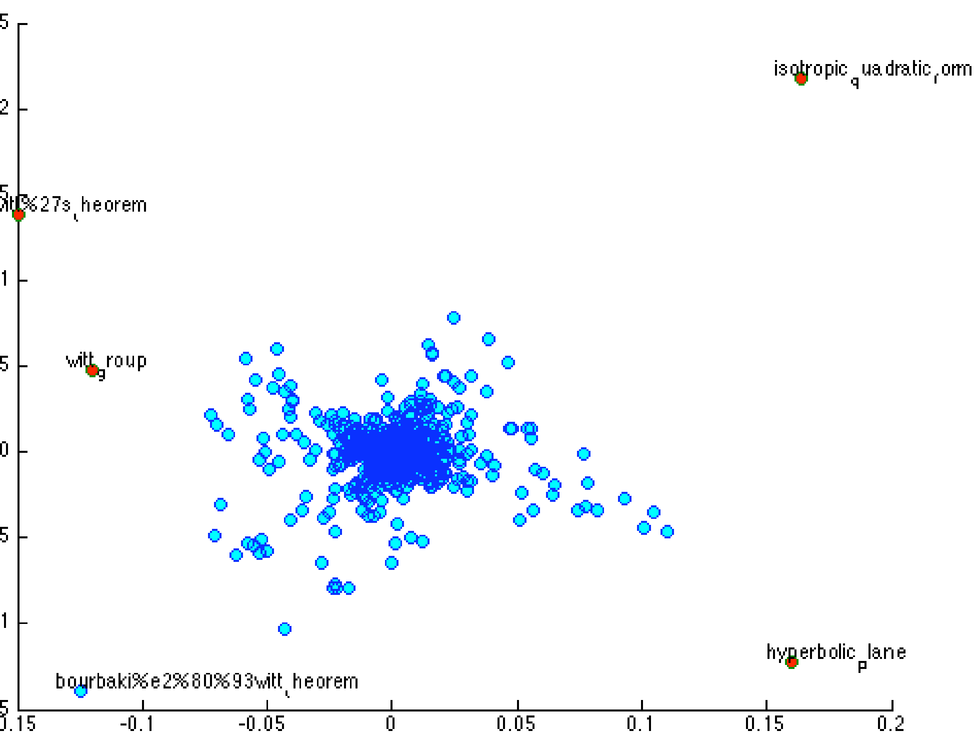
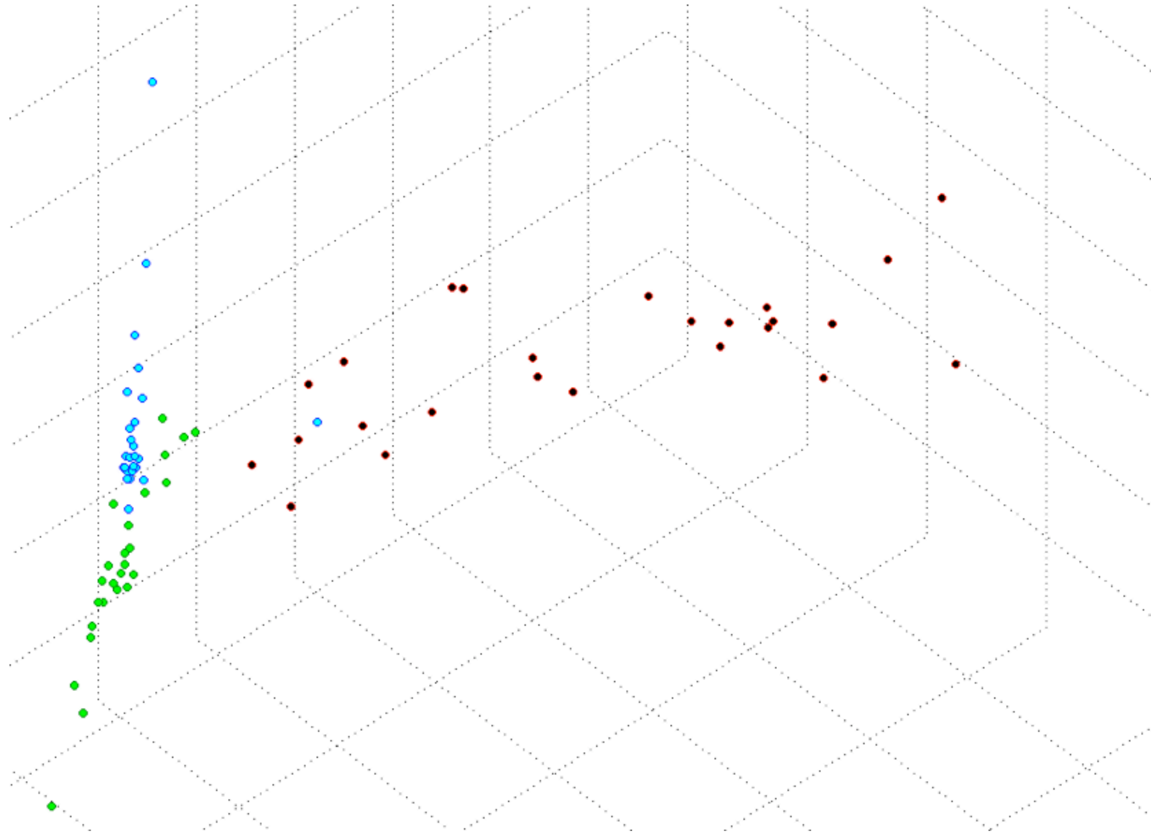
Fraud Detection, Quantum Mechanics, and Complex Systems
Lecture 3, CSSS10

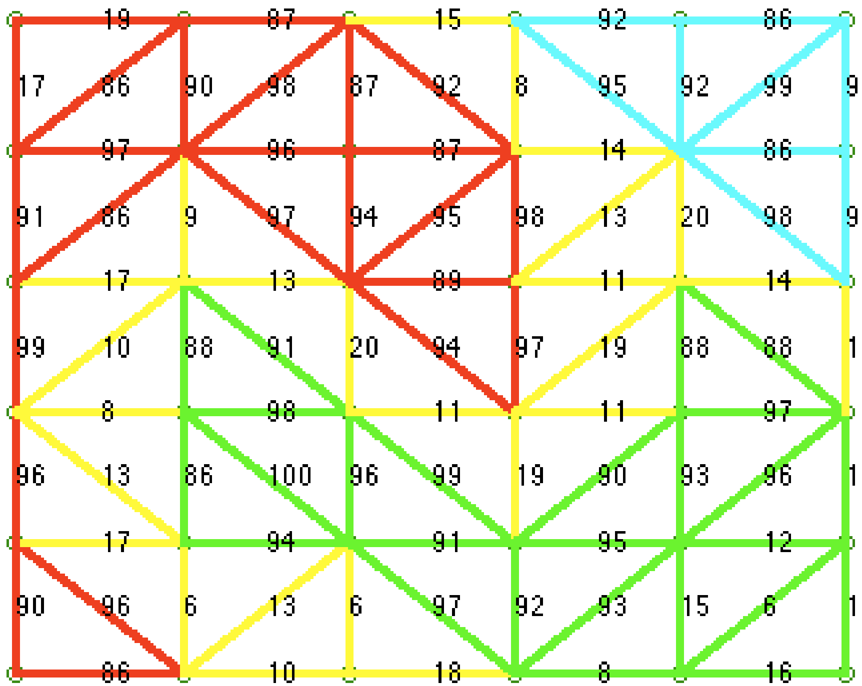




Greg Leibon
Memento, Inc
Dartmouth College

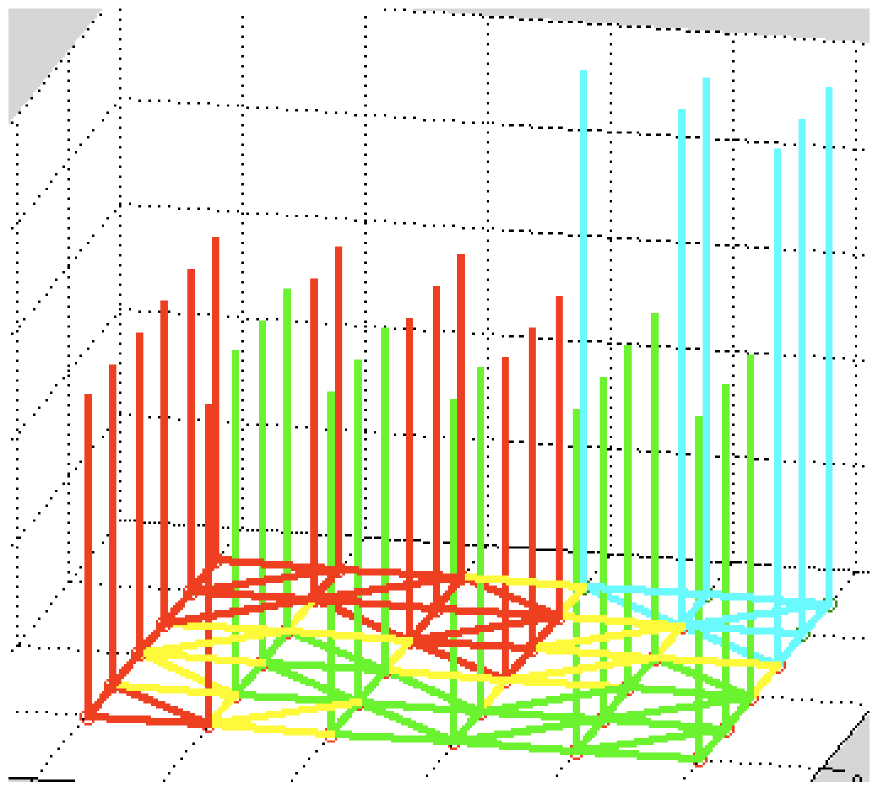
The Vagabond Problem: Finding a partition $\{R_i\}_{i=1}^{K}$ of our states into regions that minimizes

$$\sum_i P(X_{t+dt} \in \bar{R}_i \mid X_t \in R_i)$$



N=200; K=10; [states]=PlotTheState(N,K);

$$\chi_R(i) = \begin{cases} \frac{1}{\sqrt{\omega(R)}} & \text{if } i \in R \\ 0 & \text{if } i \in \bar{R} \end{cases}$$



The solution (I think!)

NP hard

$$\text{if } R_1 \text{ and } R_2 \text{ are disjoint } \langle \chi_{R_1}, \chi_{R_2} \rangle_\omega = \delta_1^2$$

Recall: $< f, g >_\omega = f^{tr}[\omega]g$

How we really solve this?

Vagabond Theorem:

$$P(X_{t+dt} \in \bar{R} \mid X_t \in R) \approx ||\chi_R||^2_{Dir}\, dt$$

proof:

$$\begin{aligned}
P(X_{t+dt} \in \bar{R} \mid X_t \in R) &= \tfrac{P(X_{t+dt} \in \bar{R}, X_t \in R)}{P(X_t \in R)} && \text{(def. conditional prob.)}\\
&= \frac{\sum_{i \in \bar{R}, j \in R} P(X_{t+dt}=i, X_t=j)}{\omega(R)} && \text{(def.)}\\
&= \frac{\sum_{i \in \bar{R}, j \in R} P(X_{t+dt}=i | X_t=j) P(X_t=i)}{\omega(R)} && \text{(def. conditional prob.)}\\
&= \frac{\sum_{i \in \bar{R}, j \in R} (1/\tau_j) p_{ij} dt \omega_i}{\omega(R)} && \text{(instantaneous transition rate)}\\
&= \frac{\sum_{i \in \bar{R}, j \in S} p_{ij} \pi_i}{\omega(R)} dt && (\omega_i = \tau_j \pi_i)\\
&= \sum_{i \in \bar{R}, j \in R} \pi_i p_{ij} (\chi_R(i) - \chi_R(j))^2 dt && \text{def.}\\
&= (1/2) \sum_{i,j} \pi_i p_{ij} (\chi_R(i) - \chi_R(j))^2 dt && \text{(summand 0 if } i, j \in S \text{ or } i, j \in \bar{R})\\
&= \langle \chi_R, \Delta_\pi \chi_R \rangle_\pi dt && \text{Green's Identity}\\
&= ||\chi_R||^2_{Dir}\, dt && \text{conformal invariance}
\end{aligned}$$

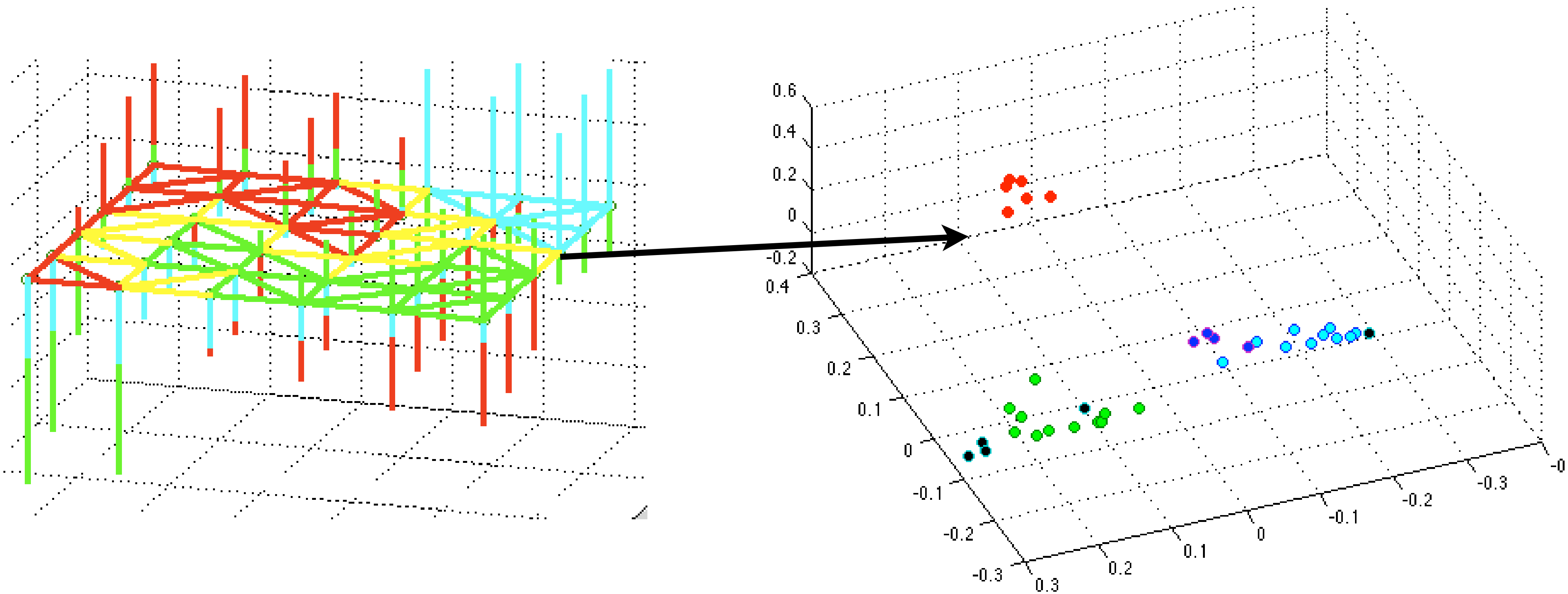
Q.E.D

$$\sum_i P(X_{t+dt} \in \bar{R}_i \mid X_t \in R_i) = \sum_i ||\chi_{R_i}||^2_{Dir}.$$

Now we can relax the Vagabond Problem and search for a ω-orthonormal set $\{f_i\}_{i=1}^{L}$ such that

$$\sum_i^L ||f_i||^2_{Dir}$$

is minimized.



Now we cluster in Euclidean space...

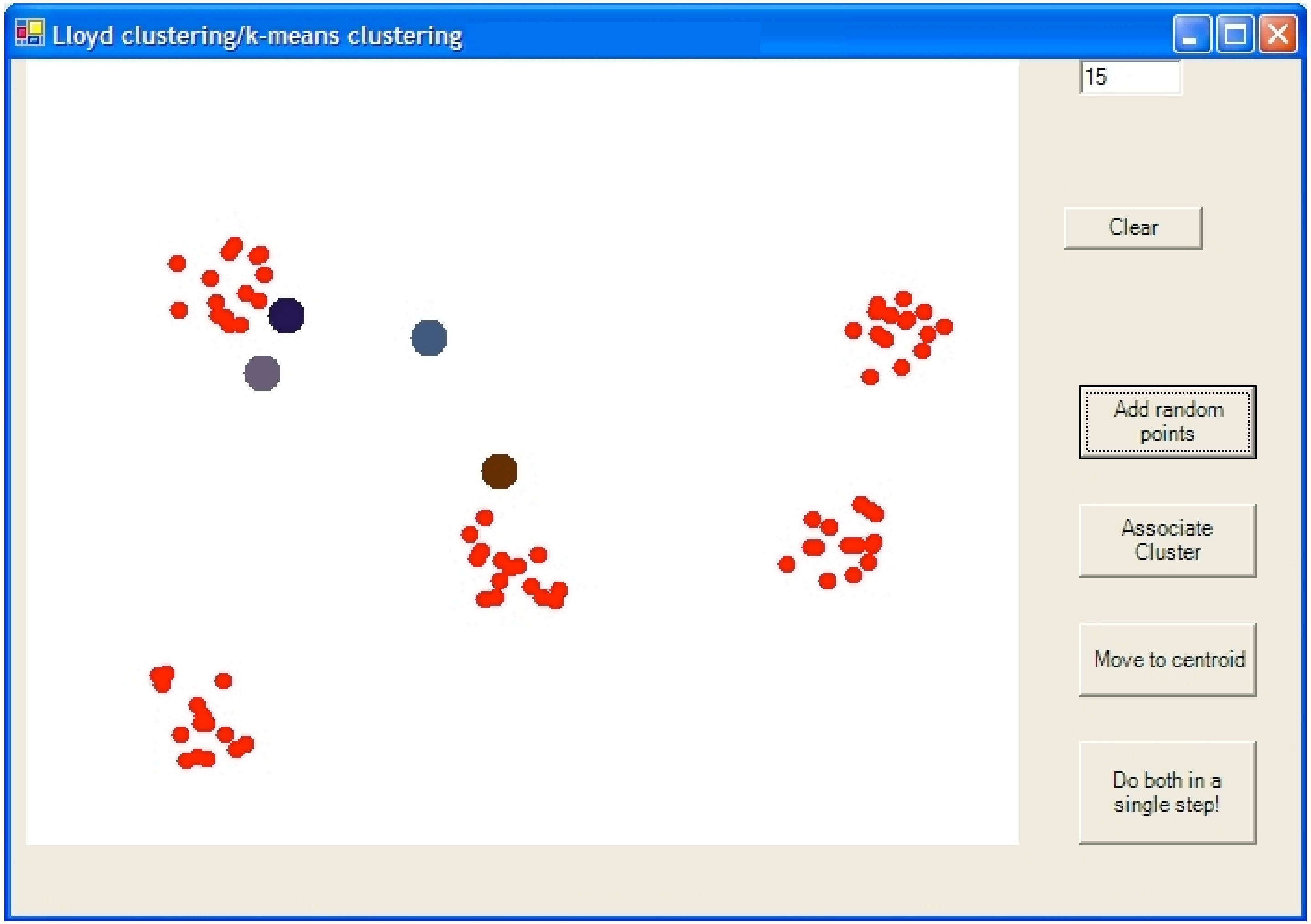
K-means

- Simplest clustering algorithm is k-means
- To run requires fixing K=#(Clusters)
- Requires an Euclidean type embedding
- We are attempting to minimizing a loss function:

$$L = \sum_{k=1}^{K} \sum_{x_i \in C_k} \left(x_i - \mu_k \right)^2$$

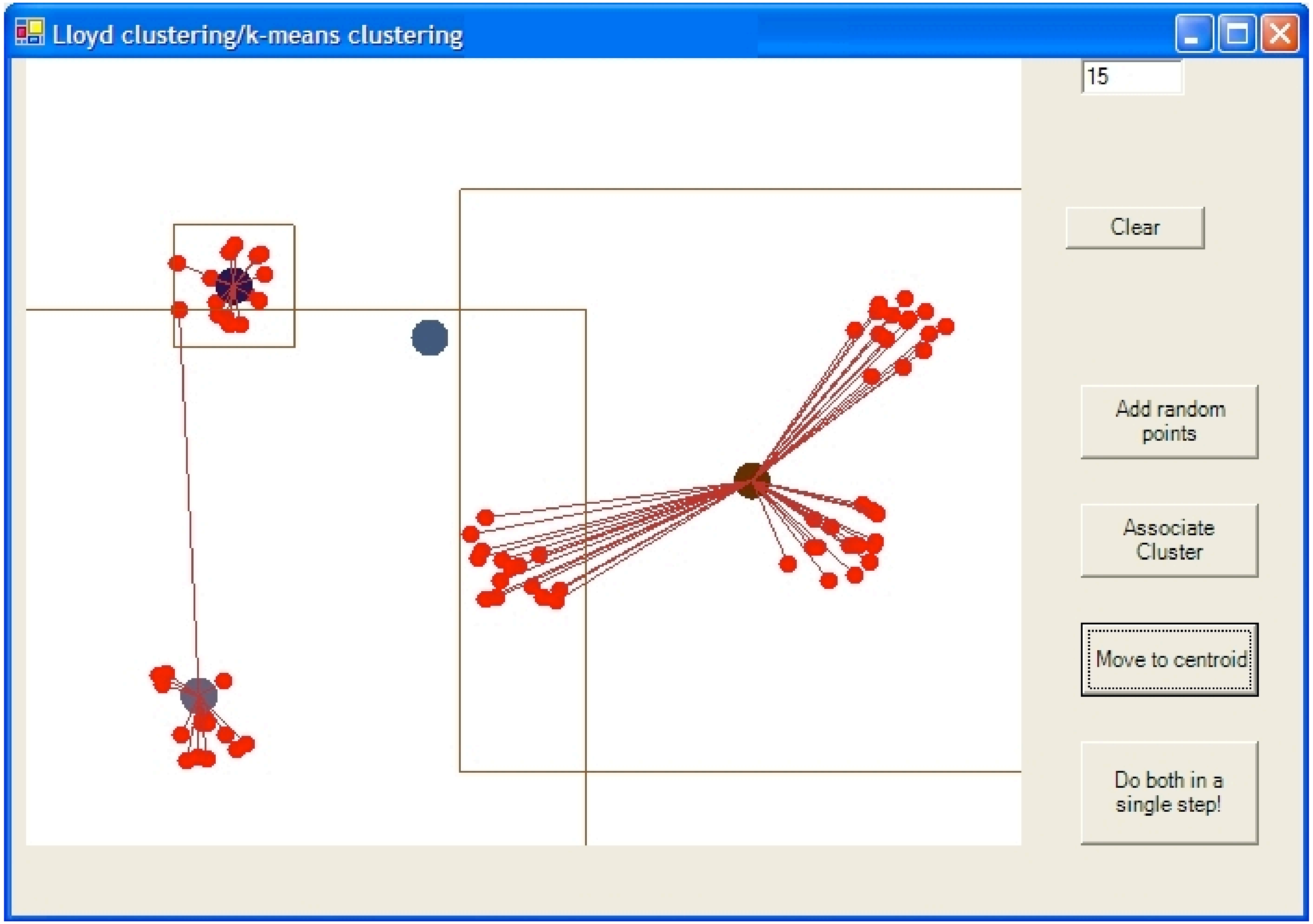
K-means algorithm:

1. Randomly choose points in each cluster and compute centroids.

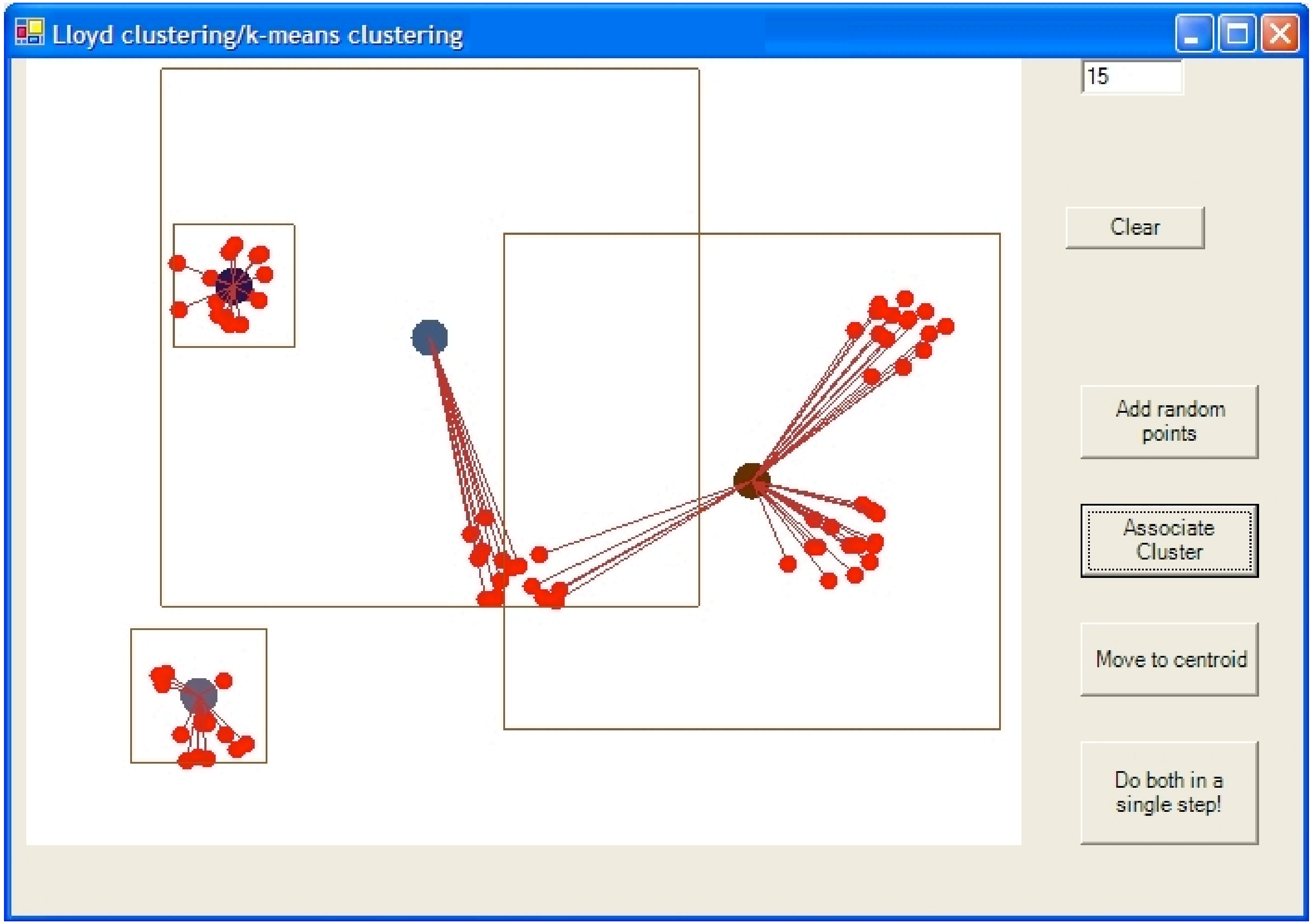


Example from: http://en.wikipedia.org/wiki/K-means_algorithm

2. Organize points by distance to the centroids.
3. Update centroids

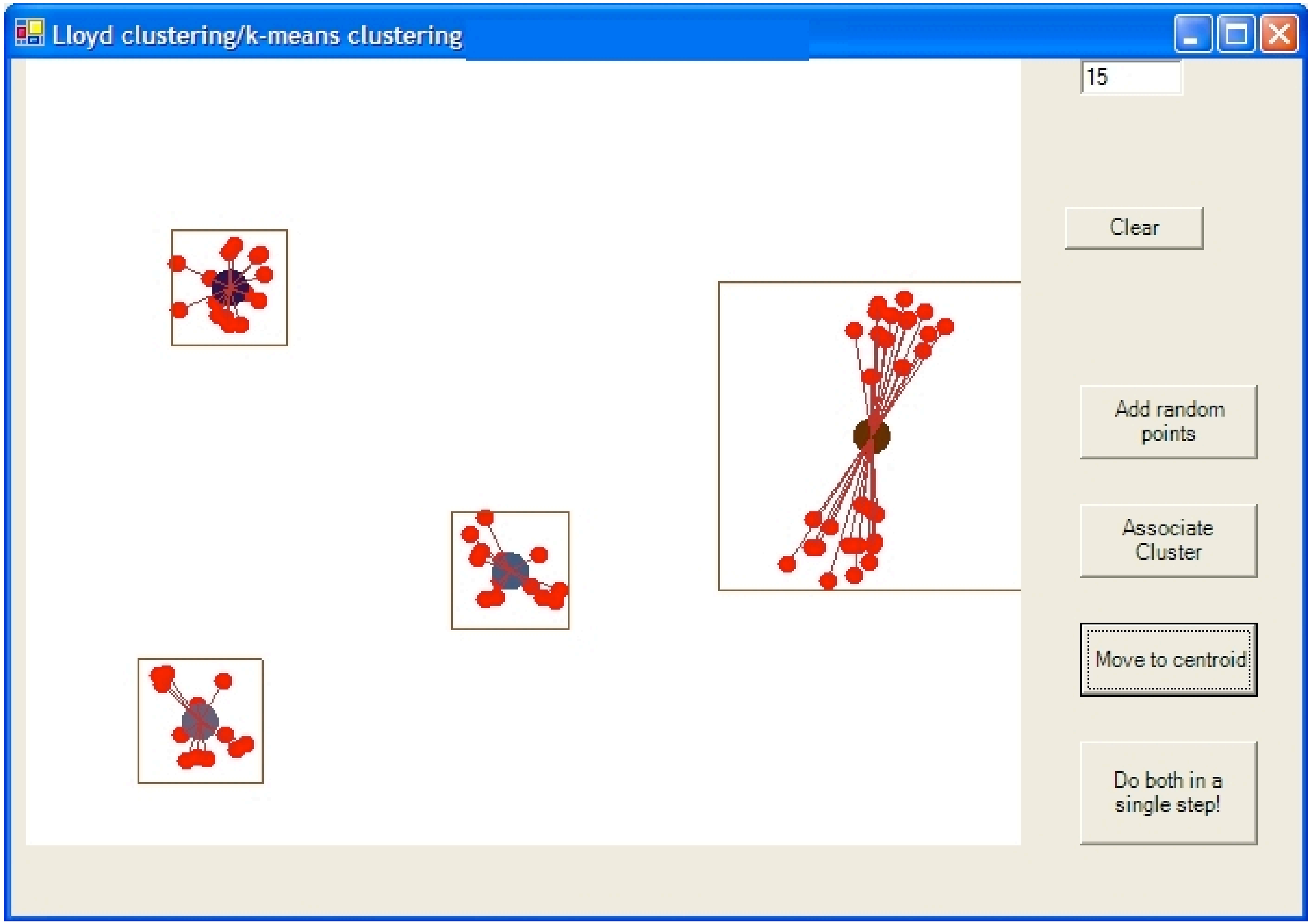


4. Repeat...



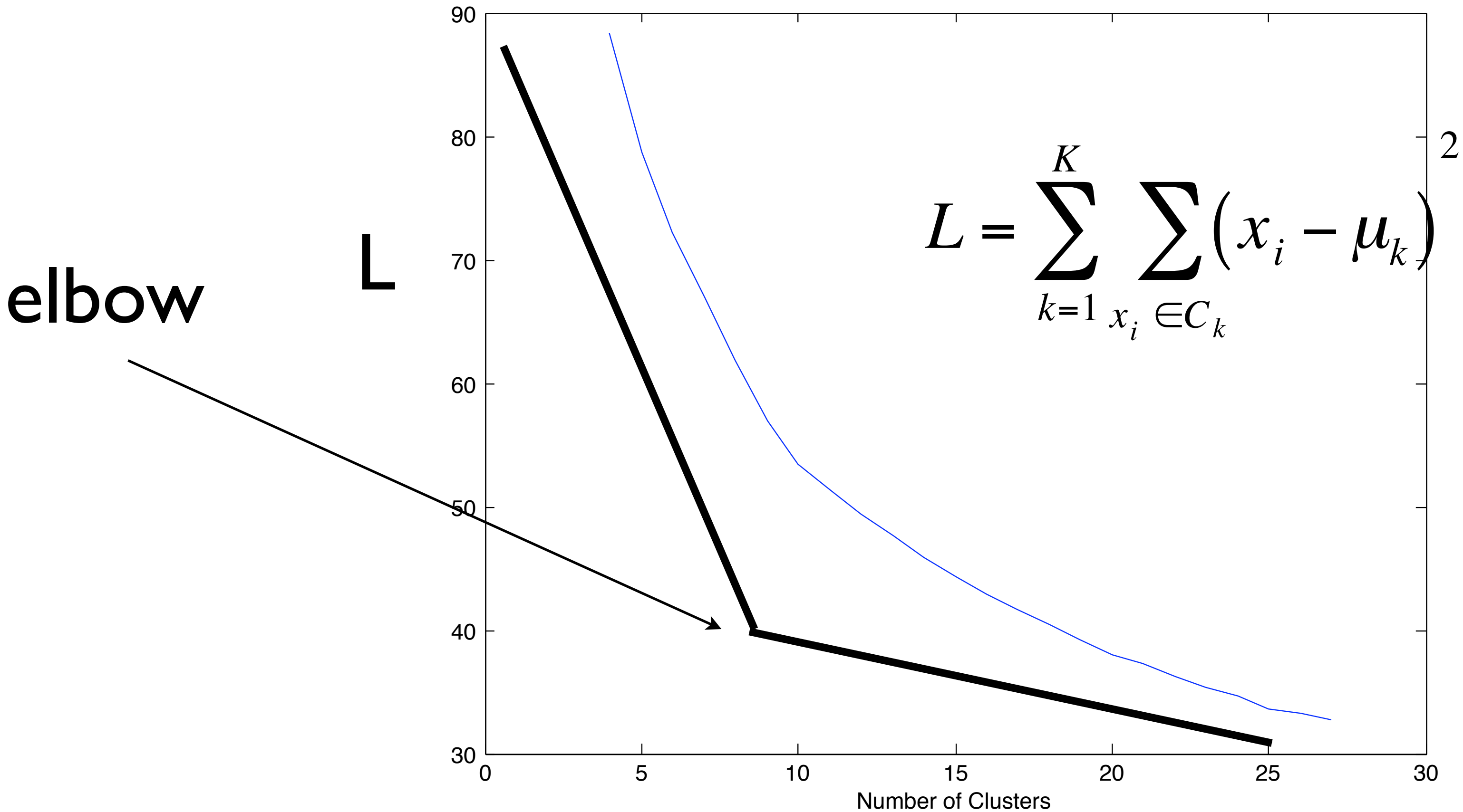
Lloyd clustering/k-means clustering
15
Clear
Add random points
Associate Cluster
Move to centroid
Do both in a single step!

...until stable.



A hard part is choosing K=#(Clusters)

Elbowlogy:



$$L = \sum_{k=1}^{K} \sum_{x_i \in C_k} \left(x_i - \mu_k \right)^2$$

Though is practice this rarely works in a complex multi-scalar system system

Now we can relax the Vagabond Problem and search for a ω-orthonormal set $\{f_i\}_{i=1}^{L}$ such that

$$\sum_i^L ||f_i||_{Dir}^2$$

is minimized.

How?

$< f_i, \Delta_\omega f_i >_\omega = ||f_i||_{Dir}^2$ is being used as a quadradic form, hence each term in the sum remains the same if we view Δ_ω an the Hermitian operator

$$C_\omega = \frac{\Delta_\omega + \Delta_\omega^*}{2}.$$

In other words,

$$< f, \Delta_\omega f >_\omega = < f, C_\omega f >_\omega$$

and we have replaced the minimization problem with minimizing

$$\sum_i \frac{< f_i, C_\omega f_i >_\omega}{< f_i, f_i >_\omega}.$$

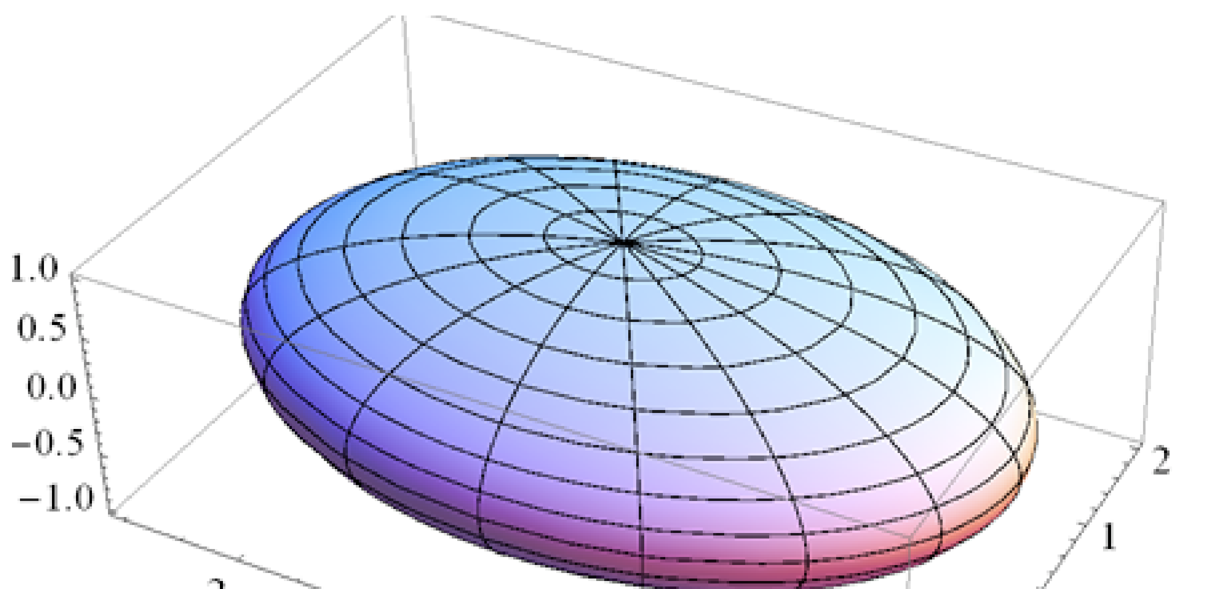
and we have replaced the minimization problem with minimizing

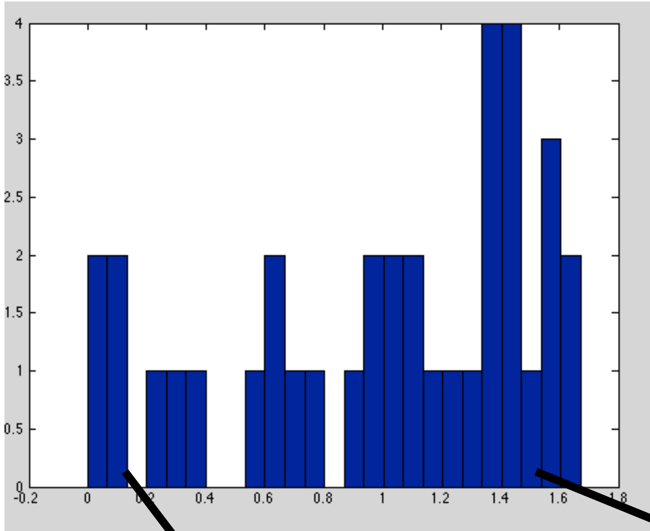
$$\sum_i \frac{< f_i, C_\omega f_i >_\omega}{< f_i, f_i >_\omega}.$$ where $$\langle f, C_\omega g \rangle = \langle C_\omega f, g \rangle$$

Why is this good, well it takes two to Tango!

C_ω is our *context operator* and being Hermitian allows us to use the spectral theorem and the Raileigh-Rtiz method to identify our Vagabond Embedding as the ω-unit length eigenfunctions of C_ω with the smallest eigenvalues.

Spectral Theorem: If <,> is a Hermitian inner product and <Av,w>=<v,Aw>, then there is an orthonormal basis of A eigenvectors.

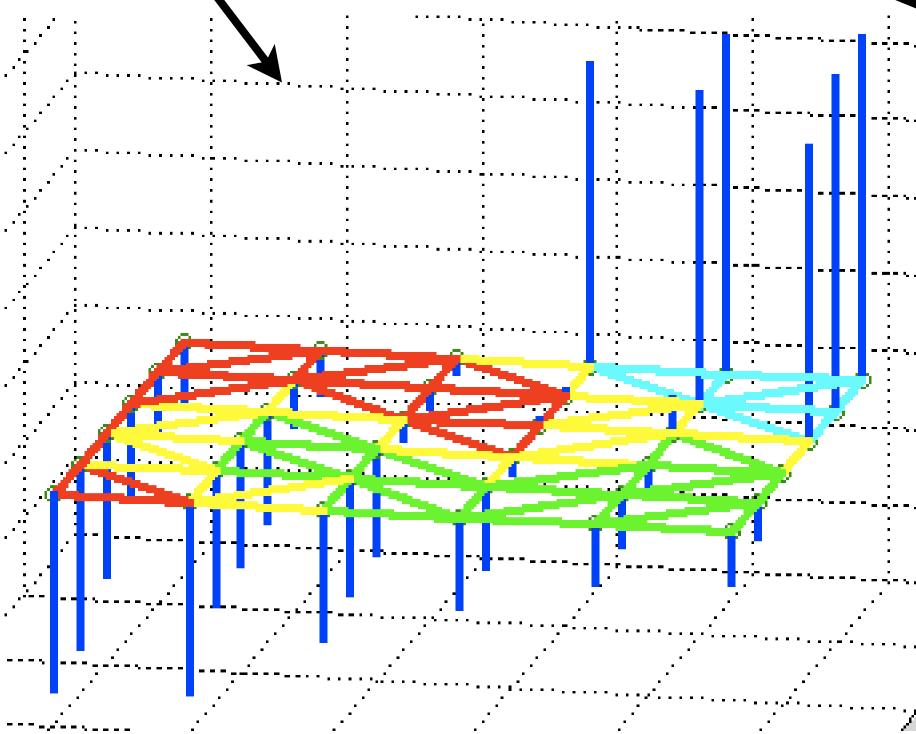




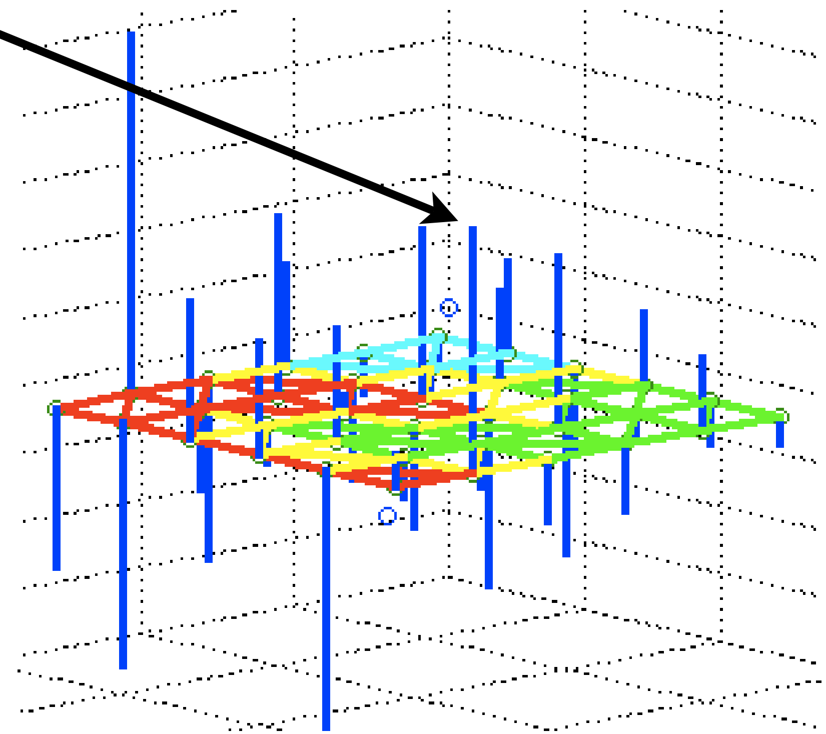
The Context Operator

$$C_\omega = \frac{\Delta_\omega + \Delta_\omega^*}{2}$$

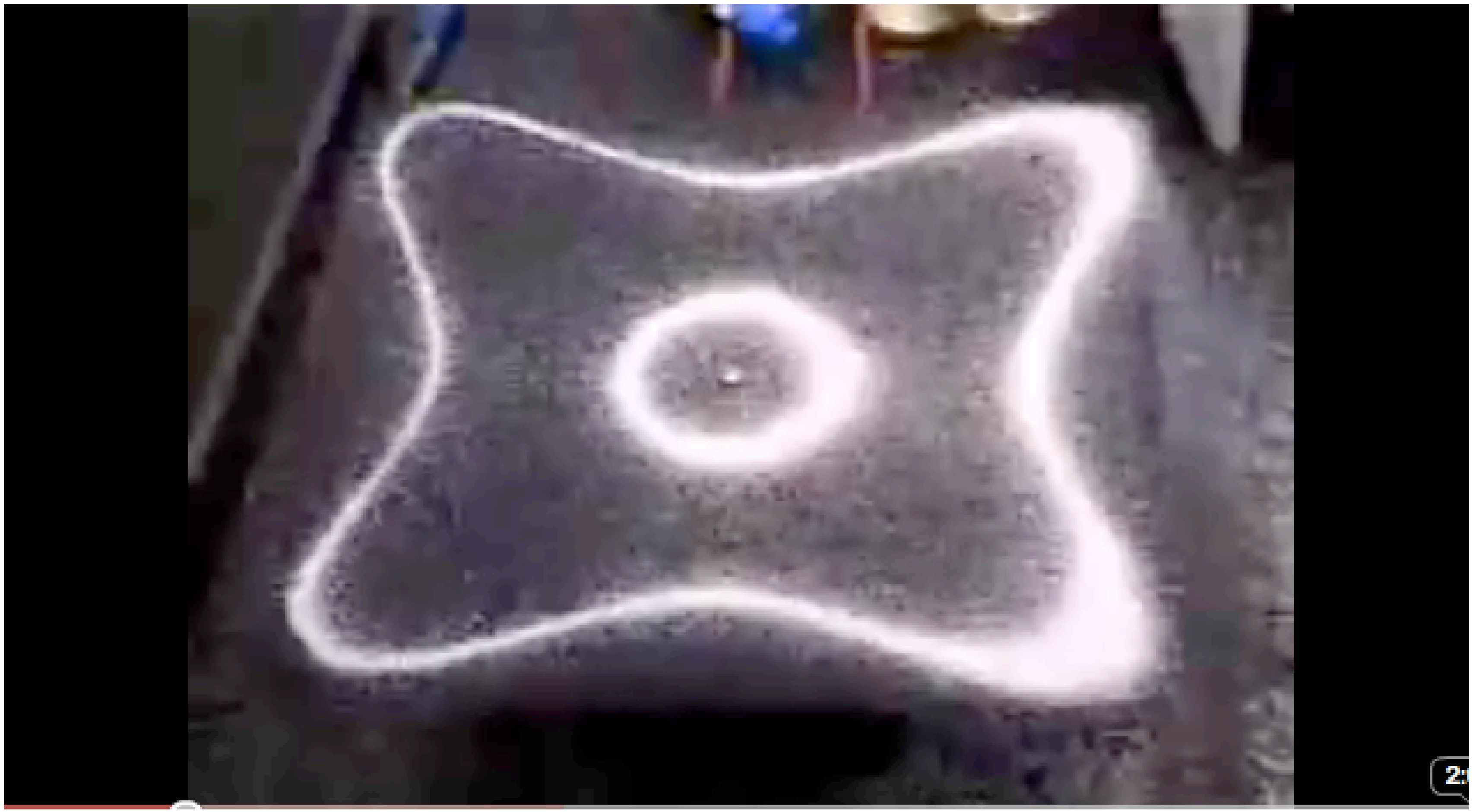
Information in Eigenfunctions Not Local



λ_2



λ_{30}



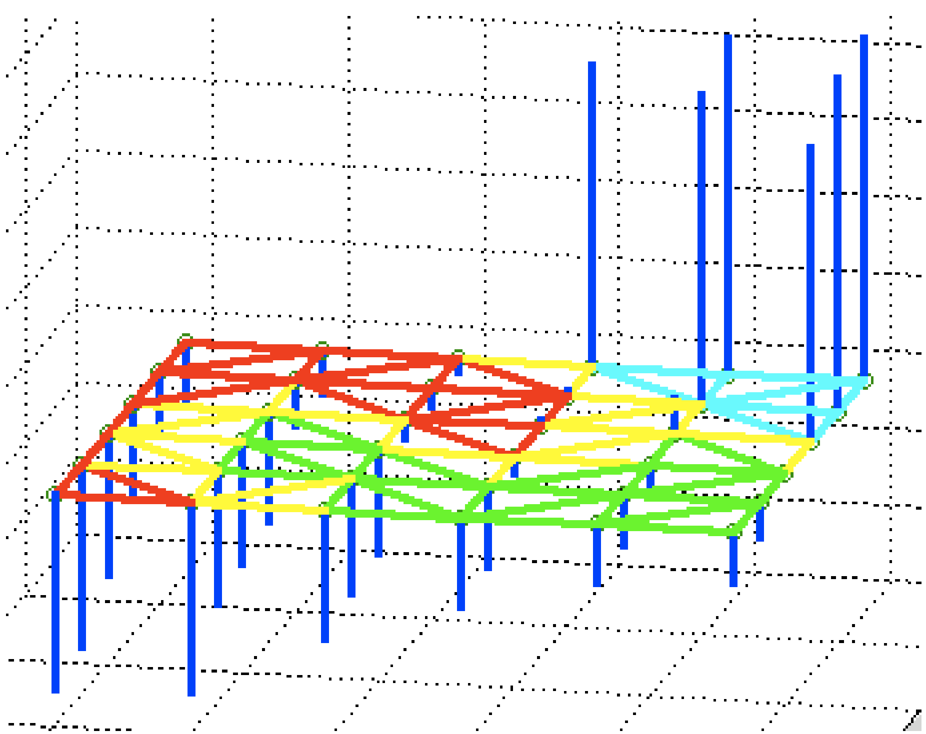
http://www.youtube.com/watch?v=Uu6Ox5LrhJg&feature=response_watch

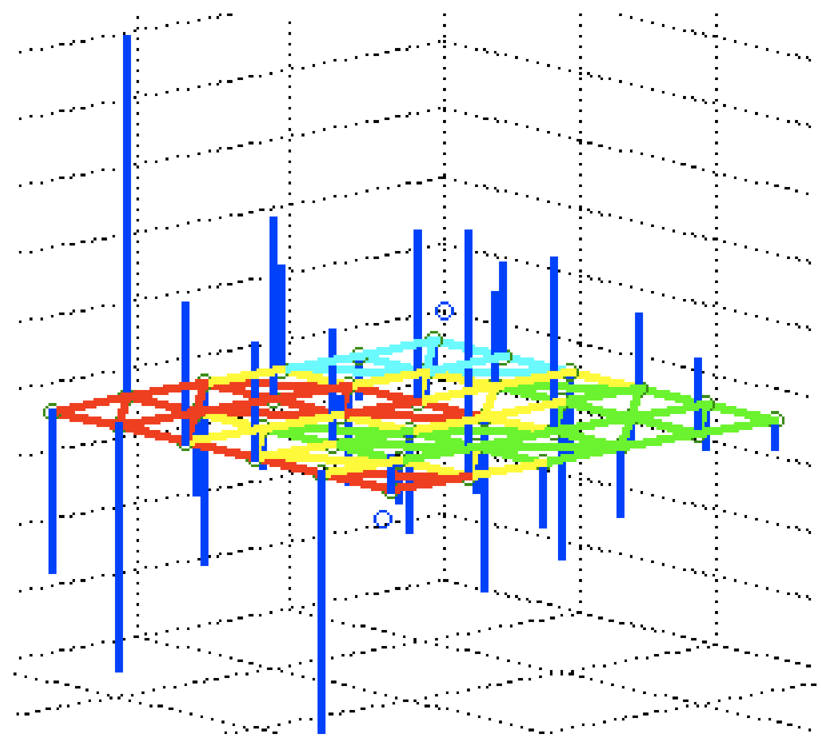
The Green-Kelvin Identity

$$\langle f, \Delta f \rangle = \int f \Delta f \, d\vec{x} = \int |\nabla f|^2 \, d\vec{x} \qquad \langle f, g \rangle = \int_M f(x) g(x) \, dVol(x)$$

$$\langle f, g \rangle = \sum_i f_i w^i g_i$$

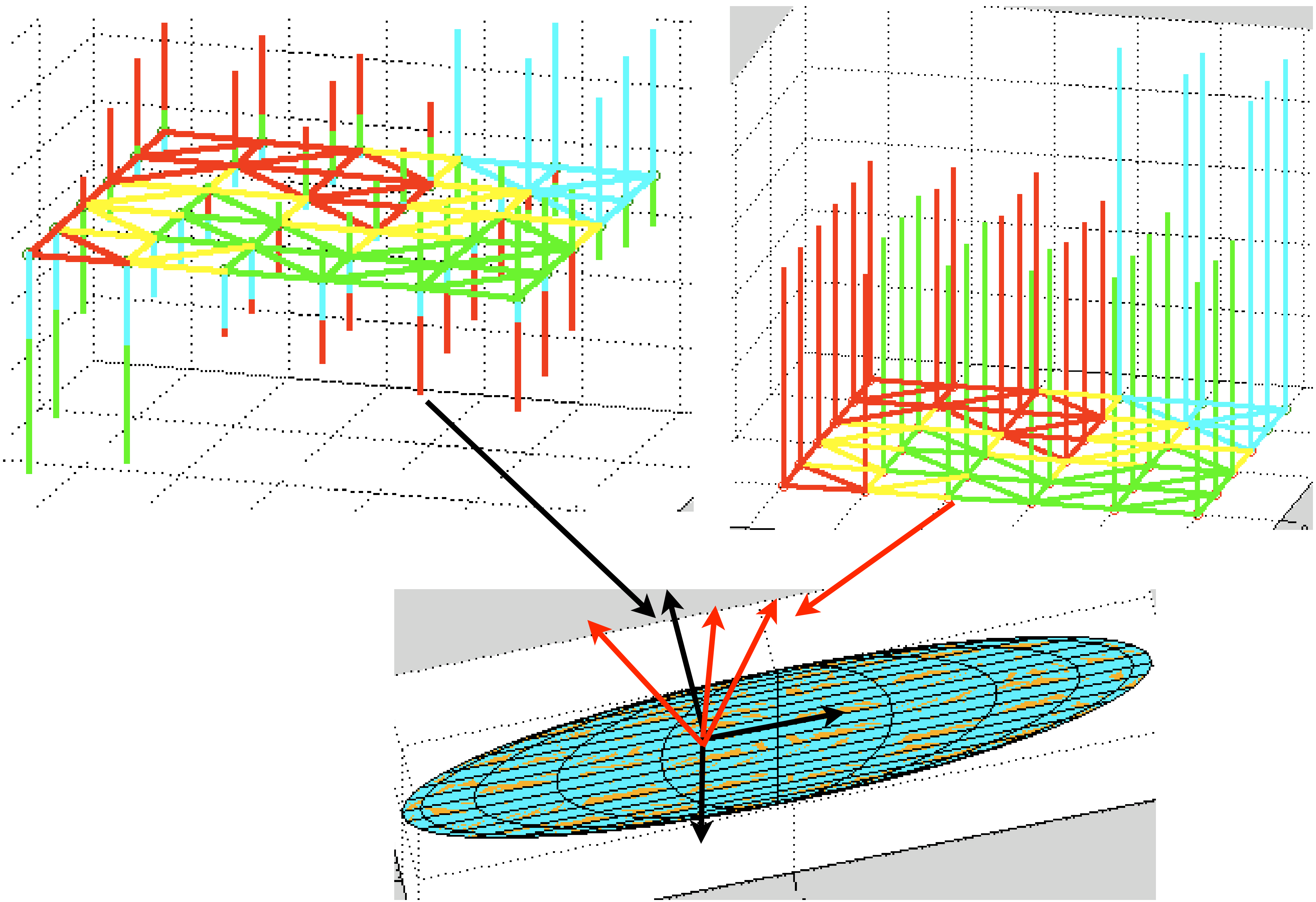
$$\langle f, \Delta f \rangle = \sum_{i,j} f_j w^i (I_i^{\ j} - P_i^{\ j}) f_j = \frac{1}{2} \sum_{i,j} w^i P_i^{\ j} \left(f_i - f_j \right)^2$$



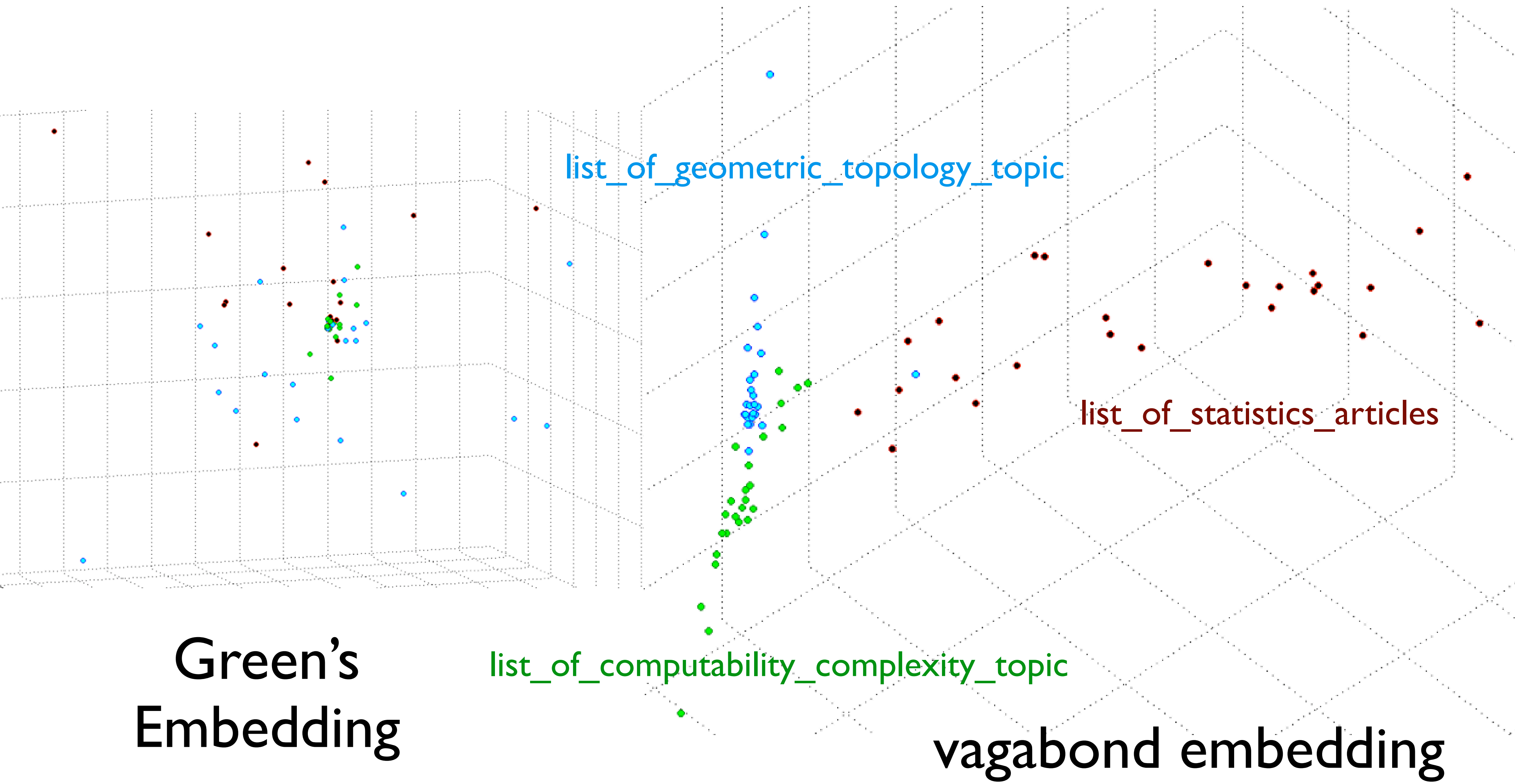


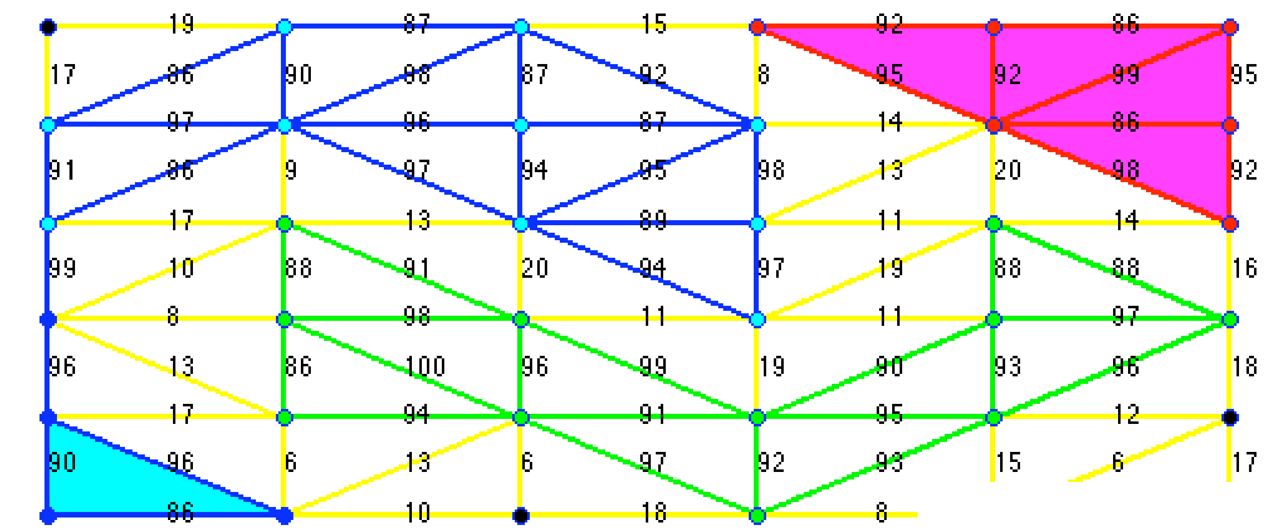
First three (non-trivial) eigenfunctions

The perfect Vagabond functions



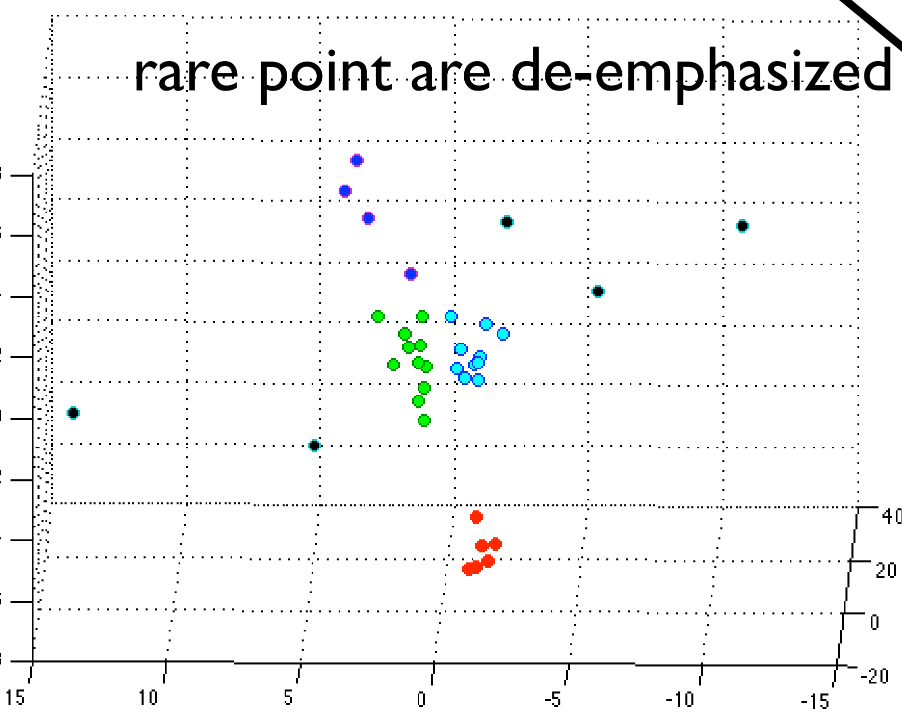
Space of mathematics

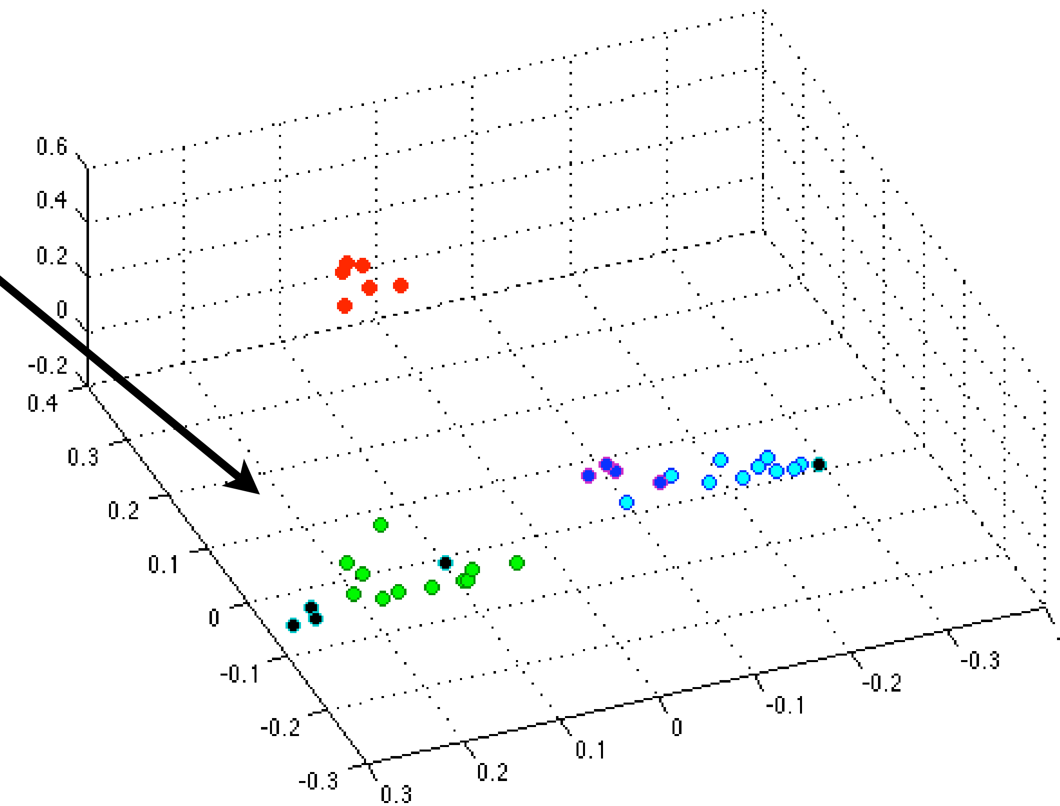




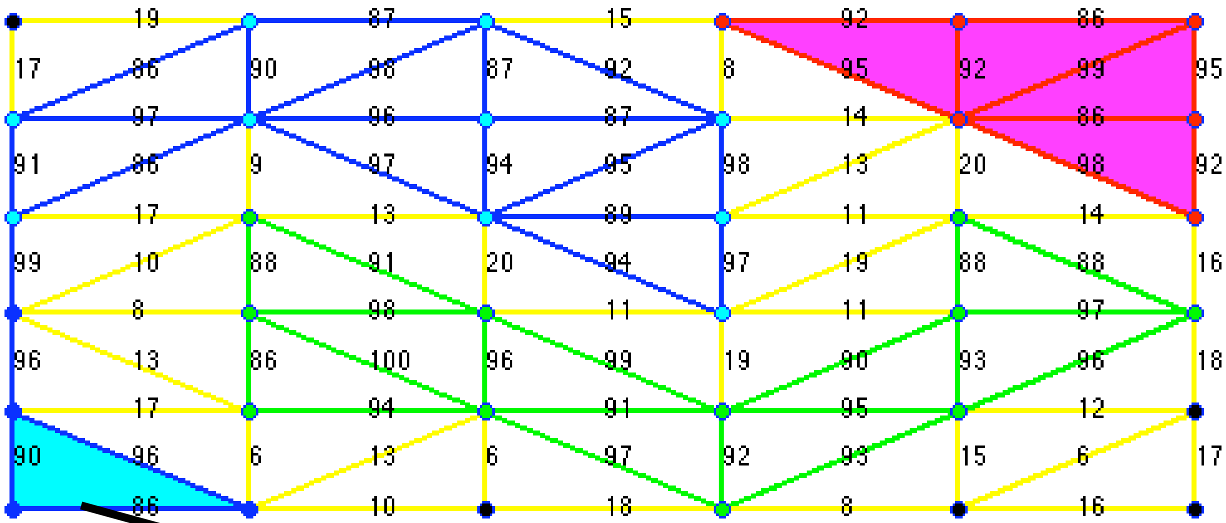
$$\tau_i = 1$$

rare point are de-emphasized



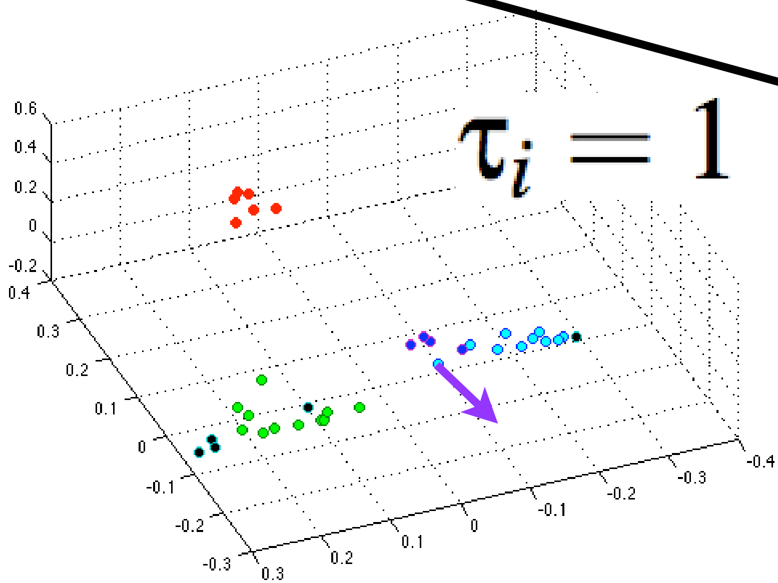


$$\tau_0(blue) = 5,\ \tau_0(magenta) = 1/5, \text{ and } \tau_0(other) = 1$$

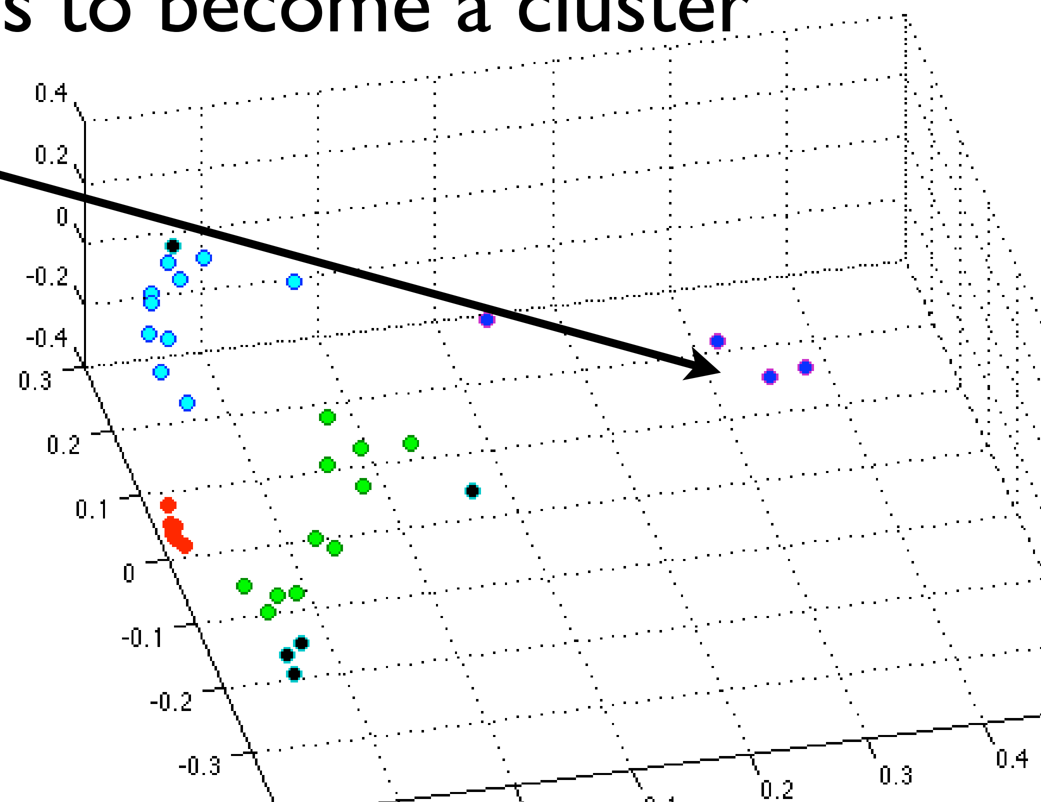


$$\tau = \frac{\tau_0}{\sum \tau_0 \pi}$$

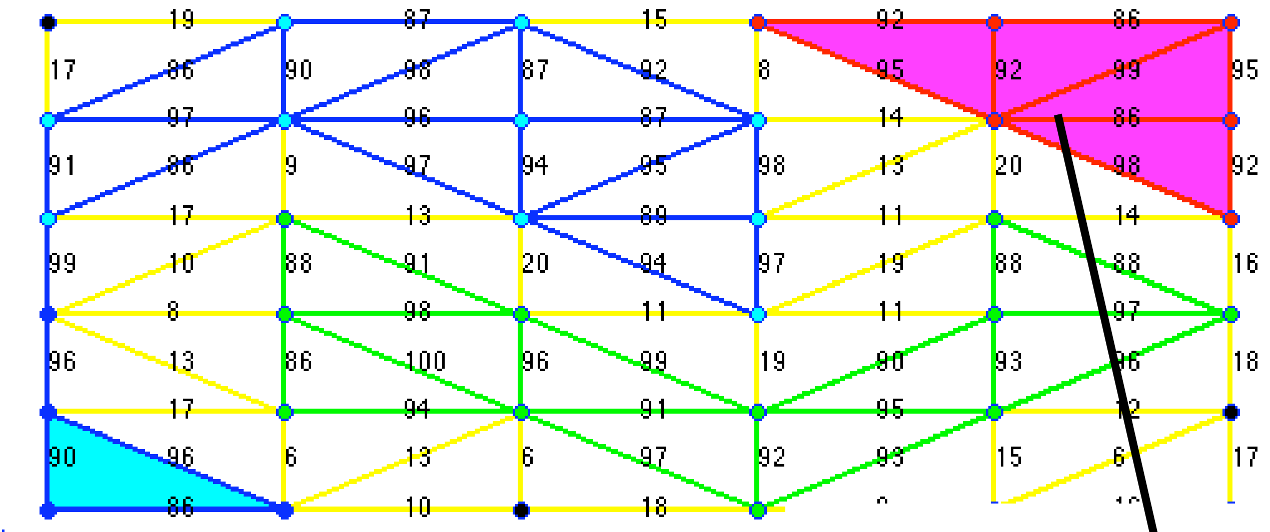
tries to become a cluster



$$\tau_i = 1$$

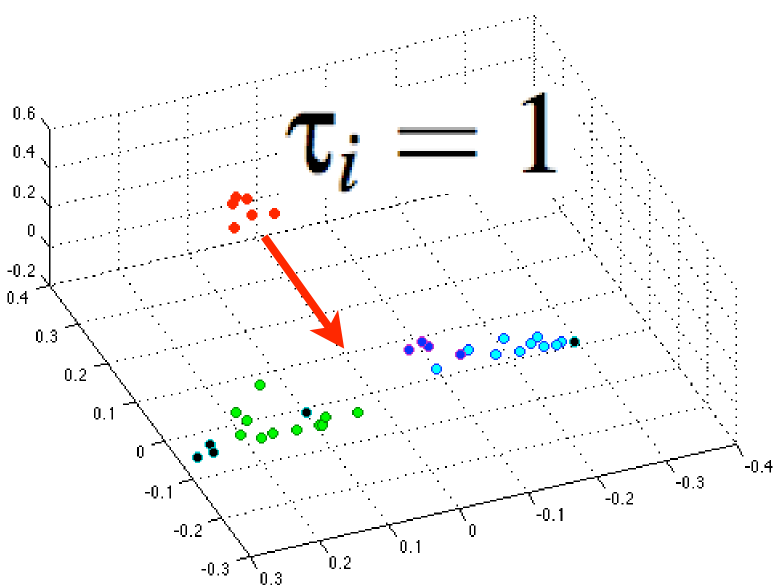


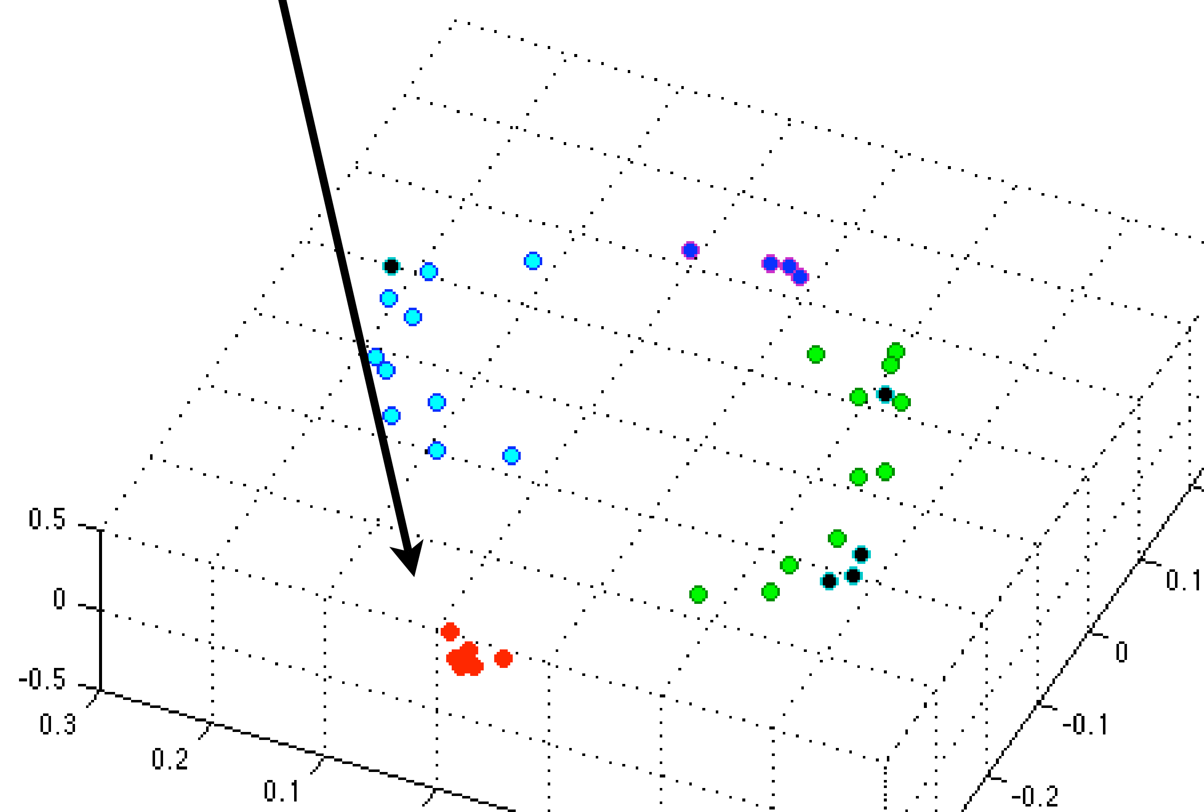
$$\tau_0(blue) = 1/5,\ \tau_0(magenta) = 1/5, \text{ and } \tau_0(other) = 1$$



$$\tau = \frac{\tau_0}{\sum \tau_0 \pi}$$

$$\tau_i = 1$$





Kakawa's Salted Carmel time! (well, except Kakawa wasn't open)

What is the difference?

figure 1

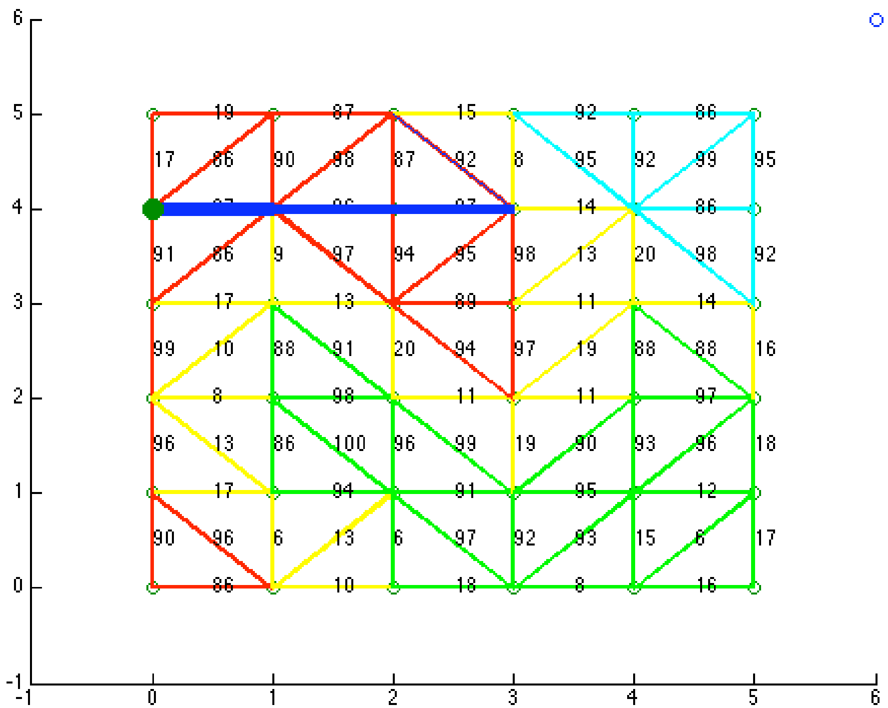
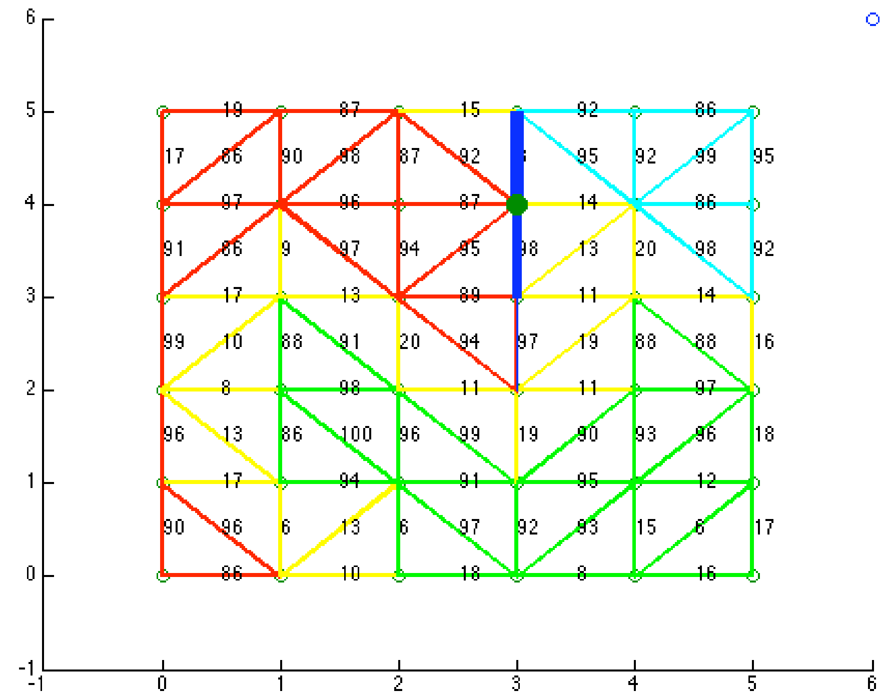


figure 2



```
close all
Show=0; N=100;
figure(1); [states WR TR]=SnakeRev(N,5,'Vg',Show,1,1,1,1,1,0,0,0,0,0,0,0);
figure(2); [states WR TR]=SnakeRev(N,5,'Vg',Show,0,1,1,1,1,0,0,0,0,0,0,0);
```

The answer:

figure 1

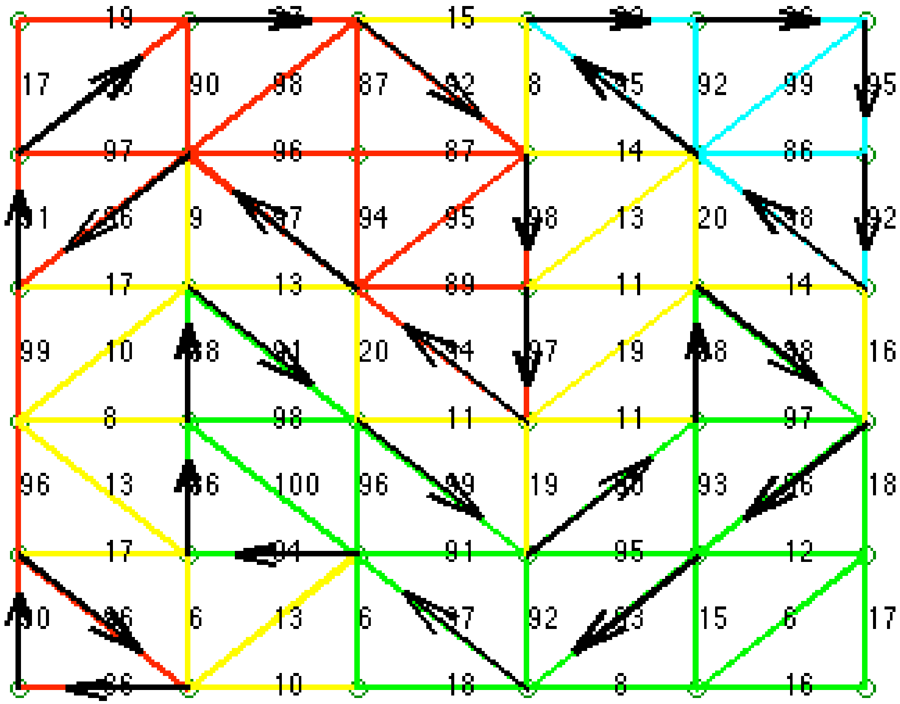
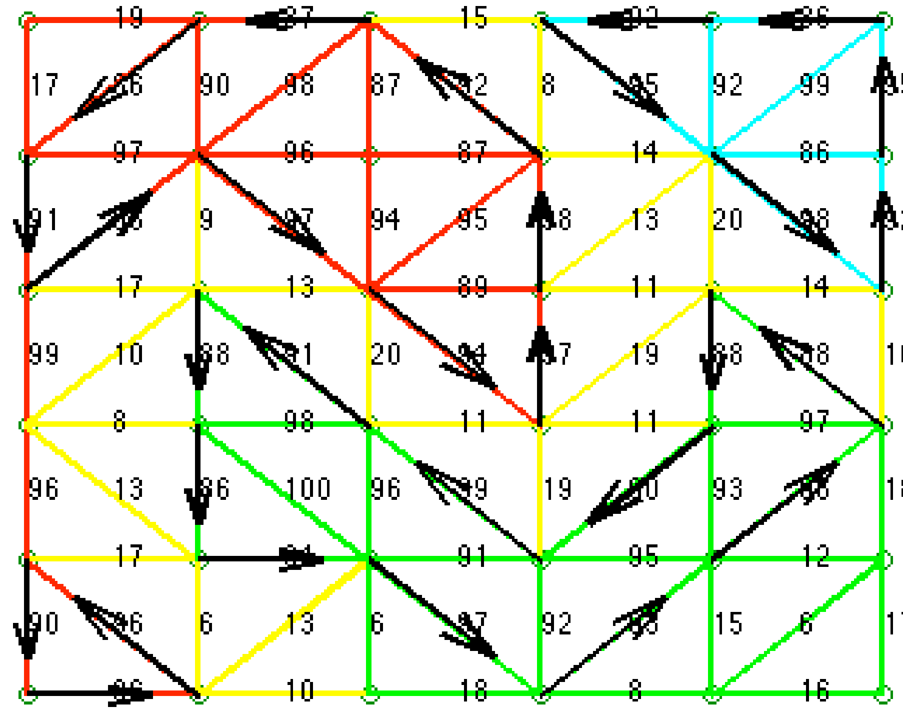


figure 2



```
close all
Show=1; N=100;
figure(1); [states WR TR]=SnakeRev(N,5,'Vg',Show,1,1,1,1,1,0,0,0,0,0,0,0);
figure(2); [states WR TR]=SnakeRev(N,5,'Vg',Show,0,1,1,1,1,0,0,0,0,0,0,0);
```

Reversibility

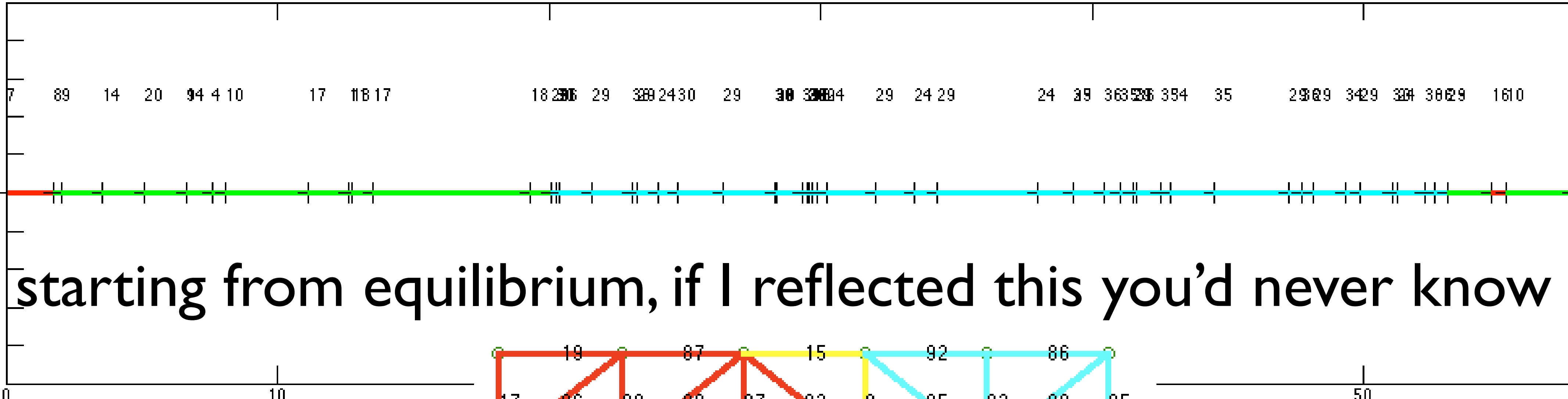
$$P^* = [\pi]^{-1} P^{tr} [\pi]$$

$$\textit{reversible} \text{ if } P = P^*$$

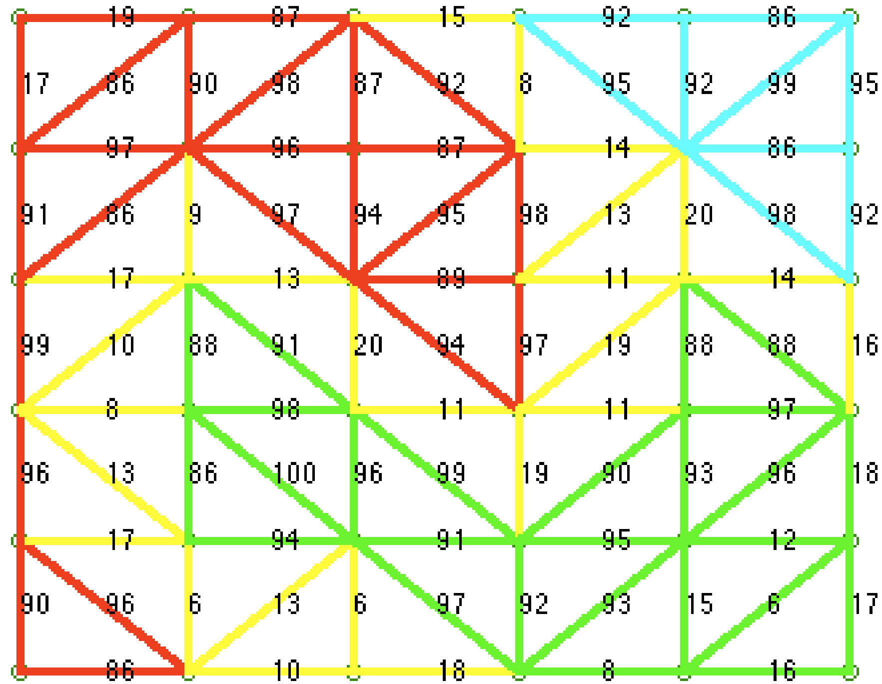
Theorem: Reversible if and only if there is symmetric conductance matrix such that

$$P = [W\mathbf{1}]^{-1} W$$

$$\pi = W\mathbf{1} / \left(\mathbf{1}^{tr} W \mathbf{1}\right)$$



starting from equilibrium, if I reflected this you'd never know



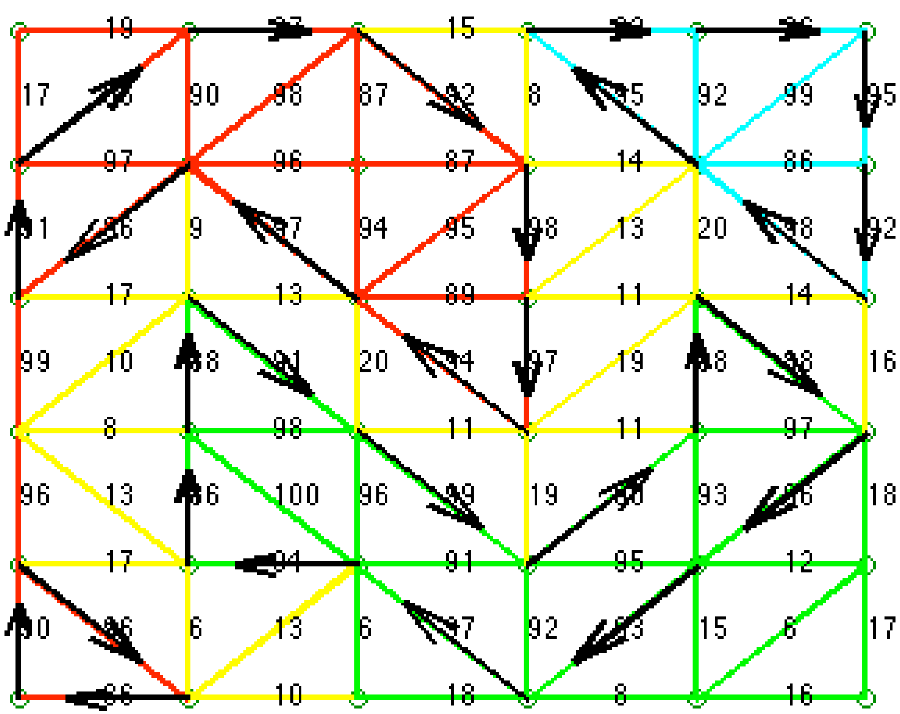
N=200; K=10; [states]=PlotTheState(N,K);

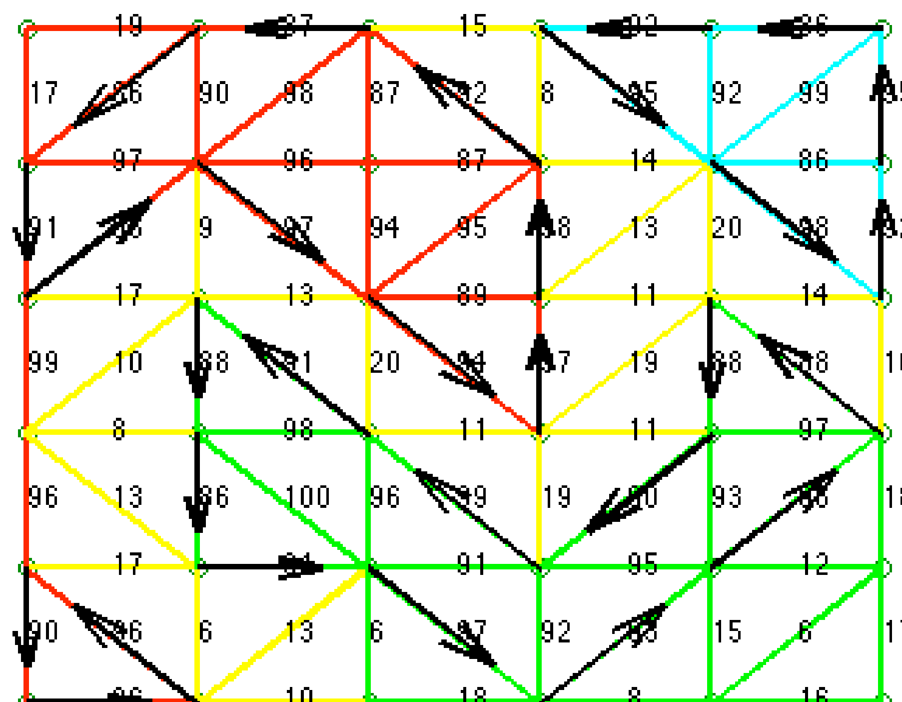
Theorem: For a reversible chain, the vagabond embedding is the PCA of the green's embedding weighted by the equilibrium vector

Theorem: For any chain P, there is a unique (up to a multiplicative constant) conductance W and divergence free, compatible flow F such that

$$P = [W\mathbf{1}]^{-1}(W+F) \qquad P^* = [W\mathbf{1}]^{-1}(W-F)$$

$$\pi = W\mathbf{1}/\left(\mathbf{1}^{tr}W\mathbf{1}\right)$$





Flow Independence Corollary: The vagabond embedding and C_ω depend only on on the Conductance W and and the equilibrium measure ω, and are independent of the flow F.

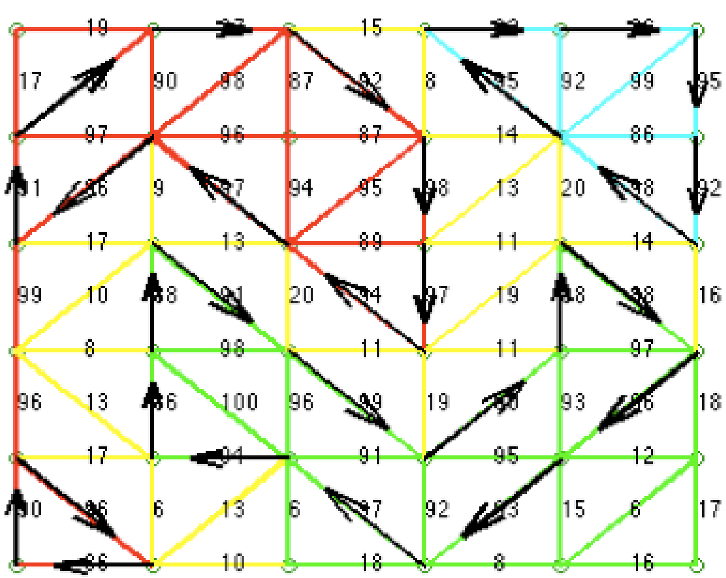
proof: Simply note

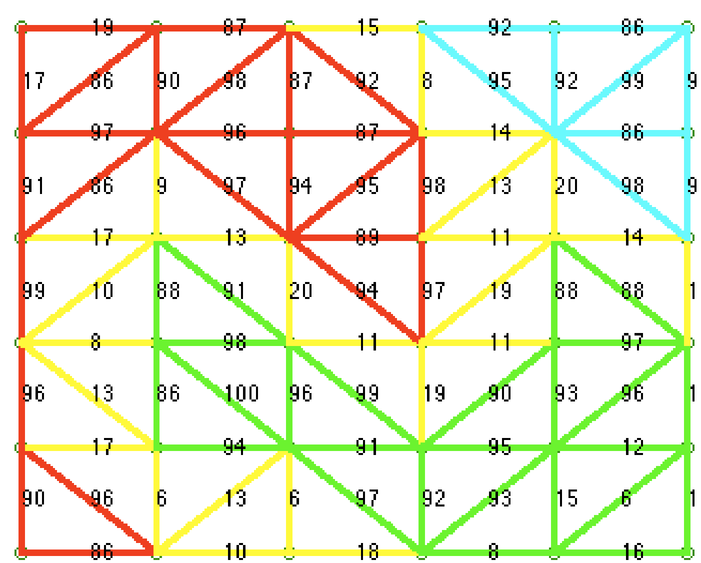
$$C_\omega = \frac{\Delta_\omega + \Delta_\omega^*}{2} = [\tau]^{-1}\left(I - \frac{P+P^*}{2}\right) = [\tau]^{-1}\left(I - [\pi]^{-1}W\right)$$

we have:

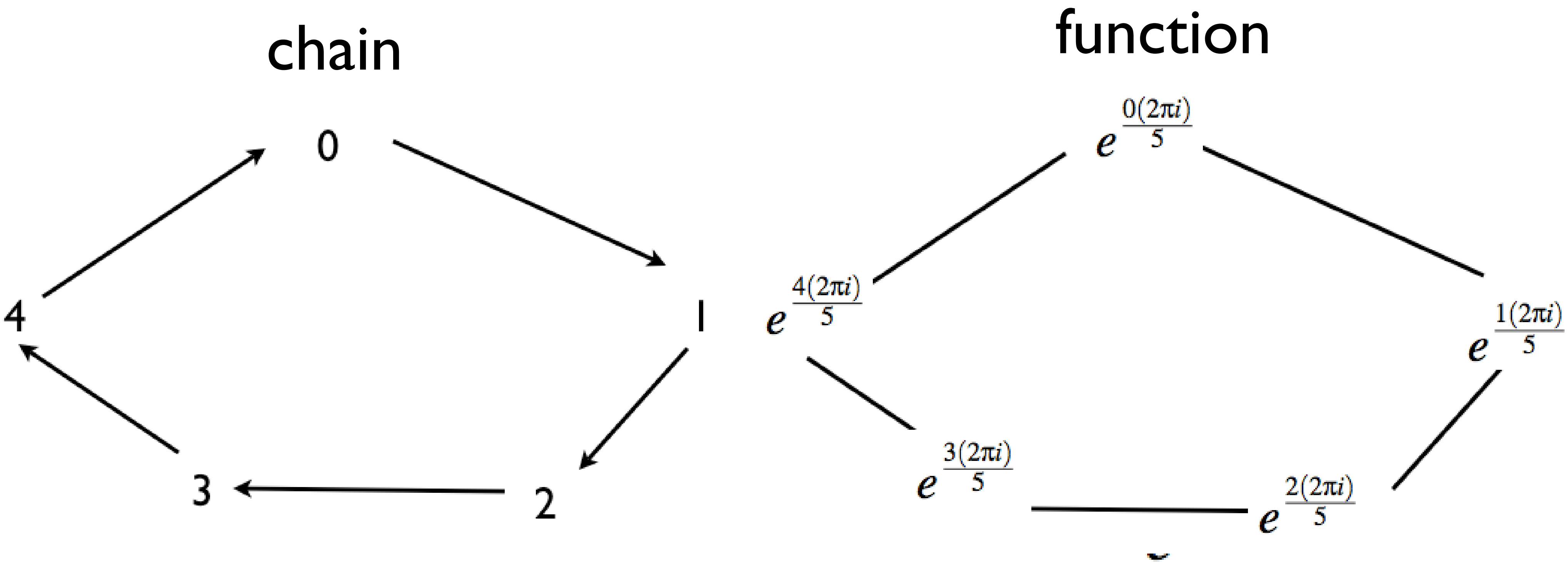
Q.E.D

Clearly the vagabond is missing half the geometry.





Here we are going explore a scenario in our space. This scenario is going to be 'cycle like behavior', and hence we must ask: what does a directed cycle look like from the point of view of the functions on the chain?



Shall we dance?

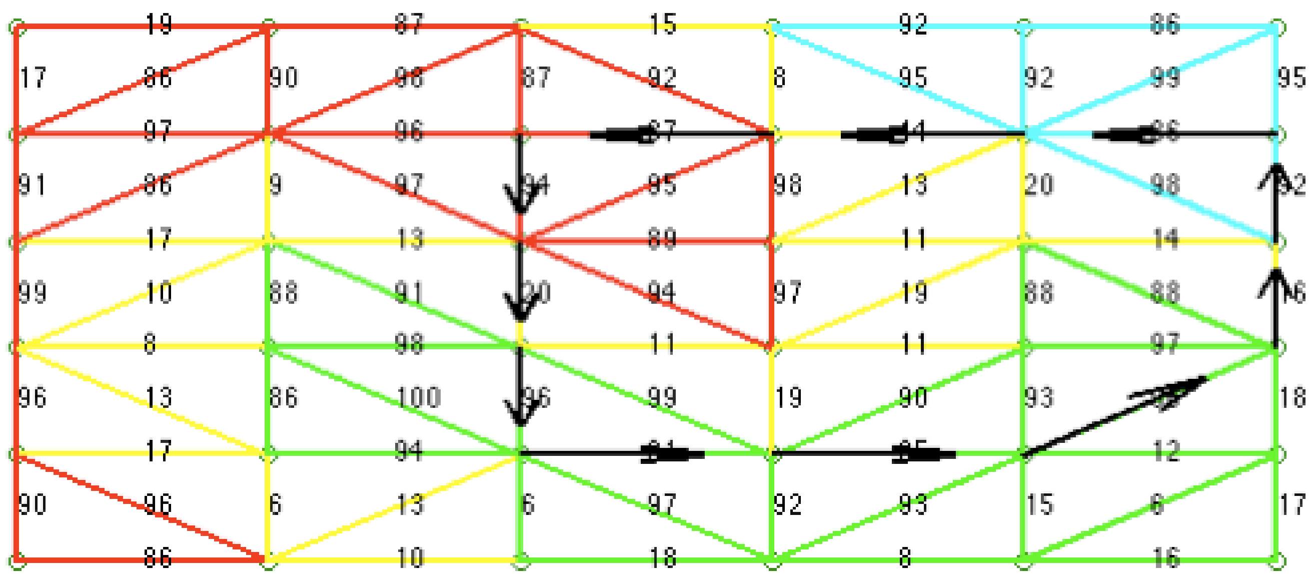
$$e^{-i\left(\frac{K2\pi}{N}\right)n}P^n f = f$$

More generally, cycles will support functions where there is a relatively large frequency κ such that

$$e^{-i\kappa t}e^{-t\Delta}f$$

is pointwise relatively stable. One way to find largish κ such the function f and κ pair locally minimize $\left|\left|\frac{d}{dt}e^{-(i\kappa+\Delta)t}f\right|\right|_{\omega}^{2}$ at $t=0$. So we are looking for critical points of

$$F(\kappa,f)=\left.\left|\left|\frac{d}{dt}e^{-(i\kappa+\Delta)t}f\right|\right|_{\omega}^{2}\right|_{t=0}.$$

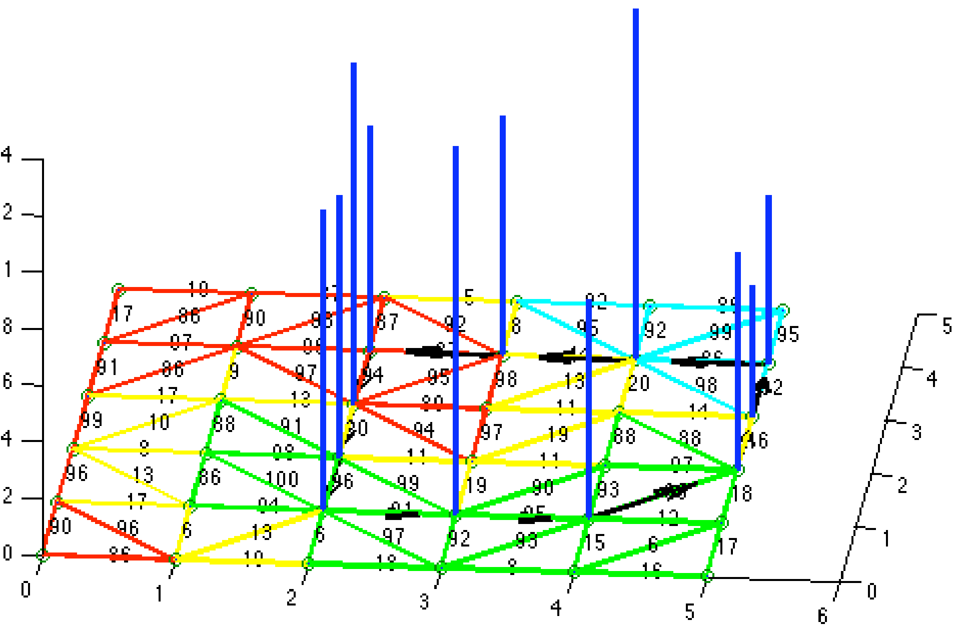


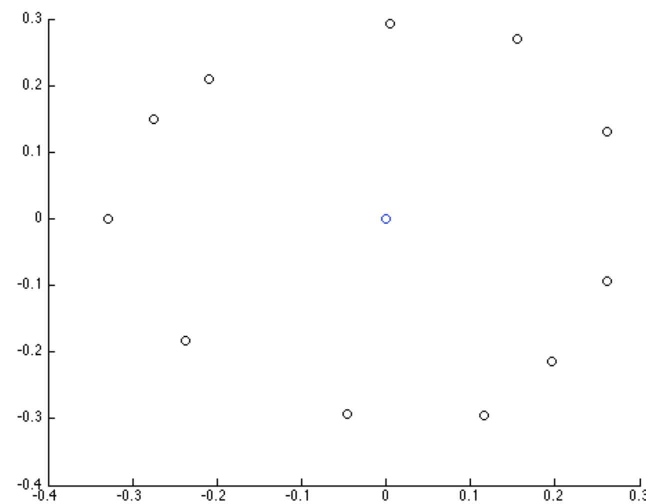
More generally, cycles will support functions where there is a frequency κ such that

$$e^{-i\kappa t}e^{-t\Delta}f$$

is pointwise relatively stable. One way to find such functions would be to minimize $\left|\left|\frac{d}{dt}e^{-(i\kappa+\Delta)t}f\right|\right|_{\omega}^{2}$ at t=0. Hence we are looking for critical points of

$$F(\kappa, f) = \left.\left|\left|\frac{d}{dt}e^{-(i\kappa+\Delta)t}f\right|\right|_{\omega}^{2}\right|_{t=0},$$

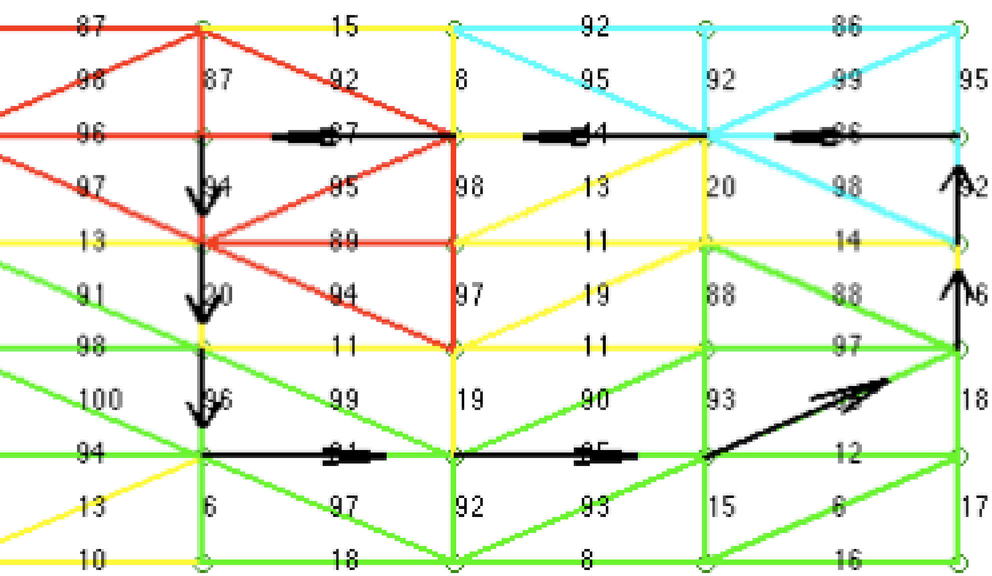




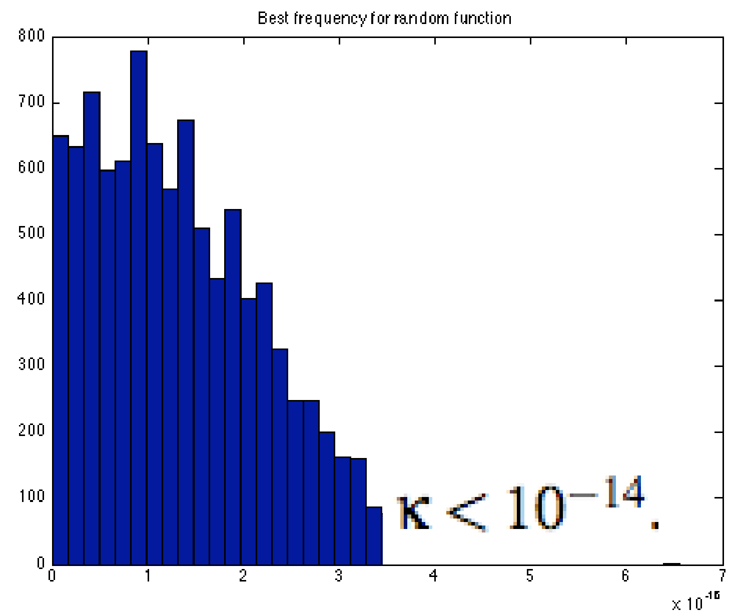
$$\kappa = 0.54112634652896\,6$$

We might call these local minima cycle functions. To get a sense for ‘largish’, putting a cycle on our toy chain as in figure 1, we can compare the function in the figure 5, which for $\kappa = 0.541126346528966$ has a nearly zero derivative, to a random function (real and imaginary parts random in $[0,1]$ normalized to have norm 1). For a random function we find its optimal κ for minimizing the derivative and after 10000 runs the largest such $\kappa < 10^{-14}$.

$$\kappa = 0.541126346528966$$



verses



How to find them...

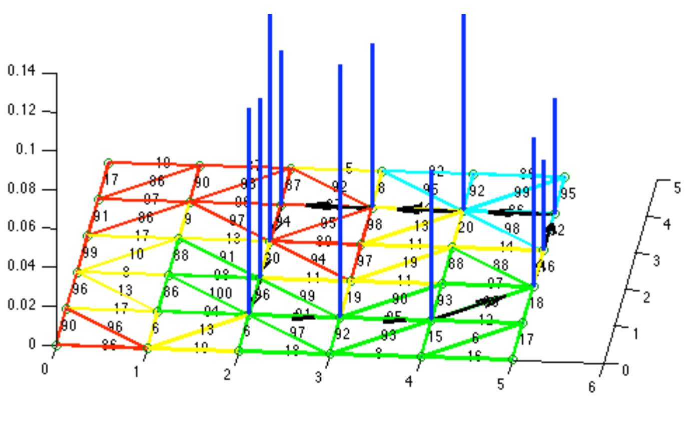
Cycle Detection Lemma: If $B_\omega f = \kappa f$ where $S_\omega = i\frac{\Delta - \Delta^*}{2}$, then $\frac{d}{d\kappa}F = 0$.

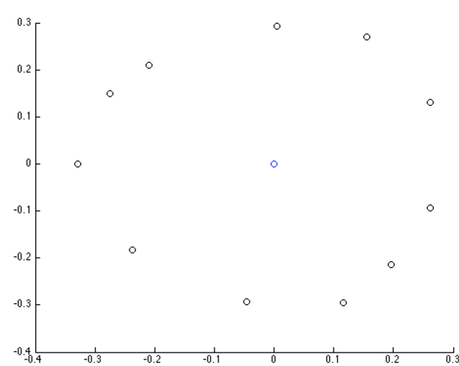
proof:

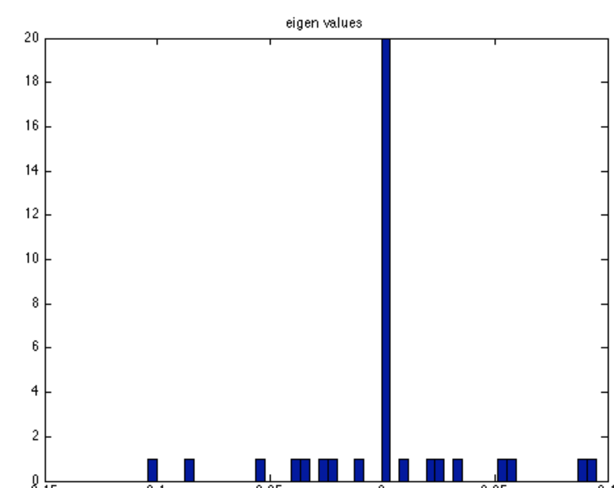
$$
\begin{aligned}
0 &= \frac{d}{d\kappa}\left.\left|\left|\frac{d}{dt}e^{-(i\kappa+\Delta)t}f\right|\right|_\omega^2\right|_{t=0} && \text{(definition)}\\
&= \frac{d}{d\kappa}\left.\left|\left|(i\kappa+\Delta)e^{-(i\kappa+\Delta)t}f\right|\right|_\omega^2\right|_{t=0} && \text{(matrix derivative, sec1.1ex??)}\\
&= \frac{d}{d\kappa}\left|\left|(i\kappa+\Delta)f\right|\right|_\omega^2 && (e^{At} = I \text{ at } t = 0 \text{)}\\
&= \frac{d}{d\kappa}\langle(i\kappa+\Delta)f,(i\kappa+\Delta)f\rangle_\omega && \text{(def. } <,>_\omega \text{)}\\
&= \langle if,(i\kappa+\Delta)f\rangle_\omega + \langle(i\kappa+\Delta)f,if\rangle_\omega && \text{(chain rule)}\\
&= \langle f,(\kappa-i\Delta)f\rangle_\omega + \langle(\kappa-i\Delta)f,f\rangle_\omega && \text{(conjugate linear)}\\
&= \langle f,((\kappa-i\Delta)+(\kappa-i\Delta)^*)f\rangle_\omega && \text{(def. adjoint)}\\
&= \langle f,2(\kappa-S_\omega)f\rangle_\omega && \text{(compute adjoint)}
\end{aligned}
$$

Q.E.D

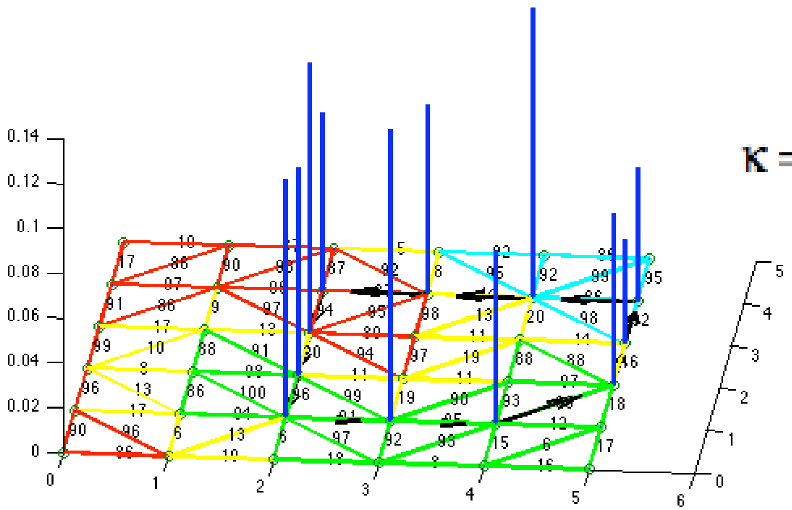
$\kappa = 0.541126346528966$



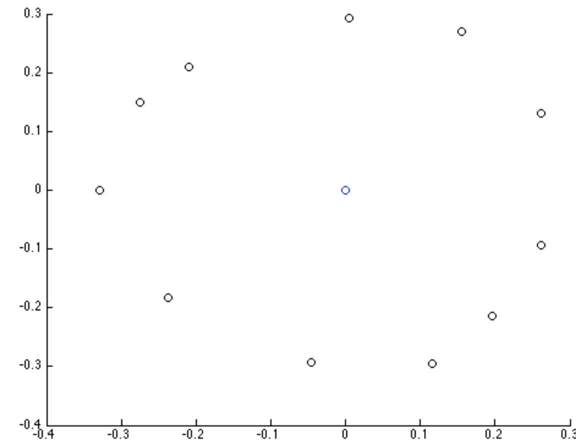




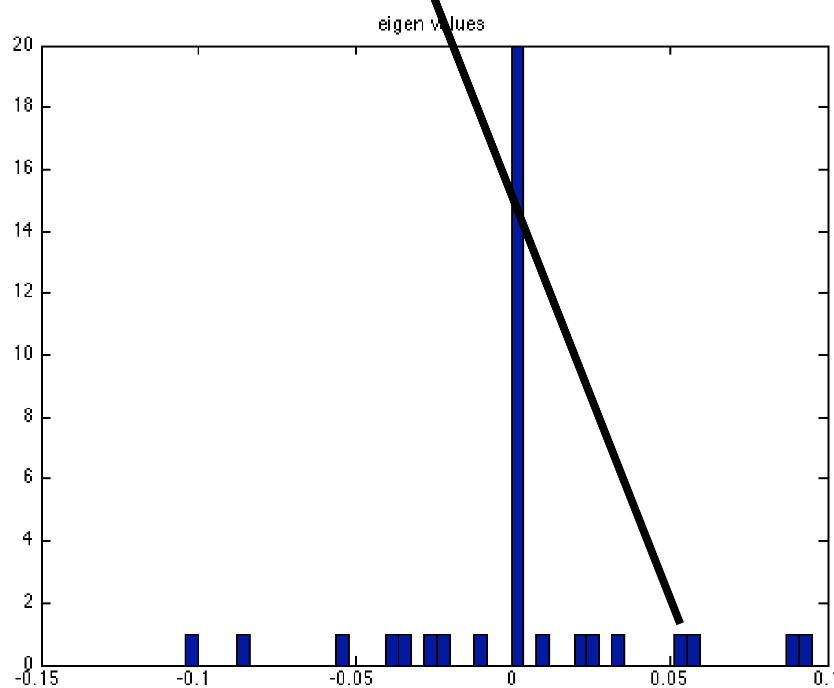
The scenario operator: $S_\omega = i\frac{\Delta - \Delta^*}{2}$



$\kappa = 0.541126346528966$



Text



Scenario operators have highly localized eigenfunctions

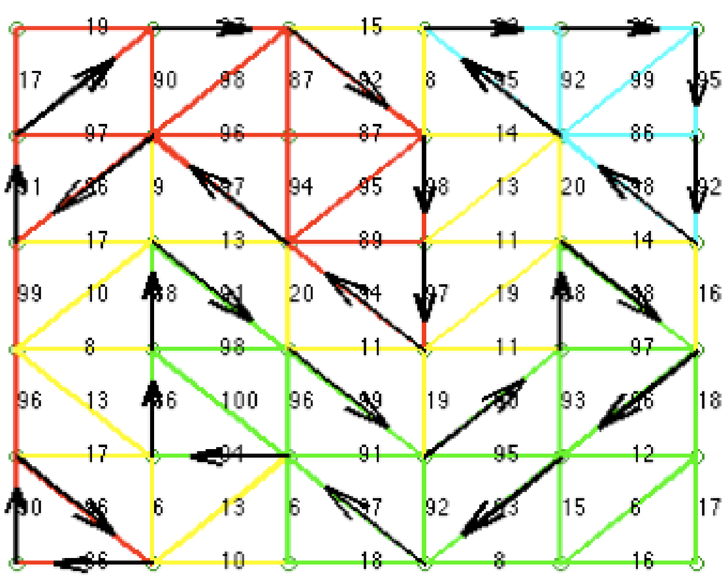
The scenario operator: $S_\omega = i\frac{\Delta - \Delta^*}{2}$

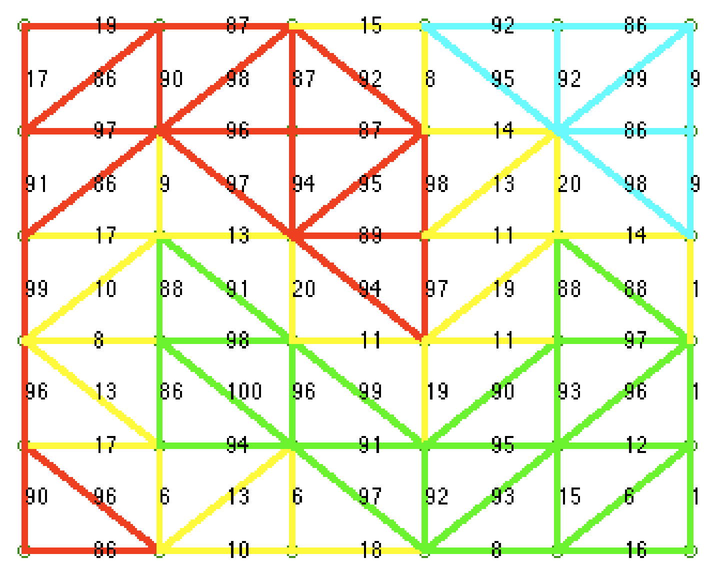
Condctance Independence Corollary: S_ω depends only on the flow F and equilibrium vector ω and is independent of conductance C. Furthermore the chain is reversible if and only if $S_\omega = 0$.

proof:

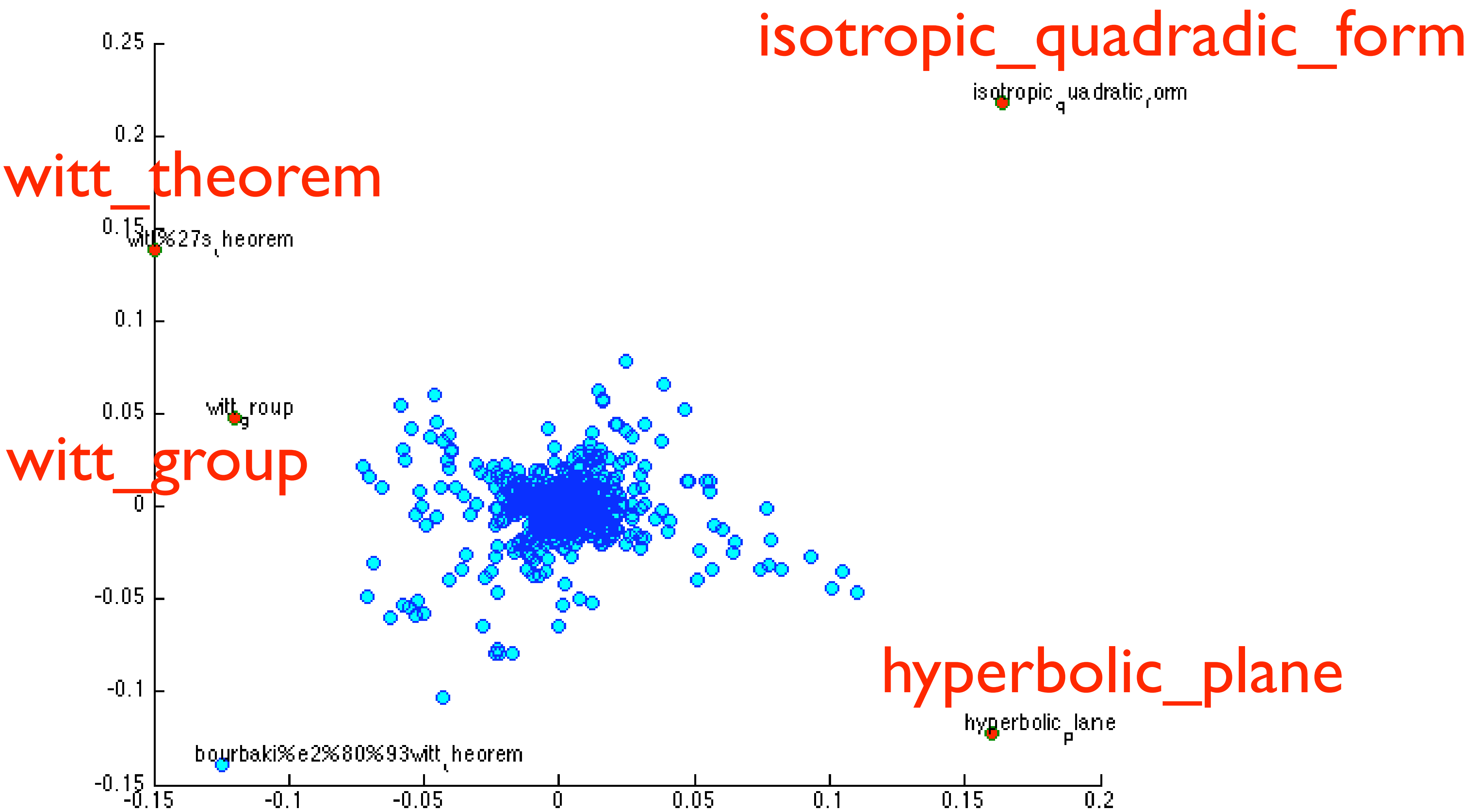
$$S_\omega = i\frac{\Delta_\omega - \Delta_\omega^*}{2} = i[\tau]^{-1}\left(\frac{P - P^*}{2}\right) = i[\tau]^{-1}[\pi]^{-1}F = i[\omega]^{-1}F$$

Q.E.D



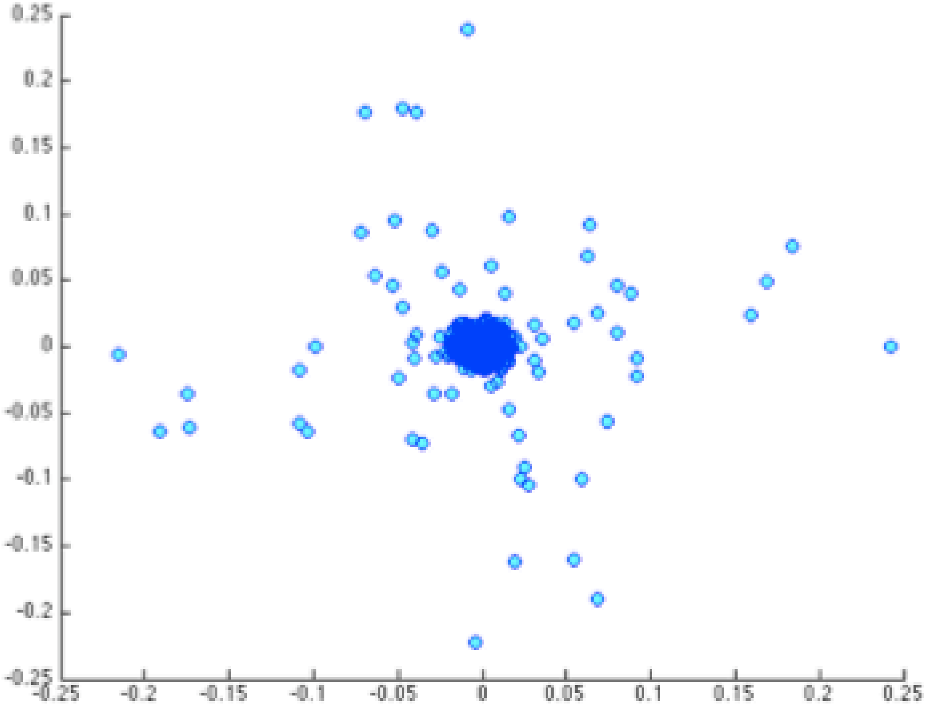


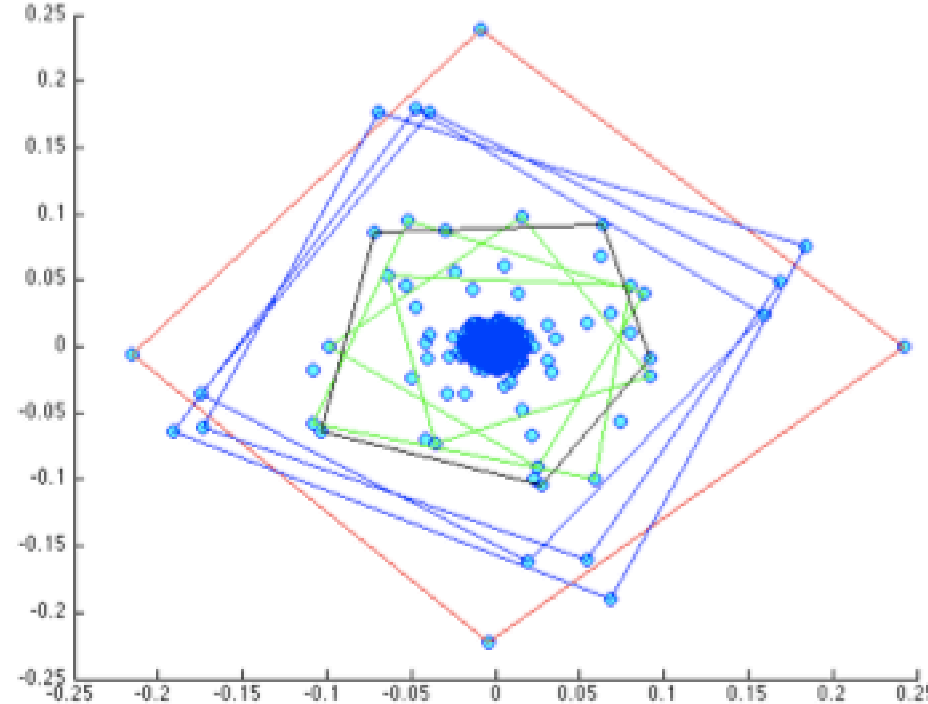
Eigenfunction in the Space of Mathematics



Scenario operators have highly localized eigenfunctions

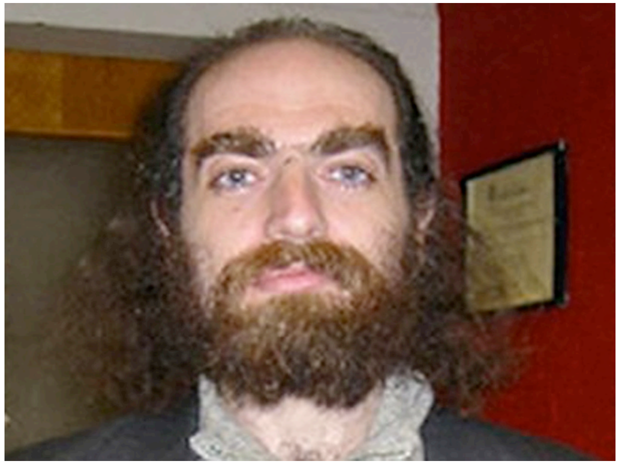
Reality....





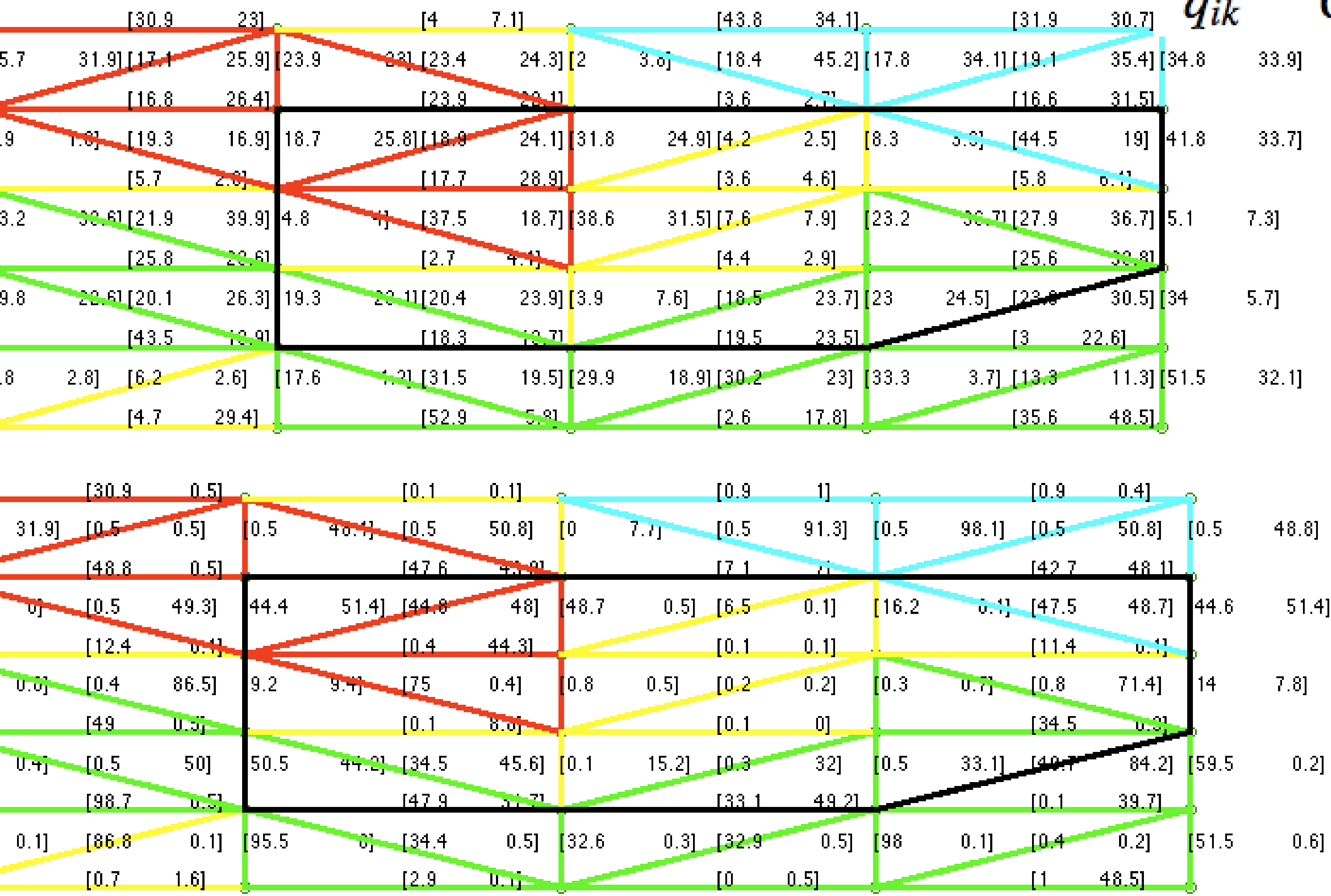
We need something like the vagabond clustering for cycles

The Co-conformal Cycle Hunt

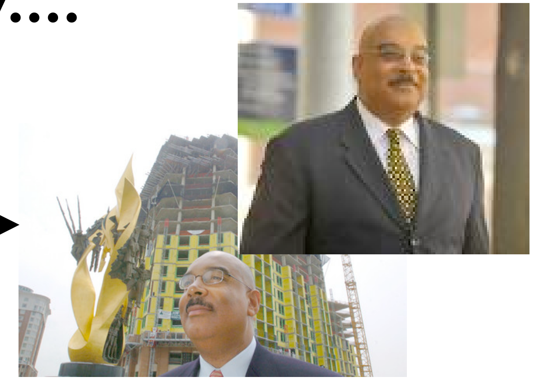


Co-conformal Magnetization

$$Q = [P[\alpha]\mathbf{1}]^{-1}P[\alpha] \qquad \frac{q_{ij}}{q_{ik}} = \frac{\alpha_j}{\alpha_k}\frac{p_{ij}}{p_{ik}}$$

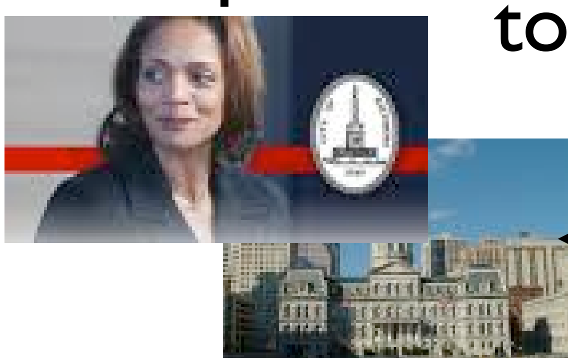


Not all cycles are created equally....



Wy do need this fellow

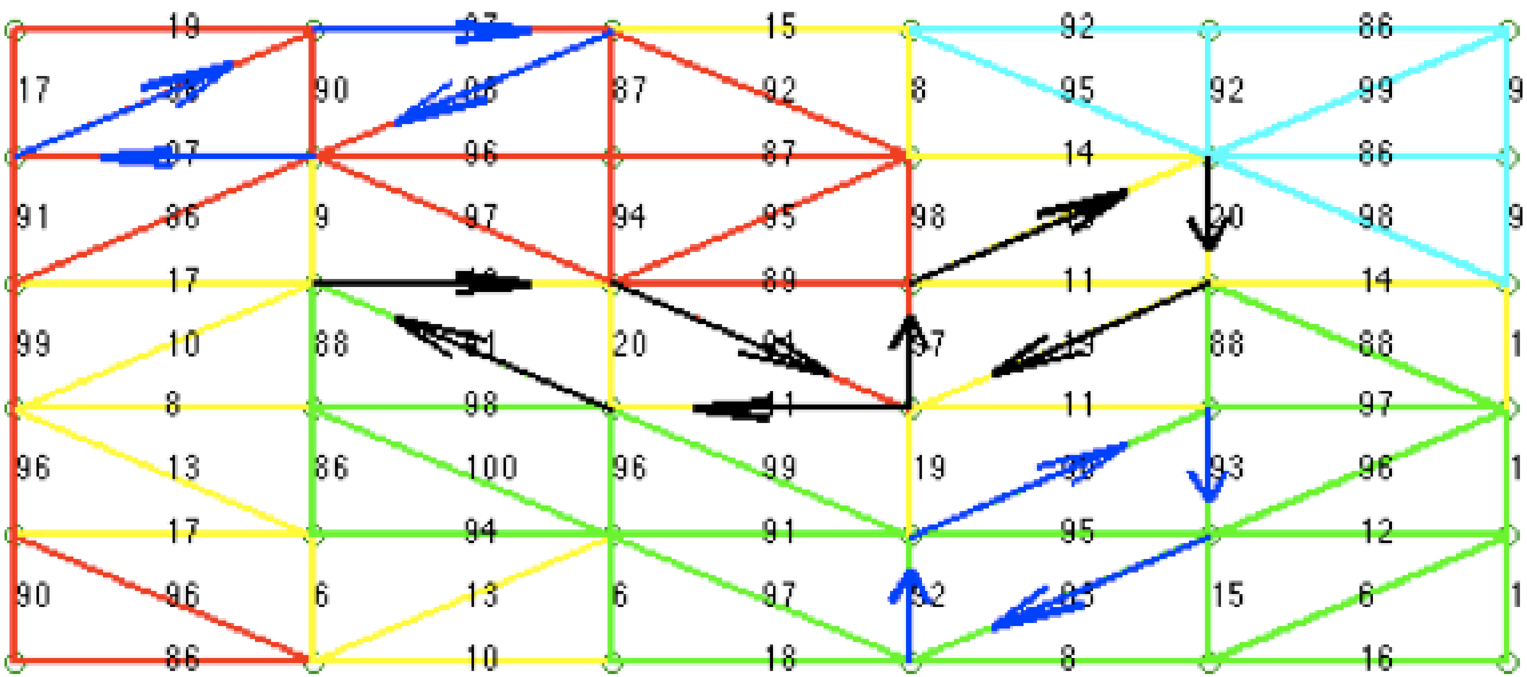
to find a naughty cycle?



Claim: The operator $i[\Delta_\omega, \Delta_\omega^*] = \Delta_\omega \Delta_\omega^* - \Delta_\omega^* \Delta_\omega$

allows us to detect scenarios which are anomalous relative to the context

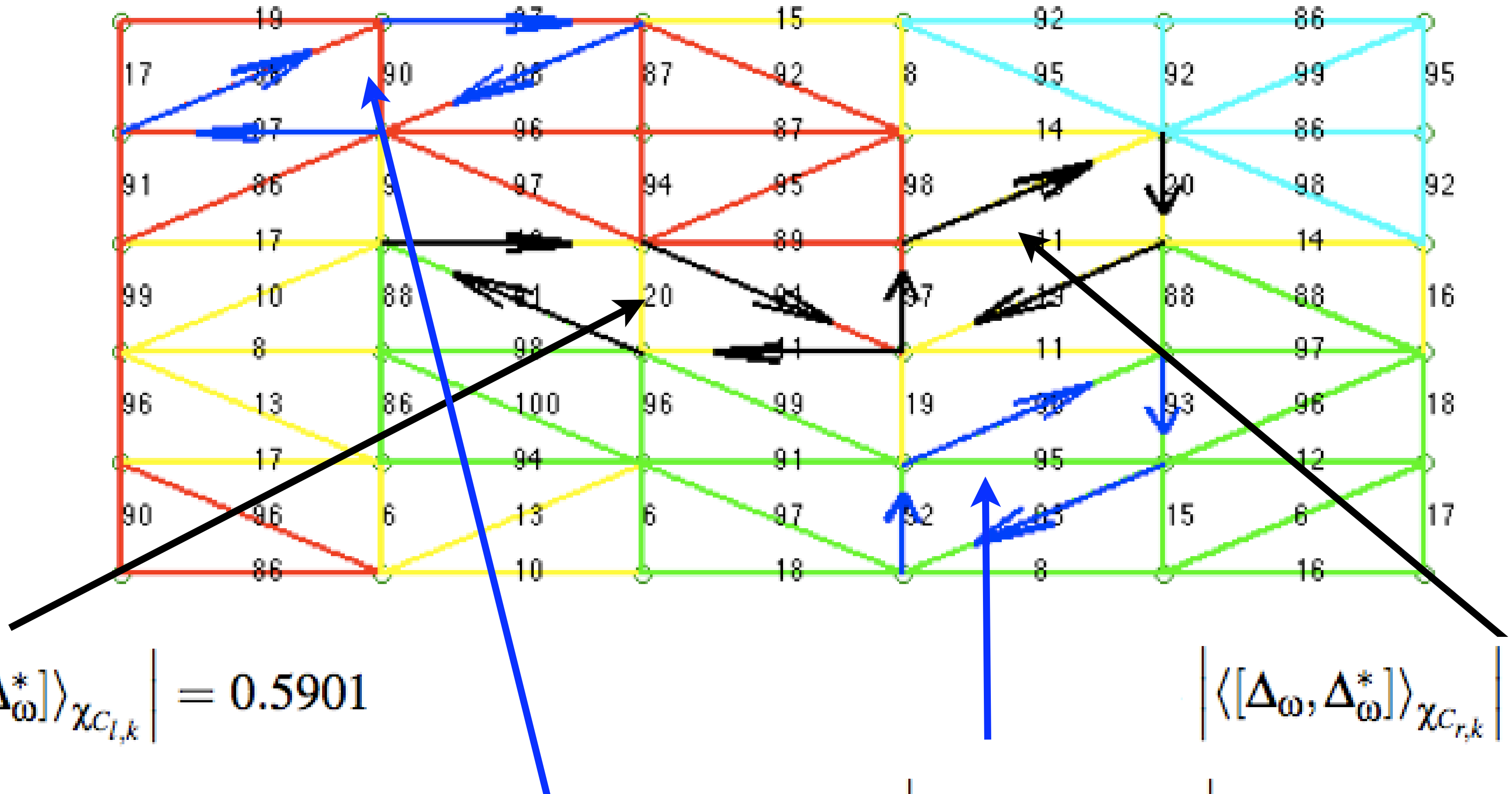
(all cycles weight 10)



Which cycles are in context and which are not?

Claim: The operator $i[\Delta_\omega, \Delta_\omega^*] = \Delta_\omega\Delta_\omega^* - \Delta_\omega^*\Delta_\omega$

allows us to detect scenarios which are anomalous relative to the context



$\left|\langle[\Delta_\omega, \Delta_\omega^*]\rangle_{\chi_{C_{l,k}}}\right| = 0.5901$

$\left|\langle[\Delta_\omega, \Delta_\omega^*]\rangle_{\chi_{C_{r,k}}}\right| = 0.3157$

$\left|\langle[\Delta_\omega, \Delta_\omega^*]\rangle_{\chi_{C_{u,b}}}\right| = 0.0141$

$\left|\langle[\Delta_\omega, \Delta_\omega^*]\rangle_{\chi_{C_{l,b}}}\right| = 0.0039$

http://www.youtube.com/watch?v=KT7xJ0tjB4A

Mathematics of Quantum Mechanics

$$< f, g >_\omega = \bar{f}^{tr}[\omega] g$$

Hilbert Space

$$\langle \psi, \psi \rangle = 1$$

States

$$\langle \psi, A\phi \rangle = \langle A\psi, \phi \rangle$$

Measurements

$$\langle \psi, A\psi \rangle_\omega = \langle A \rangle_\psi$$

Expectation

$$A_\omega - \langle A_\omega \rangle_\psi$$

$$\langle \hat{A}^2 \rangle_\psi$$

$$\sigma_\psi A = \sqrt{\langle \hat{A}^2 \rangle_\psi}$$

Deviation

Why does this work?

The Uncertainty Principle for Markov Chains Lemma:

$$(\sigma_\psi C_\omega)\,(\sigma_\psi S_\omega) \geq \left|\frac{1}{2}\,\langle[\Delta_\omega,\Delta^*_\omega]\rangle_\psi\right|$$

proof: Notice

$$[C_\omega, S_\omega] = C_\omega S_\omega - S_\omega C_\omega = i[\Delta_\omega, \Delta^*_\omega],$$

so

$$\begin{aligned}
\left|\tfrac{1}{2}\,\langle[\Delta_\omega,\Delta^*_\omega]\rangle_\psi\right| = \left|\tfrac{1}{2}\,\langle[C_\omega,C_\omega]\rangle_\psi\right| & \quad \text{(def. plus foil)} \\
= \left|\tfrac{1}{2}\,\left\langle[\hat{C}_\omega,\hat{S}_\omega]\right\rangle_\psi\right| & \quad \text{(differ by constants)} \\
= \tfrac{1}{2}\left|\left\langle\psi, \hat{C}_\omega\hat{S}_\omega - \hat{C}_\omega\hat{S}_\omega\psi\right\rangle\right| & \quad \text{(definition)} \\
= \tfrac{1}{2}\left|\left\langle\hat{C}_\omega\psi, \hat{S}_\omega\psi\right\rangle - \left\langle\hat{S}_\omega\psi, \hat{C}_\omega\psi\right\rangle\right| & \quad \text{(self adjoint)} \\
= \left|Im\left\langle\hat{C}_\omega\psi, \hat{S}_\omega\psi\right\rangle\right| & \quad \text{(Hemitian Inner-product)} \\
\leq \left|\left\langle\hat{C}_\omega\psi, \hat{S}_\omega\psi\right\rangle\right| & \quad (|a+ib| = \sqrt{a^2+b^2}) \\
\leq \sqrt{\left\langle\hat{C}_\omega\psi, \hat{C}_\omega\psi\right\rangle}\sqrt{\left\langle\hat{S}_\omega\psi, \hat{S}_\omega\psi\right\rangle} & \quad \text{(Cauchy-Schwarz inequality)} \\
= \sqrt{\left\langle\psi, \hat{C}^2_\omega\psi\right\rangle}\sqrt{\left\langle\psi, \hat{S}^2_\omega\psi\right\rangle} & \quad \text{(Self adjoint)} \\
= (\sigma_\psi C_\omega)\,(\sigma_\psi S_\omega) & \quad \text{(defintion)}
\end{aligned}$$

Q.E.D