum_y \frac{(p(y,t|x_i) - p(y,t|x_j))^2}{\phi_0(y)}$$

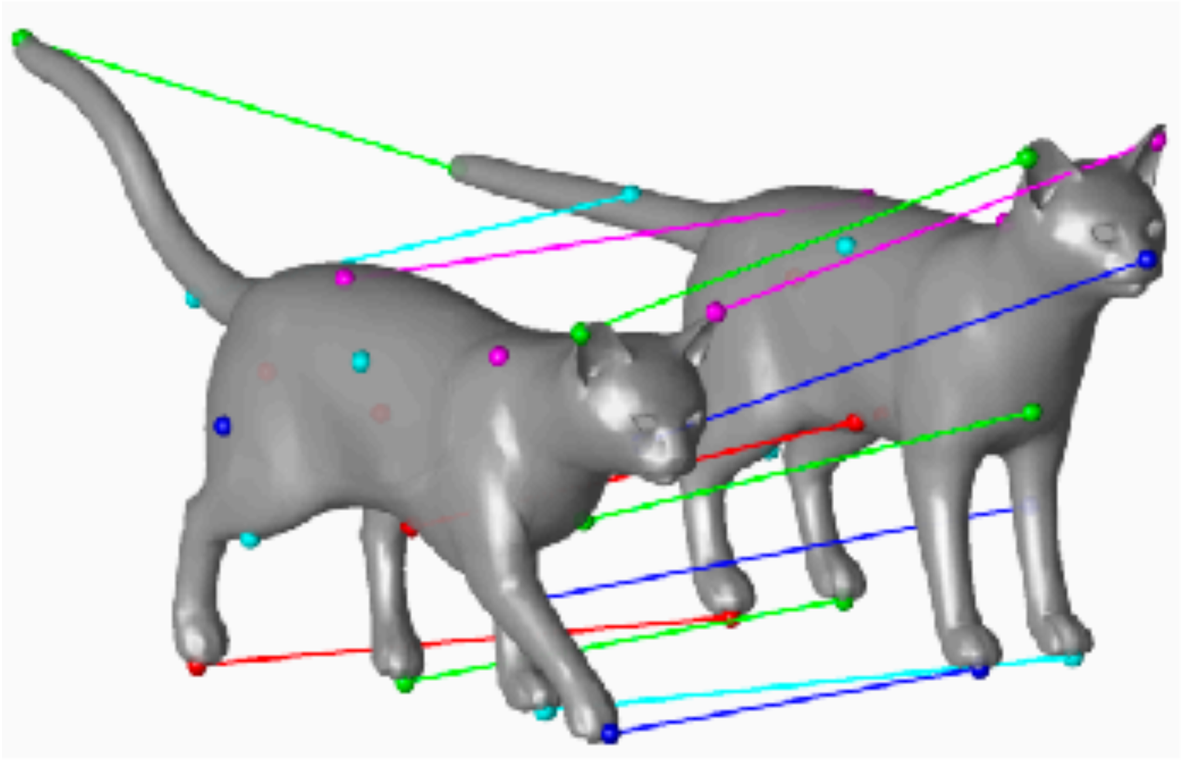
$$\Psi_t(x) = (\lambda_1^t \psi_1(x), \lambda_2^t \psi_2(x), \ldots, \lambda_k^t \psi_k(x))$$

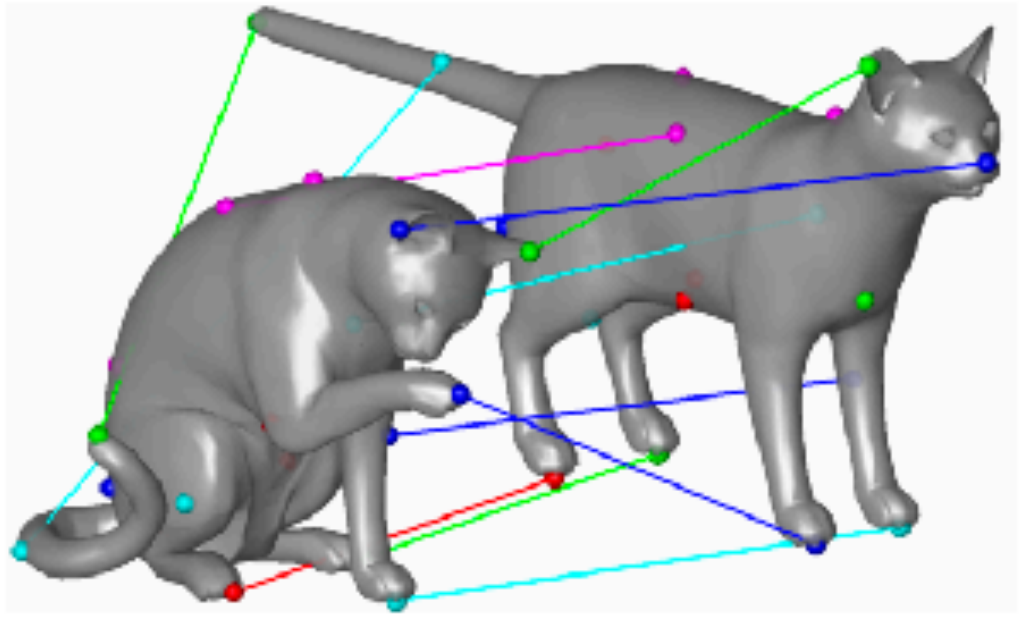
$$D_t(x_i, x_j)^2 \approx ||\Psi_t(x_i) - \Psi_t(x_j)||^2$$

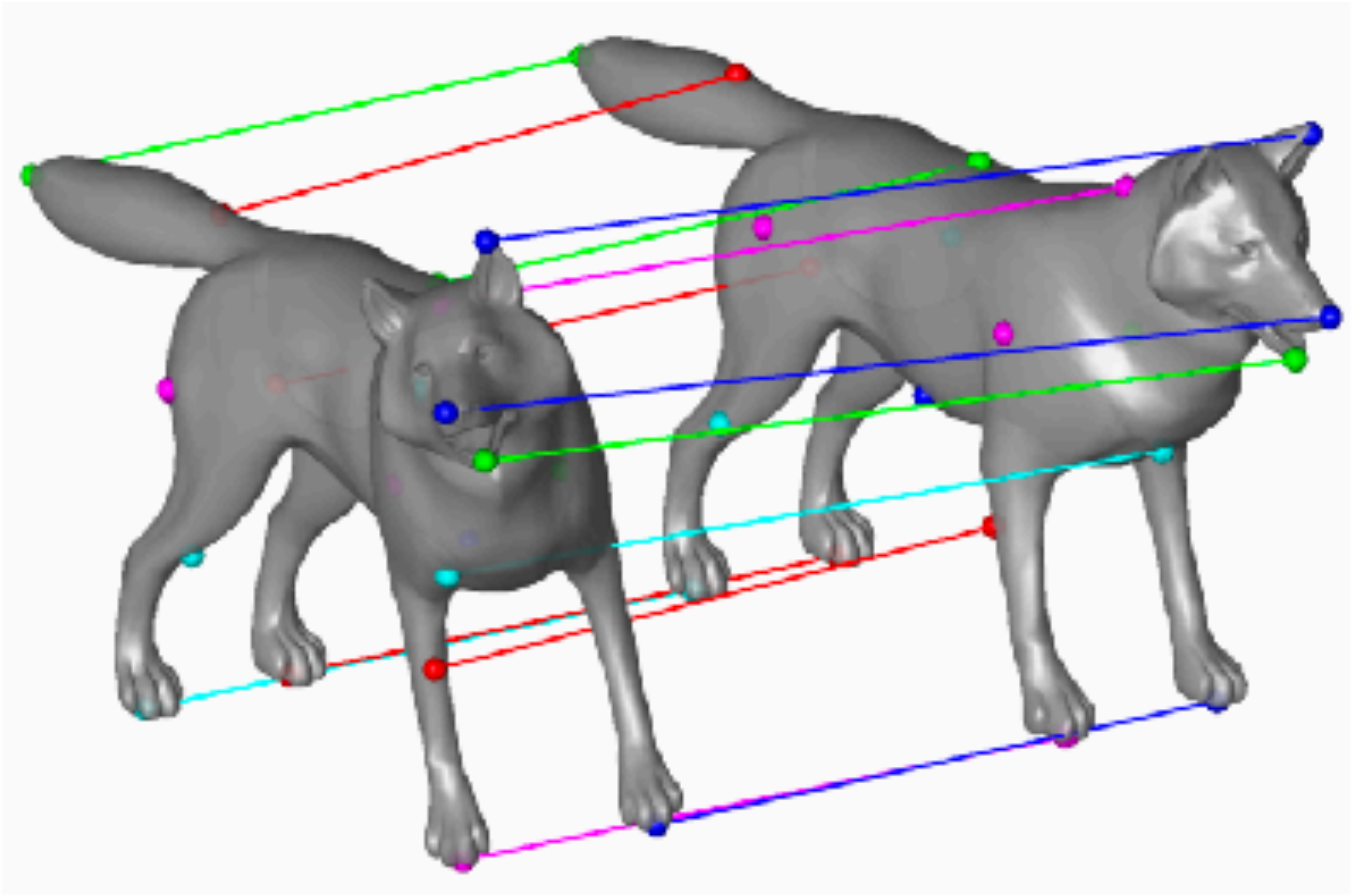


$s(x) = (\varphi_1(x), \varphi_2(x), \ldots, \varphi_N(x))$ for vertex x.

Spectral shape matching

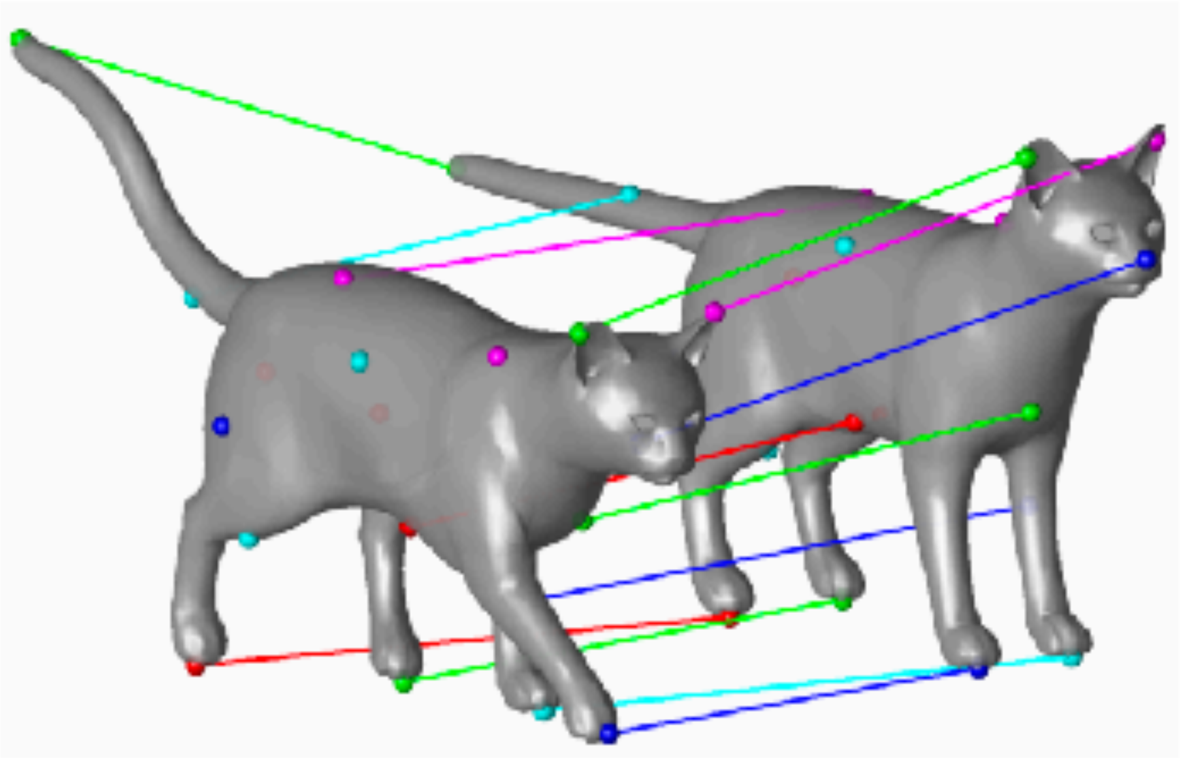


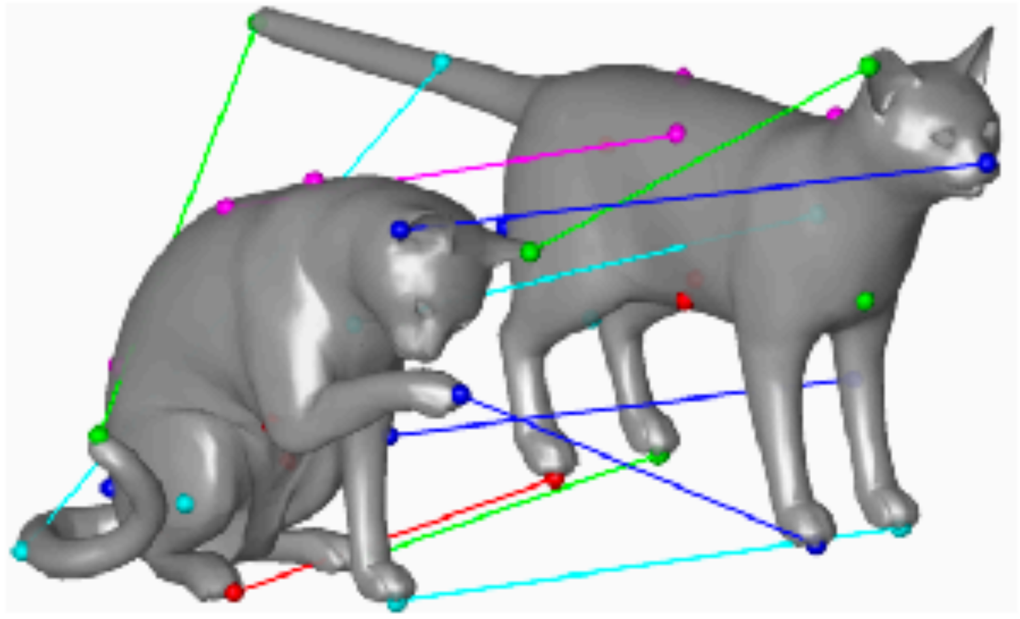


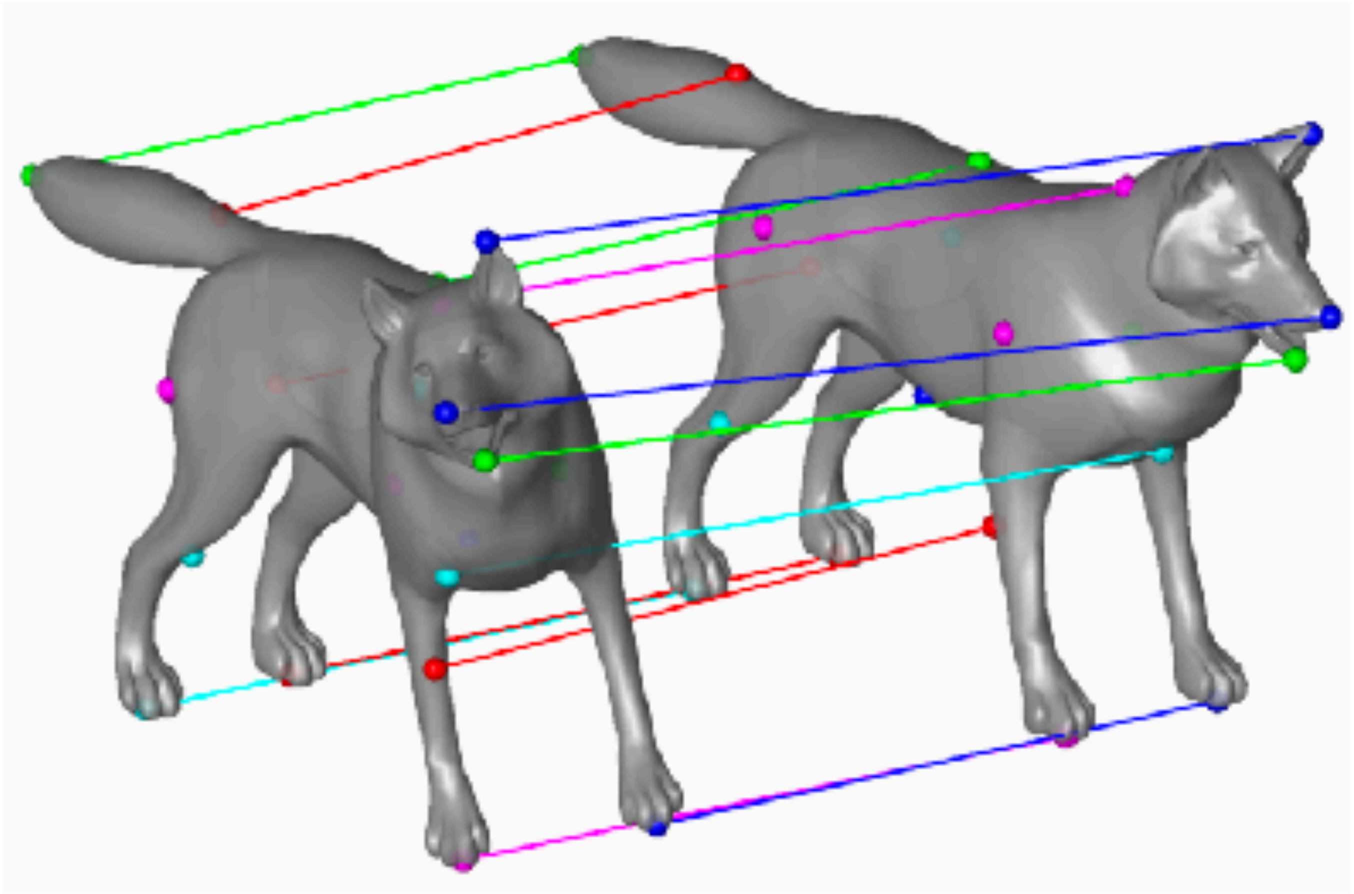


Spectral shape matching

$s(x) = (\varphi_1(x), \varphi_2(x), \ldots, \varphi_N(x))$ for vertex x.

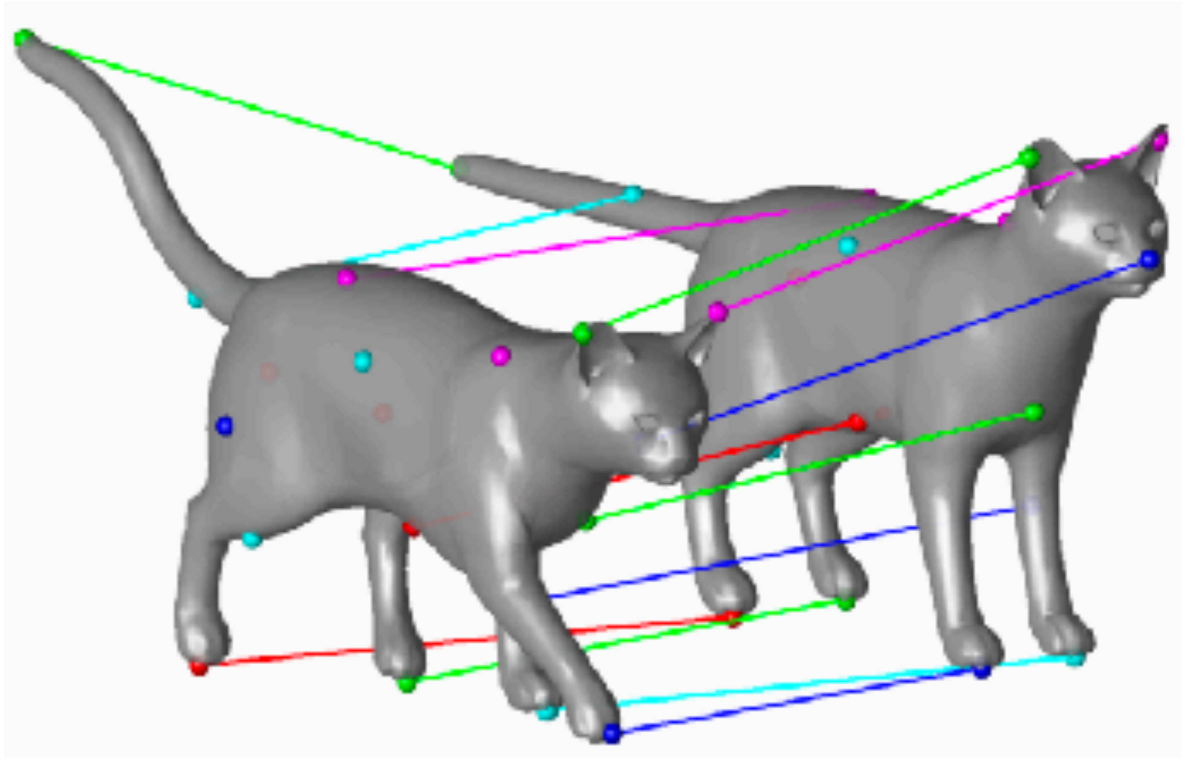


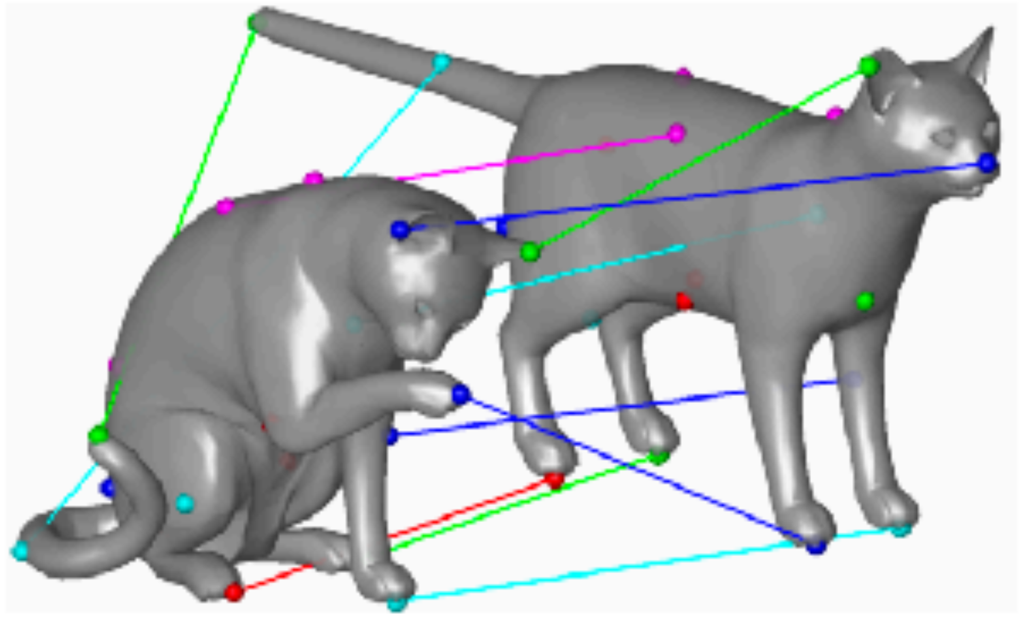


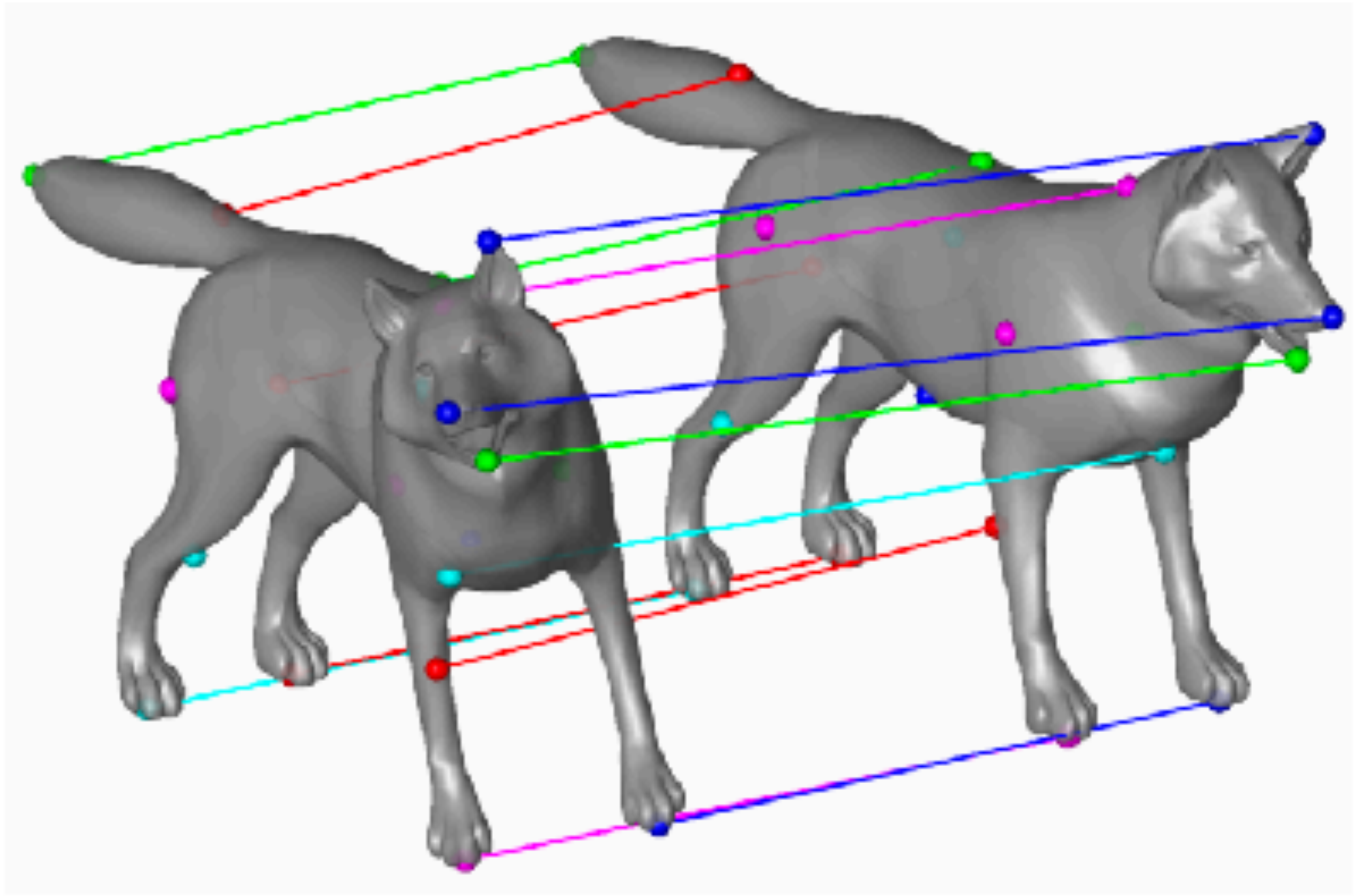


Spectral shape matching

$s(x) = (\varphi_1(x), \varphi_2(x), \ldots, \varphi_N(x))$ for vertex x





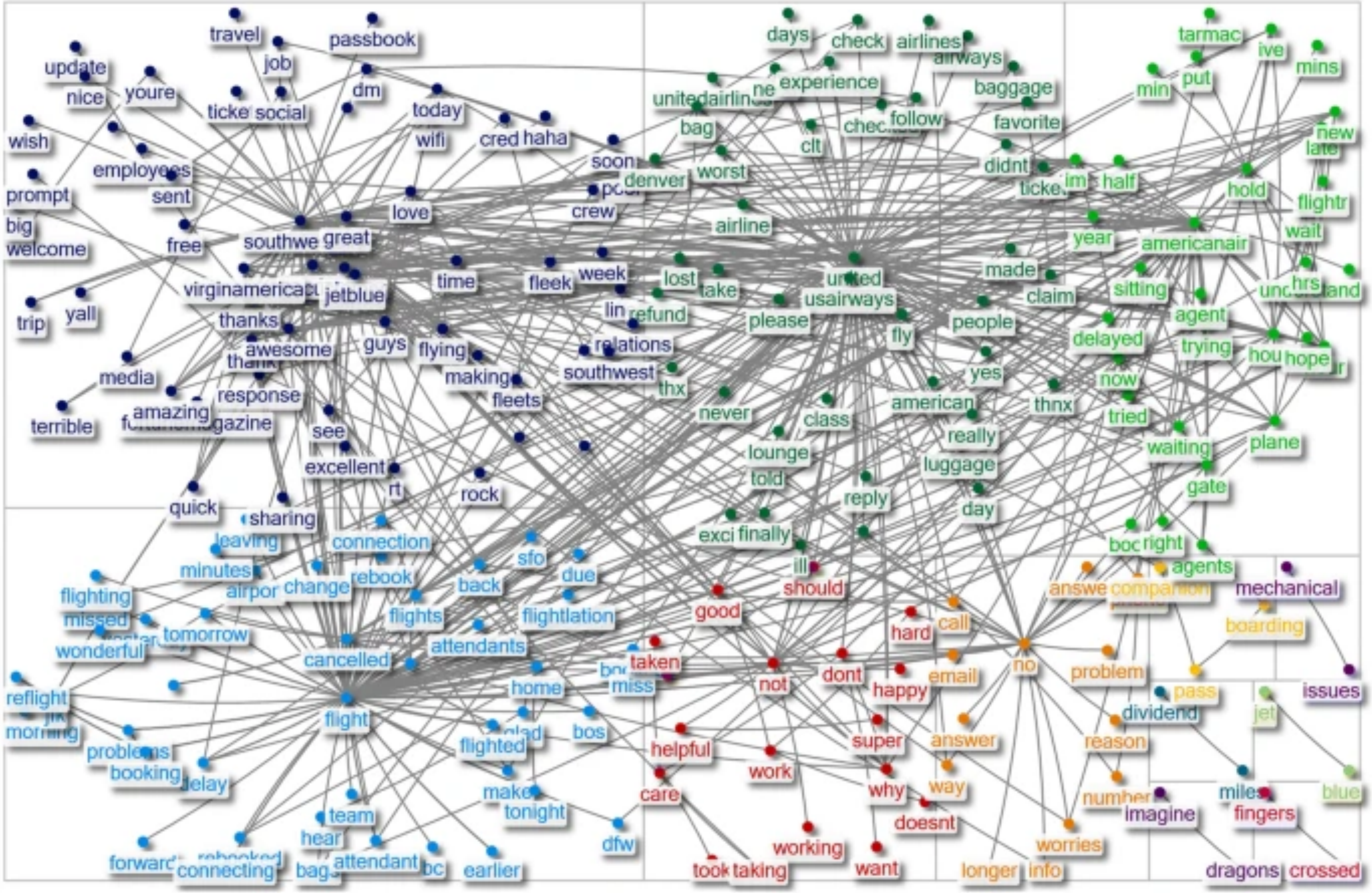


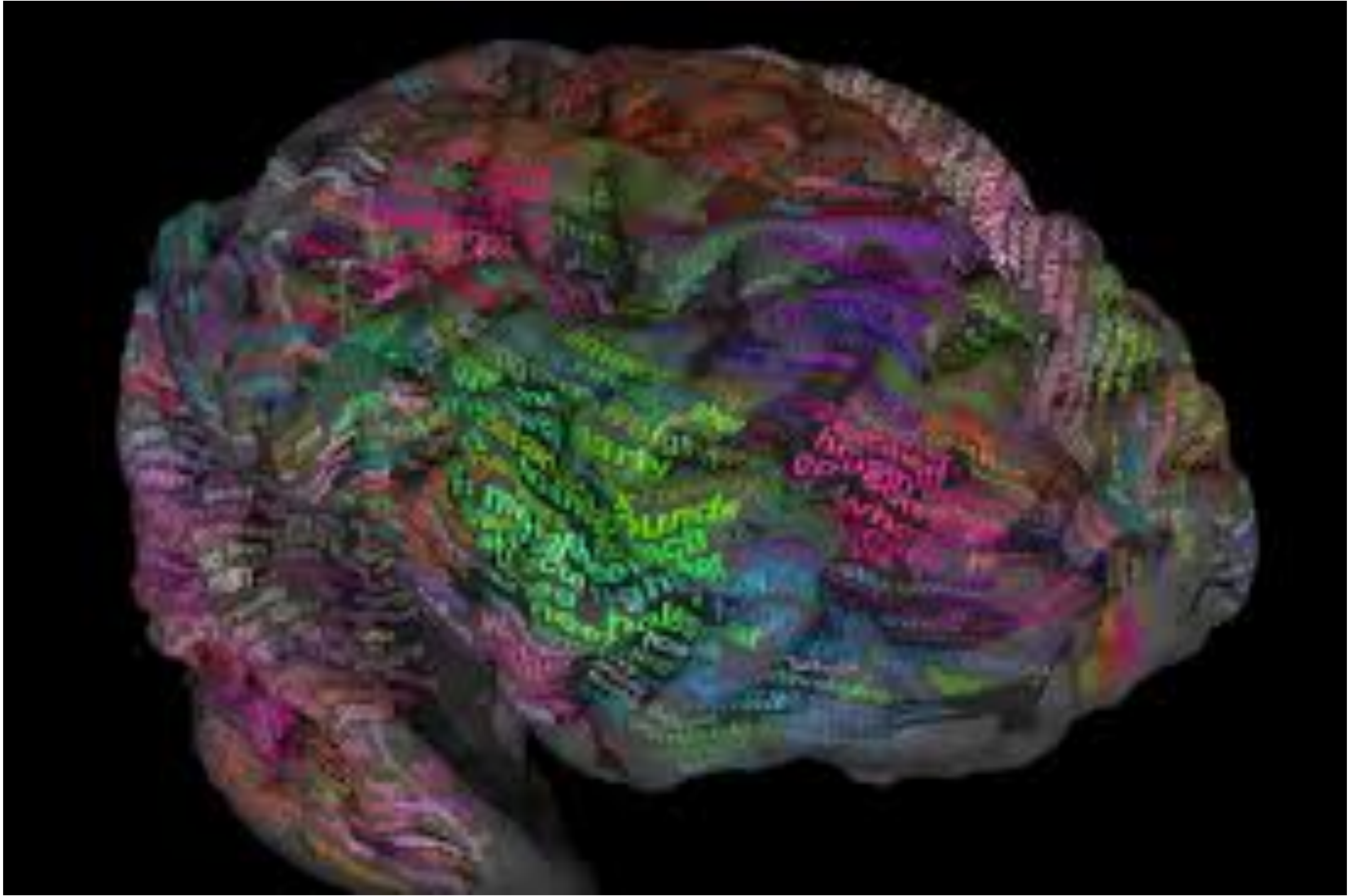
- Low-dimensional representations based on connectivity

- Diffusion distance shows invariance to non-stretching or non-topology changing transformations

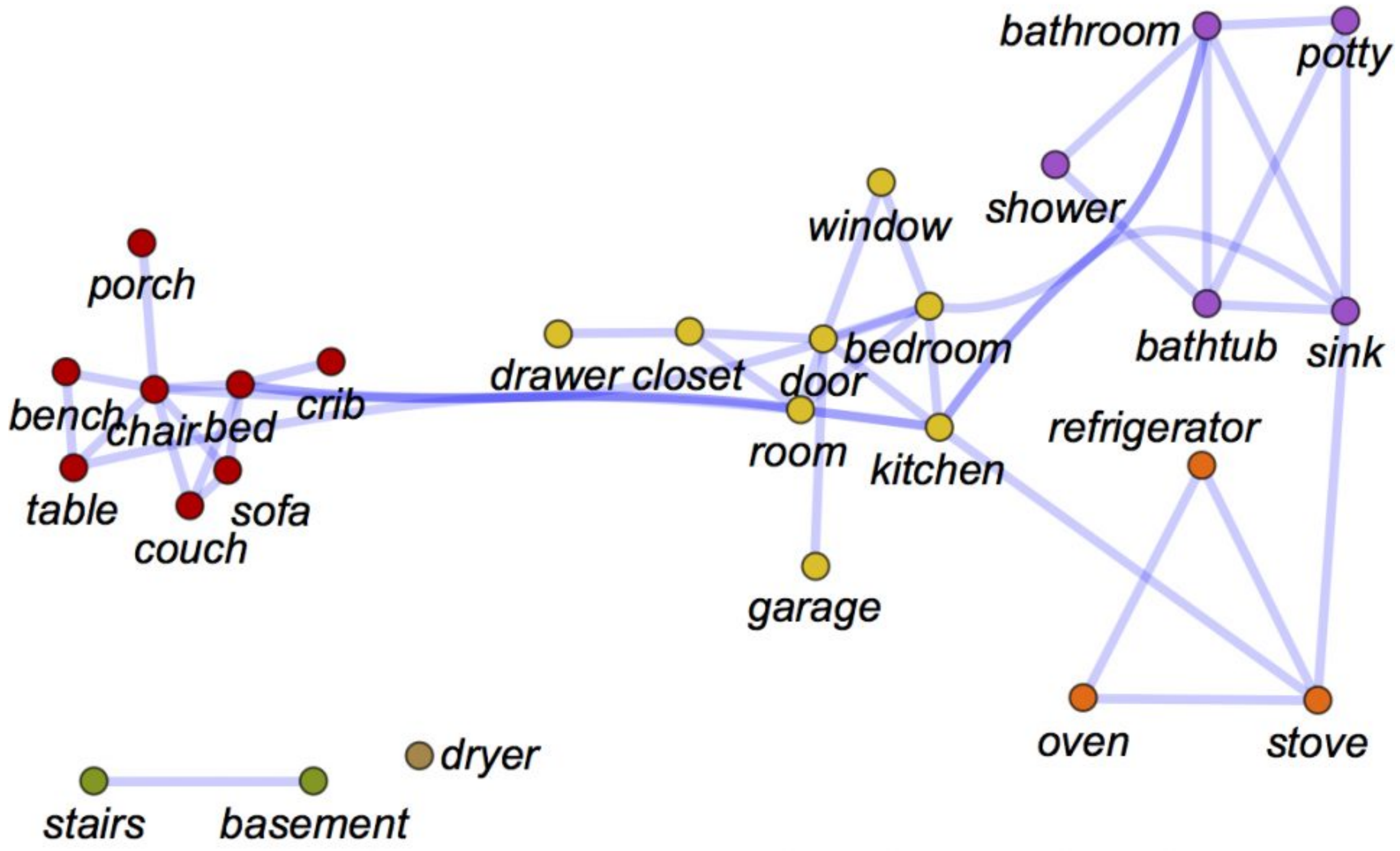
- Correspondence problem

- Analogy-based inferences?


travel
passbook
update
job
nice
youre
dm
social
today
wish
wifi
cred
haha
employees
soon
prompt
sent
love
crew
big
welcome
free
great
time
fleek
week
virginamerica
jetblue
trip
yall
thanks
lin
relations
awesome
guys
flying
making
southwest
media
response
fleets
terrible
amazing
see
excellent
rt
rock
quick
sharing
leaving
connection
minutes
flighting
missed
change
rebook
sfo
due
back
flights
flightlation
tomorrow
wonderful
cancelled
attendants
home
reflight
miss
morning
flight
bos
problems
flighted
booking
delay
make
tonight
team
hear
forward
connecting
attendant
bc
earlier
dfw
days
check
airlines
airways
experience
baggage
unitedairlines
follow
favorite
bag
clt
worst
denver
didnt
half
airline
year
americanair
hold
flightr
wait
lost
take
united
usairways
made
claim
sitting
refund
please
people
agent
fly
delayed
trying
hope
thx
yes
now
never
class
american
thnx
tried
really
waiting
plane
lounge
luggage
told
reply
day
gate
finally
right
agents
should
good
mechanical
call
boarding
hard
taken
not
dont
email
no
problem
happy
pass
dividend
issues
jet
helpful
super
answer
reason
work
why
way
care
number
miles
blue
imagine
fingers
doesnt
working
worries
took
taking
want
longer
info
dragons
crossed
tarmac
ive
mins
min
put
new
late



Semantic networks

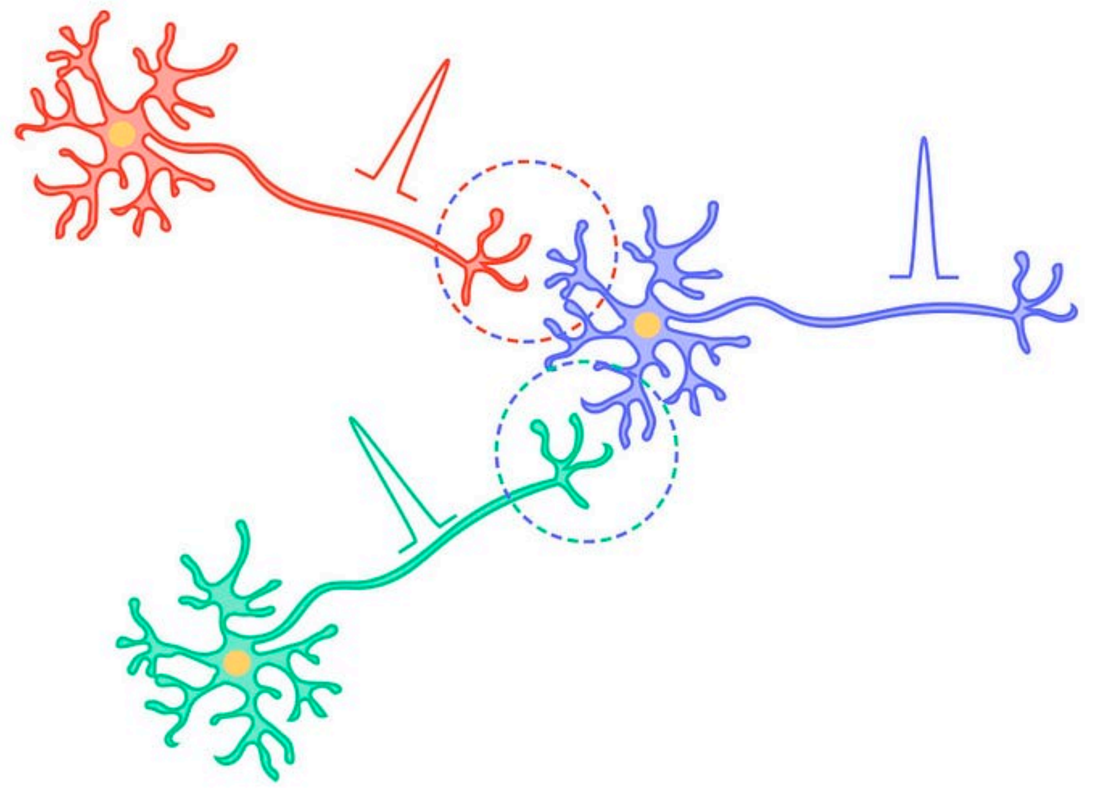


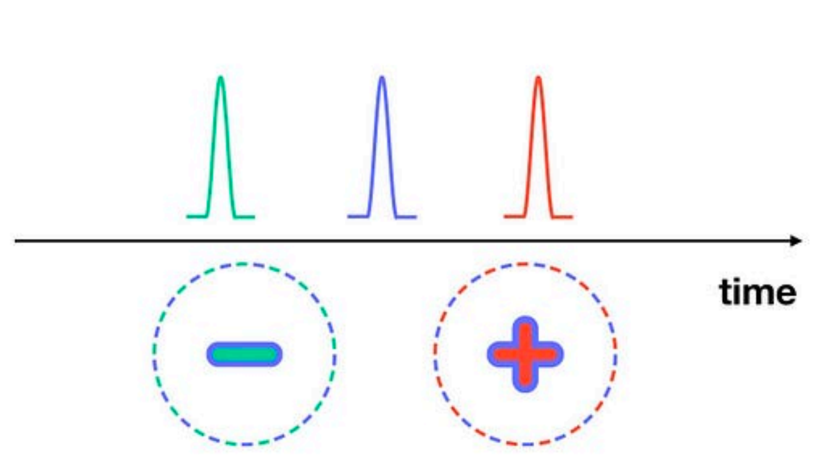
Associative learning

$s, s' \quad s, s' \quad s, s'$ $\qquad$ $W_{s,s'} \uparrow \quad W_{s',s} \downarrow$

Temporal sequences
$s, s', s'', s''', \cdots$

Sculpting of neural circuits
by weight changes





Hebbian learning (e.g. STDP)

Search in semantic networks: Remote Associates Test

Which word relates to all three cue words?

Word 1	Word 2	Word 3	Answer
man	glue	star	
dew	comb	bee	
rain	test	stomach	

Search in semantic networks: Remote Associates Test

Which word relates to all three cue words?

Word 1	Word 2	Word 3	Answer
man	glue	star	super
dew	comb	bee	
rain	test	stomach	

Search in semantic networks: Remote Associates Test

Which word relates to all three cue words?

Word 1	Word 2	Word 3	Answer
man	glue	star	super
dew	comb	bee	honey
rain	test	stomach	

Search in semantic networks: Remote Associates Test

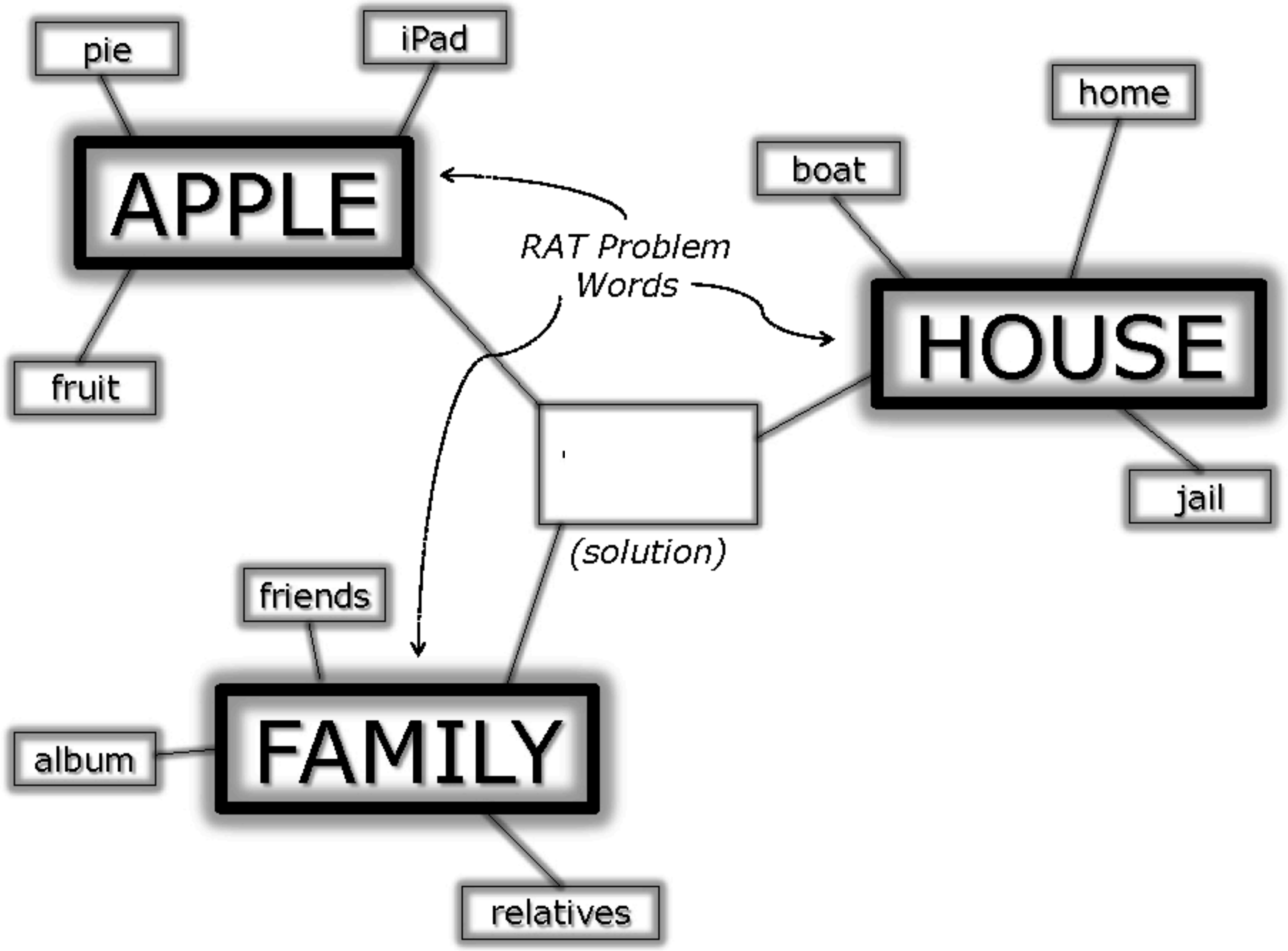
Which word relates to all three cue words?

Word 1	Word 2	Word 3	Answer
man	glue	star	super
dew	comb	bee	honey
rain	test	stomach	acid

Search in semantic networks: Remote Associates Test

Which word relates to all three cue words?

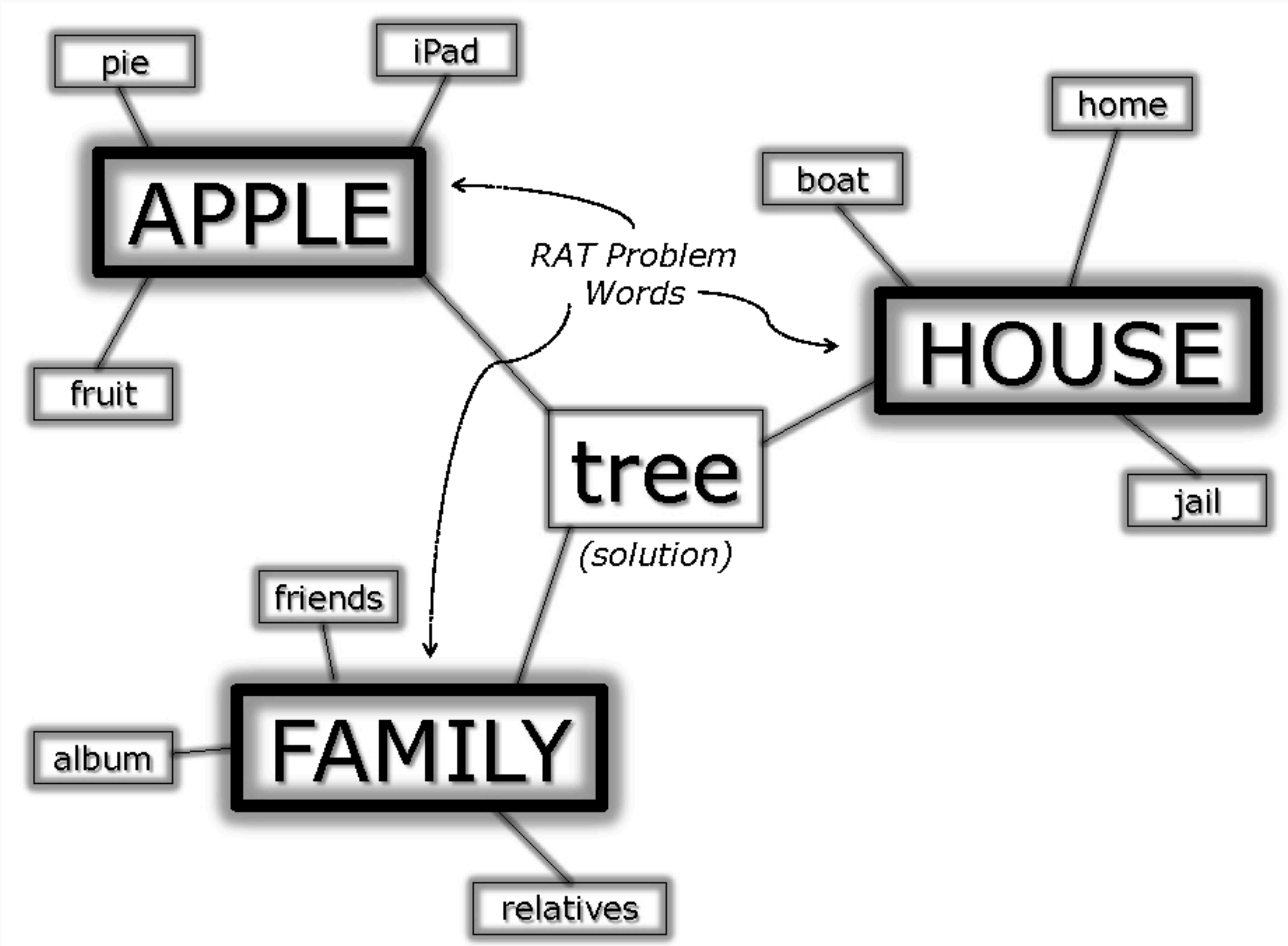
Word 1	Word 2	Word 3	Answer
man	glue	star	super
dew	comb	bee	honey
rain	test	stomach	acid



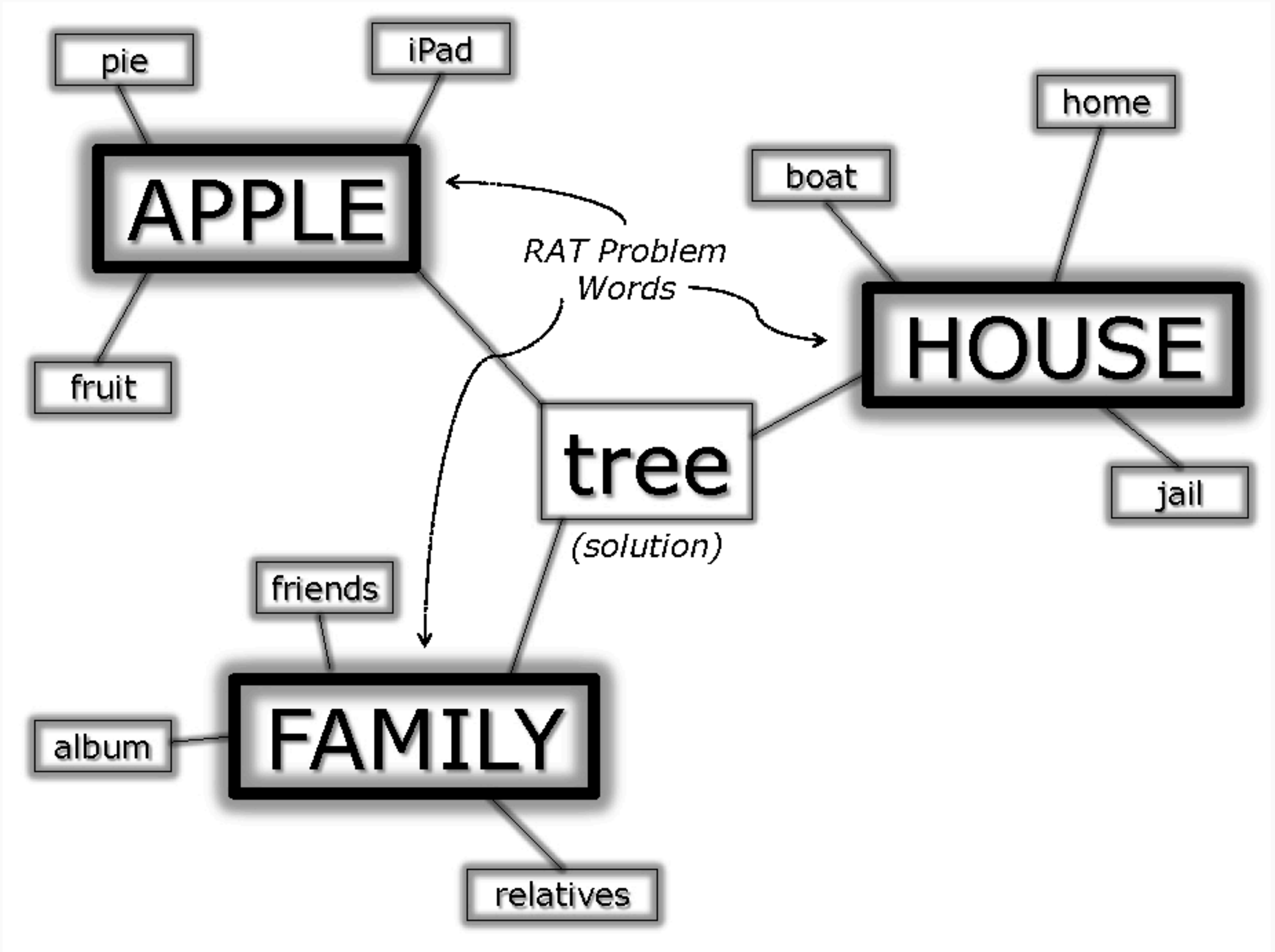
Search in semantic networks: Remote Associates Test

Which word relates to all three cue words?

Word 1	Word 2	Word 3	Answer
man	glue	star	super
dew	comb	bee	honey
rain	test	stomach	acid



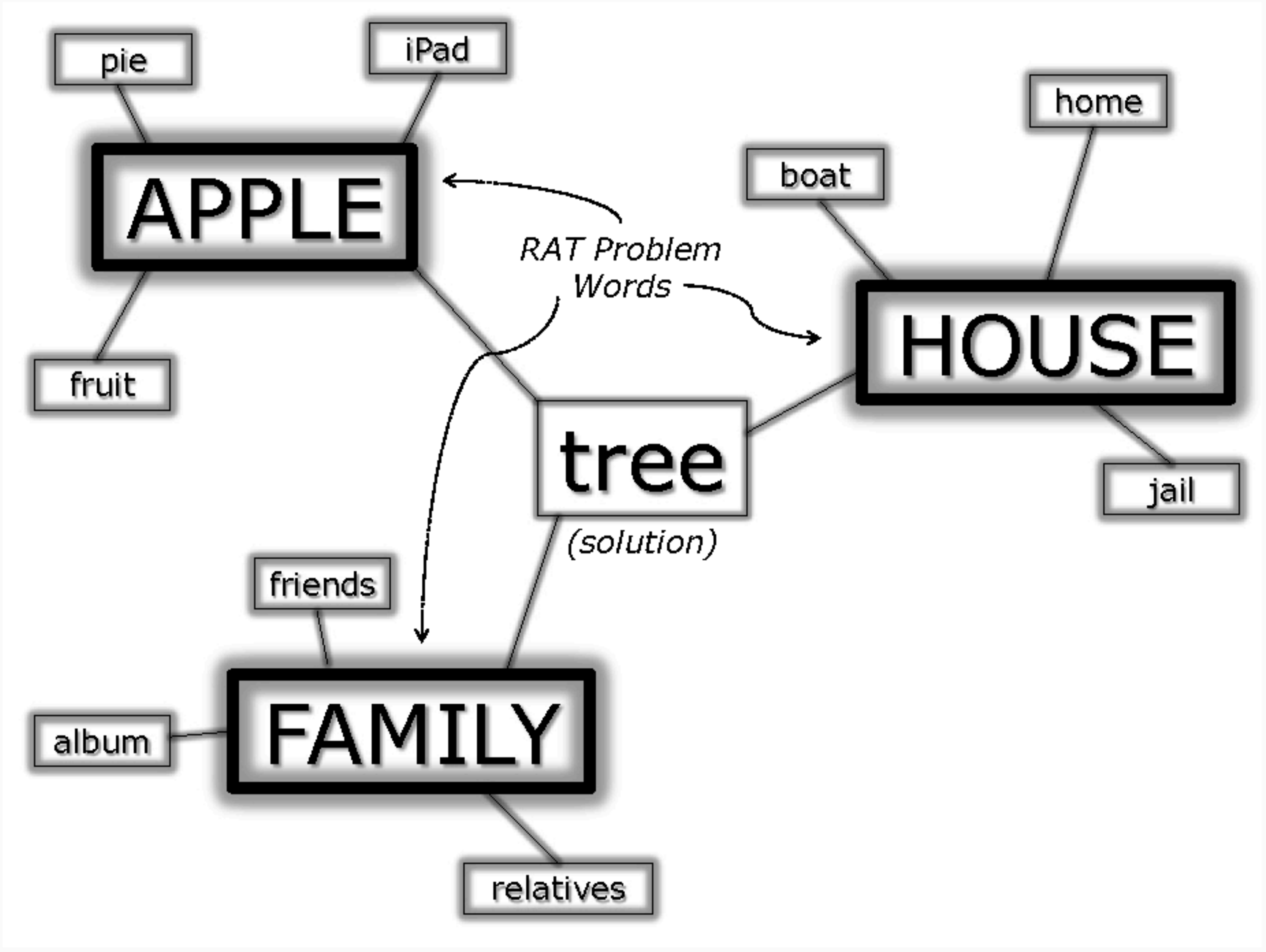
RAT: how does the brain solves it?



Candidates: common nearest neighbor to all three cue words

Constraints from humans: local search, memory constraints, difficulty

RAT: how does the brain solves it?

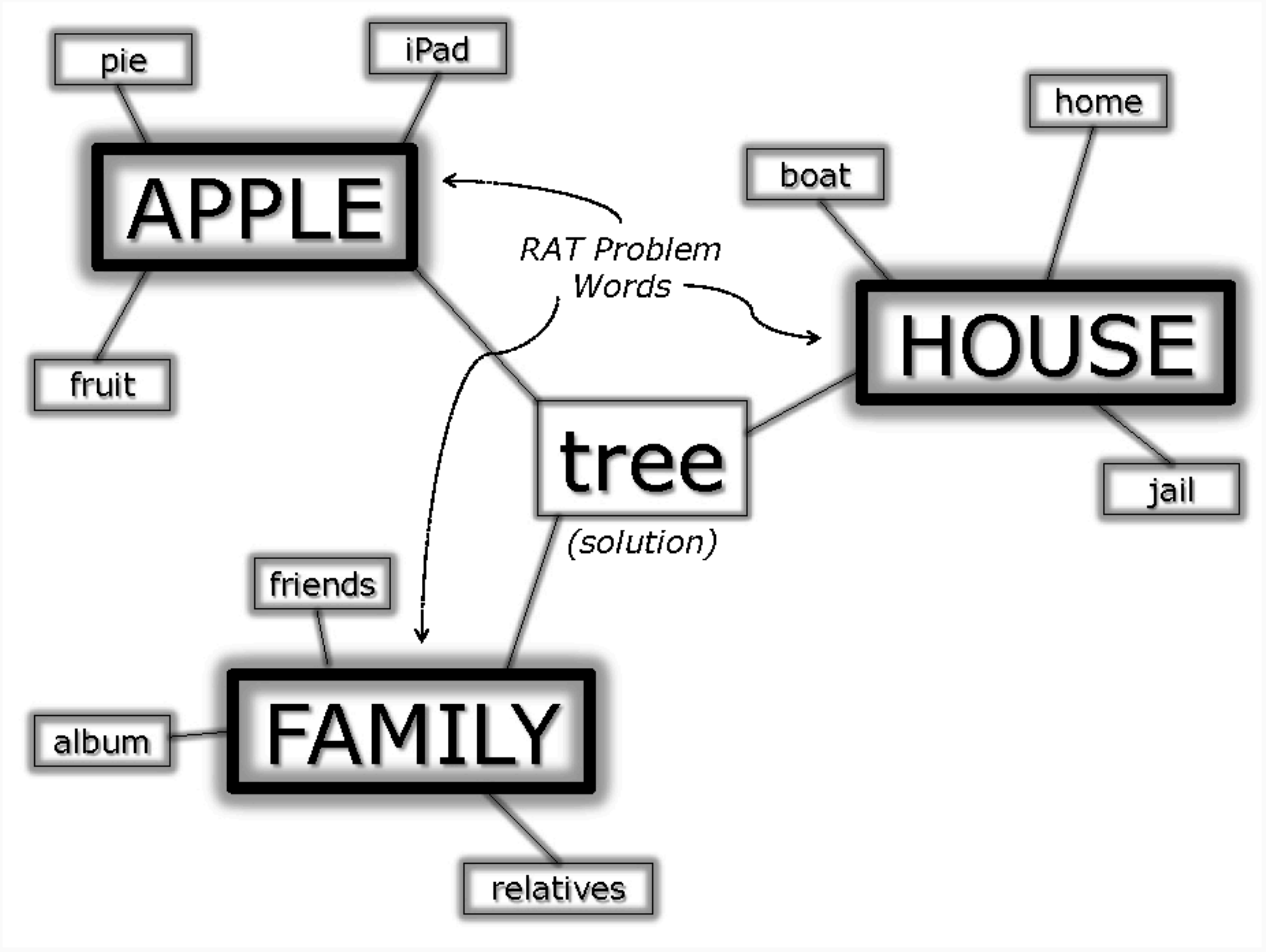


Candidates: common nearest neighbor to all three cue words

Constraints from humans: local search, memory constraints, difficulty

Naive algorithm 1. Launch random walks from 3 initial cue words
Stopping criteria: 1st commonly visited word

RAT: how does the brain solves it?



Candidates: common nearest neighbor to all three cue words

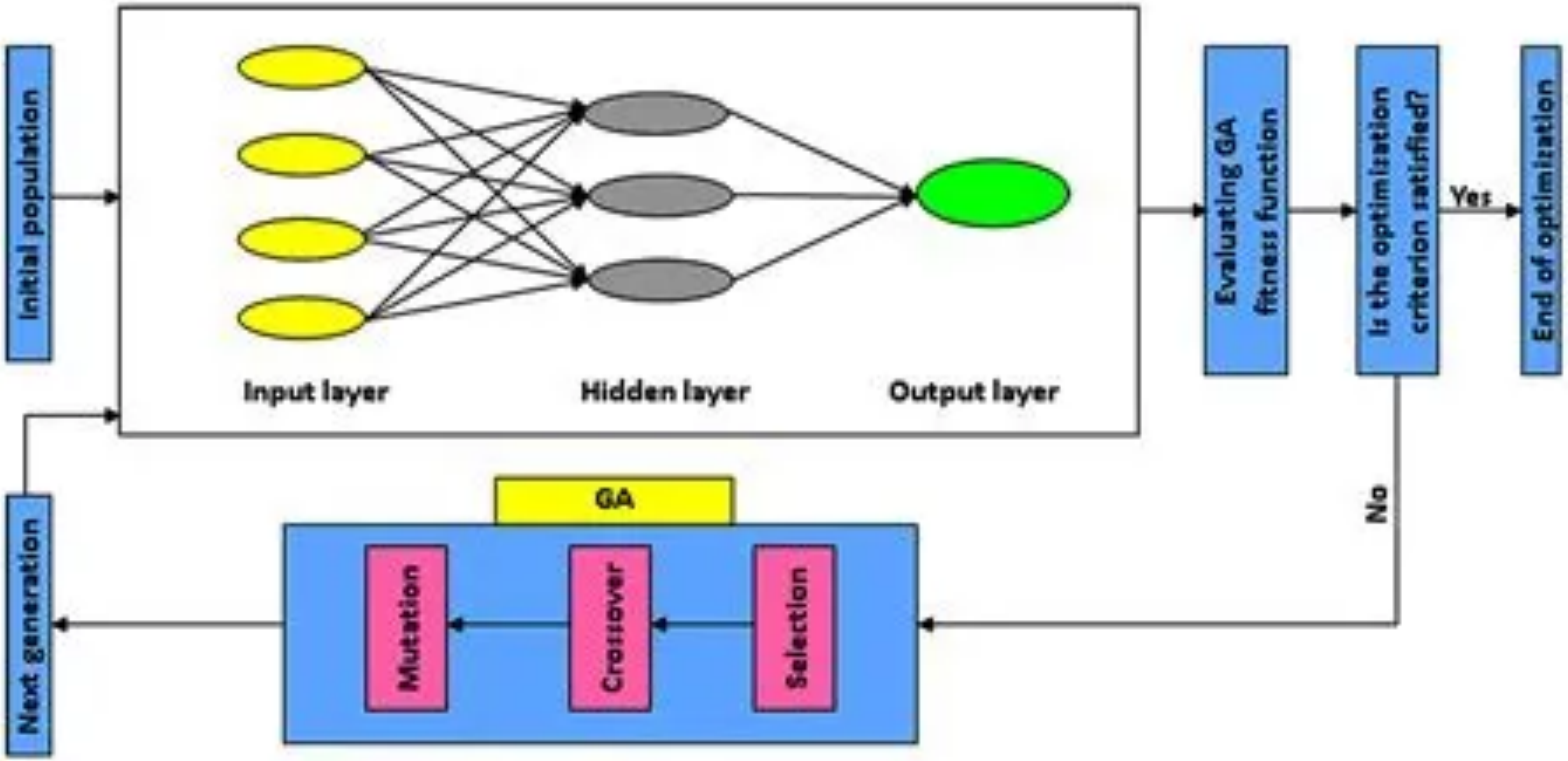
Constraints from humans: local search, memory constraints, difficulty

Naive algorithm 1. Launch random walks from 3 initial cue words
Stopping criteria: 1st commonly visited word

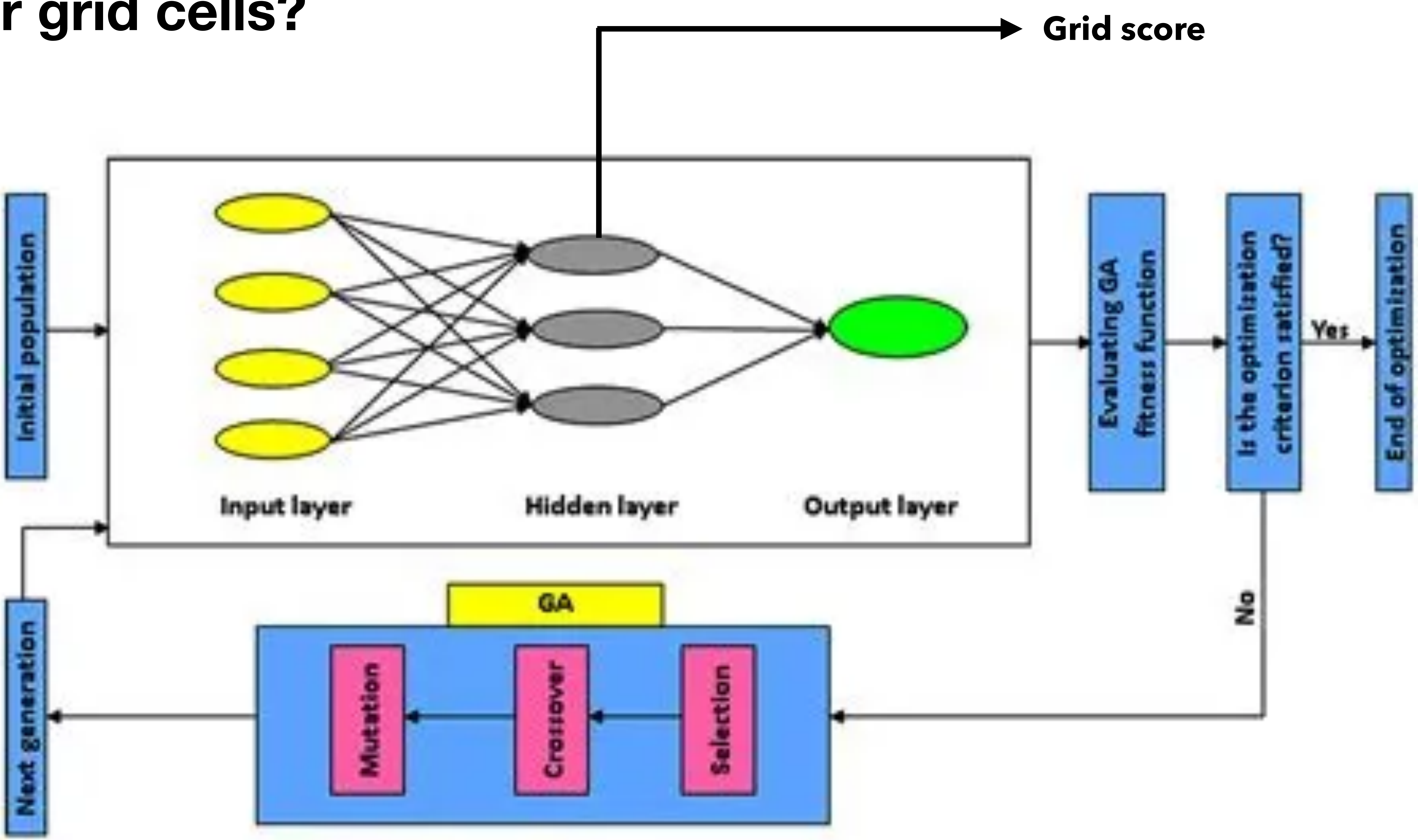
Naive algorithm 2. Represent each word with spectral coordinates
Output: Word with MinAvg Euclidean distance to the 3 cues

Functional role of grid cells?
More tasks involving grid cells…

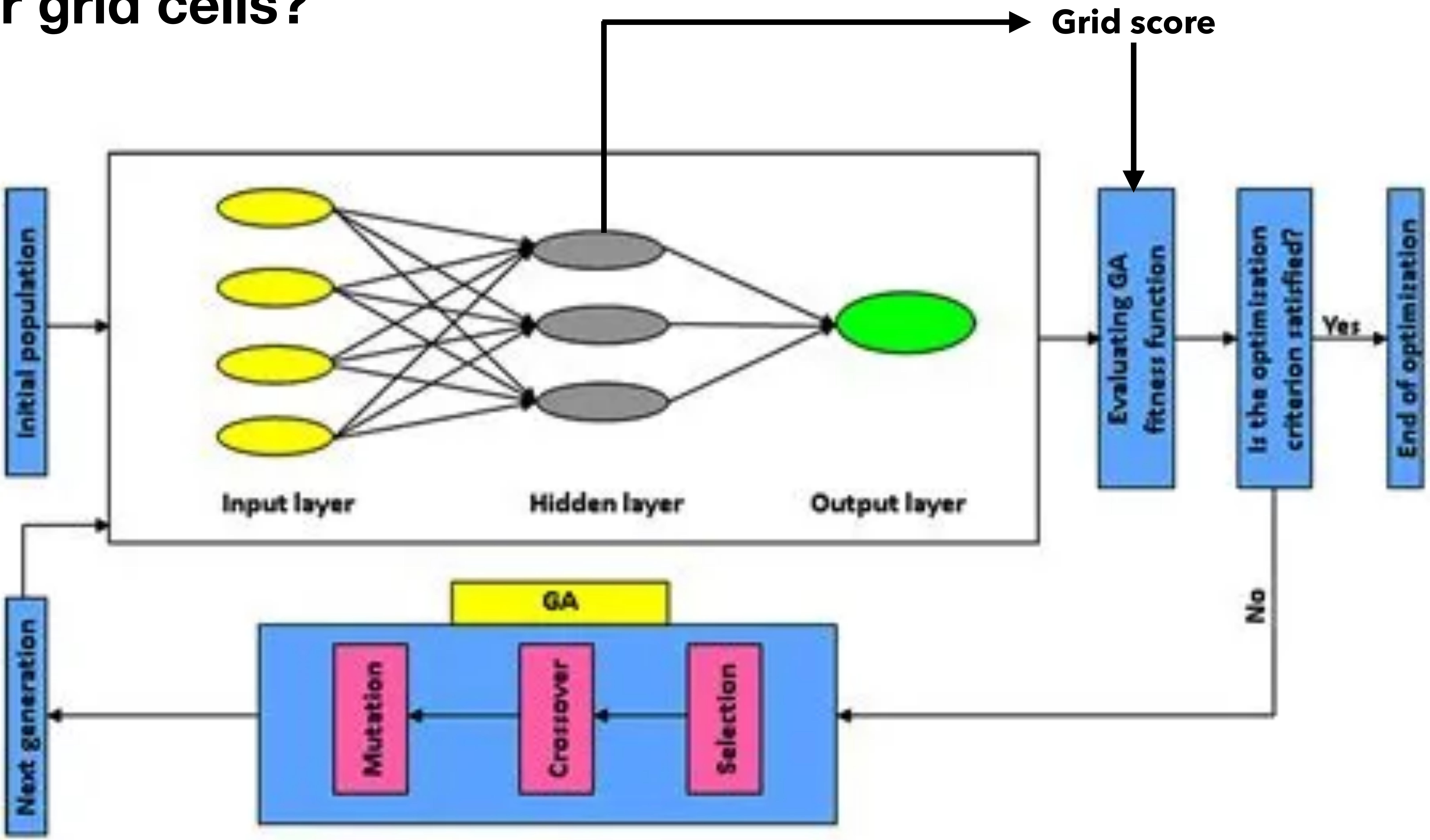
GP for grid cells?

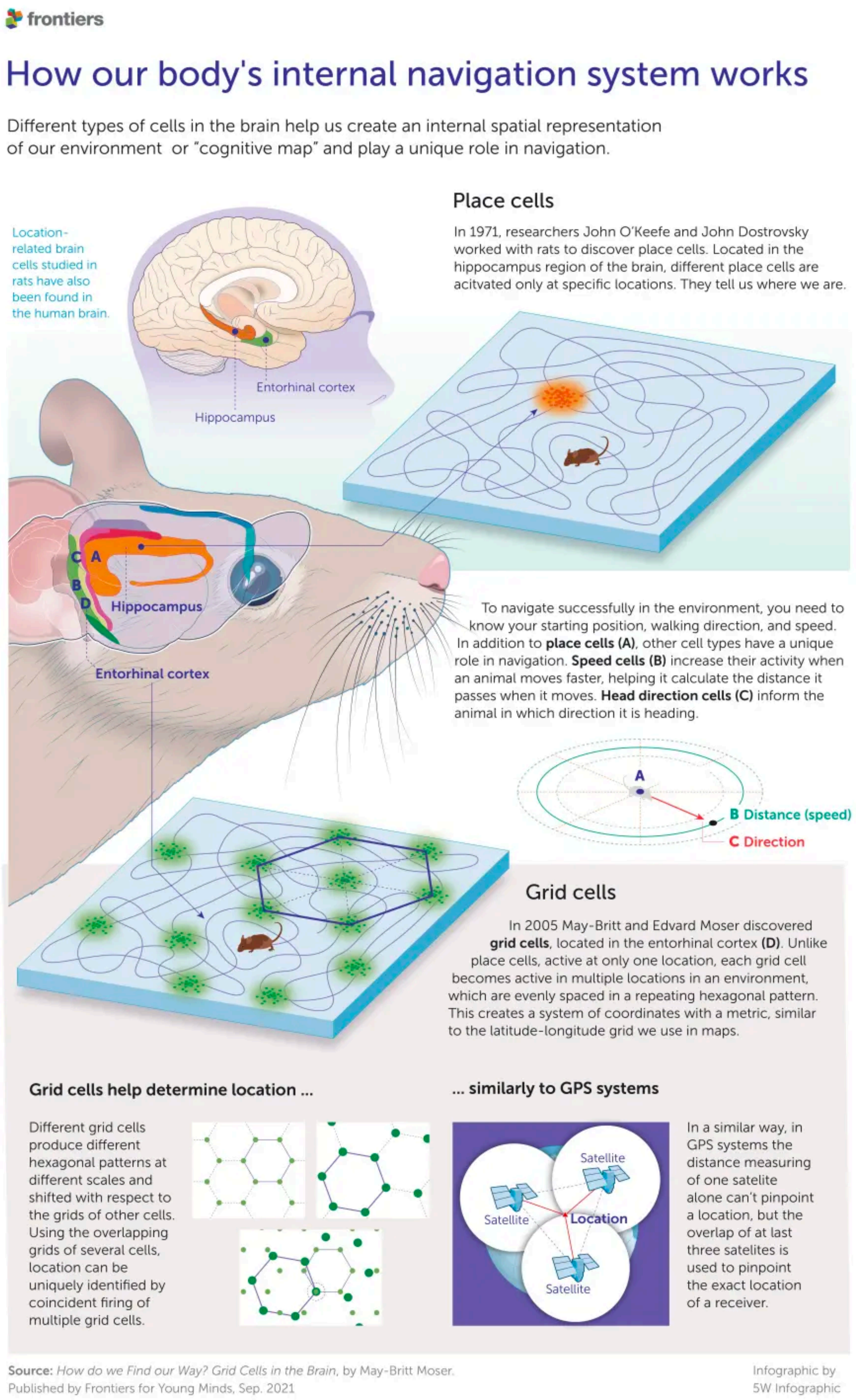


GP for grid cells?



GP for grid cells?





- **Grid cells periodic pattern may** allow the brain to form cognitive maps in physical space (coordinate system for distance and orientation in space; how things are related to each other independently of the observer position)

- They are thought to allow us to engage **mental navigation** in abstract spaces

- Did the brain discovered Fourier analysis?