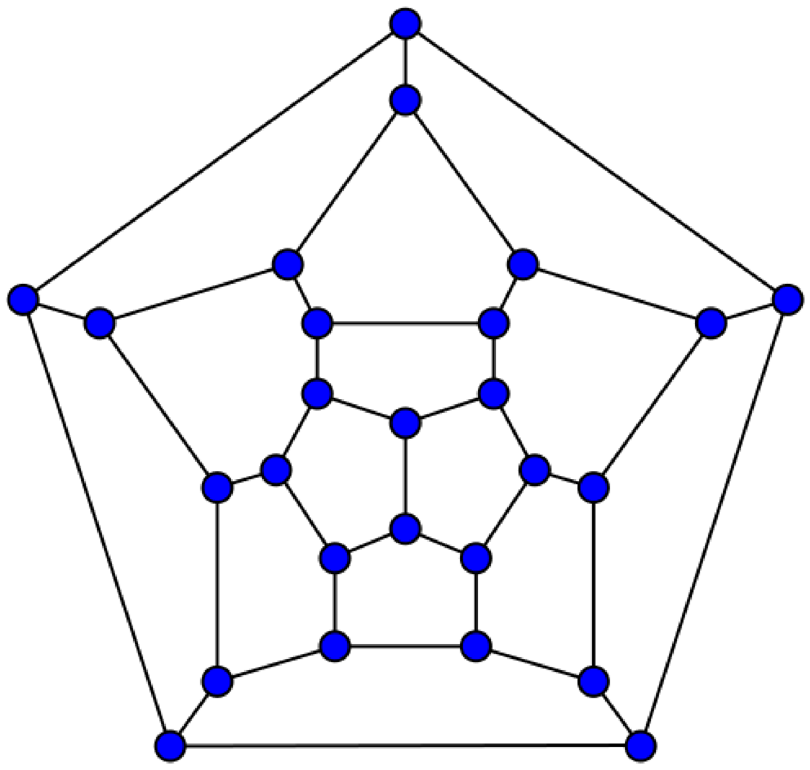
Introduction to Graph Spectral Analysis

Umberto Castellani, Simone Melzi, Riccardo Marin

Tuesday 28th July 2020
Verona, Italy

Context

Graphs



Eigenvalues & Eigenvectors

$$\boxed{A v = \lambda v}$$

where $A \in \mathbb{R}^{m \times m}$ (Square Matrix)

eigenvectors $\longrightarrow$ $v \in \mathbb{R}^{m \times 1}$ (Column Vector)

eigenvalues $\longrightarrow$ $\lambda \in \mathbb{R}^{m \times m}$ (Diagonal Matrix)

Context

1973

W. E. Donath
A. J. Hoffman

Lower Bounds for the Partitioning of Graphs

Abstract: Let a k-partition of a graph be a division of the vertices into k disjoint subsets containing $m_1 \geq m_2 \cdots \geq m_k$ vertices. Let E_c be the number of edges whose two vertices belong to different subsets. Let $\lambda_1 \geq \lambda_2 \cdots \geq \lambda_k$ be the k largest eigenvalues of a matrix, which is the sum of the adjacency matrix of the graph plus any diagonal matrix U such that the sum of all the elements of the sum matrix is zero. Then

$$E_c \geq \frac{1}{2} \sum_{r=1}^{k} -m_r \lambda_r.$$

A theorem is given that shows the effect of the maximum degree of any node being limited, and it is also shown that the right-hand side is a concave function of U. Computational studies are made of the ratio of upper bound to lower bound for the two-partition of a number of random graphs having up to 100 nodes.

1973

Czechoslovak Mathematical Journal, 23 (98) 1973, Praha

ALGEBRAIC CONNECTIVITY OF GRAPHS*)

MIROSLAV FIEDLER, Praha

(Received April 14, 1972)

1. INTRODUCTION

Let $G = (V, E)$ be a non-directed finite graph without loops and multiple edges. Having chosen a fixed ordering $w_1, w_2, \ldots, w_n$ of the set V, we can ... n-rowed matrix $A(G)$ whose off-diagonal entries are $a_{ik} = a_{ki} = -1$ ... and $a_{ik} = 0$ otherwise and whose diagonal entries a_{ii} are equal to the ... the vertices w_i. This matrix $A(G)$, which is frequently used to enumerate ... trees of the graph G, is symmetric, singular (all the row sums are zero) ... semidefinite ($A(G) = UU^T$ where U is the $(0, 1, -1)$ vertex-edge adj... of arbitrarily directed graph G). Let $n \geq 2$ and $0 = \lambda_1 \leq \lambda_2 = a(G) \ldots \leq \lambda_n$ be the eigenvalues of the matrix $A(G)$. From the Perron-Frobenius ... applied to the matrix $(n-1) I - A(G)$ it follows that $a(G)$ is zero if and only if the graph G is not connected. We shall call the second smallest eigenvalue $a(G)$ of the matrix $A(G)$ algebraic connectivity of the graph G. It is the purpose of this paper to find its relation to the usual vertex and edge connectivities.

We recall that many authors, e.g. A. J. HOFFMAN, M. DOOB, D. K. RAY-CHAUDHURI, J. J. SEIDEL have characterized graphs by means of the spectra of the (0, 1) and (0, 1, −1) adjacency matrices.

2. NOTATION AND CONVENTIONS

The notation introduced above is used throughout the present paper. All matrices and vectors are considered real. The transpose of a matrix M is denoted by M^T, the identity matrix by I, the vector $(1, 1, \ldots, 1)^T$ by e, the universal matrix ee^T by J, the cardinality of a set S by $|S|$.

For our purpose it is convenient to denote by W the set of all column vectors x such that $x^T x = 1$, $x^T e = 0$. Any square matrix M with all zero row sums has an

*) Presented at the Graph Theory Meeting in Zlatá Idka, May 1971.

298

2000

IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 22, NO. 8, AUGUST 2000

Normalized Cuts and Image Segmentation

Jianbo Shi and Jitendra Malik, *Member, IEEE*

Abstract—We propose a novel approach for solving the perceptual grouping problem in vision. Rather than focusing on local features and their consistencies in the image data, our approach aims at extracting the global impression of an image. We treat image segmentation as a graph partitioning problem and propose a novel global criterion, the *normalized cut*, for segmenting the graph. The *normalized cut* criterion measures both the total dissimilarity between the different groups as well as the total similarity within the groups. We show that an efficient computational technique based on a generalized eigenvalue problem can be used to optimize this criterion. We have applied this approach to segmenting static images, as well as motion sequences, and found the results to be very encouraging.

Index Terms—Grouping, image segmentation, graph partitioning.

2018

SPECTRALNET: SPECTRAL CLUSTERING USING DEEP NEURAL NETWORKS

Uri Shaham, Kelly Stanton, Henry Li
Yale University
New Haven, CT, USA

Boaz Nadler, Ronen Basri
Weizmann Institute of Science
Rehovot, Israel

arXiv:1801.01587v6 [stat.ML] 4 Apr 2018

2014

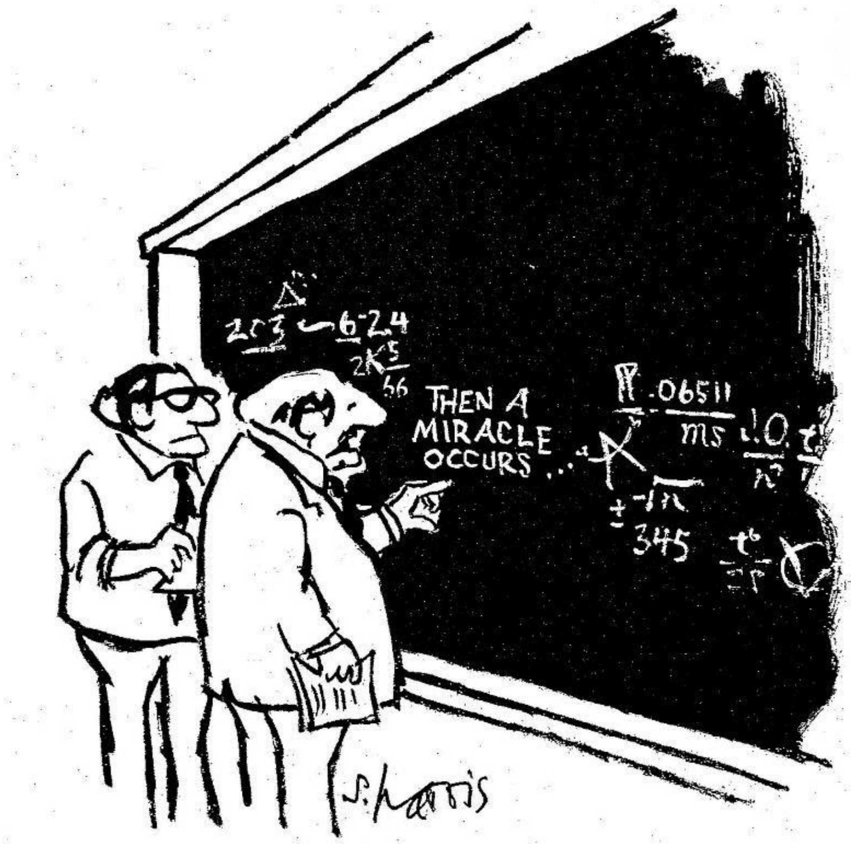
Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence

Learning Deep Representations for Graph Clustering

Fei Tian, University of Science and Technology of China
Bin Gao, Microsoft Research
Qing Cui, Tsinghua University
Enhong Chen, University of Science and Technology of China
Tie-Yan Liu, Microsoft Research

Index

- Basic concepts (linear algebra, function analysis)
- Dirichlet energy and smoothness
- Laplacian
- Graph Laplacian
- Eigen-decomposition of the Laplacian
- Spectral properties
- Applications (mincut, clustering, Fourier-like analysis)



"I THINK YOU SHOULD BE MORE EXPLICIT HERE IN STEP TWO."

Basis

Given a vector space $\mathbf{V}$, a subset $\mathbf{B}$ is a *basis* iif*:*

- Its elements are linear indipendent
- And they span all the vectors in $\mathbf{V}$

The canonical base for $\mathbf{v} \in \mathbb{R}^3$

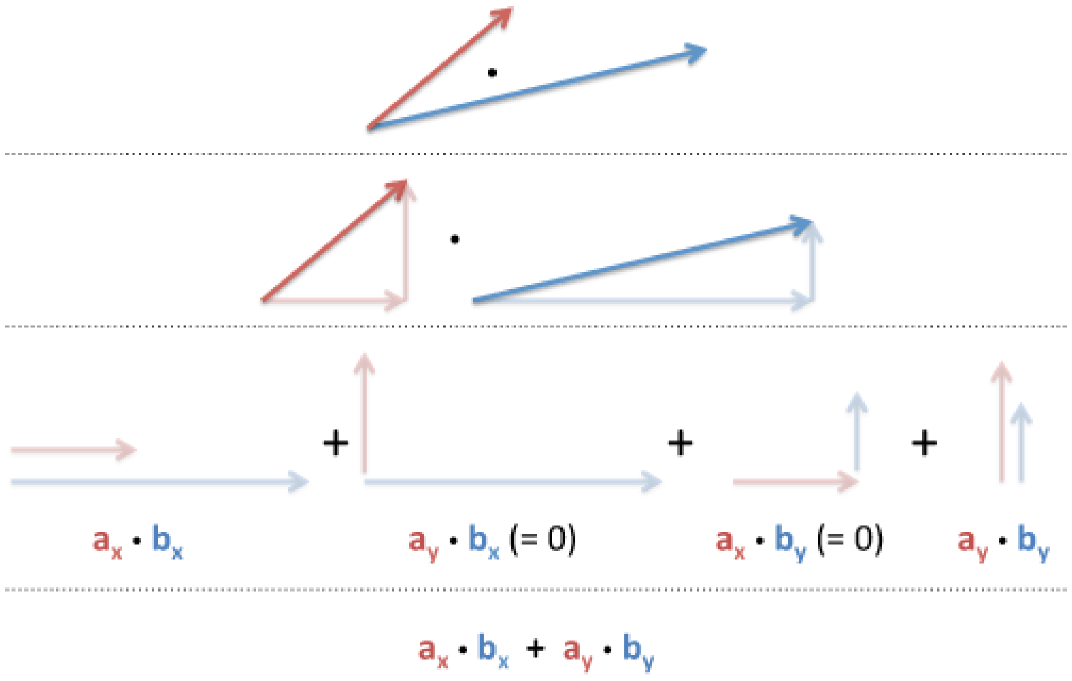
$$\alpha_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + \alpha_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \mathbf{v}$$

Every $\mathbf{v}$ can be written as a weighted sum of the elements of the basis.

Basis - Properties

Defining an inner product:

$$\langle \mathbf{a}, \mathbf{b} \rangle = a_0 b_0 + a_1 b_1 + \cdots + a_n b_n = \sum_{i=0}^{n} a_i b_i$$



It is useful to see it as a projection operation

We say our basis $\mathbf{B} = \{\mathbf{b}_0, \mathbf{b}_1, \ldots, \mathbf{b}_n\}$ is:

Orthogonal → $\forall i, j \in [0, n]$, if $i \neq j$ then $\langle \mathbf{b}_i, \mathbf{b}_j \rangle = 0$ → Projection gives nothing

Normal → $\forall i, ||\mathbf{b}_i||_2 = \sqrt{\langle \mathbf{b}_i, \mathbf{b}_i \rangle} = 1$ → Projection preserves length

[1] https://betterexplained.com/articles/vector-calculus-understanding-the-dot-product/

Basis - Coefficients

$$\mathbf{v} = \mathbf{b}_1 c_1 + \mathbf{b}_2 c_2 + \mathbf{b}_3 c_3 + \dots$$

$$\mathbf{v} \cdot \mathbf{b}_2 = (\mathbf{b}_1 c_1 + \mathbf{b}_2 c_2 + \mathbf{b}_3 c_3 + \dots) \cdot \mathbf{b}_2$$

$$\frac{\mathbf{v} \cdot \mathbf{b}_2 - \mathbf{b}_2 \mathbf{b}_1 c_1 - \mathbf{b}_2 \mathbf{b}_3 c_3 - \dots}{||\mathbf{b}_2||_2} = c_2$$

If $\mathbf{B}$ is orthonormal $\Rightarrow$ $\langle \mathbf{b}_i, \mathbf{b}_j \rangle = 0$, $||\mathbf{b}_2||_2 = 1$

$$\implies \mathbf{v} \cdot \mathbf{b}_2 = c_2$$

Each coefficent is indipendent and comes from a multiplication.

[1]: Fourier Analysis for Beginners, L.N. Thibos, 2017,https://scholarworks.iu.edu/dspace/bitstream/handle/2022/21365/Fourier%20analysis%202014.pdf?sequence=1

Basis - Functions

For functions:

$$f(x) = \alpha_1 g_1(x) + \alpha_2 g_2(x) + \alpha_3 g_3(x) + \dots$$

Weirestrass Theorem (of approximation):

Suppose f is a continuous real-valued function defined on the real interval $[a, b]$. For every $\epsilon > 0$, there exists a polynomial p such that for all x in $[a, b]$, we have $|f(x) - p(x)| < \epsilon$

i.e., Given a function, there exist a polynomial arbitrary close to it.

Monomials are a basis for the functional space:

$$G(x) = \{g_i(x)\} = \{x^i\} = \{1, x, x^2, x^3, \dots\}$$

Basis – Coefficients for functions

$$G(x) = \{g_i(x)\} = \{x^i\} = \{1, x, x^2, x^3, \ldots\}$$

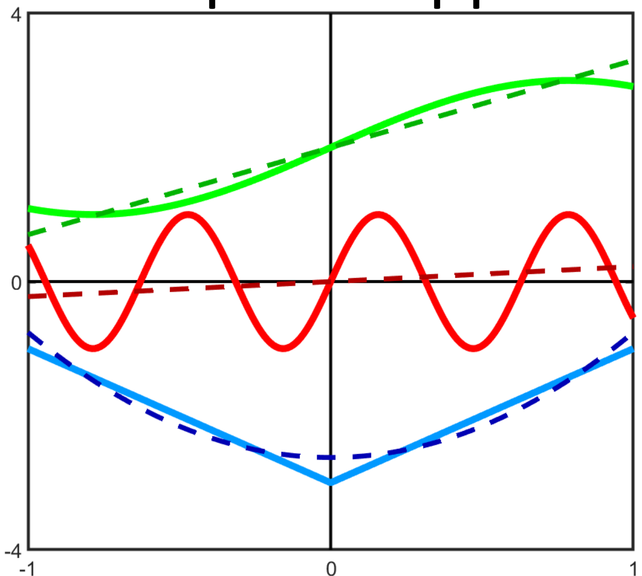
We use polynomial interpolation by sampling the function:

$$\begin{cases} y_1 = \alpha_1 g_1(x_1) + \alpha_2 g_2(x_1) + \alpha_3 g_3(x_1) + \ldots \\ y_2 = \alpha_1 g_1(x_2) + \alpha_2 g_2(x_2) + \alpha_3 g_3(x_2) + \ldots \\ y_3 = \alpha_1 g_1(x_3) + \alpha_2 g_2(x_3) + \alpha_3 g_3(x_3) + \ldots \\ \ldots \end{cases} \Rightarrow \begin{cases} y_1 = \alpha_1 \cdot 1 + \alpha_2 \cdot x_1 + \alpha_3 \cdot x_1^2 + \ldots \\ y_2 = \alpha_1 \cdot 1 + \alpha_2 \cdot x_2 + \alpha_3 \cdot x_2^2 + \ldots \\ y_3 = \alpha_1 \cdot 1 + \alpha_2 \cdot x_3 + \alpha_3 \cdot x_3^2 + \ldots \\ \ldots \end{cases}$$

$$\begin{bmatrix} y1 \\ y2 \\ y3 \\ \ldots \end{bmatrix} = \begin{bmatrix} 1 & x_1 & x_1^2 & \ldots \\ 1 & x_2 & x_2^2 & \ldots \\ 1 & x_3 & x_3^2 & \ldots \\ \ldots & \ldots & \ldots & \ldots \end{bmatrix} \cdot \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \ldots \end{bmatrix}$$

$$\mathbf{y} = \mathbf{Ga} \quad \text{with } \mathbf{a} \text{ unknown} \qquad \Longrightarrow \mathbf{G}^{-1}\mathbf{y} = \mathbf{a}$$

Example of approx



Coefficients are informative on the function behavior

Basis – Coefficients for functions (infinite sampling)

Considering infinite sampling

$$\mathbf{x} = \begin{bmatrix} -1 \\ -1+dx \\ -1+2dx \\ \dots \\ 1-dx \\ 1 \end{bmatrix} \quad \mathbf{y} = \begin{bmatrix} f(-1) \\ f(-1+dx) \\ f(-1+2dx) \\ \dots \\ f(1-dx) \\ f(1) \end{bmatrix}$$

Defining an inner product

$$\langle \mathbf{f}, \mathbf{g} \rangle = \begin{bmatrix} f(-1) \\ f(-1+dx) \\ f(-1+2dx) \\ \dots \\ f(1-dx) \\ f(1) \end{bmatrix} \cdot \begin{bmatrix} g(-1) \\ g(-1+dx) \\ g(-1+2dx) \\ \dots \\ g(1-dx) \\ g(1) \end{bmatrix}$$

$$= f(-1)g(-1) + f(-1+dx)g(-1+dx) + \cdots + f(1)g(1)$$

$$= \int_{-1}^{1} f(x)g(x)dx$$

Basis – Properties for functions

$$\langle \mathbf{f}, \mathbf{g} \rangle = \boxed{\int_{-1}^{1} f(x)g(x)dx}$$

	Orthogonality	Normality
Vectors	$\langle v, w \rangle = 0$	$\langle v, v \rangle = 1$
Functions	$\langle f, g \rangle = \int_{-1}^{1} f(x)g(x)dx = 0$	$\langle f, f \rangle = \int_{-1}^{1} f(x)^2 = 1$

$$\langle 1, x \rangle = \int_{-1}^{1} 1 \cdot x dx = \frac{x^2}{2}\Big|_{-1}^{1} = \frac{1}{2} - \frac{1}{2} = 0$$

$$\langle x, x^2 \rangle = \int_{-1}^{1} x \cdot x^2 dx = \frac{x^4}{4}\Big|_{-1}^{1} = \frac{1}{4} - \frac{1}{4} = 0$$

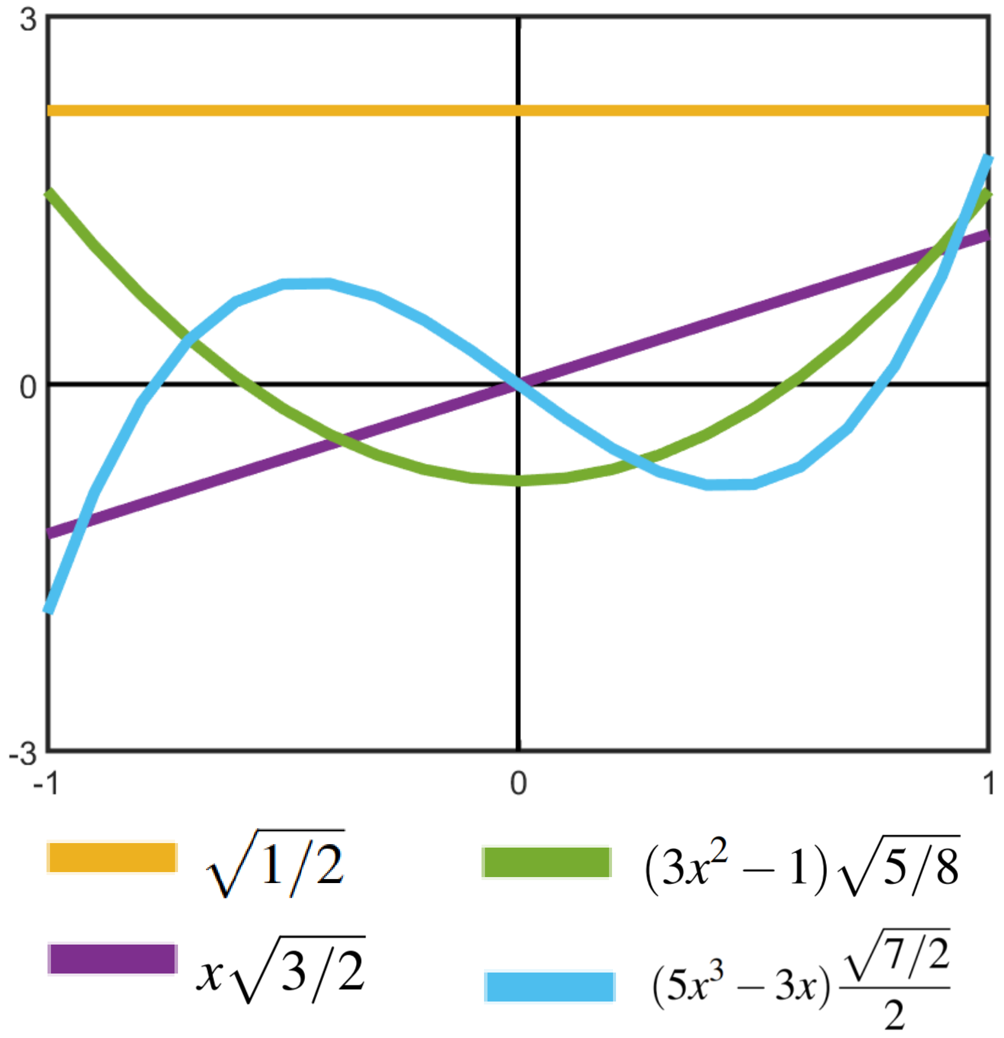
$$\langle 1, x^2 \rangle = \int_{-1}^{1} 1 \cdot x^2 dx = \frac{x^3}{3}\Big|_{-1}^{1} = \frac{1}{3} + \frac{1}{3} = \frac{2}{3}$$

$$\langle 1, 1 \rangle = \int_{-1}^{1} 1 dx = x\Big|_{-1}^{1} = 1 + 1 = 2$$

Basis – Orthonormalization for functions

Gram-Schmidt process over the monomial basis produces the **Legendre Polynomials**, that are orthonormals.

There are other useful properties for functional basis?



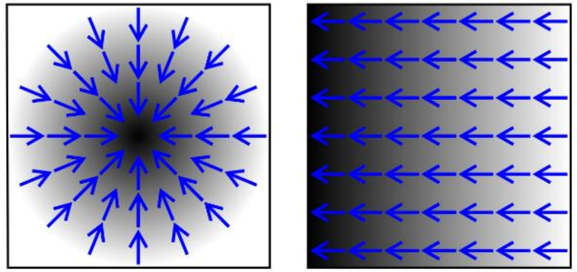
Gram-Schmidt for functions: Legendre polynomials, S. G. Johnson, MIT course 18.06, Spring 2009

Gradient, Divergence, Laplacian

Important tools from analysis:

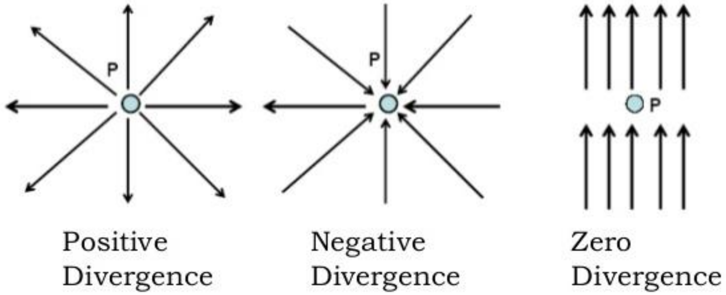
The gradient

$$\nabla: \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}, \ldots\right)$$



The divergence

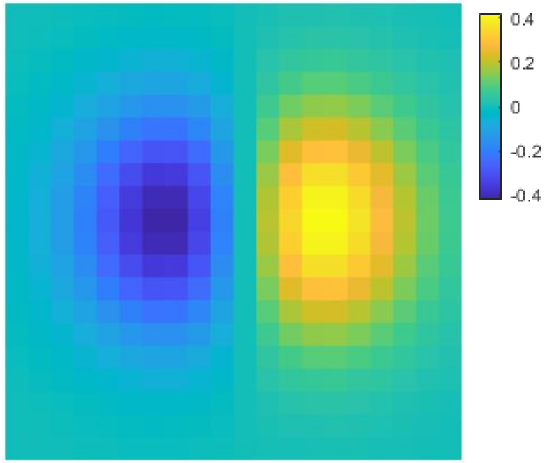
$$div: \frac{\partial}{\partial x} + \frac{\partial}{\partial y} + \frac{\partial}{\partial z} + \ldots.$$

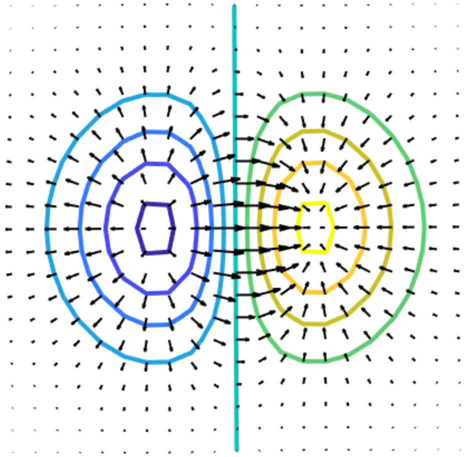


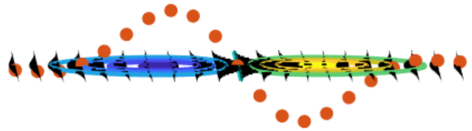
The Laplacian

$$\Delta: div\, \nabla = \frac{\partial^2}{\partial^2 x} + \frac{\partial^2}{\partial^2 y} + \frac{\partial^2}{\partial^2 z} + \ldots$$

The change of a rate of change







Note: in 1D, the Laplacian is equal to the second order derivative.

Smoothness – Dirichlet Energy

Dirichlet Energy:

$$E(f) = \int_X ||\nabla f(x)||^2 dx$$

It measures the variability of the function.

A miracle occours

First Green Identity:

$$E(f) = \int_X ||\nabla f(x)||^2 dx = \cdots = \int_X f(x) \Delta f(x) dx$$

We want the orthonormal basis that minimizes this energy.

Geometric deep learning: going beyond euclidean data. Bronstein et al., IEEE Signal Processing Magazine (2017)

Smoothness – (Discrete) Dirichlet Energy

$$E(f) = \int_X f(x) \Delta f(x) dx \qquad E(\phi_i) = \int_X \phi_i(x) \Delta \phi_i(x) dx = \langle \phi_i, \Delta \phi_i \rangle$$

$$\mathbf{\Phi}_k = \begin{bmatrix} \phi_0 & \phi_1 & \phi_2 & \dots & \phi_k \end{bmatrix}$$

$$\mathbf{\Phi}_k \in \mathbb{R}^{n \times k}$$

$$\min_{\phi_0} E(\phi_0) \quad \text{s.t.} \quad ||\phi_0|| = 1$$

$$\min_{\phi_i} E(\phi_i) \quad \text{s.t.} \quad ||\phi_i|| = 1, \quad \phi_i \perp \text{span}\{\phi_0, \dots, \phi_{i-1}\}$$

$$\boxed{\min_{\mathbf{\Phi}_k} \quad \text{trace}(\mathbf{\Phi}_k^T \Delta \mathbf{\Phi}_k) \qquad \text{s.t} \qquad \mathbf{\Phi}_k^T \mathbf{\Phi}_k = \mathbf{I}}$$

Smoothness – (Discrete) Dirichlet Energy

$$\boxed{\min_{\mathbf{\Phi}_k} \quad \mathrm{trace}(\mathbf{\Phi}_k^T \Delta \mathbf{\Phi}_k) \qquad \mathrm{s.t} \qquad \mathbf{\Phi}_k^T \mathbf{\Phi}_k = \mathbf{I}}$$

$$\Delta \mathbf{\Phi_k} = \mathbf{\Phi}_k \Lambda_k$$

The set of eigenvectors (also called *eigenfunctions*) of the Laplacian is an orthonormal basis, optimal for the smoothness.

In the 1D case they are the standard Fourier basis:

$$\Delta \sin(n\theta) = -|n|^2 \sin(n\theta)$$

$$\Delta e^{n2\pi i\theta} = -|n|^2 4\pi^2 e^{n2\pi i\theta}$$

Fourier Transform, Terence Tao, 2015,https://www.math.ucla.edu/~tao/preprints/fourier.pdf

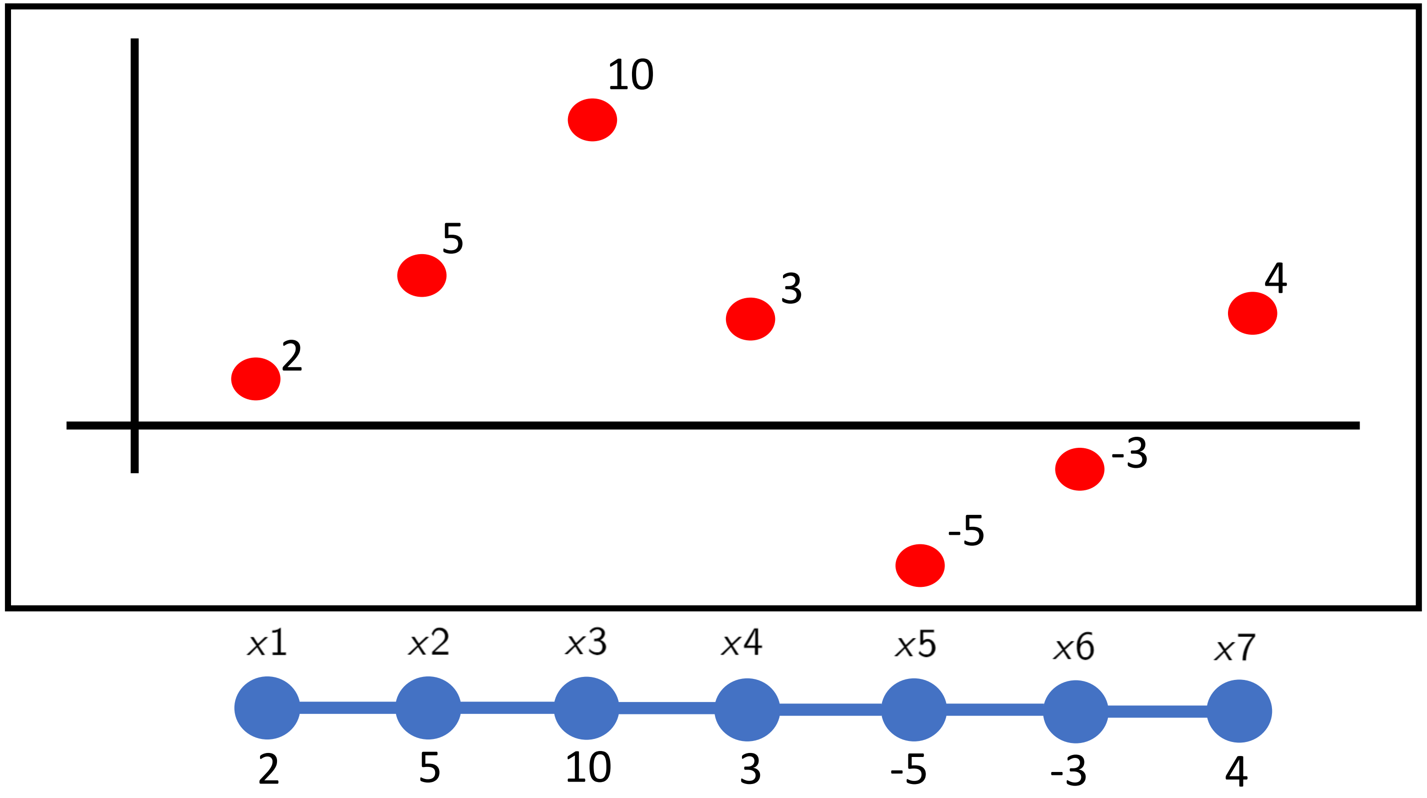
First part summary

We have seen so far:

- The concept of basis
- The properties of basis (Orthogonality and Normality)
- How to extend these to the functions
- Some main tools: gradient, divergence, Laplacian
- The Dirichlet Energy to add a new property (Smoothness)
- The link with Fourier Basis

Key idea: given a domain, we need a Laplacian.

Discrete Setting



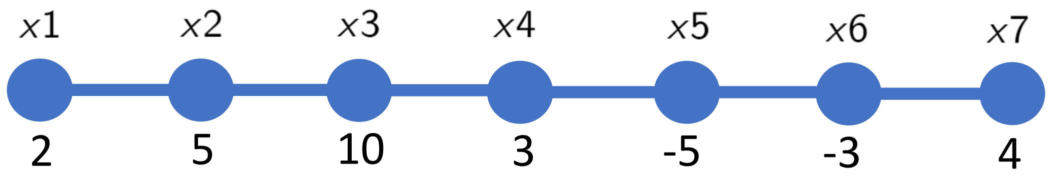
Vertices $\{x1, x2, x3, x4, x5, x6, x7\}$

Edges connect consecutive points

The y-value is a function on the graph $F : V \to \mathbb{R}$

Discrete Setting

$$f(x) = \boxed{2}\boxed{5}\boxed{10}\boxed{3}\boxed{-5}\boxed{-3}\boxed{4}$$



$$f'(x_i) = \frac{f(x_{i+1}) - f(x_i)}{h}$$

$$f'(x) = \boxed{3}\boxed{5}\boxed{-7}\boxed{-8}\boxed{2}\boxed{7}\boxed{\ldots}$$

$$f''(x_i) = \frac{f'(x_i) - f'(x_i - 1)}{h} = \cdots = \frac{f(x_{i+1}) + f(x_{i-1}) - 2f(x_i)}{h^2}$$

assuming $h = 1$:

$$f''(x_i) = f(x_{i+1}) + f(x_{i-1}) - 2f(x_i)$$

$$f''(x) = \boxed{\ldots}\boxed{2}\boxed{-12}\boxed{-1}\boxed{10}\boxed{5}\boxed{\ldots}$$

Encode distance of a point from the mean of the neighboor values.

Digital Geometry Processing, Maks Ovsjanikov, 2015,http://www.lix.polytechnique.fr/~maks/Verona_MPAM/schedule.html

1D Discrete Laplacian

$f(x) =$	2	5	10	3	-5	-3	4

$$f''(x_i) = f(x_{i+1}) + f(x_{i-1}) - 2f(x_i)$$

$$f''(x_i) = Lf = \begin{bmatrix} -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} 2 \\ 5 \\ 10 \\ 3 \\ -5 \\ -3 \\ 4 \end{bmatrix} = \begin{bmatrix} \dots \\ 2 \\ -12 \\ -1 \\ 10 \\ 5 \\ \dots \end{bmatrix}$$

$$L_{ij} = \begin{cases} 1 & \text{if } i \neq j, i \sim j \\ -d_{ij} & \text{if } i = j \\ 0 & \text{otherwise} \end{cases} \quad \text{or more generally} \quad \begin{cases} w_{ij} & \text{if } i \neq j, i \sim j \\ -\sum_{j \neq i, j \sim i} w_{ij} & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

$L = W - D$, where D is the diagonal matrix of the degrees, and W is the weighted adjacency matrix.

Laplacian Eigenvectors

$$E(f) = \int_X ||\nabla f(x)||^2 dx = \int_X f(x) \Delta f(x) dx$$

$$E(f) = \sum \langle \nabla f, \nabla f \rangle = \sum_i f_i (\Delta f)_i$$

$$\sum_{(i,j) \in E} (f_i - f_j)^2 = f^T L f$$

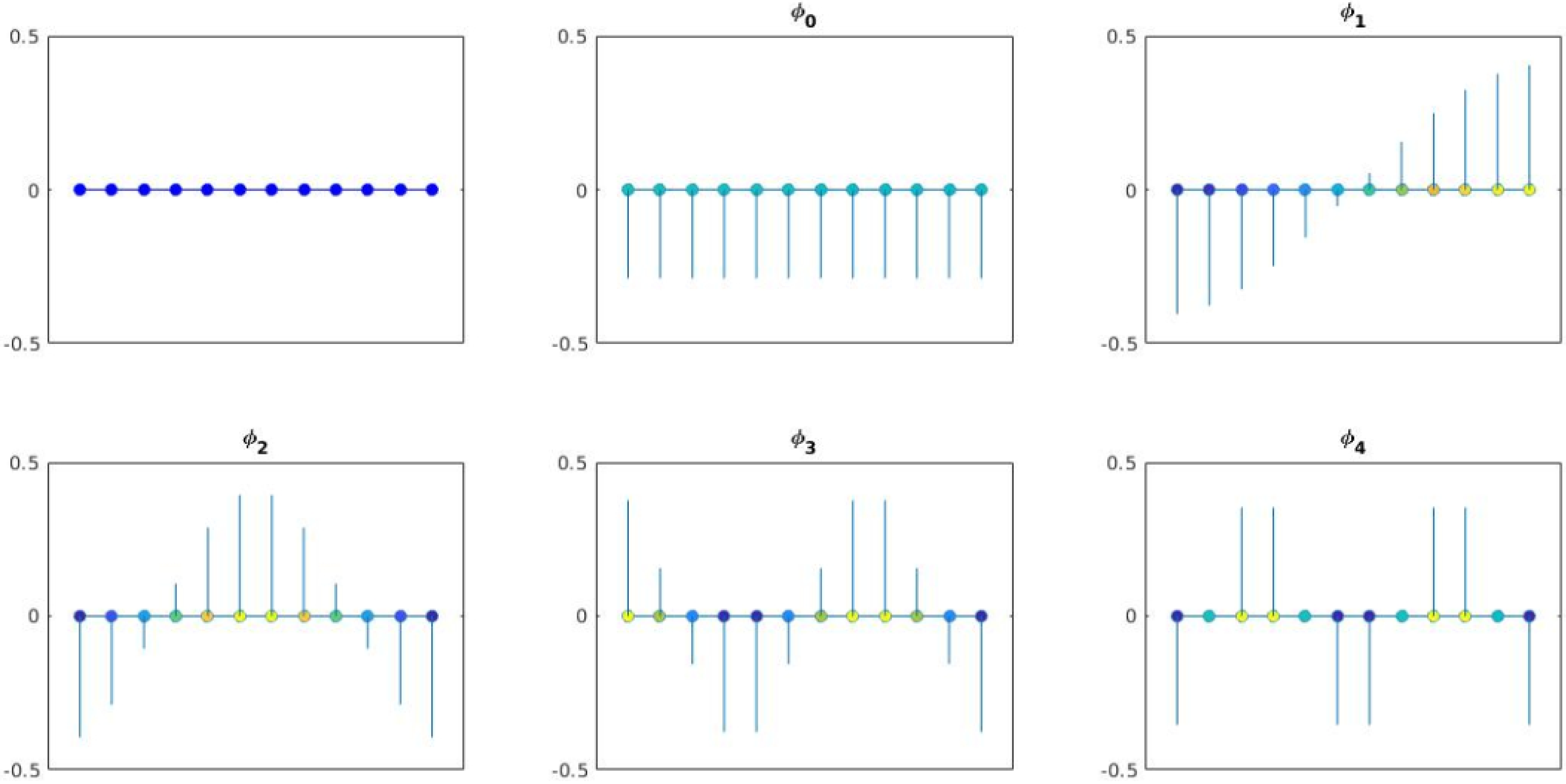
$$\lambda = f^T L f$$

$$\lambda f = L f$$

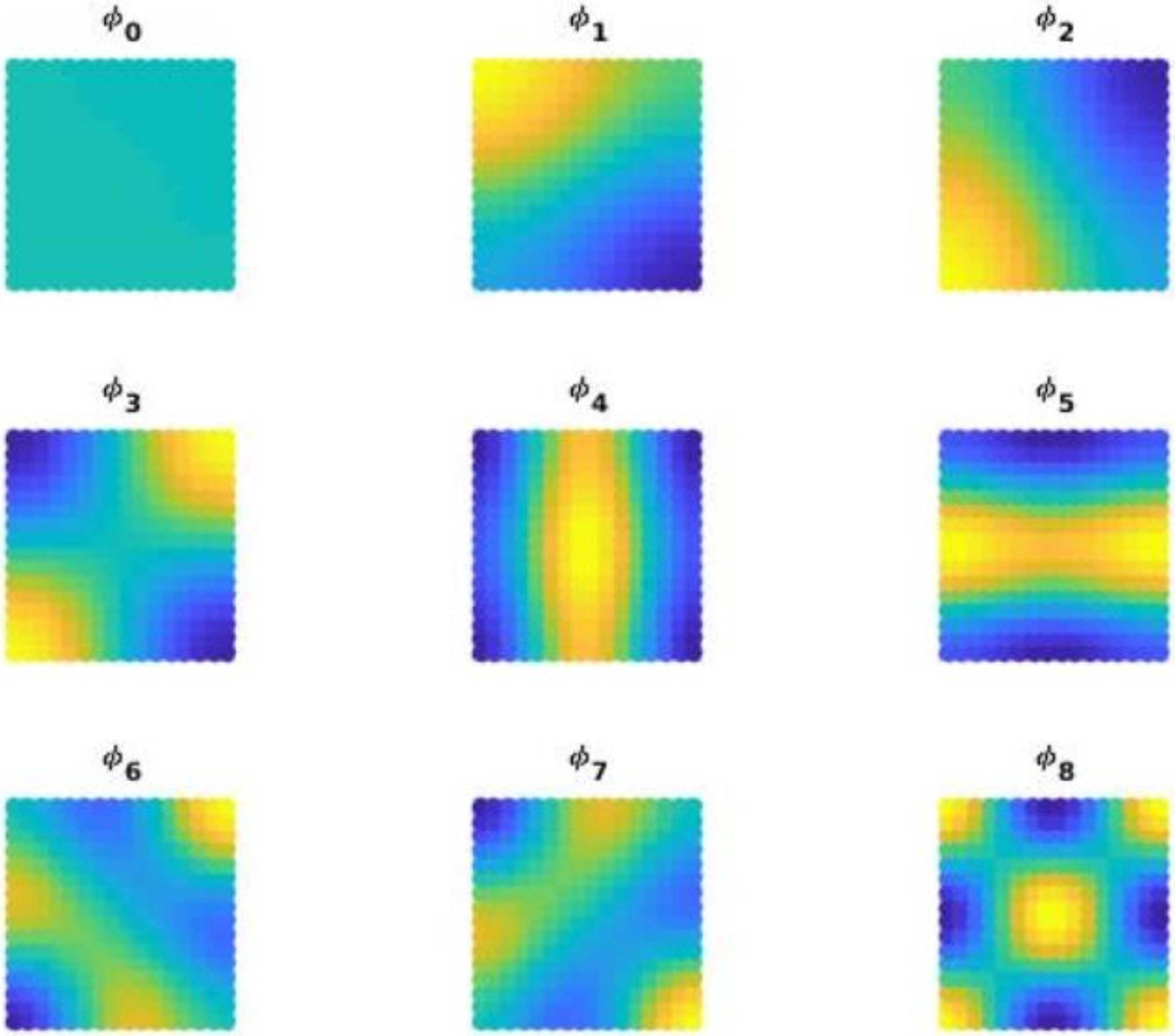
L is symmetric $\Rightarrow$ its eigenvalues are real and non-negative

$$\lambda_i = \phi_i^T L \phi_i \geq 0$$

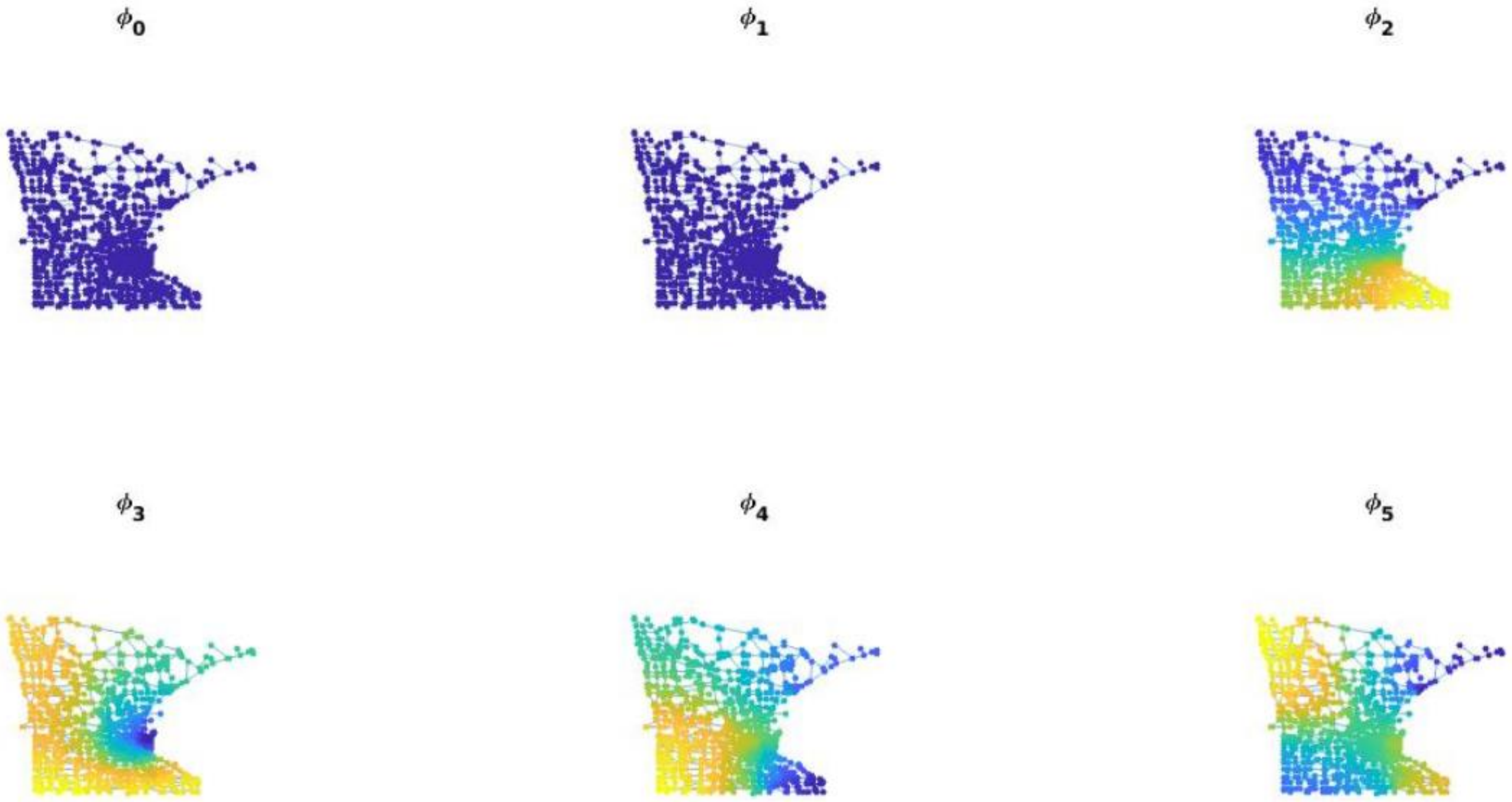
Laplacian Eigenvectors – 1D



Laplacian Eigenvectors – 2D



Laplacian Eigenvectors – Generic Graph



Minnesota Road Map, D. Gleich, 2010,https://www.cise.ufl.edu/research/sparse/matrices/Gleich/minnesota.html

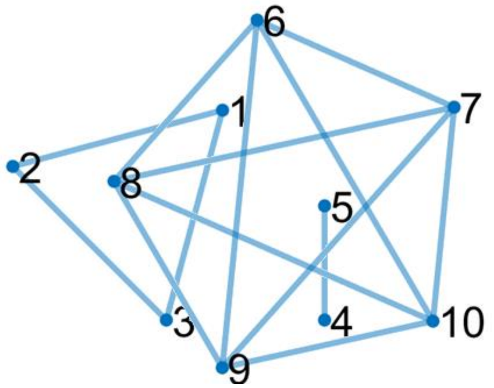
Disconnected components

Consider the eigenvalues 0 on a connected graph:

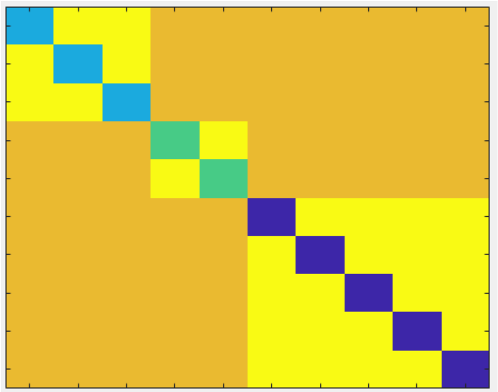
$$\sum_{(i,j)\in E} (\phi_i - \phi_j)^2 = \phi^T L \phi = 0 \implies (\phi_i - \phi_j)^2 = 0, \forall (i,j) \implies \phi_i = \phi_j$$

ϕ_0 is the constant function.

Consider the adjency matrix on a disconnected graph $\implies$ Block matrix



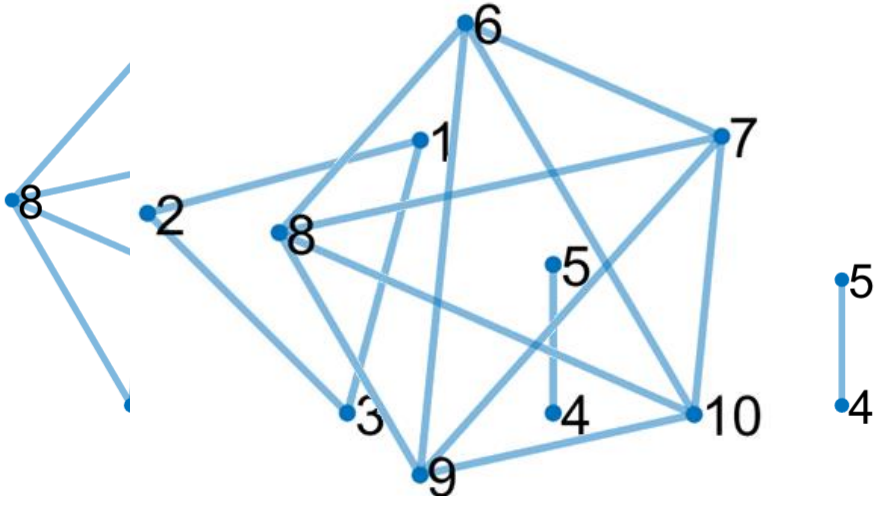
$$W = \begin{bmatrix} A_1 & & & \\ & A_2 & & \\ & & \cdots & \\ & & & A_k \end{bmatrix} \qquad L = \begin{bmatrix} L_1 & & & \\ & L_2 & & \\ & & \cdots & \\ & & & L_k \end{bmatrix}$$

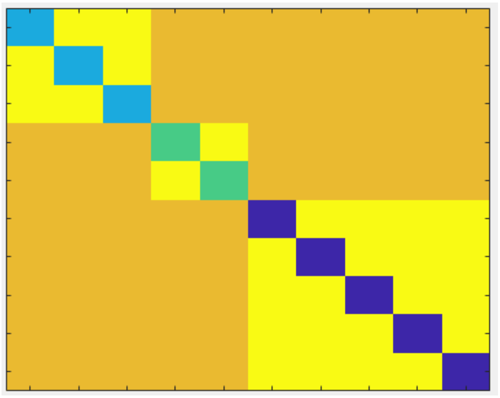


The eigenfunctions of a block matrix are the union of the eigenfunctions of the blocks $\implies$ 0s eigenvalues as many as connected components.

A Tutorial on Spectral Clustering, Ulrike von Luxburg, 2006,https://arxiv.org/pdf/0711.0189.pdf

Disconnected components

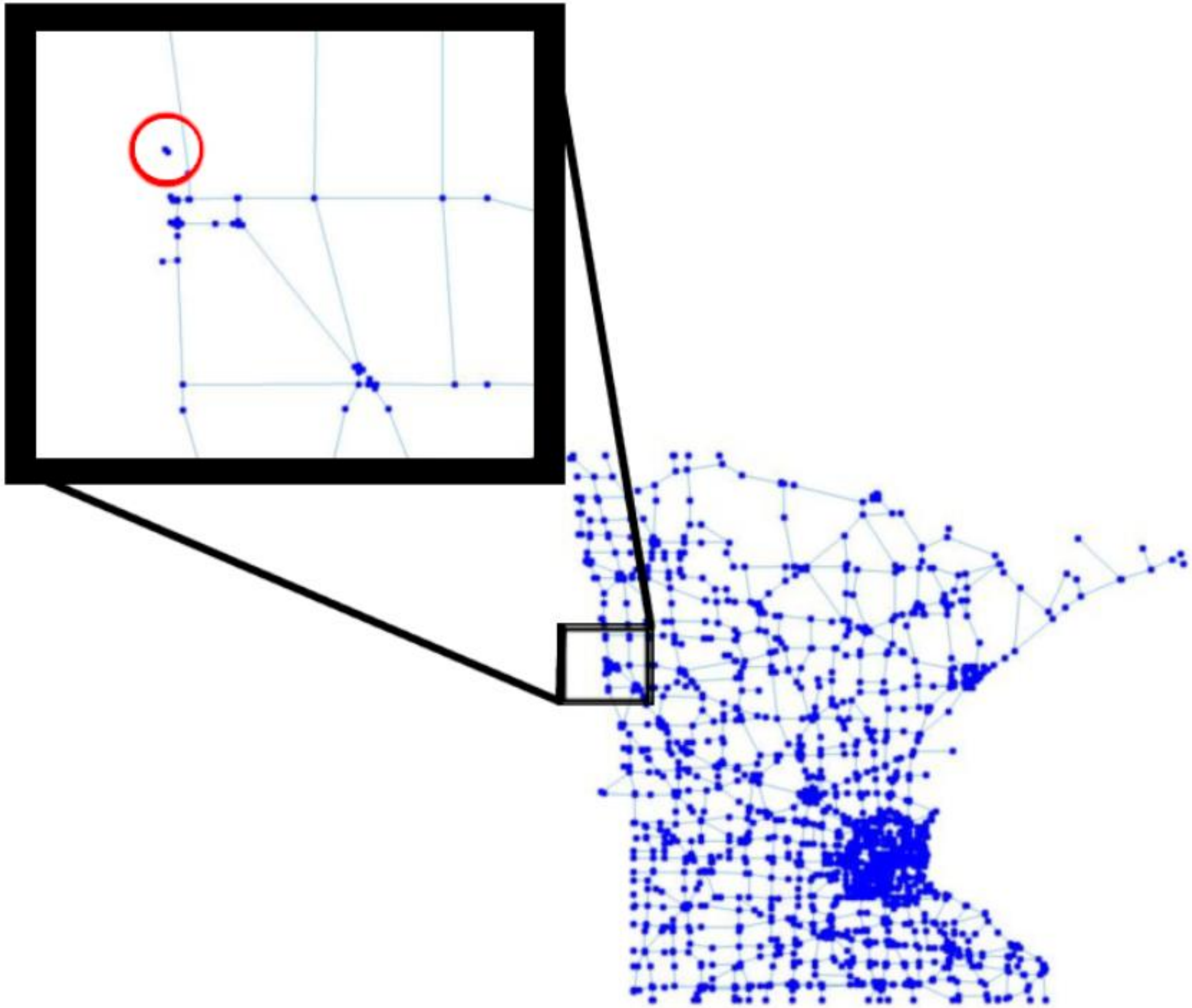




$$\lambda_0 = 0$$
$$\lambda_1 = 0$$
$$\lambda_2 = 0$$
$$\lambda_3 = 2$$
$$\lambda_4 = 3$$
$$\dots$$

$$\phi_0 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0.4472 \\ 0.4472 \\ 0.4472 \\ 0.4472 \\ 0.4472 \end{bmatrix} \quad \phi_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0.7071 \\ 0.7071 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \phi_2 = \begin{bmatrix} 0.5774 \\ 0.5774 \\ 0.5774 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \phi_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0.7071 \\ -0.7071 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \dots$$

Disconnected components



Min-cut

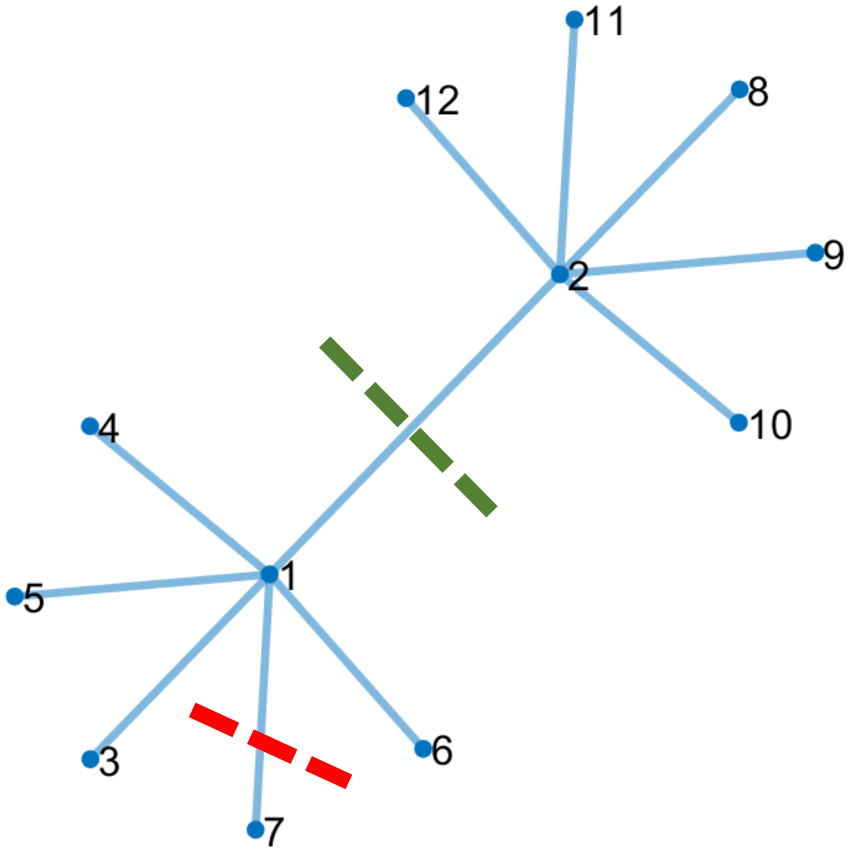
Minimum number of edges to remove to disconnect the graph.

Given a Graph (V,E), find the minimum cut s.t. the two disjoint sets $A, B \subset V$

$$\min \operatorname{cut}(A, B) = \sum_{i \in A, j \in B} w_{ij}$$

Green and red cuts have the same cost.
In general we prefer a *balanced* cut.

$$|A_1| = |\bar{A}_1| = 0.5$$



Ratiocut

Approximation of a balanced cut

$$\min \text{Ratiocut}(A_1, \dots, A_2) = \sum_{i=1}^{k} \frac{cut(A_i, \bar{A}_i)}{|A_i|}$$

Case of split in two

$$\min \text{Ratiocut}(A_1, \bar{A}_1) = \frac{cut(A_1, \bar{A}_1)}{|A_1|} + \frac{cut(\bar{A}_1, A_1)}{|\bar{A}_1|} = \epsilon$$

$$|\bar{A}_1| cut + |A_1| cut = \epsilon |A_1||\bar{A}_1|$$

$$\frac{cut}{|A_1||\bar{A}_1|} = \frac{\epsilon}{|V|}$$

$$\min \text{Ratiocut}(A_1, \bar{A}_1) = \frac{cut}{|A_1||\bar{A}_1|}$$

maximize $|A_1||\bar{A}_1| \implies |A_1| = |\bar{A}_1| = 0.5$

minimize $cut \implies$ few «bridge» edges

Ratiocut

$$\sum_{(i,j)\in E} (\phi_{1_i} - \phi_{1_j})^2 = \lambda_1$$

Minimize Dirichlet Energy

$$\phi_0 \perp \phi_1 \implies \langle \phi_0, \phi_1 \rangle = 0 \implies \sum_i \phi_{0_i}\phi_{1_i} = 0 \implies \phi_{1_1} + \phi_{1_2} + \cdots = 0$$

First eigenfunction is constant

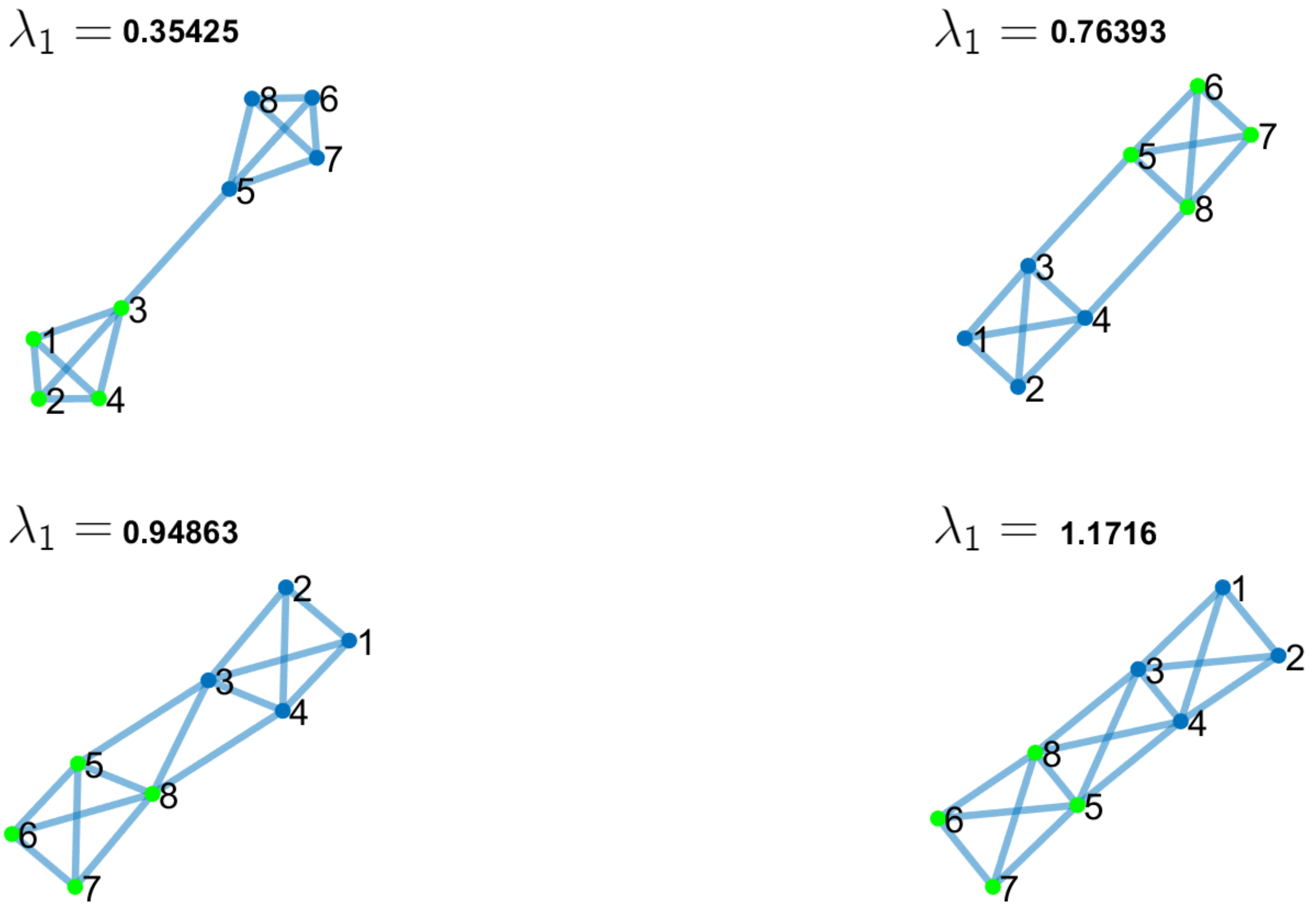
Also:

$$||\phi_1|| = 1$$

$$\phi_{1_1} + \phi_{1_2} + \cdots = 0$$

Split the graph eavenly by sign → Relaxed version of RatioCut

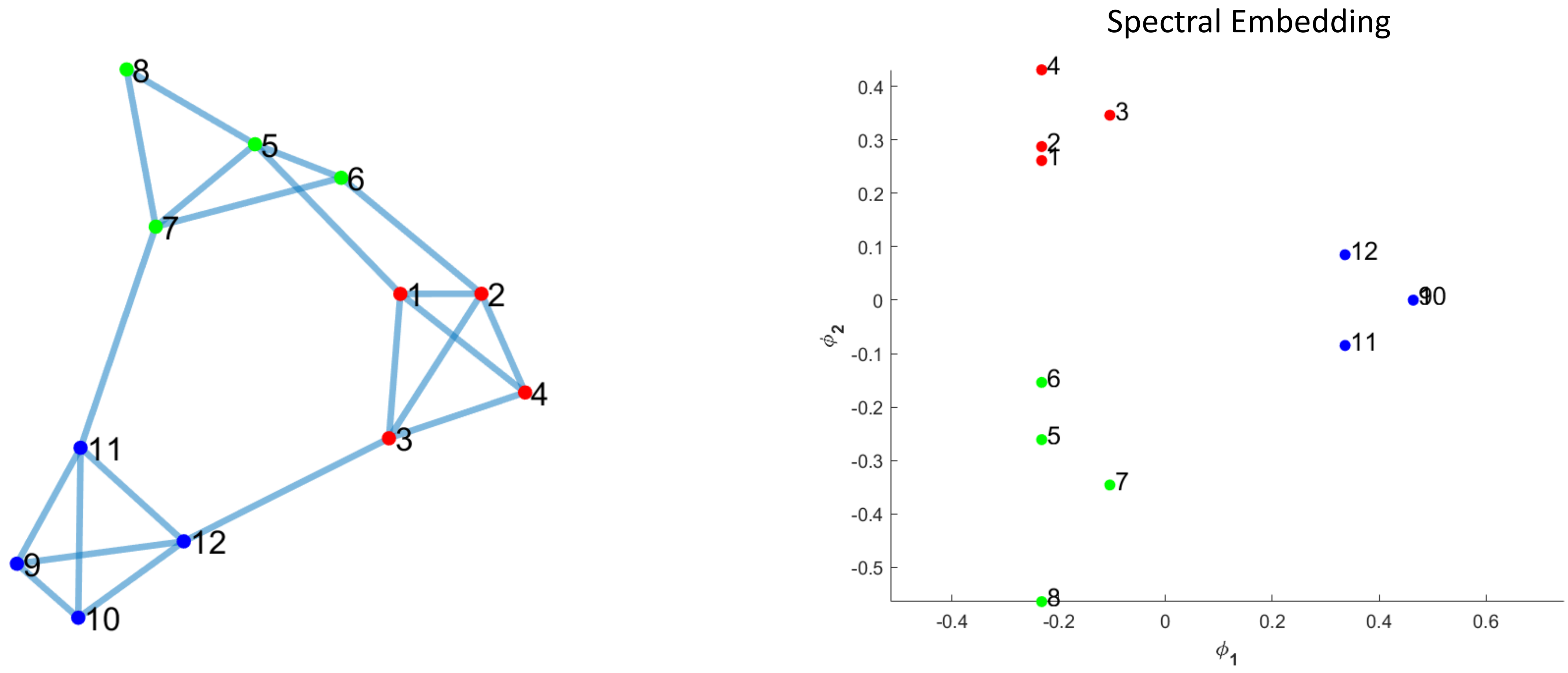
Fiedler value



Spectral Clustering

$$\phi_1, \phi_2, \ldots, \phi_d$$

$$\{\phi_1(v), \phi_2(v), \ldots, \phi_d(v)\}$$

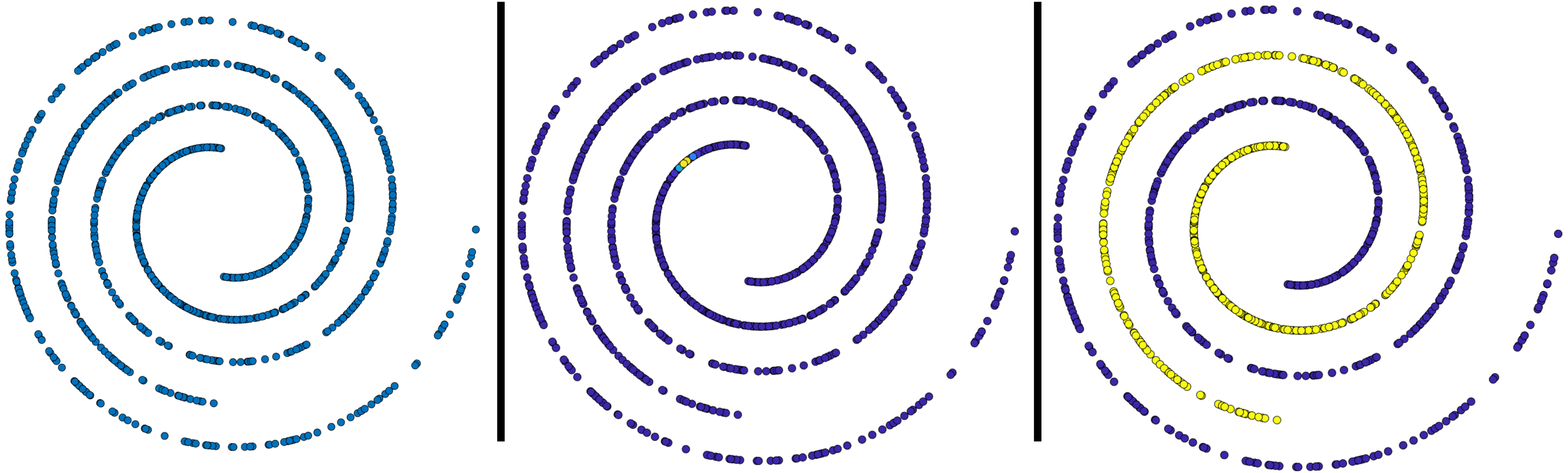


Spectral Clustering – 2D

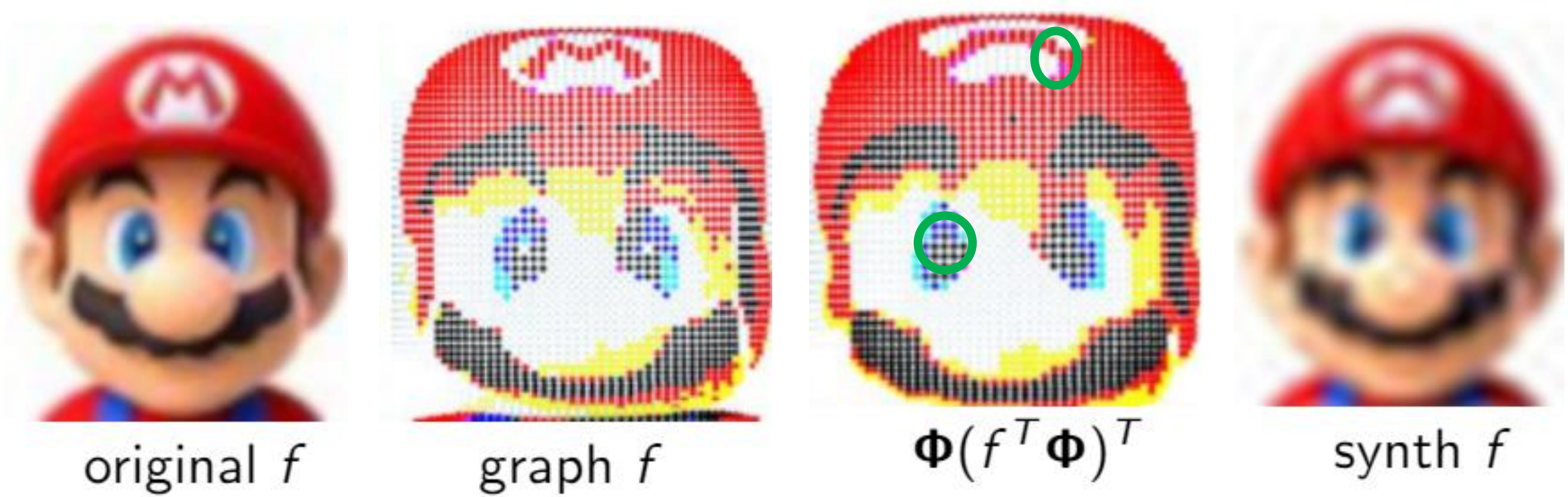
We can build a fully-connected adjancy matrix setting the weight as:

$$d(x_i, x_j) = e^{\frac{-|x_i - x_j|}{2\sigma^2}}$$

Other strategies are possible (e.g. K-NN).



Graph Signal Processing



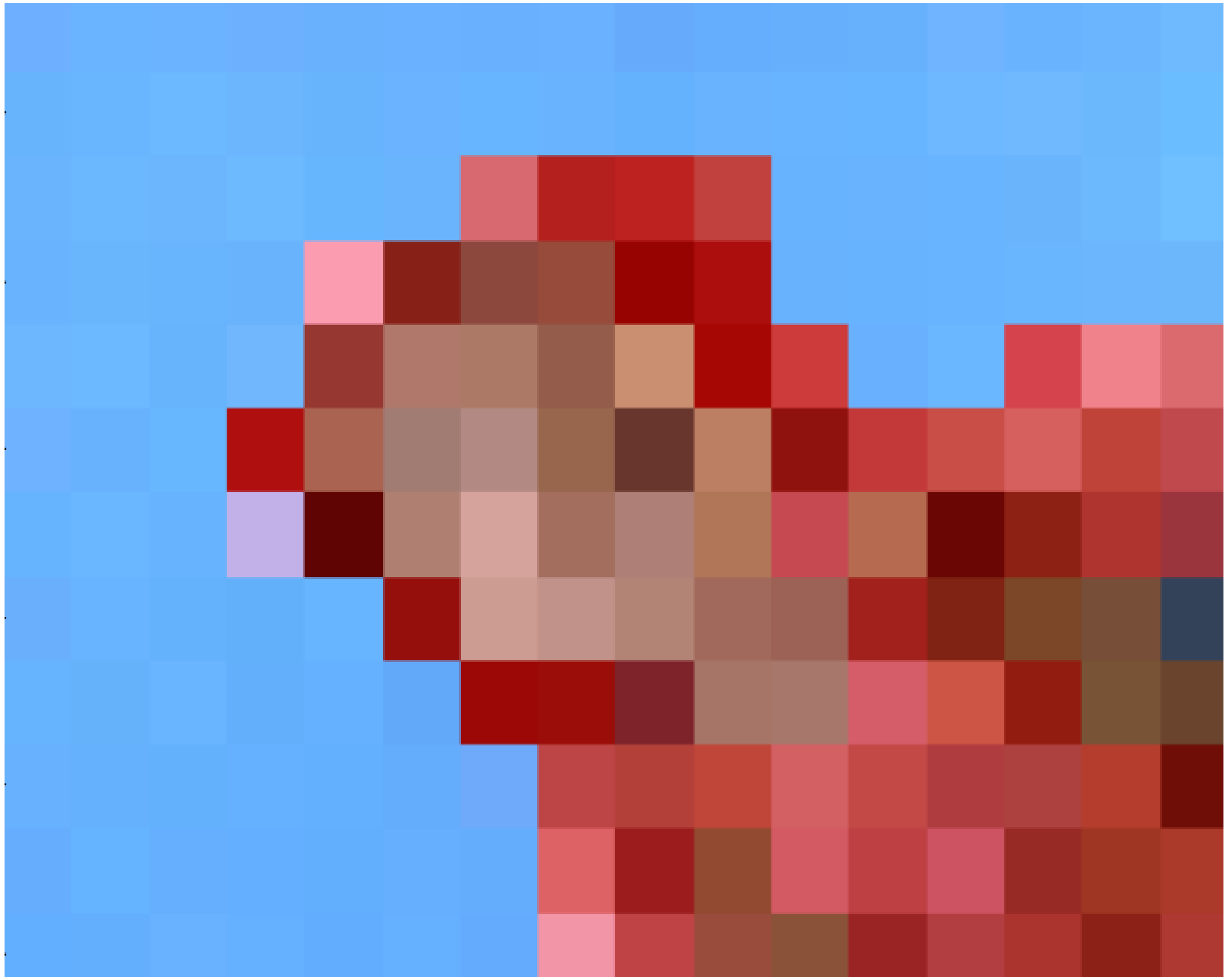
Denoising and compression

This image is $64 \times 64 = 4096$ nodes, for the 3 RGB channels: 12288 values. We have represented 3 signals with the first 500 eigenfunctions $\implies$ 1500 values.

Segmentation (Shi-Malik '00)

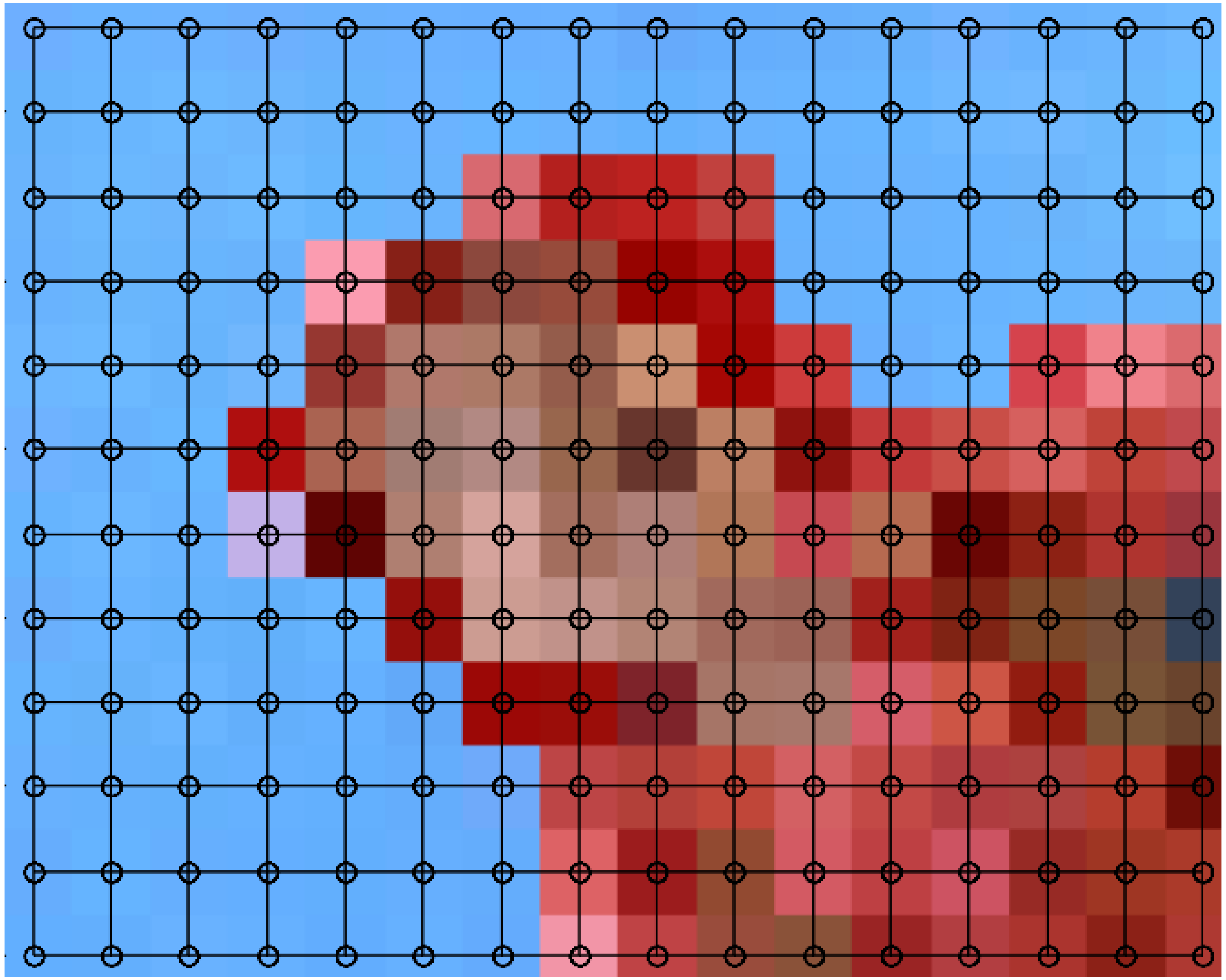
https://www.cs.yale.edu/homes/spielman/sgta/SpectTut.pdf

Segmentation (Shi-Malik ’00)



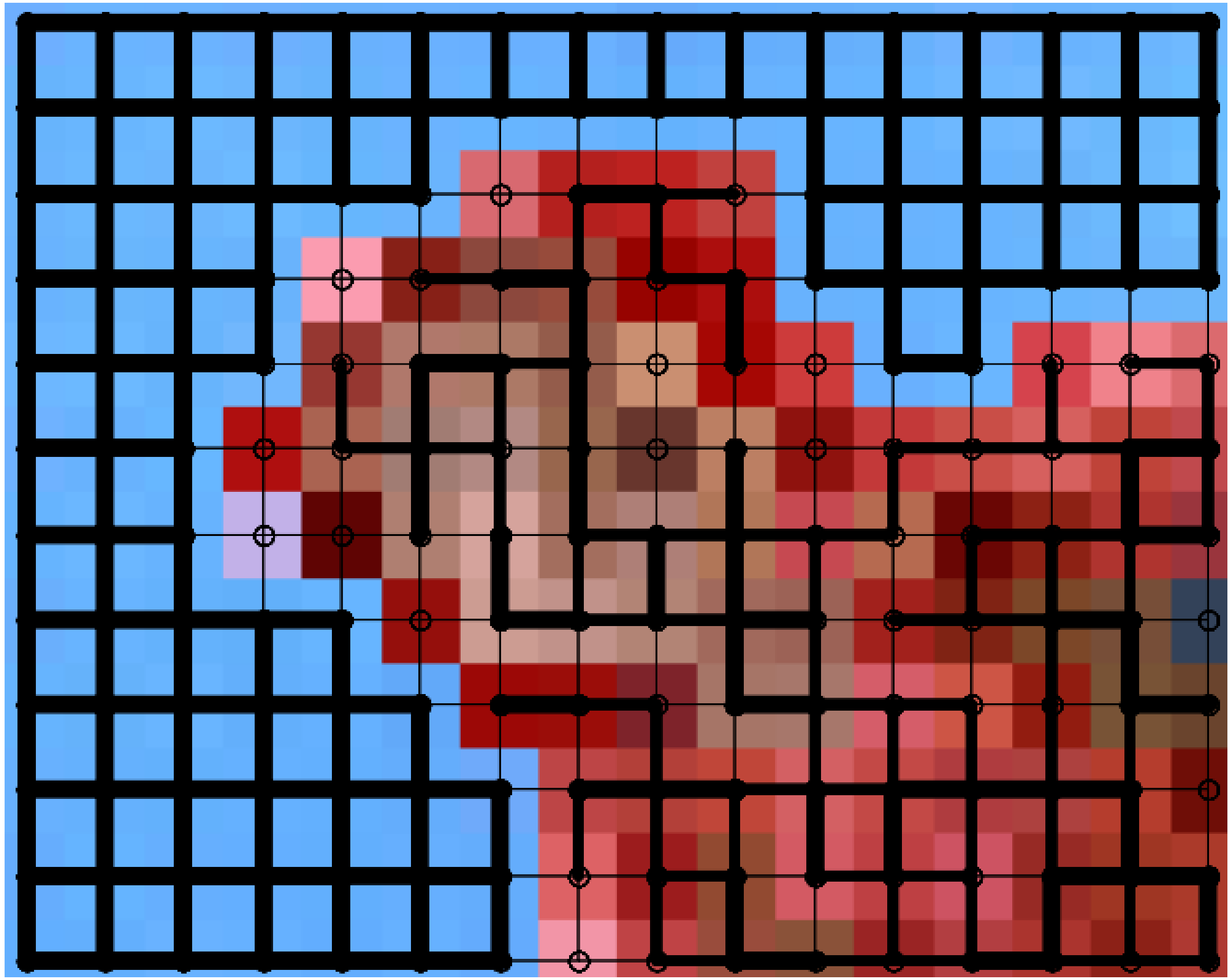
https://www.cs.yale.edu/homes/spielman/sgta/SpectTut.pdf

Segmentation (Shi-Malik ’00)



https://www.cs.yale.edu/homes/spielman/sgta/SpectTut.pdf

Segmentation (Shi-Malik '00)



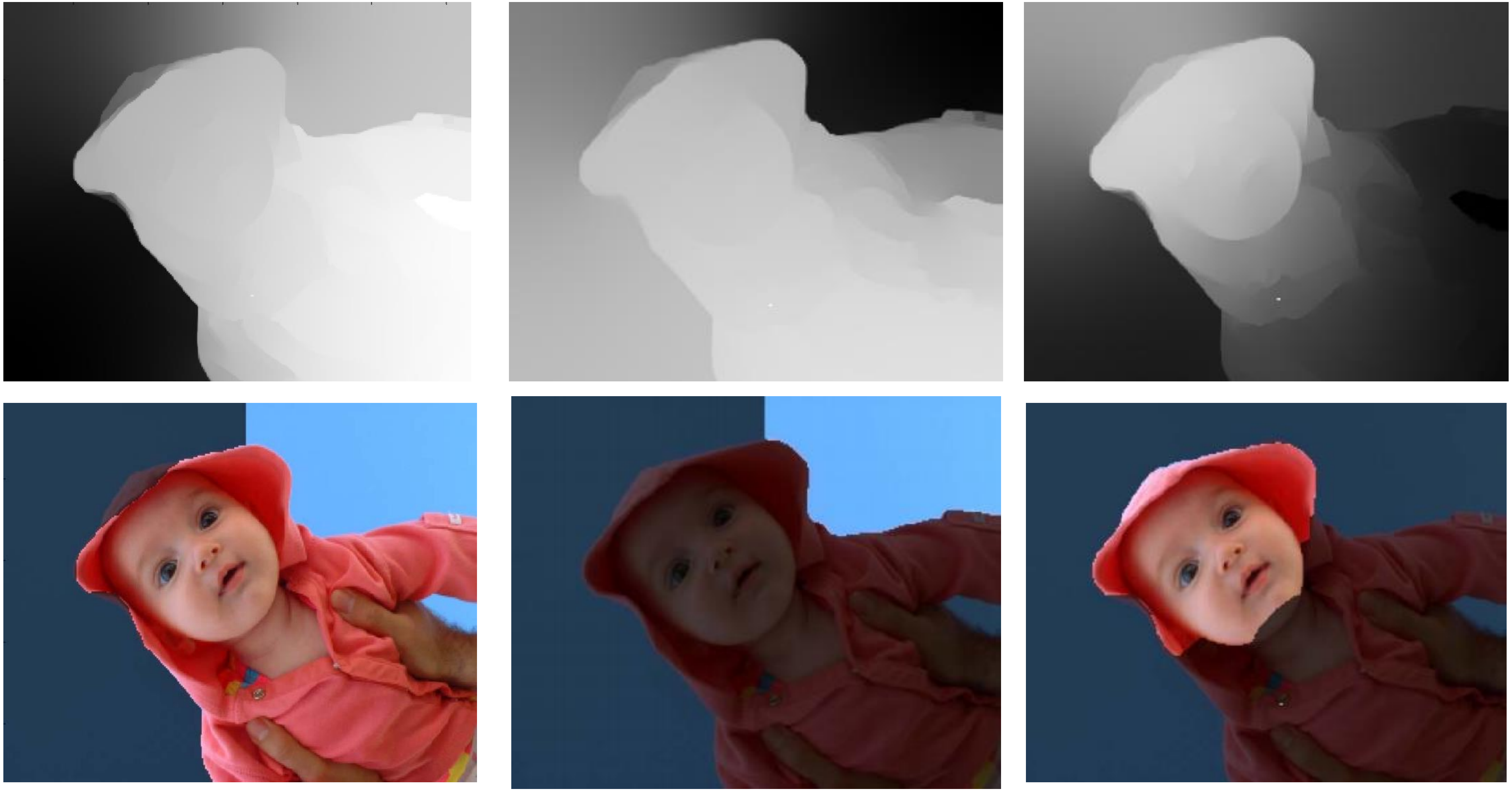
edge weight $e^{-\mathrm{diff}(\mathrm{pixel}\,i,\mathrm{pixel}\,j)^2/t^2}$

https://www.cs.yale.edu/homes/spielman/sgta/SpectTut.pdf

Segmentation (Shi-Malik '00)

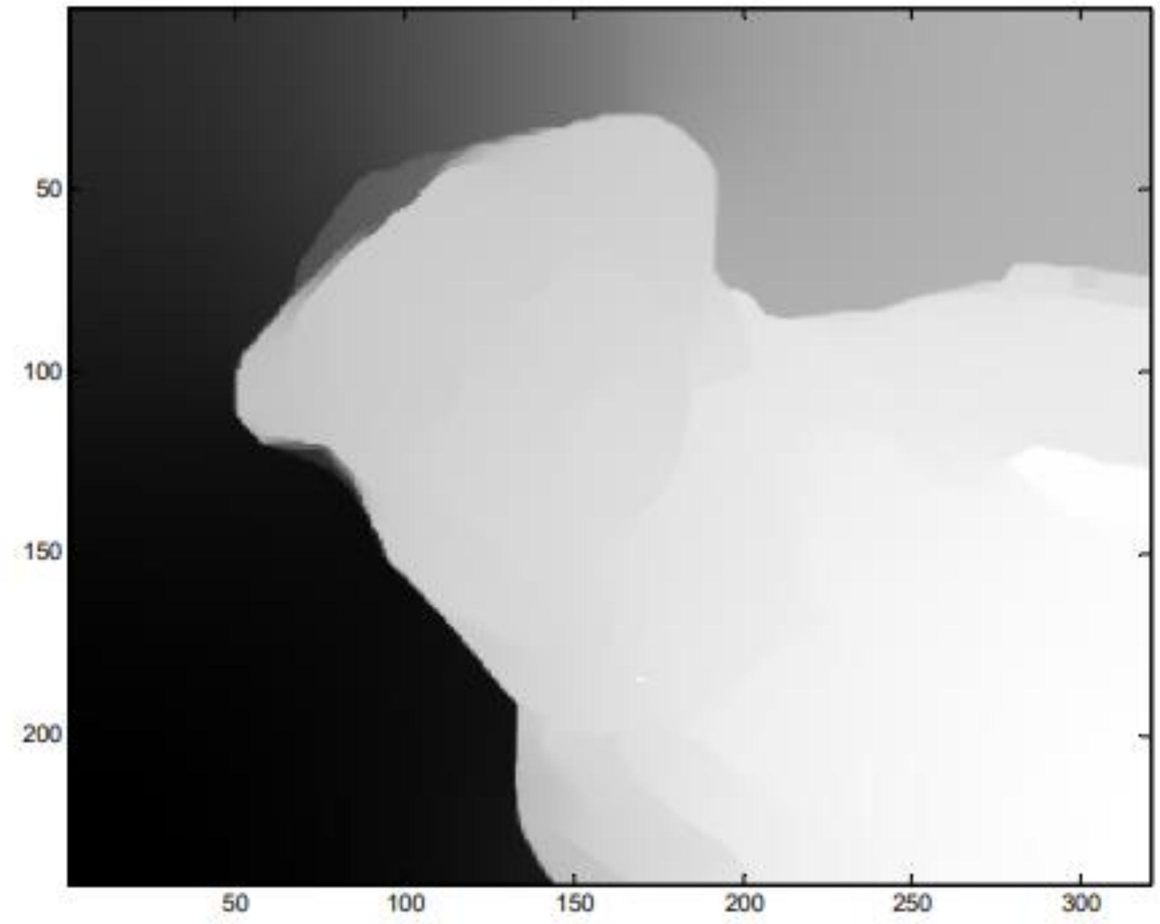
Eigen

Segm.



https://www.cs.yale.edu/homes/spielman/sgta/SpectTut.pdf

Segmentation (Miller and Tolliver ’06)



For a graph $G = (V, E, w^0)$ prescribe the number of partitions k that the edge cut is to yield. Given a valid **reweighting scheme**, iteration of the `SR-Step` produces a sequence of N weightings $\{w^{(N)}\}$ such that the graph $G^N = (V, E, w^N)$ is disconnected into k components by the weighting w^N.

https://www.cs.yale.edu/homes/spielman/sgta/SpectTut.pdf

Segmentation (Neto and Felzenszwalb '20)

Both graphs are defined over the same set of vertices, corresponding to the pixels in an image.

1. The graph G_{grid} is a grid over the image pixels, where each pixel is connected to the four neighboring pixels with an edge of weight 1. This graph encourages neighboring pixels to be grouped together, independent of their appearance.
2. The graph G_{data} is a fully connected graph that encourages pixels with similar appearance to be grouped together, independent of their location. The weights in G_{data} are based on appearance similarity of pixels, and do not depend on pixel locations,

$$w(i,j) = \exp\left(-\frac{||I(i) - I(j)||^2}{2\sigma^2}\right). \tag{6}$$

In our experiments, we use a Lanczos Process to compute the second largest eigenvector of P.

Spectral Image Segmentation with Global Appearance Modeling; Neto, Felzenszwalb; 2020 (NIPS submission)

Summary

We have learned:

- How to extend Laplacian to non-euclidean domain
- Why their eigenfunctions are so important
- How this operator and its eigendecomposition touch the connectivity
- Several applications of its eigendecomposition

! Exercises !

https://github.com/riccardomarin/SpectralShapeAnalysis

It could take some time to configure the environment

Lunch it in advance