Twice-Ramanujan Sparsifiers

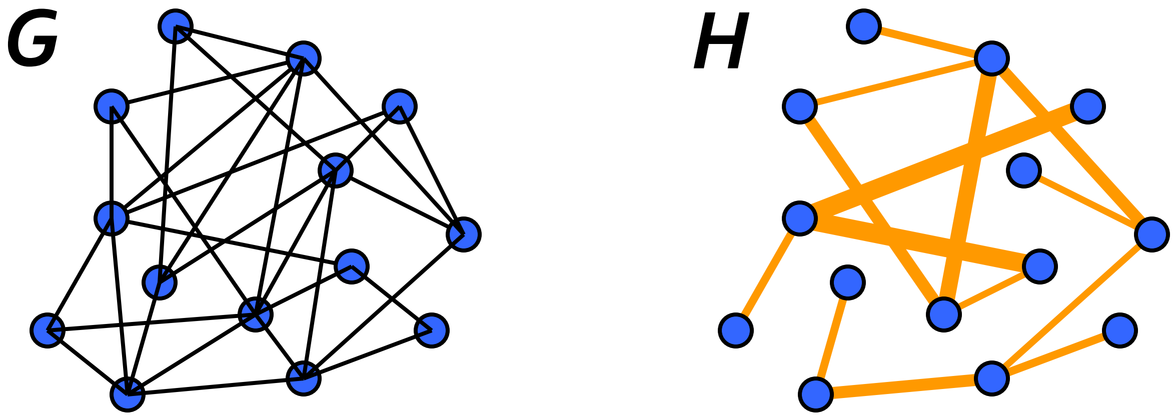
Nikhil Srivastava

MSR India / Yale

with Josh Batson and Dan Spielman

Sparsification

Approximate any graph ***G*** by a sparse graph ***H***.



– Nontrivial statement about ***G***

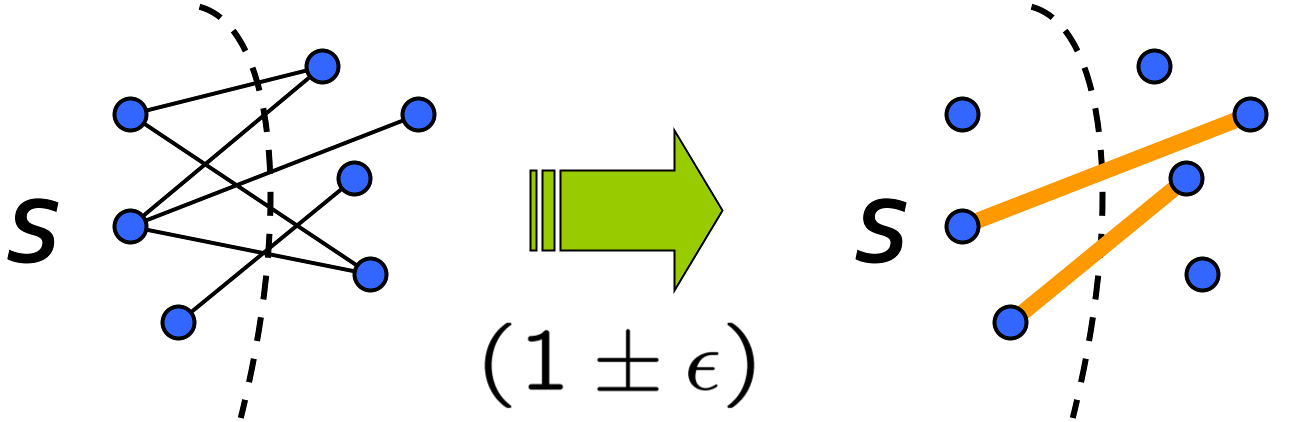
– ***H*** is faster to compute with than ***G***

Cut Sparsifiers [Benczur-Karger'96]

H approximates ***G*** if

for every cut ***S***$\subset$***V***

sum of weights of edges leaving ***S*** is preserved



Can find ***H*** with O(nlogn/ε^2) edges in $\tilde{O}(m)$ time

The Laplacian (quick review)

$$L_G = D_G - A_G = \sum_{ij \in E} c_{ij}(\delta_i - \delta_j)(\delta_i - \delta_j)^T$$

Quadratic form

$$x : V \to \mathbb{R}$$

$$x^T L_G x = \sum_{ij \in E} c_{ij}(x(i) - x(j))^2$$

Positive semidefinite

Ker(L_G)=span(**1**) if ***G*** is connected

Cuts and the Quadratic Form

For characteristic vector $x_S \in \{0,1\}^n$ of $S \subseteq V$

$$
\begin{aligned}
x_S^T L_G x_S &= \sum_{ij \in E} c_{ij}(x(i) - x(j))^2 \\
&= \sum_{ij \in (S,\overline{S})} c_{ij} \\
&= wt_G(S, \overline{S})
\end{aligned}
$$

So BK says:

$$1 - \epsilon \leq \frac{x^T L_H x}{x^T L_G x} \leq 1 + \epsilon \quad \forall x \in \{0,1\}^n$$

A Stronger Notion [ST'04]

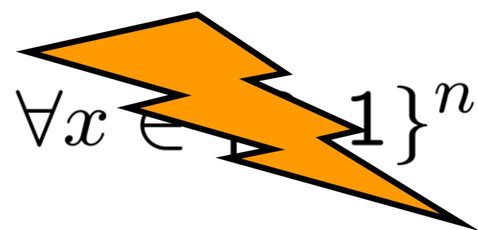
For characteristic vector $x_S \in \{0,1\}^n, S \subseteq V$

$$x_S^T L_G x_S = \sum_{ij \in E} c_{ij} (x(i) - x(j))^2$$
$$= \sum_{ij \in (S, \overline{S})} c_{ij}$$
$$= wt_G(S, \overline{S})$$

So BK says:

$$\forall x \in \mathbb{R}^n$$

$$1 - \epsilon \leq \frac{x^T L_H x}{x^T L_G x} \leq 1 + \epsilon \quad \forall x \in \{0,1\}^n$$



Why?

1. All eigenvalues are preserved

By Courant-Fischer,

$$(1-\epsilon)\lambda_i(G) \leq \lambda_i(H) \leq (1+\epsilon)\lambda_i(G)$$

G and ***H*** have similar eigenvalues.

For spectral purposes, ***G*** and ***H*** are equivalent.

($x^T Lx$ says a lot)

2. Behavior of electrical flows.

   (***$x^T Lx$*** = "energy" for potentials ***x***:V->R)

3. Behavior of random walks: commute times, mixing time, etc.

4. 'Relative condition number' in lin-alg.

5. Fast linear system solvers.

strong notion of approximation.

Examples

Example: Sparsify Complete Graph by Ramanujan Expander [LPS,M]

G is complete on n vertices. $\lambda_i(L_G) = n$

H is d-regular Ramanujan graph. $\lambda_i(L_H) \sim d$

$$\lambda_i(\tfrac{n}{d} L_H) \sim n$$

Example: Sparsify Complete Graph by Ramanujan Expander [LPS,M]

G is complete on n vertices. $\lambda_i(L_G) = n$

H is d-regular Ramanujan graph. $\lambda_i(L_H) \sim d$

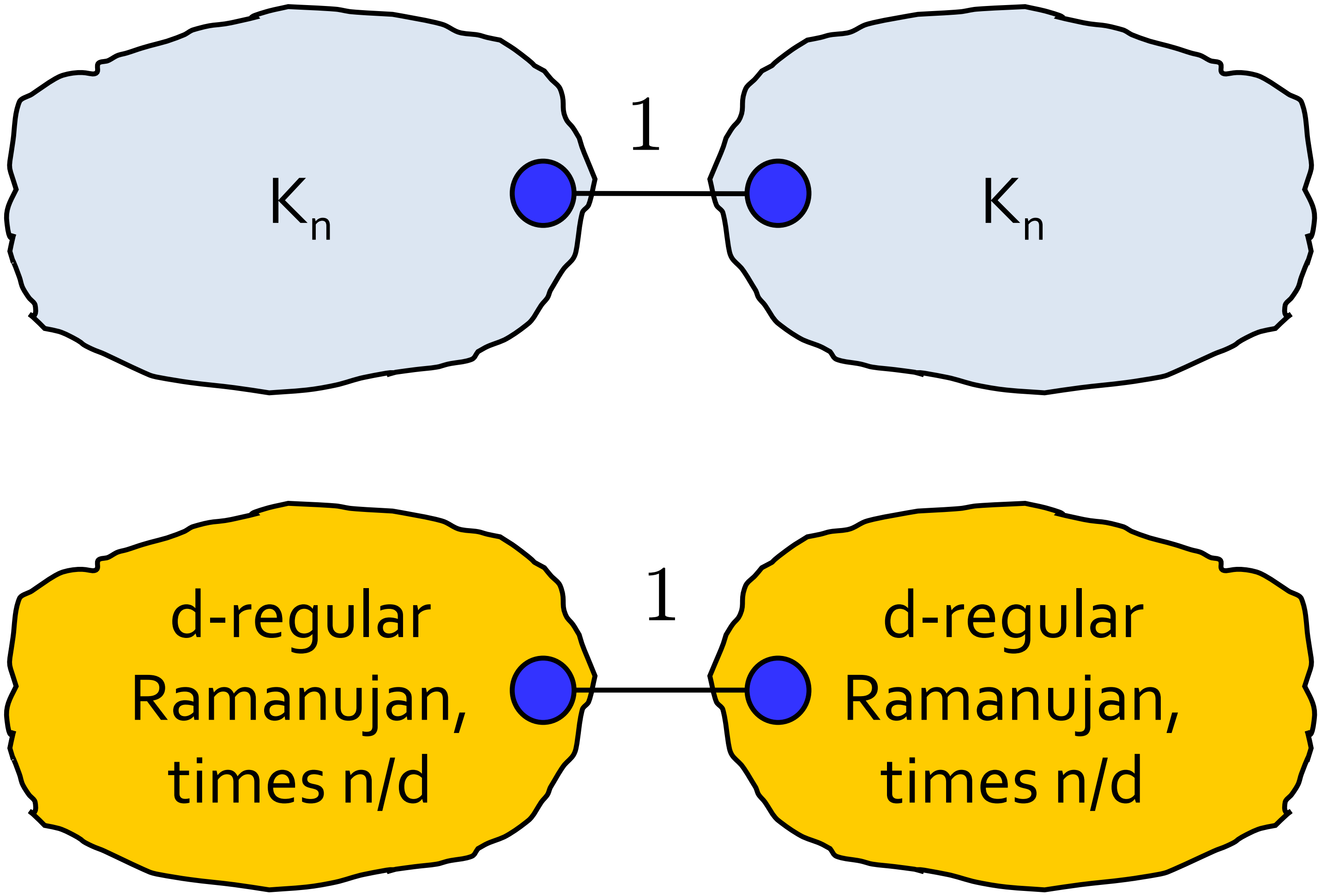
$$\lambda_i(\tfrac{n}{d}L_H) \sim n$$

$$\frac{x^T(\frac{n}{d}L_H)x}{x^T L_G x} \sim 1$$

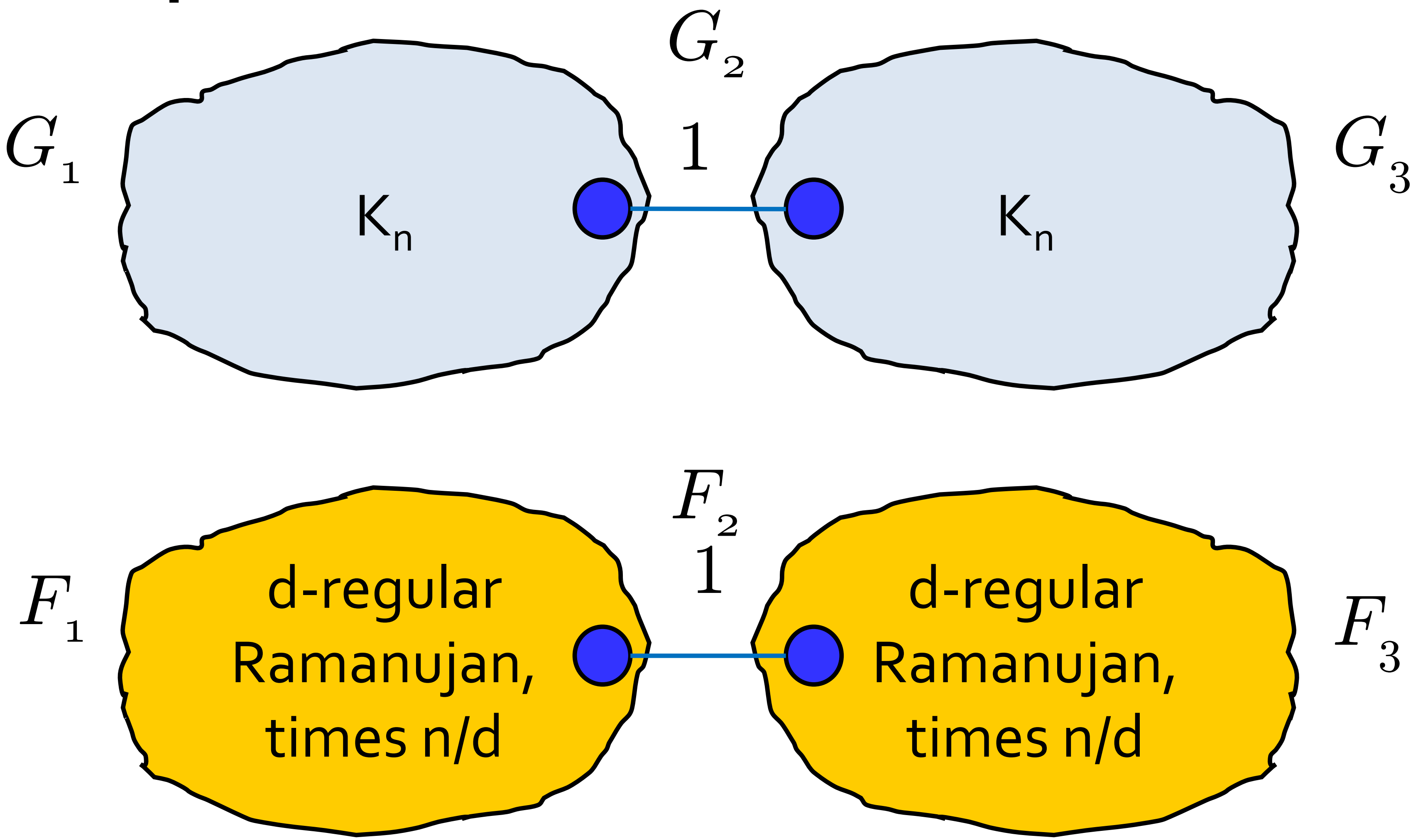
Each edge has weight (n/d)

So, $\frac{n}{d}H$ is a good sparsifier for ***G***.

Example: Dumbell



Example: Dumbell



$$G = G_1 + G_2 + G_3$$

$$x^T G x = x^T G_1 x + x^T G_2 x + x^T G_3 x$$

Results

Results

*We can do this well for every **G**.*

(upto a factor of 2)

Previously Known

Expanders/Ramanujan graphs exist:

"There are very sparse ***H*** that look like $\mathbf{K_n}$"

Previously Known

Expanders/Ramanujan graphs exist:

"There are very sparse ***H*** that look like $\mathbf{K_n}$"

degree *d*

$$1 \leq \frac{x^T L_H x}{x^T L_{K_n} x} = \frac{d+2\sqrt{d-1}}{d-2\sqrt{d-1}}$$

New Result

Expanders/Ramanujan graphs exist:

"There are very sparse ***H*** that look like $\mathbf{K_n}$"

degree *d*

$$1 \leq \frac{x^T L_H x}{x^T L_G x} = \frac{d+2\sqrt{d-1}}{d-2\sqrt{d-1}}$$

Sparsifiers exist:

"There are very sparse ***H*** that look like *any graph* **G**."

avg. degree *2d*

New Result

Expanders/Ramanujan graphs exist:

"There are very sparse ***H*** that look like $\mathbf{K_n}$"

degree *d*

$$1 \leq \frac{x^T L_H x}{x^T L_G x} = \frac{d+2\sqrt{d-1}}{d-2\sqrt{d-1}}$$

Sparsifiers exist:

"There are very sparse ***H*** that look like *any graph* **G**."

avg. degree *2d*

weighted subgraph

New Result

deterministic
O(dmn³)
algorithm

Expanders/Ramanujan graphs exist:

"There are very sparse ***H*** that look like $\mathbf{K_n}$"

degree *d*

$$1 \leq \frac{x^T L_H x}{x^T L_G x} = \frac{d+2\sqrt{d-1}}{d-2\sqrt{d-1}}$$

Sparsifiers exist:

"There are very sparse ***H*** that look like *any graph* **G**."

avg. degree *2d*

weighted subgraph

New Result

deterministic
O(dmn³)
algorithm

Expanders/Ramanujan graphs exist:

"There are very sparse ***H*** that look like $\mathbf{K_n}$"

degree *d*

$$1 \le \frac{x^T L_H x}{x^T L_G x} \le \frac{1+\epsilon}{1-\epsilon}$$

Sparsifiers exist:

"There are very sparse ***H*** that look like *any graph* **G**."

avg. degree $8/\epsilon^2$

weighted subgraph

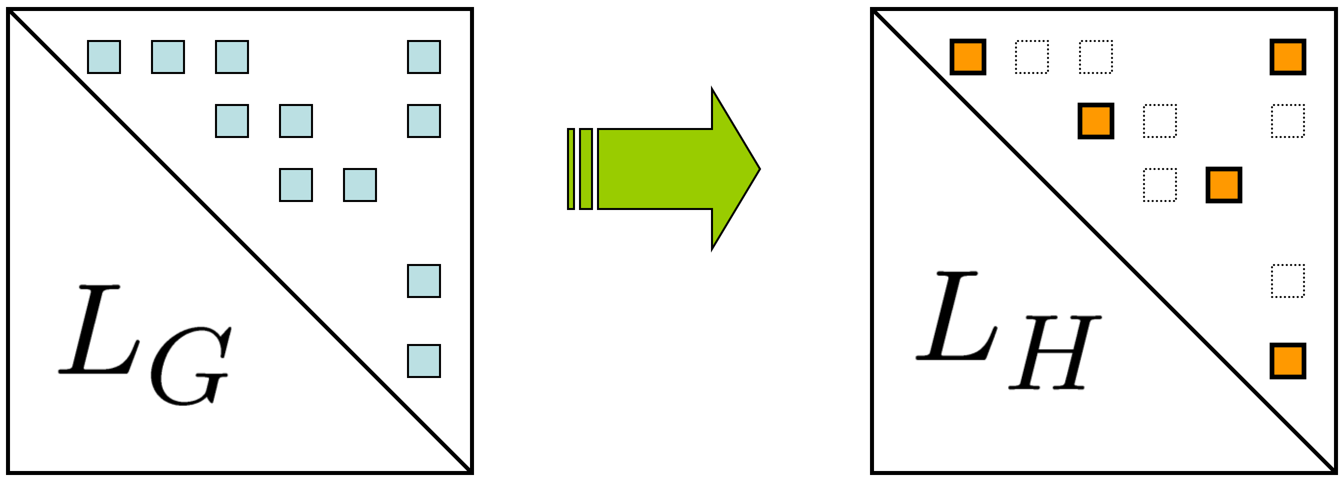
The Method

The Method

(13-approximation with 6n edges.)

Step 1: Reduction to Linear Algebra

Goal



$$1 \leq \frac{x^T L_H x}{x^T L_G x} \leq 13 \qquad \forall x \in \mathbb{R}^n$$

Outer Product Expansion

Recall:

$$L_G = \sum_{ij \in E} (\delta_i - \delta_j)(\delta_i - \delta_j)^T = \sum_{e \in E} b_e b_e^T .$$

Outer Product Expansion

Recall:

$$L_G = \sum_{ij \in E} (\delta_i - \delta_j)(\delta_i - \delta_j)^T = \sum_{e \in E} b_e b_e^T .$$

For a weighted subgraph ***H***:

$$L_H = \sum_{e \in E} s_e b_e b_e^T$$

where s_e=wt(e) in H.

$$1 \leq \frac{x^T L_H x}{x^T L_G x} \leq 13 \quad \forall x \in \mathbb{R}^n$$

$$1 \leq \frac{x^T L_H x}{x^T L_G x} \leq 13 \quad \forall x \in \mathbb{R}^n$$

$$\boxed{1 \leq \lambda(L_G^{-1/2} L_H L_G^{-1/2}) \leq 13.}$$

$$1 \leq \frac{x^T L_H x}{x^T L_G x} \leq 13 \quad \forall x \in \mathbb{R}^n$$

$$1 \leq \lambda(L_G^{-1/2} L_H L_G^{-1/2}) \leq 13.$$

$$\boxed{1 \leq \lambda\left(\sum_{e \in E} s_e L_G^{-1/2} b_e b_e^T L_G^{-1/2}\right) \leq 13.}$$

$$1 \leq \frac{x^T L_H x}{x^T L_G x} \leq 13 \quad \forall x \in \mathbb{R}^n$$

$$1 \leq \lambda(L_G^{-1/2} L_H L_G^{-1/2}) \leq 13.$$

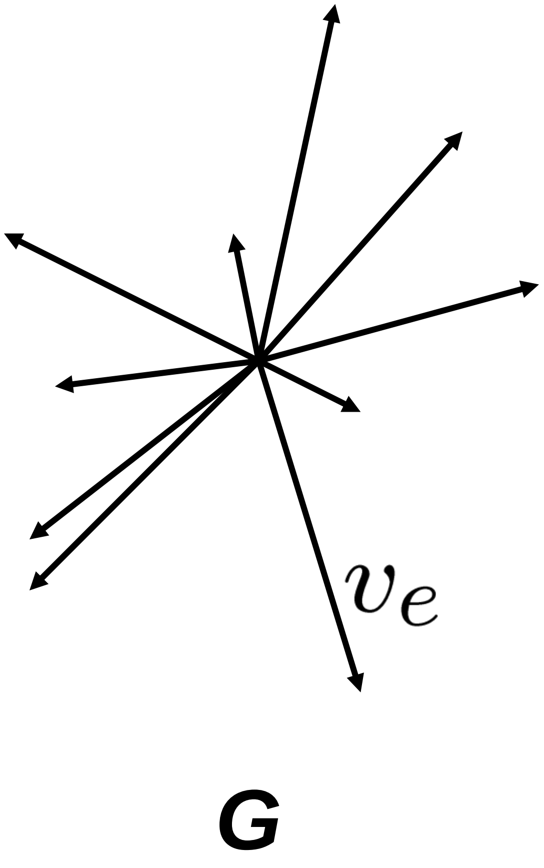
$$1 \leq \lambda \left(\sum_{e \in E} s_e L_G^{-1/2} b_e b_e^T L_G^{-1/2} \right) \leq 13.$$

$$1 \leq \lambda \left(\sum_{e \in E} s_e v_e v_e^T \right) \leq 13$$

with $v_e = L_G^{-1/2} b_e$.

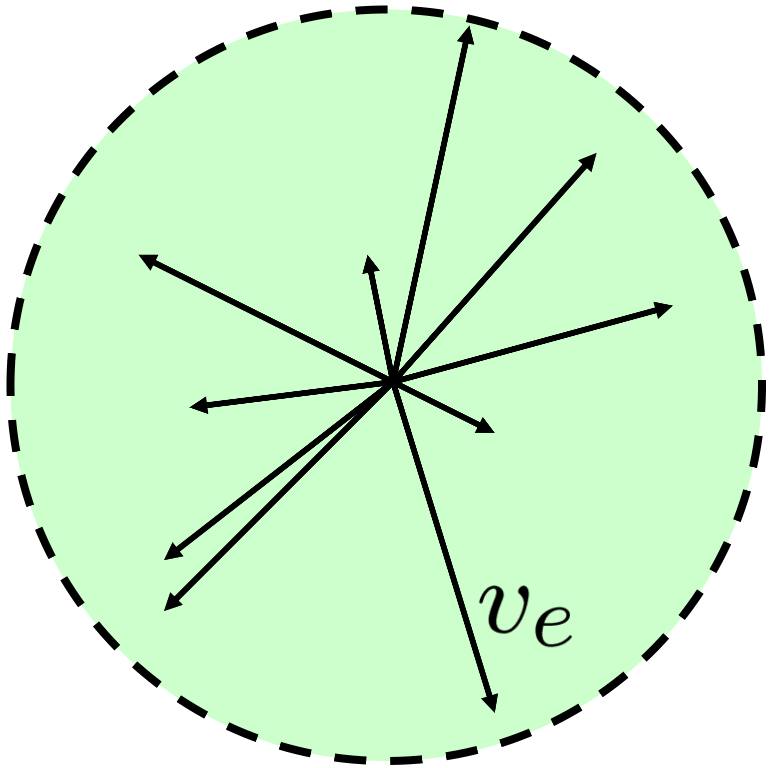
A closer look at v_e

$$v_e = L_G^{-1/2} b_e.$$



m vectors in $\boldsymbol{R^{n-1}}$

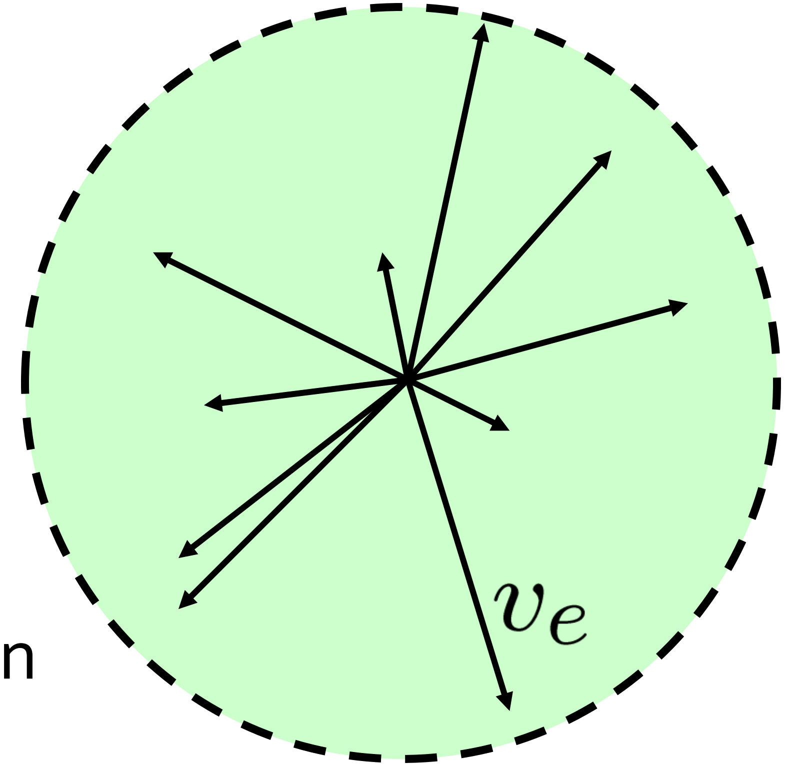
A closer look at v_e



m vectors
in $\boldsymbol{R}^{n-1}$

$$\sum_e v_e v_e^T = L_G^{-1/2} \left(\sum_e b_e b_e^T \right) L_G^{-1/2} = I$$

A closer look at v_e

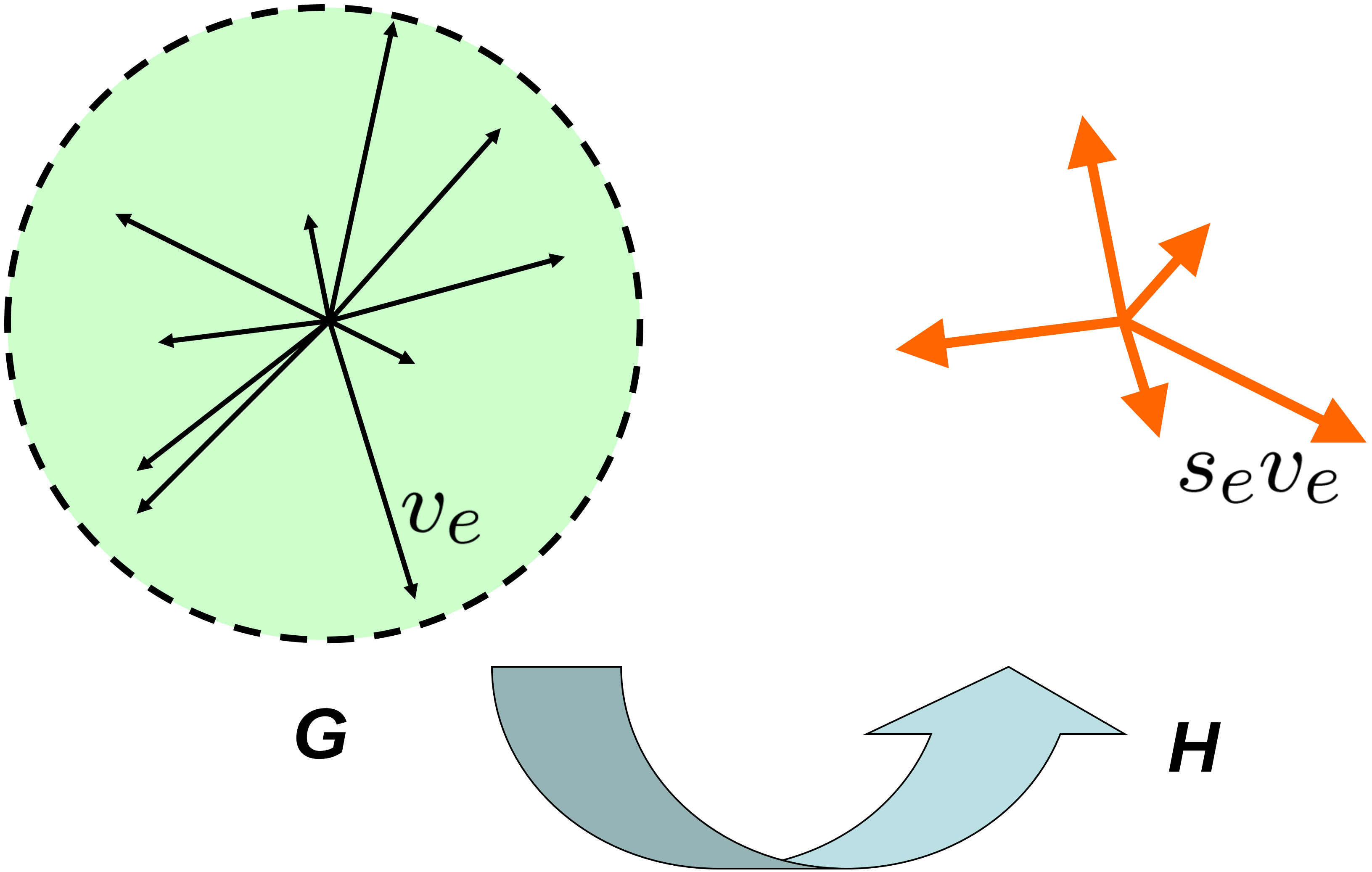


m vectors in $\boldsymbol{R}^{n-1}$

"decomposition of identity"

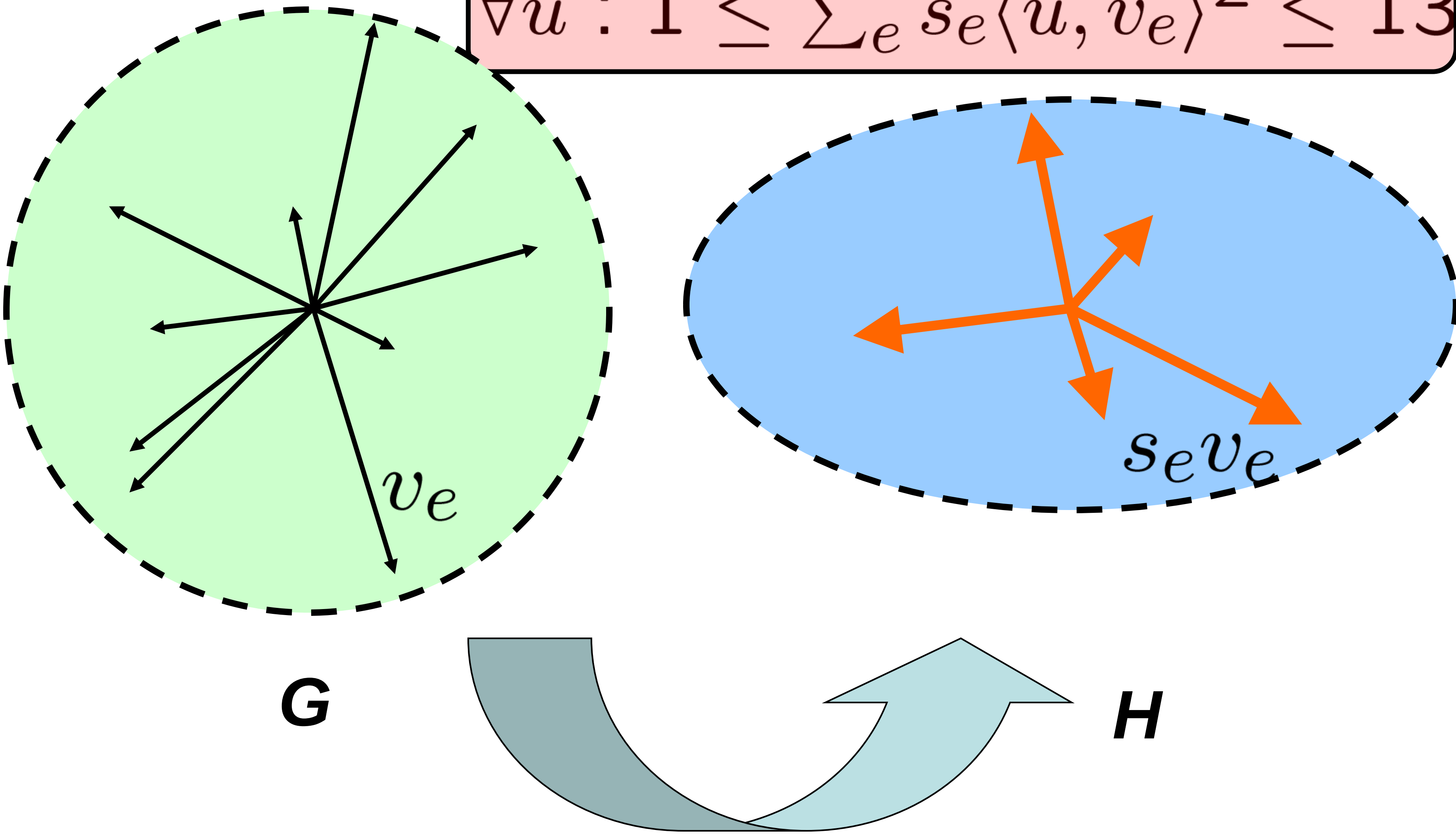
$$\forall u \quad \sum_e \langle u, v_e \rangle^2 = 1$$

Choosing a Subgraph

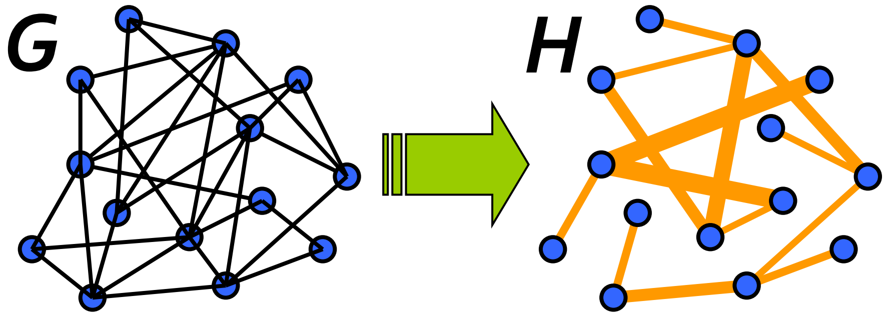


New Goal

$$\forall u : 1 \leq \sum_e s_e \langle u, v_e \rangle^2 \leq 13$$

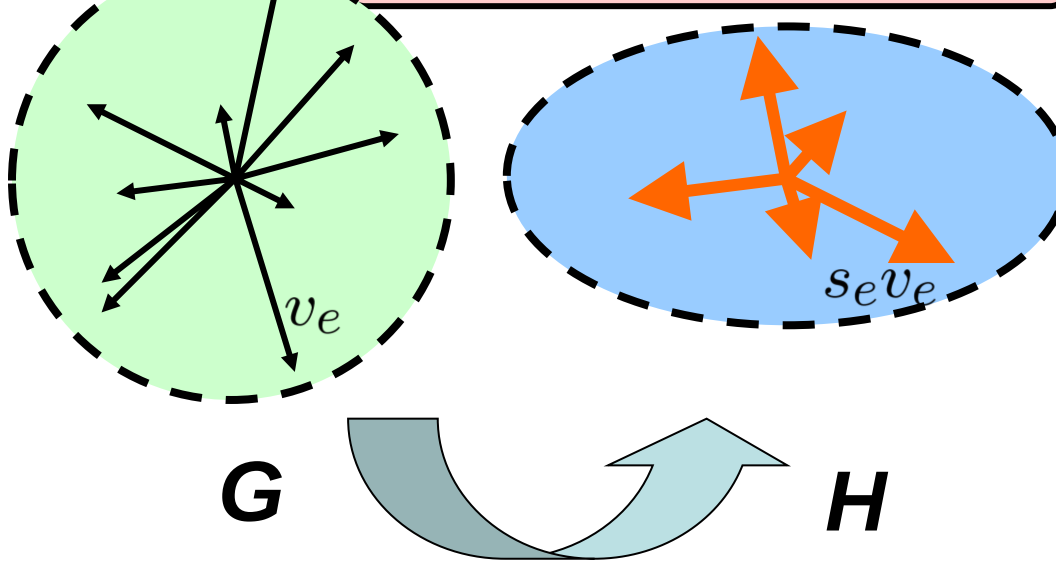


New Goal



$$1 \le \frac{x^T L_H x}{x^T L_G x} \le 13 \quad \forall x \in \mathbb{R}^n$$

$$\forall u : 1 \le \sum_e s_e \langle u, v_e \rangle^2 \le 13$$



Main theorem

If

$$\sum_e v_e v_e^T = I_n$$

then there are scalars $s_e \geq 0$ with

$$1 \leq \lambda(\sum_e s_e v_e v_e^T) \leq 13$$

and $|\{s_e \neq 0\}| \leq 6n$.

Main theorem

If

$$\sum_e v_e v_e^T = I_n$$

then there are scalars $s_e \geq 0$ with

$$1 \leq \lambda(\sum_e s_e v_e v_e^T) \leq \frac{d+2\sqrt{d-1}}{d-2\sqrt{d-1}}$$

and $|\{s_e \neq 0\}| \leq dn$

Main theorem

If

$$\sum_e v_e v_e^T = I_n$$

then there are scalars $s_e \geq 0$ with

$$1 \leq \lambda(\sum_e s_e v_e v_e^T) \leq 13$$

and $|\{s_e \neq 0\}| \leq 6n$.

Step 2: Intuition for the proof

Main theorem

If

$$\sum_e v_e v_e^T = I_n$$

then there are scalars $s_e \geq 0$ with

$$1 \leq \lambda(\sum_e s_e v_e v_e^T) \leq 13$$

and $|\{s_e \neq 0\}| \leq 6n$.

Main theorem

If

$$\sum_e v_e v_e^T = I_n$$

then there are scalars $s_e \geq 0$ with

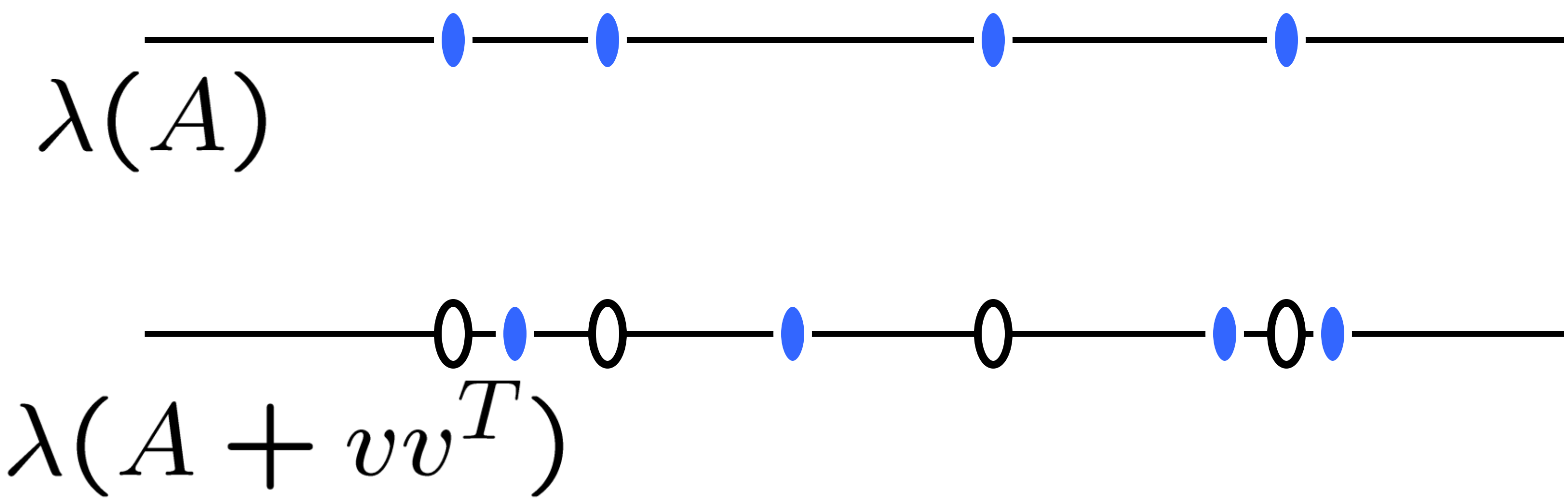
$$1 \leq \lambda\left(\sum_e s_e v_e v_e^T\right) \leq 13$$

and $|\{s_e \neq 0\}| \leq 6n$

will build this one vector at a time.

What happens when we add a vector?

Interlacing



More precisely

Characteristic Polynomial:

$$p_A(x) = \det(xI - A)$$

More precisely

Characteristic Polynomial:

$$p_A(x) = \det(xI - A)$$

Matrix-Determinant Lemma:

$$p_{A+vv^T} = p_A\left(1 + \sum_i \frac{\langle v, u_i \rangle^2}{\lambda_i - x}\right)$$

More precisely

Characteristic Polynomial:

$$p_A(x) = \det(xI - A)$$

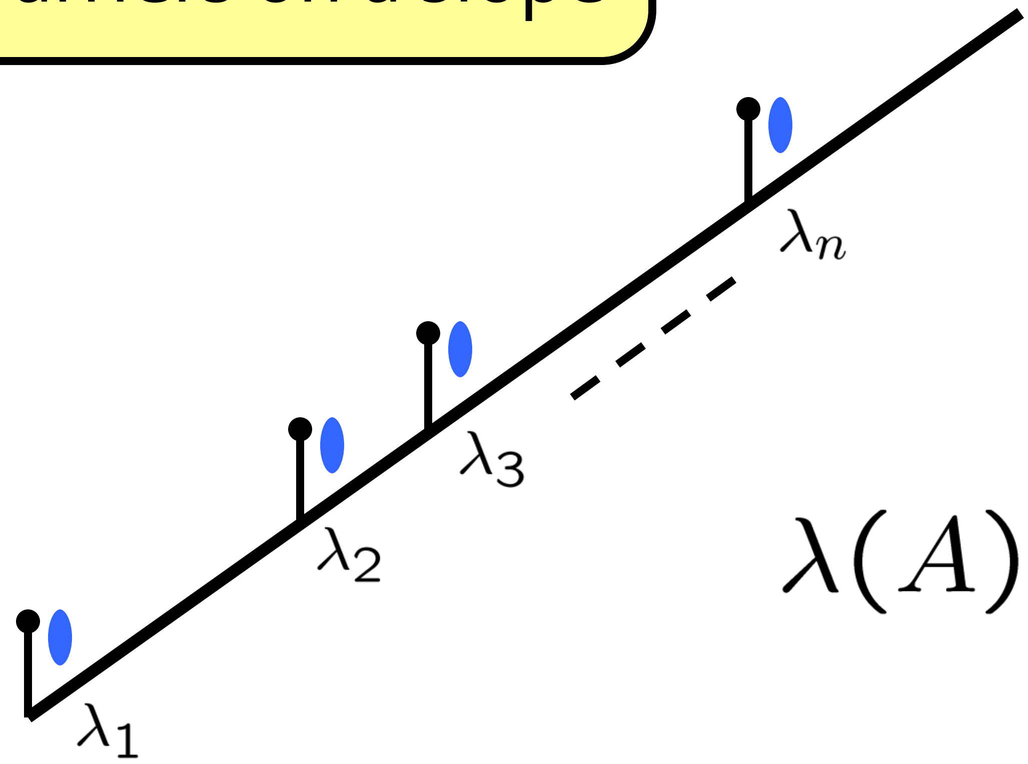
Matrix-Determinant L

$\lambda(A + vv^T)$
are zeros of this.

$$p_{A+vv^T} = p_A \left(1 + \sum_i \frac{\langle v, u_i \rangle^2}{\lambda_i - x}\right)$$

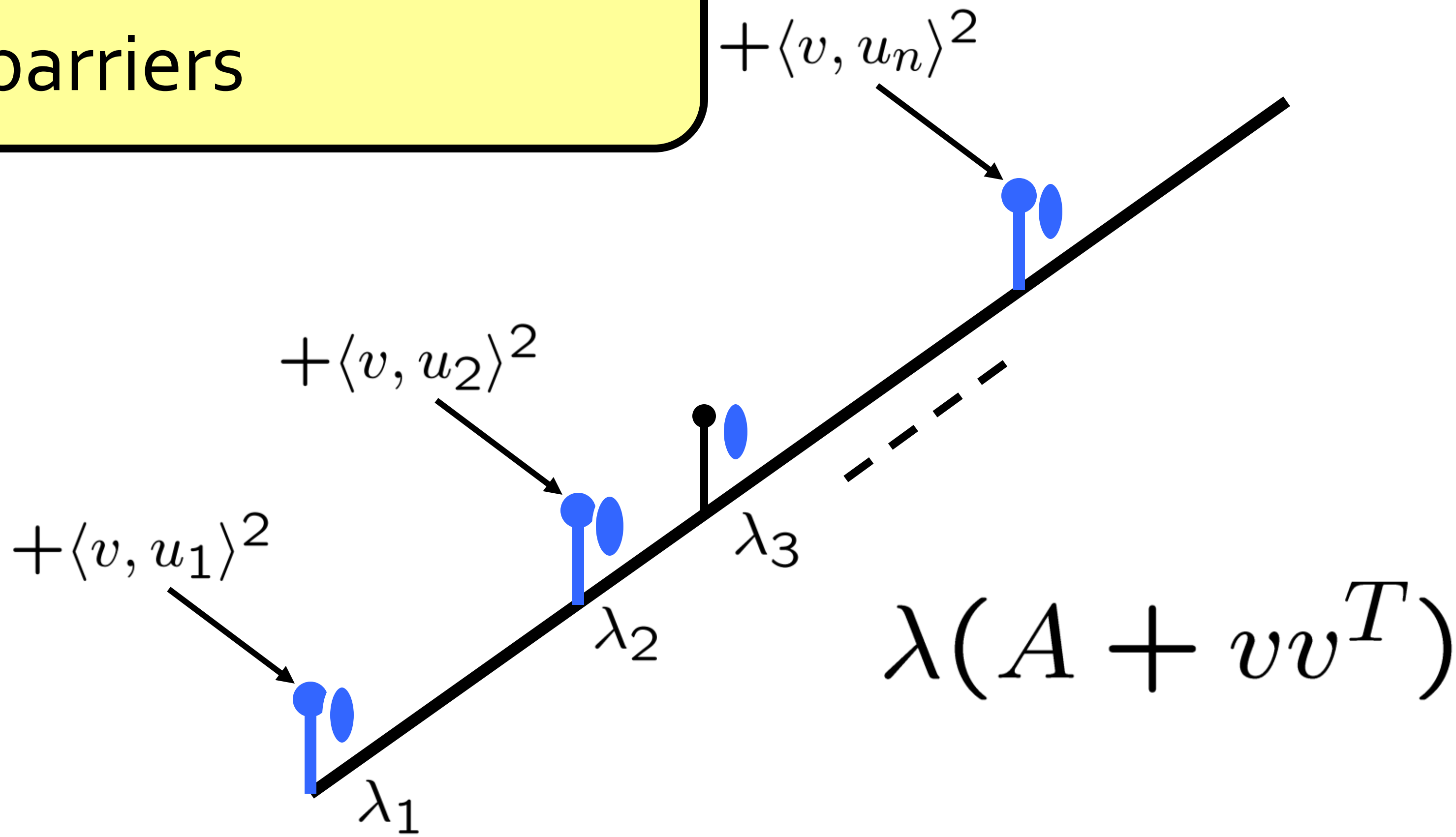
Physical model of interlacing

λ_i = positive unit charges resting at barriers on a slope



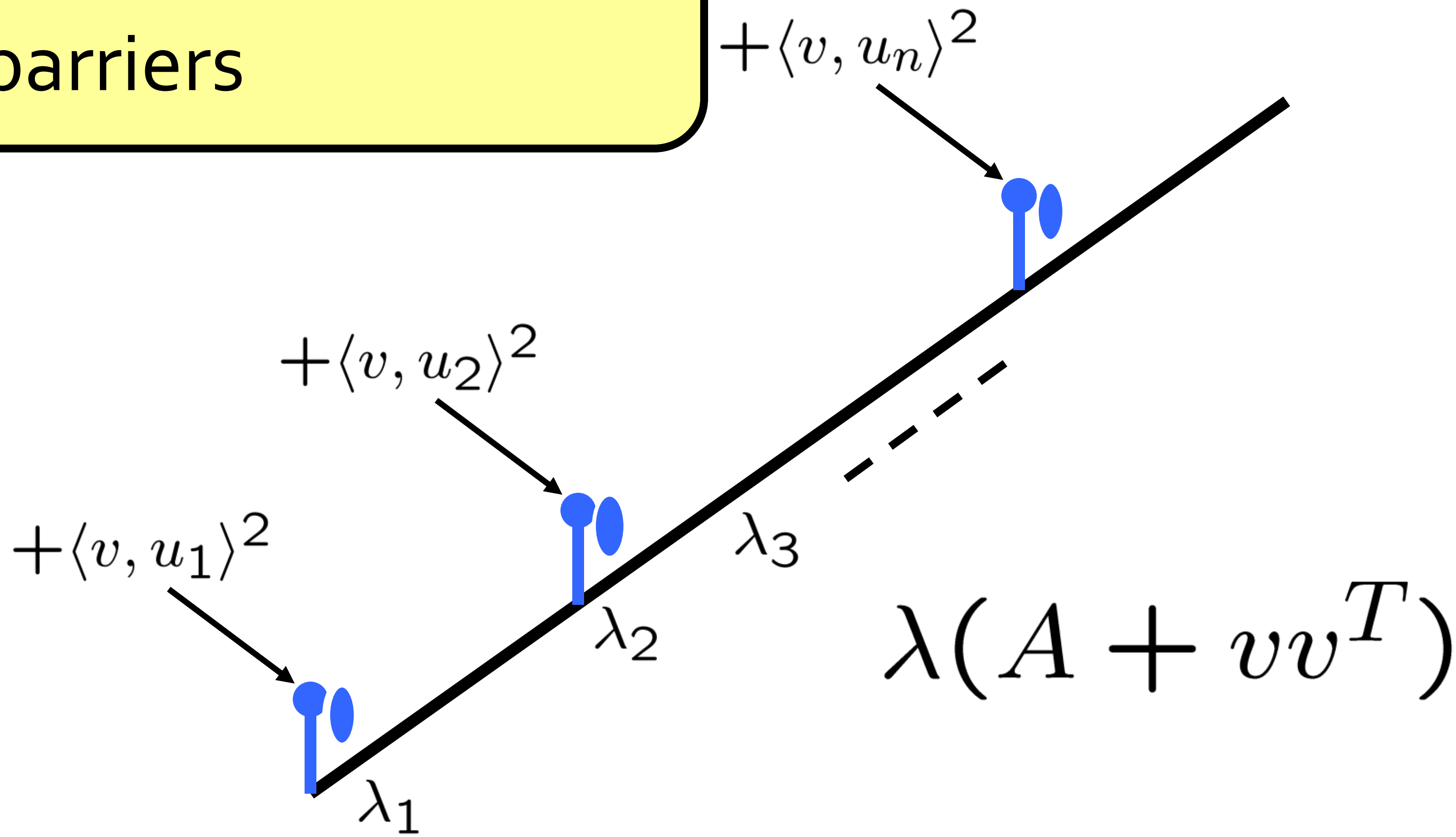
Physical model of interlacing

$<v,u_i>^2$ =charges added to barriers



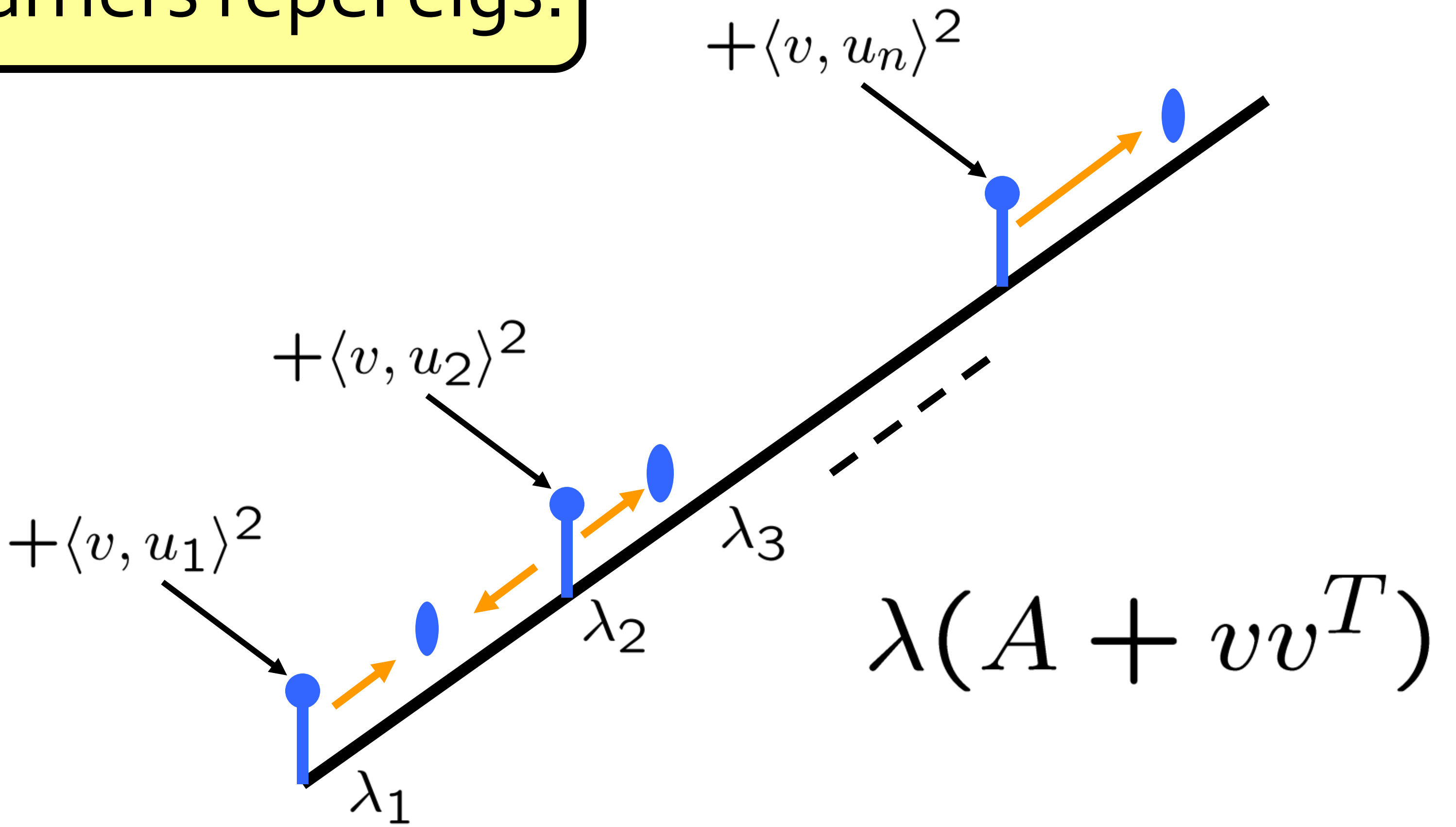
Physical model of interlacing

$\langle v, u_i \rangle^2$ =charges added to barriers

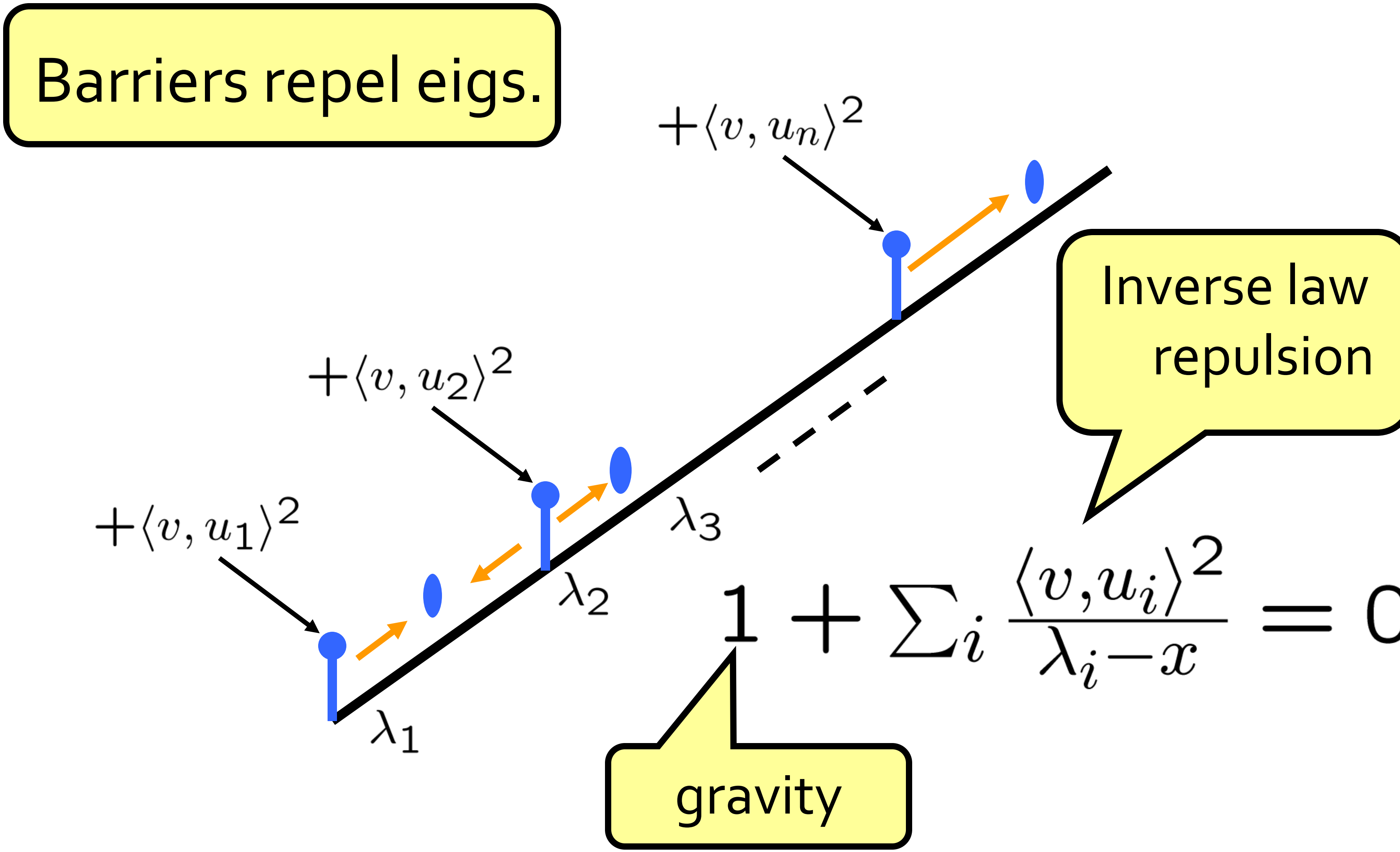


Physical model of interlacing

Barriers repel eigs.

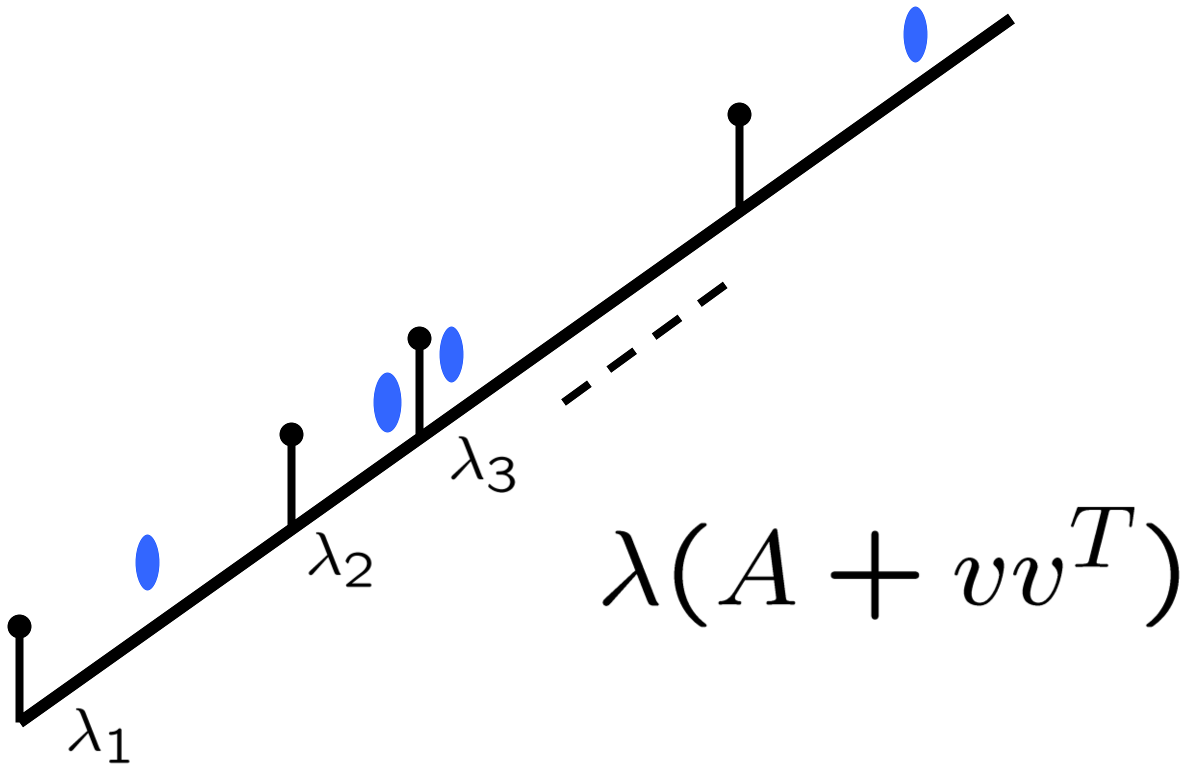


Physical model of interlacing

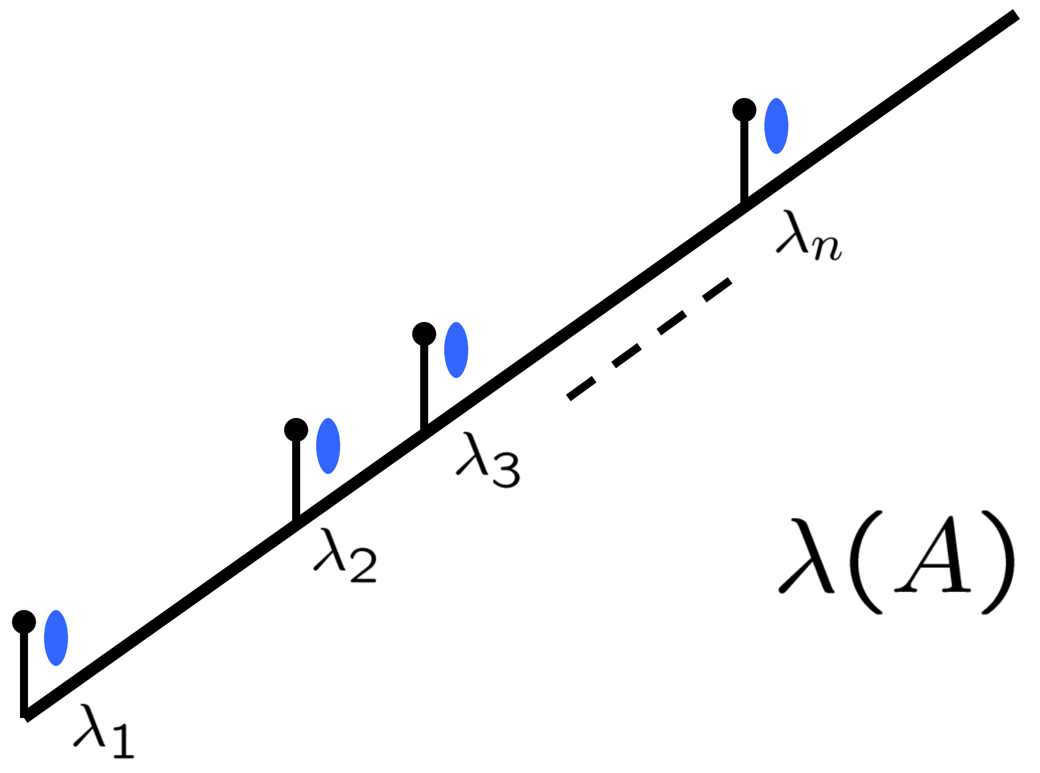


Physical model of interlacing

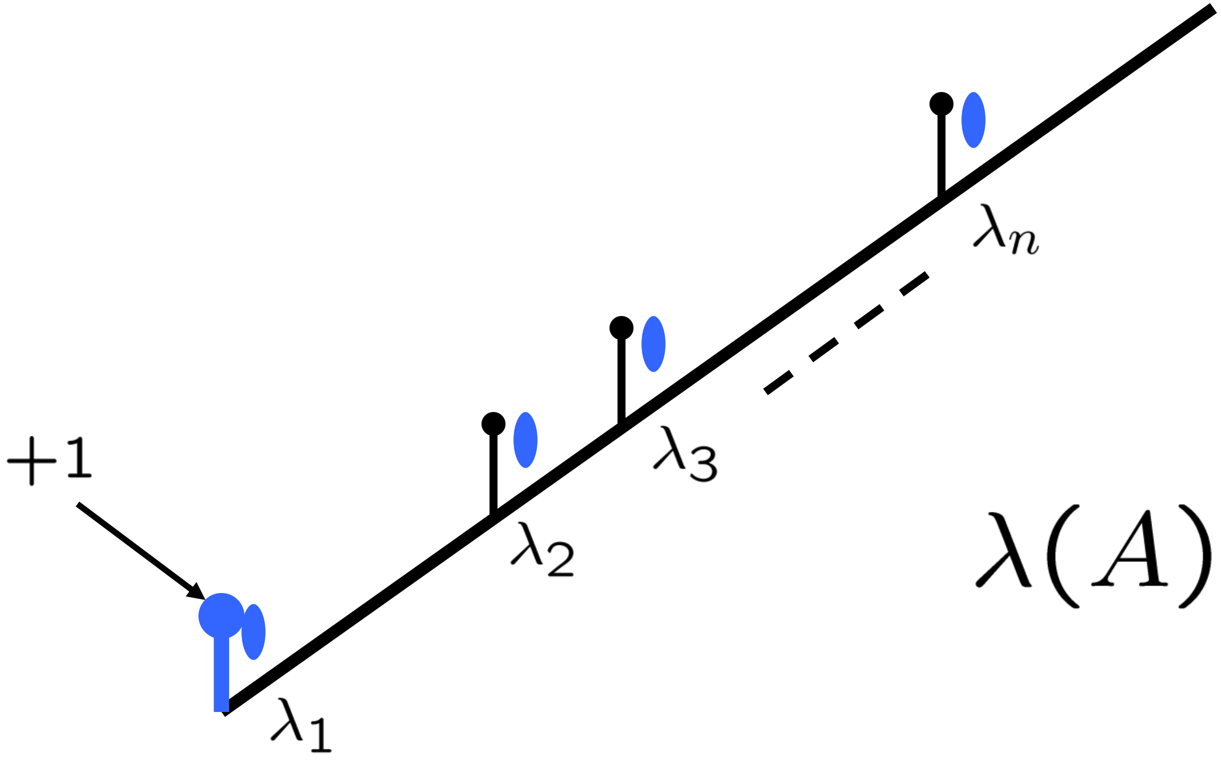
Barriers repel eigs.



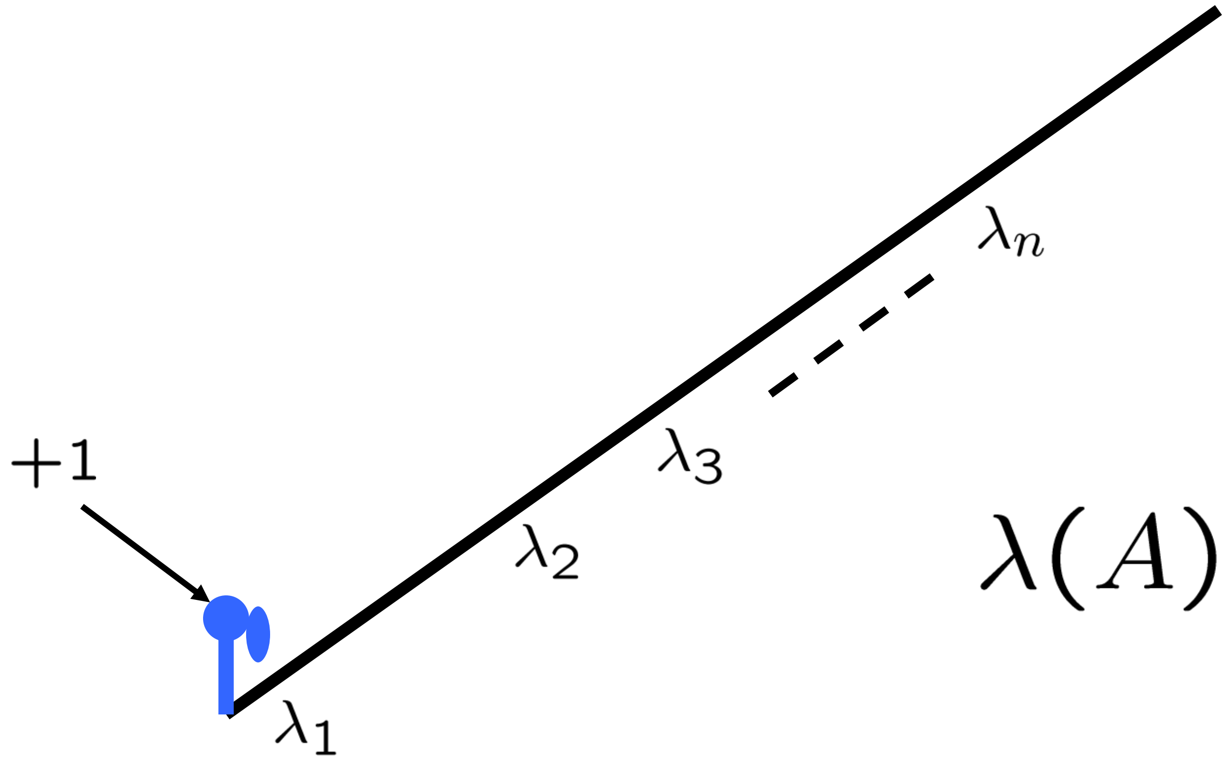
Examples



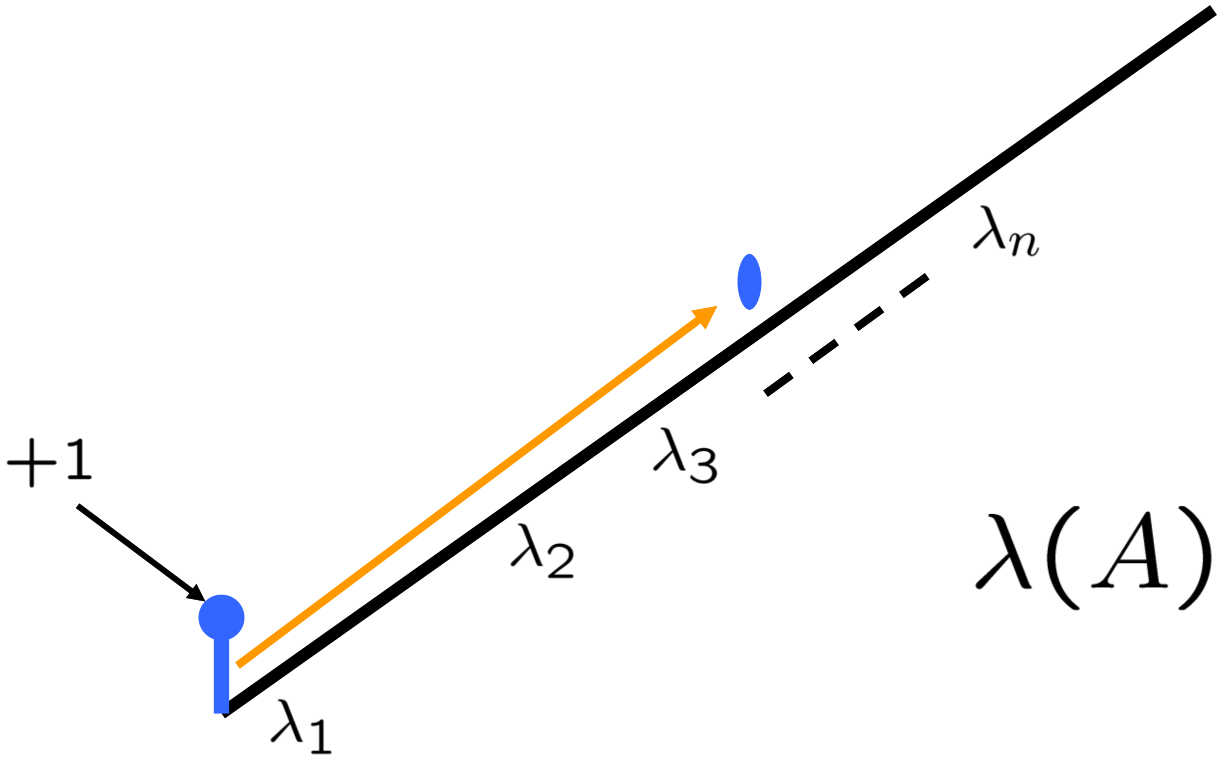
Ex1: All weight on u_1



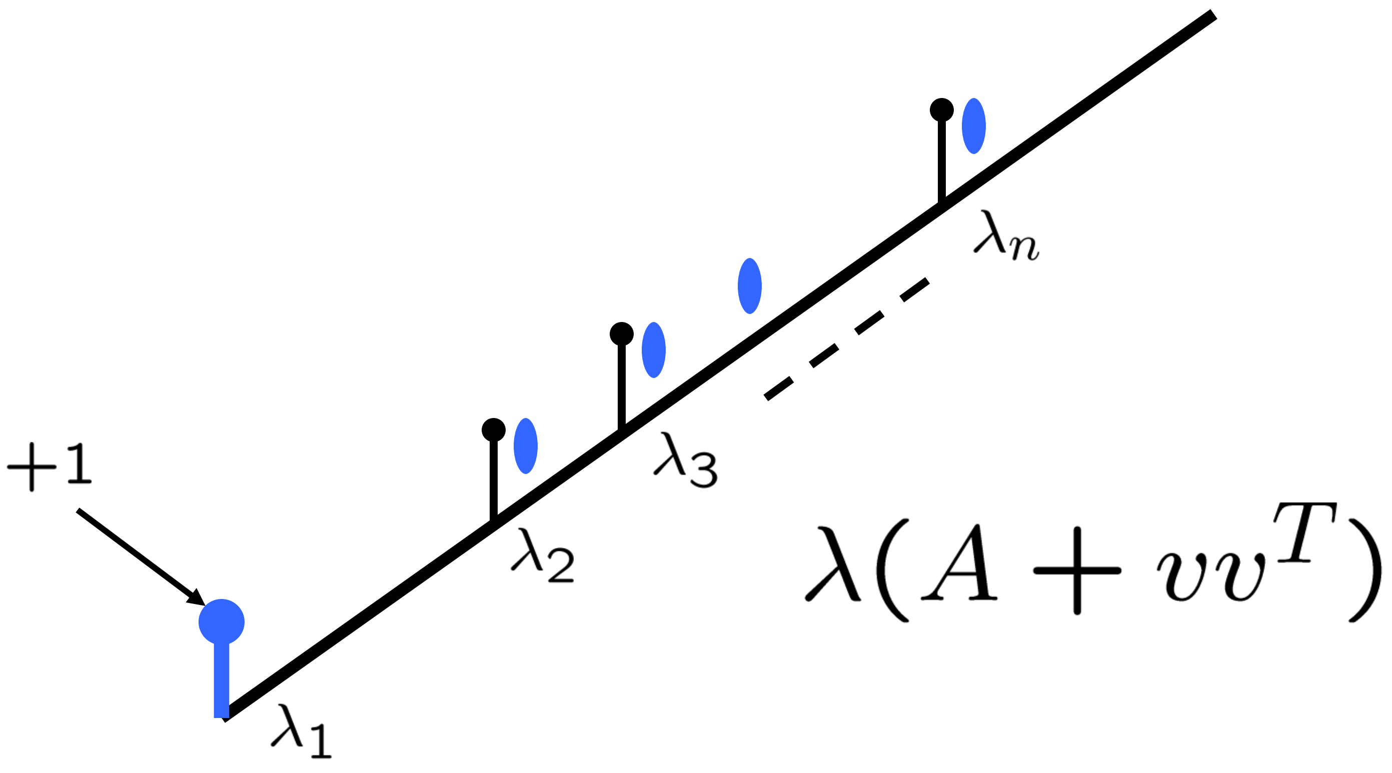
Ex1: All weight on u_1



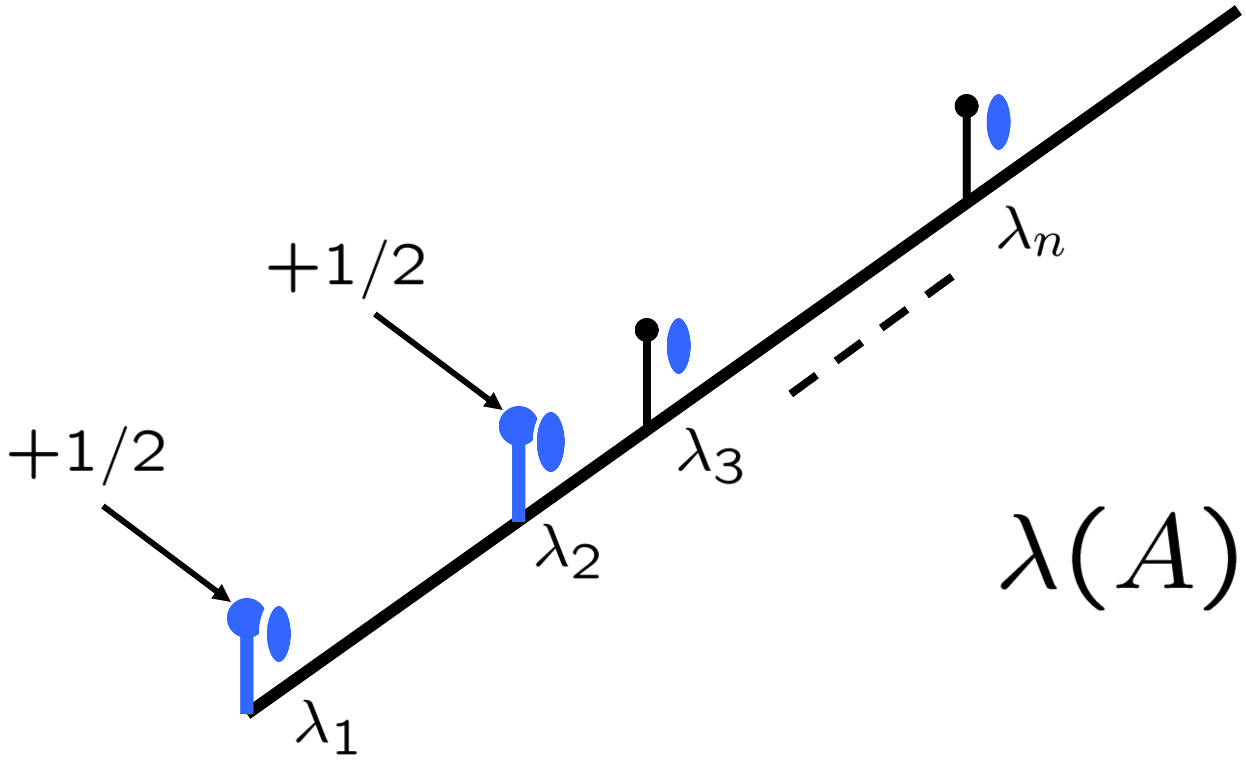
Ex1: All weight on u_1



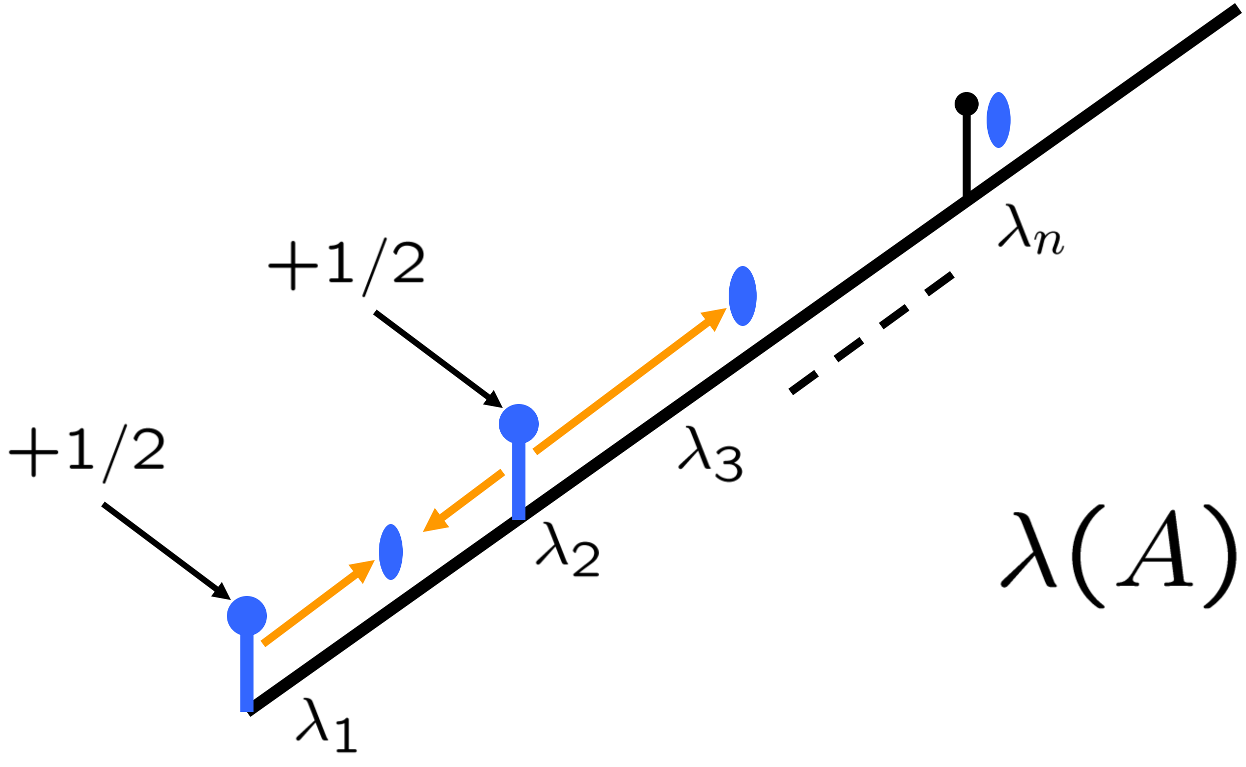
Ex1: All weight on u_1



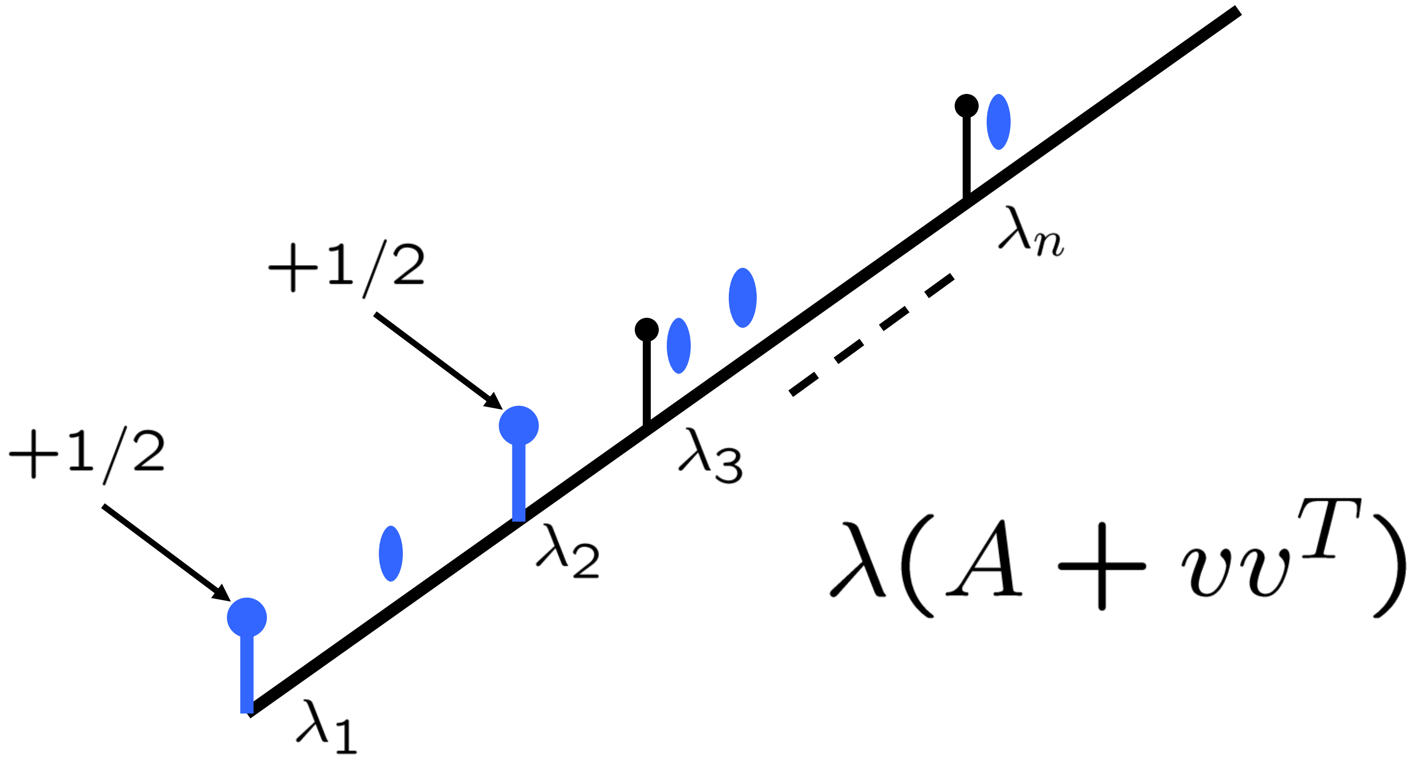
Ex2: Equal weight on u_1, u_2



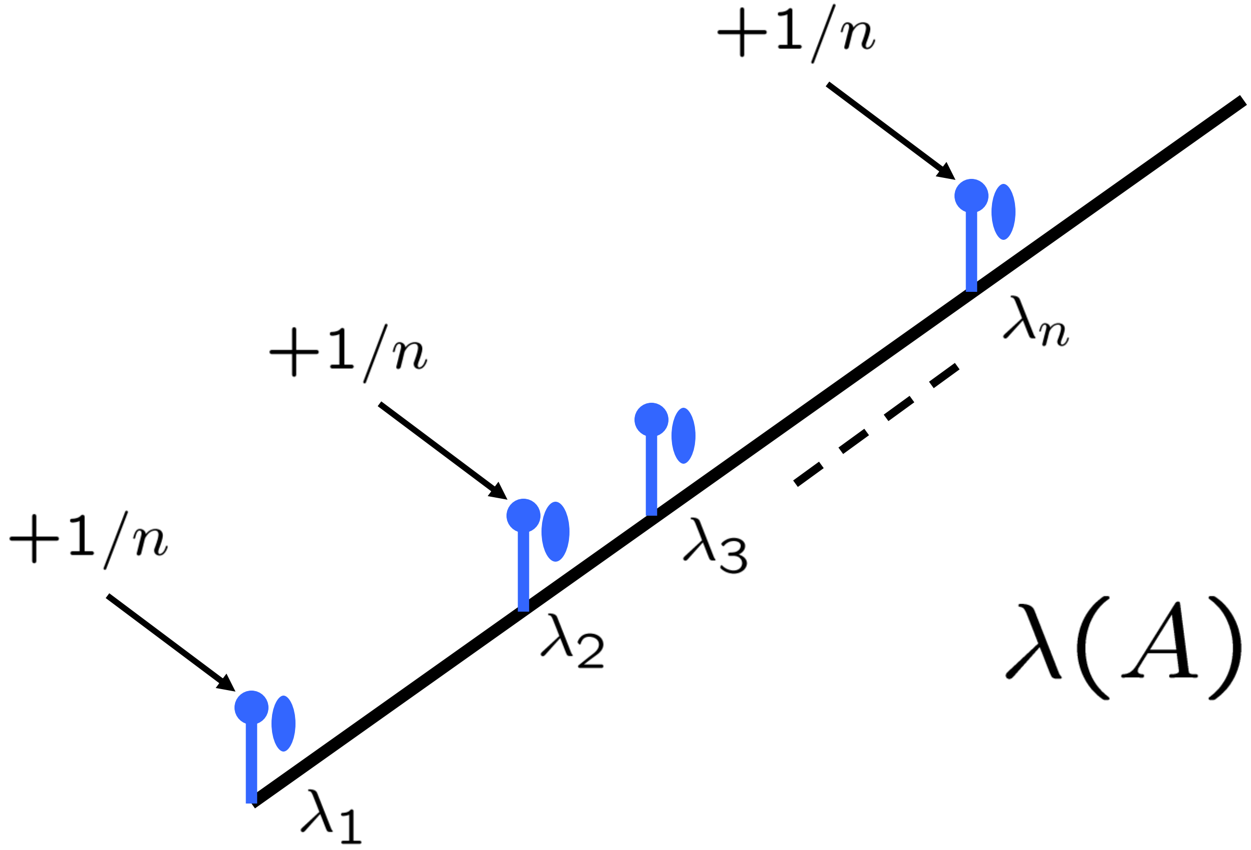
Ex2: Equal weight on u_1, u_2



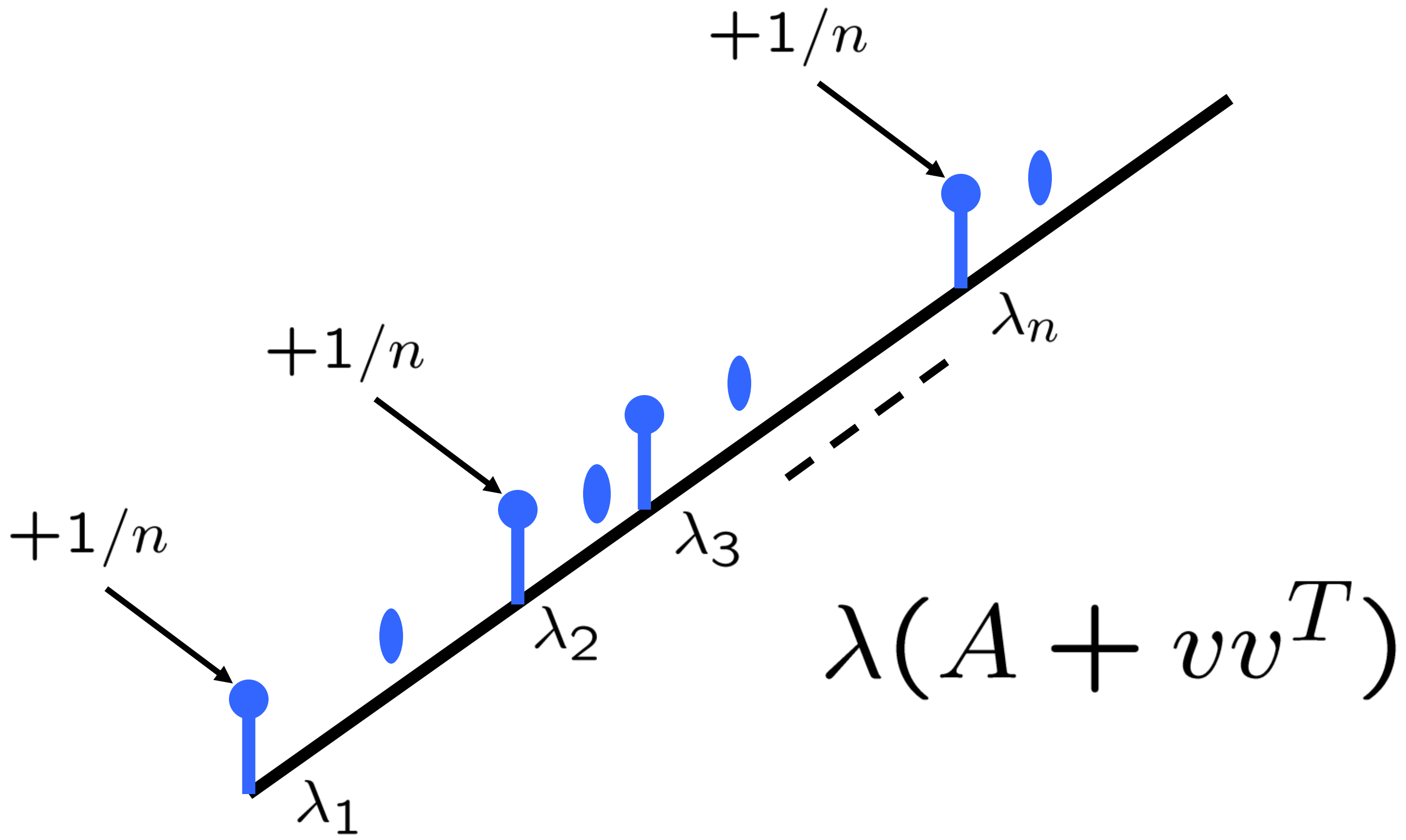
Ex2: Equal weight on u_1, u_2



Ex3: Equal weight on all u_1, u_2, …u_n



Ex3: Equal weight on all $u_1, u_2, \ldots u_n$



Adding a balanced vector

$$p_{A+vv^t} = p_A \left(1 + \sum_i \frac{\langle v, u_i \rangle^2}{\lambda_i - x}\right)$$

$$= p_A \left(1 + \sum_i \frac{1}{\lambda_i - x}\right)$$

$$= p_A - p'_A$$

Consider a random vector

If

$$\sum_e v_e v_e^T = I$$

For every u_i : $\sum_e \langle v_e, u_i \rangle^2 = 1.$

Consider a random vector

If

$$\sum_e v_e v_e^T = I$$

For every u_i : $\sum_e \langle v_e, u_i \rangle^2 = 1.$

thus a 'random' vector has the same expected projection in *every* direction ***i*** :

$$\mathbb{E}_e \langle v_e, u_i \rangle^2 = 1/m$$

Ideal proof

$$A^{(0)} = 0$$

$$p^{(0)} = x^n$$

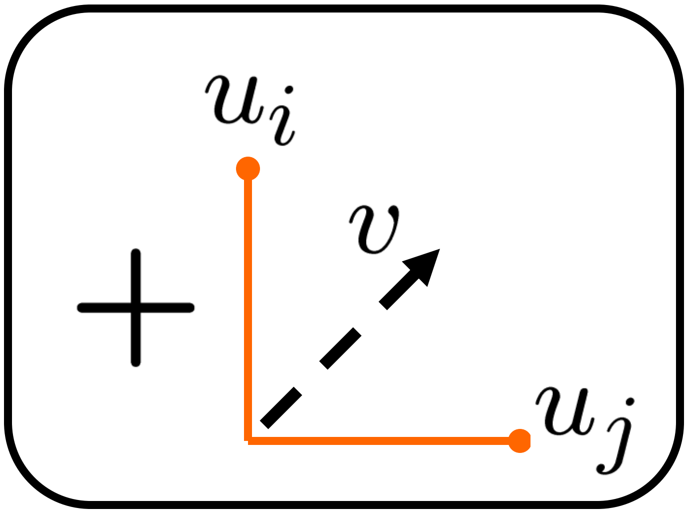
Ideal proof

$$\mathbb{E}_e \langle v_e, u_i \rangle^2 = 1/m$$

$$A^{(0)} = 0$$

$$p^{(0)} = x^n$$

Ideal proof

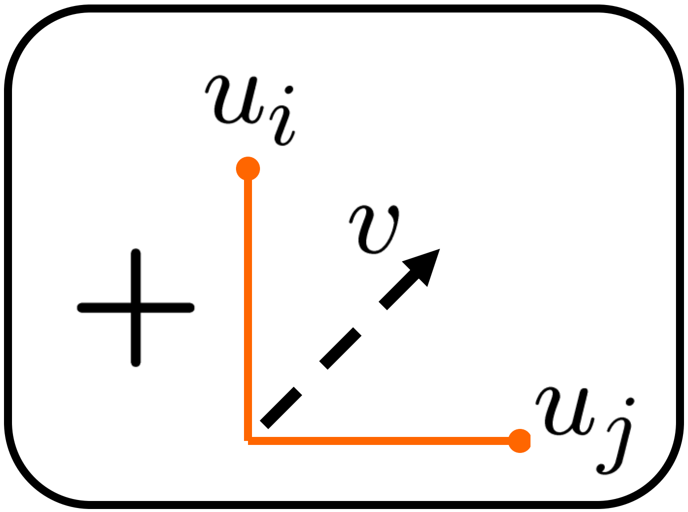


$$\mathbb{E}_e \langle v_e, u_i \rangle^2 = 1/m$$

$$A^{(1)} = 0 + vv^T$$

$$p^{(1)} = x^n - nx^{n-1}$$

Ideal proof

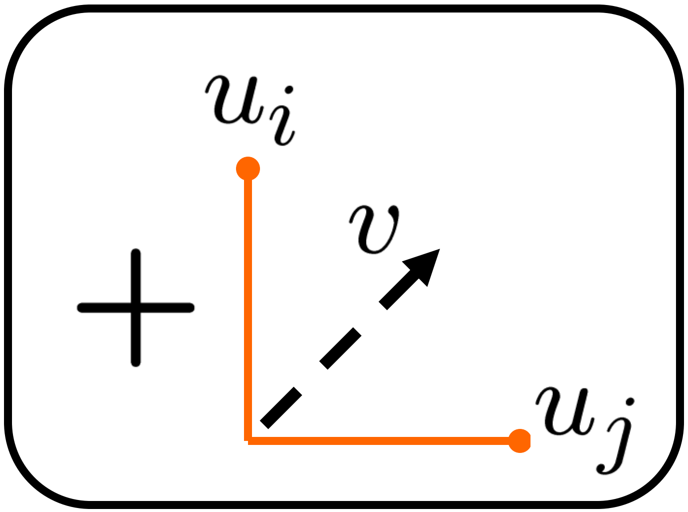


$$\mathbb{E}_e \langle v_e, u_i \rangle^2 = 1/m$$

$$A^{(2)} = A^{(1)} + vv^T$$

$$p^{(2)} = x^n - 2nx^{n-1} + n(n-1)x^{n-2}$$

Ideal proof

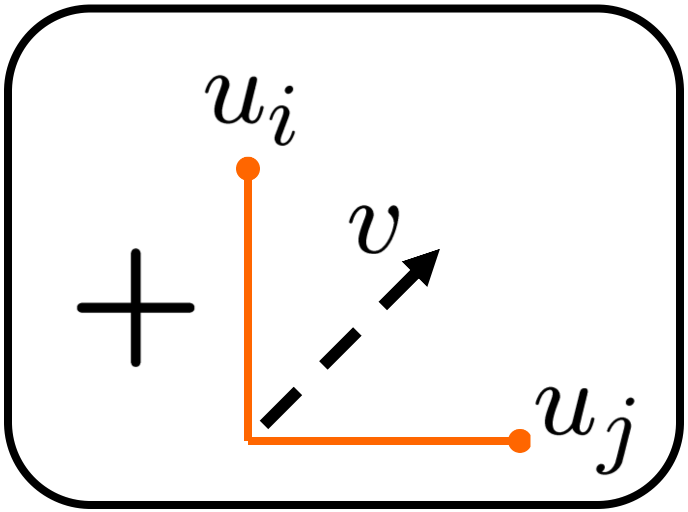


$$\mathbb{E}_e \langle v_e, u_i \rangle^2 = 1/m$$

$$A^{(3)} = A^{(2)} + vv^T$$

$$p^{(3)} = p^{(2)} - p^{(2)'}$$

Ideal proof

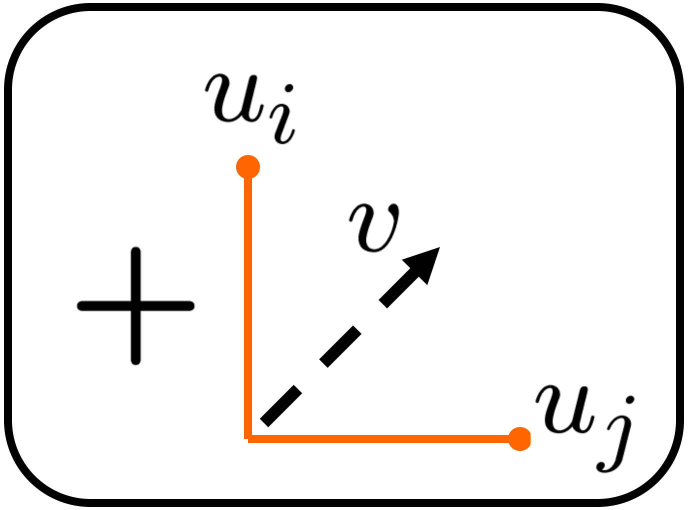


$$\mathbb{E}_e \langle v_e, u_i \rangle^2 = 1/m$$

$$A^{(i+1)} = A^{(i)} + vv^T$$

$$p^{(i+1)} = p^{(i)} - p^{(i)\prime}$$

Ideal proof

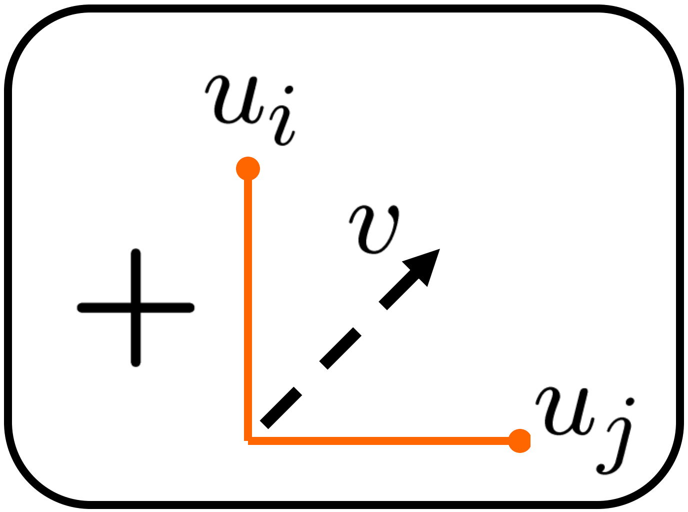


$$\mathbb{E}_e \langle v_e, u_i \rangle^2 = 1/m$$

$$A^{(i+1)} = A^{(i)} + vv^T$$

$$p^{(i+1)} = p^{(i)} - p^{(i)\prime}$$

Ideal proof

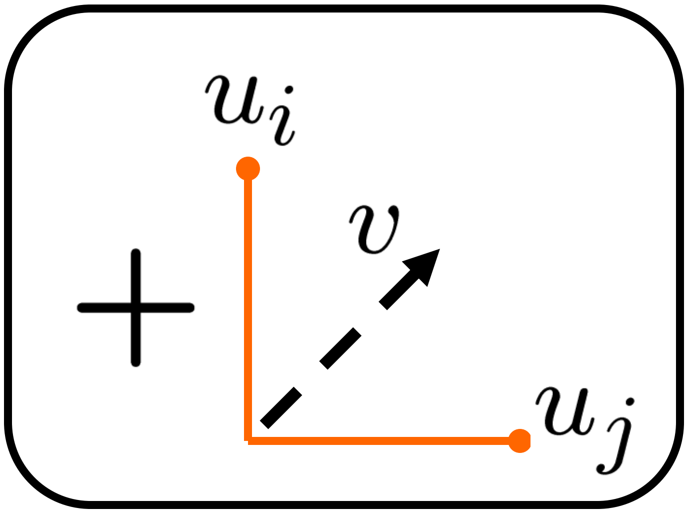


$$\mathbb{E}_e \langle v_e, u_i \rangle^2 = 1/m$$

$$A^{(i+1)} = A^{(i)} + vv^T$$

$$p^{(i+1)} = p^{(i)} - p^{(i)\prime}$$

Ideal proof

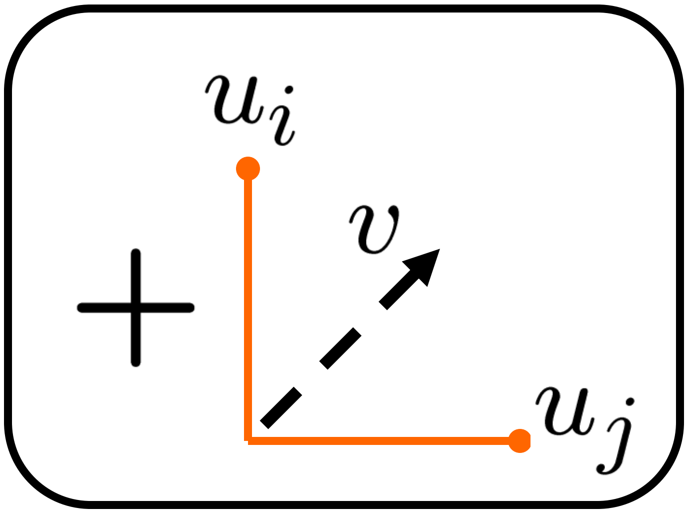


$$\mathbb{E}_e \langle v_e, u_i \rangle^2 = 1/m$$

$$A^{(i+1)} = A^{(i)} + vv^T$$

$$p^{(i+1)} = p^{(i)} - p^{(i)\prime}$$

Ideal proof

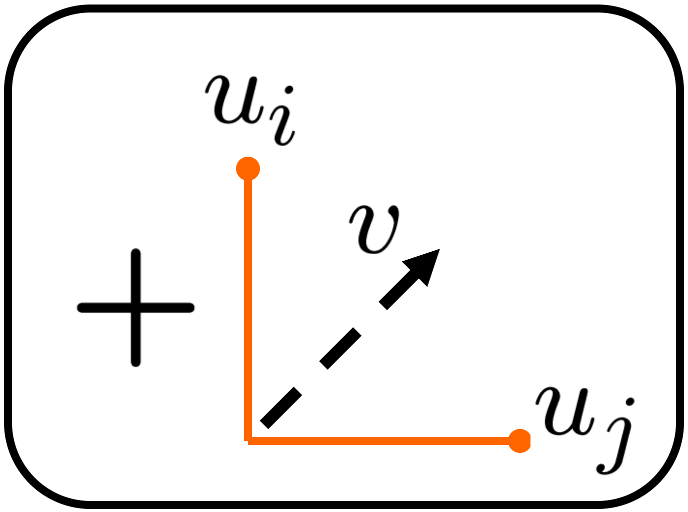


$$\mathbb{E}_e \langle v_e, u_i \rangle^2 = 1/m$$

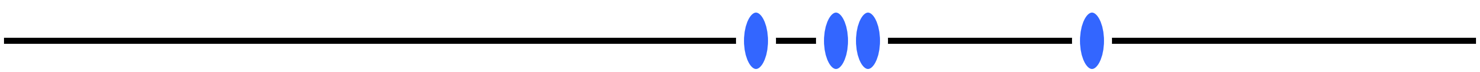
$$A^{(i+1)} = A^{(i)} + vv^T$$

$$p^{(i+1)} = p^{(i)} - p^{(i)\prime}$$

Ideal proof



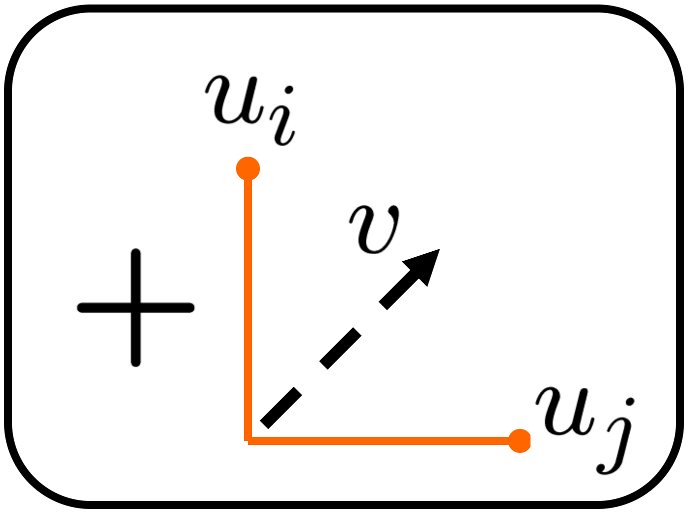
$$\mathbb{E}_e \langle v_e, u_i \rangle^2 = 1/m$$



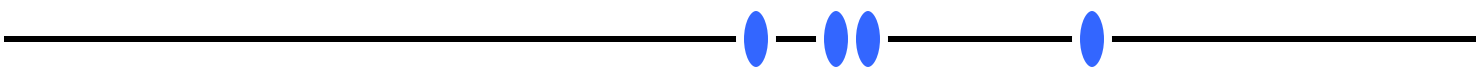
$$A^{(i+1)} = A^{(i)} + vv^T$$

$$p^{(i+1)} = p^{(i)} - p^{(i)\prime}$$

Ideal proof



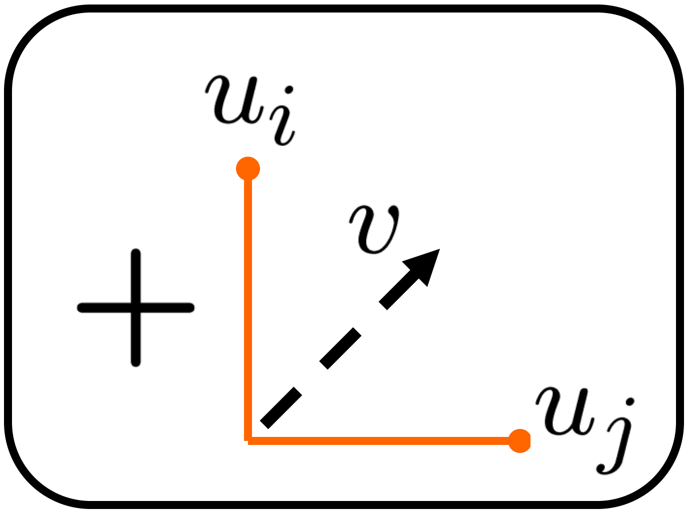
$$\mathbb{E}_e \langle v_e, u_i \rangle^2 = 1/m$$



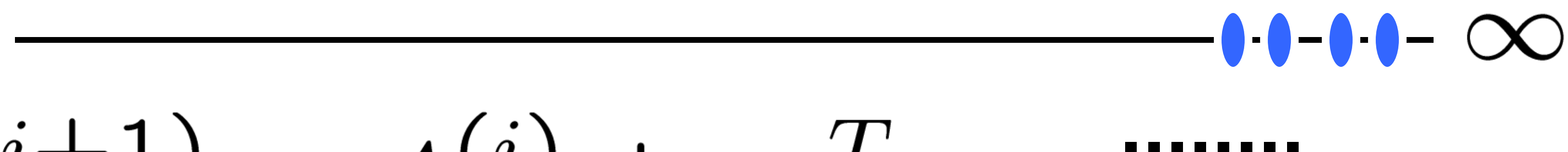
$$A^{(i+1)} = A^{(i)} + vv^T$$

$$p^{(i+1)} = p^{(i)} - p^{(i)\prime}$$

Ideal proof



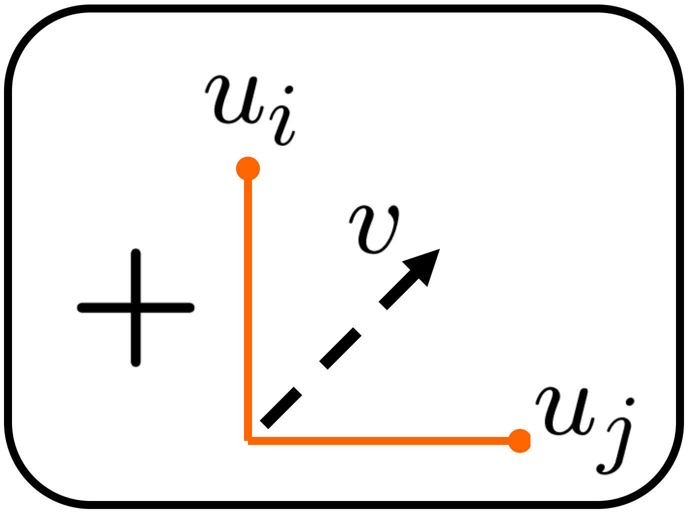
$$\mathbb{E}_e \langle v_e, u_i \rangle^2 = 1/m$$



$$A^{(i+1)} = A^{(i)} + vv^T$$

$$p^{(i+1)} = p^{(i)} - p^{(i)\prime}$$

Ideal proof



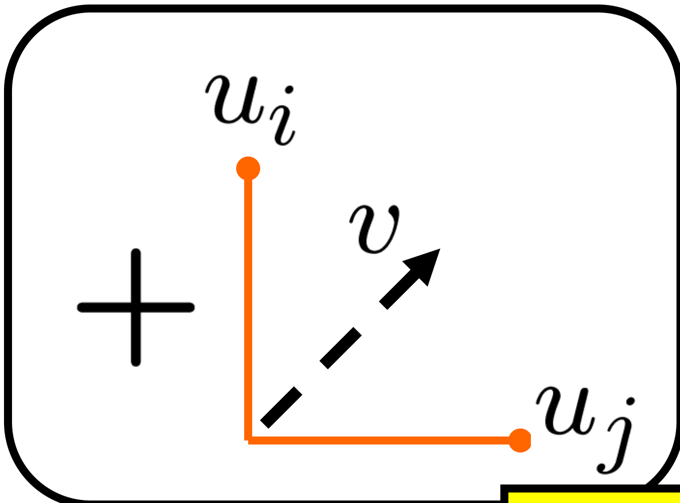
$$\mathbb{E}_e \langle v_e, u_i \rangle^2 = 1/m$$

$$A^{(i+1)} = A^{(i)} + vv^T$$

$$p^{(i+1)} = p^{(i)} - p^{(i)\prime}$$

$$\frac{\lambda_n(A)}{\lambda_1(A)} \leq 13?$$

Punch Line



$$\mathbb{E}_e\langle v_e, u_i\rangle^2 = 1/m$$

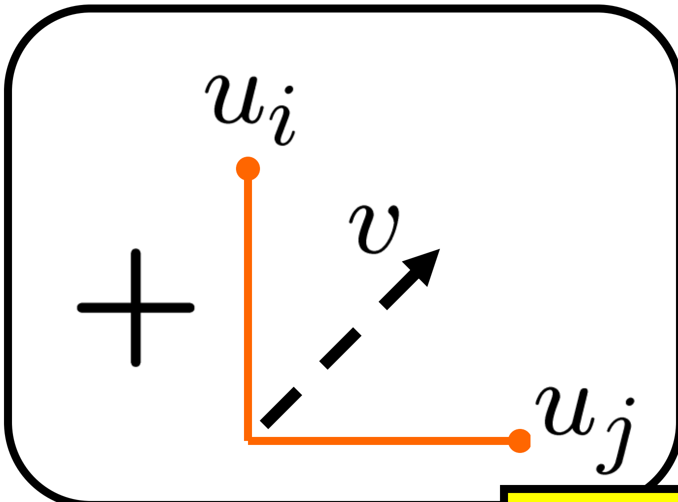
$$p^{(i)} = \text{Laguerre}^{(i)}$$

$$A^{(i+1)} =$$

$$p^{(i+1)} = p^{(i)} - p^{(i)\prime}$$

$$\frac{\lambda_n(A)}{\lambda_1(A)} \leq 13?$$

Punch Line



$$\mathbb{E}_e\langle v_e, u_i\rangle^2 = 1/m$$

$A^{(i+1)} =$

$p^{(i)} =$ Laguerre$^{(i)}$

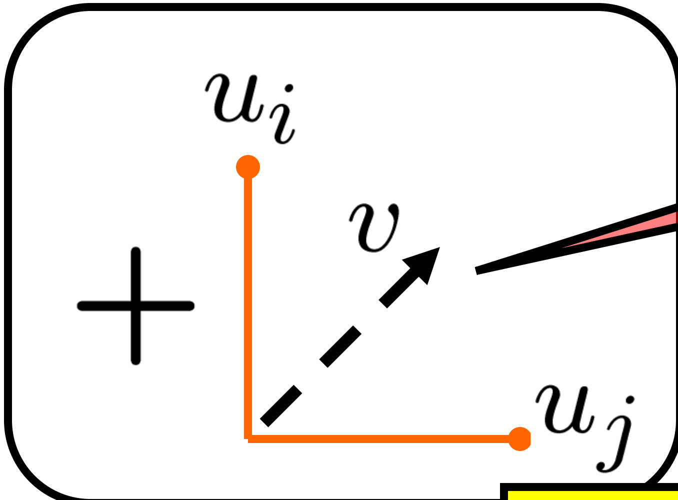
$p^{(i+1)} = p^{(i)} - p^{(i)\prime}$

$\lambda_n(A)$

In dn steps: $\frac{\lambda_n(A)}{\lambda_1(A)} \le \frac{d+2\sqrt{d-1}}{d-2\sqrt{d-2}}$

Punch Line

find actual vectors that realize this ideal behavior.



$$\mathbb{E}_e \langle v_e, u_i \rangle^2 = 1/m$$

$$p^{(i)} = \text{Laguerre}^{(i)}$$

$$A^{(i+1)} =$$

$$p^{(i+1)} = p^{(i)} - p^{(i)\prime}$$

$$\lambda_n(A) \le 13?$$

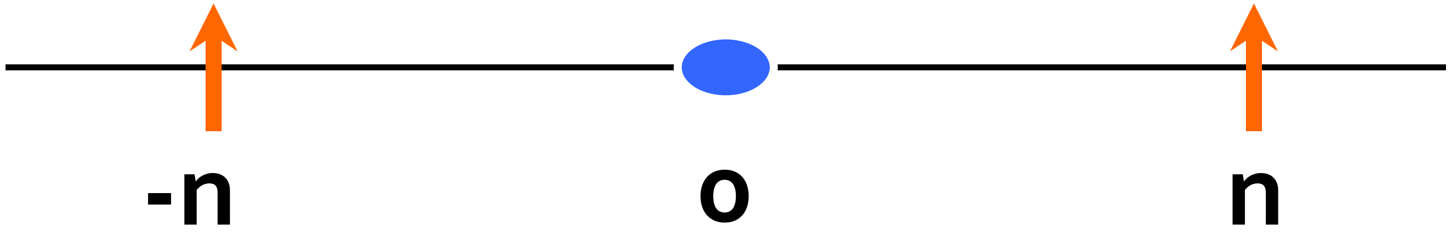
In dn steps: $\frac{\lambda_n(A)}{\lambda_1(A)} \leq \frac{d+2\sqrt{d-1}}{d-2\sqrt{d-2}}$

Step 3: Actual Proof

(for 6n vectors, 13-approx)

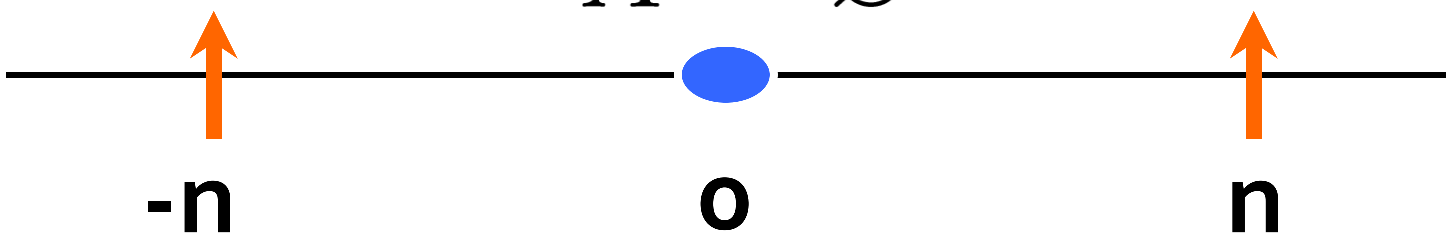
Broad outline: moving barriers

$$A = \varnothing$$



Step 1

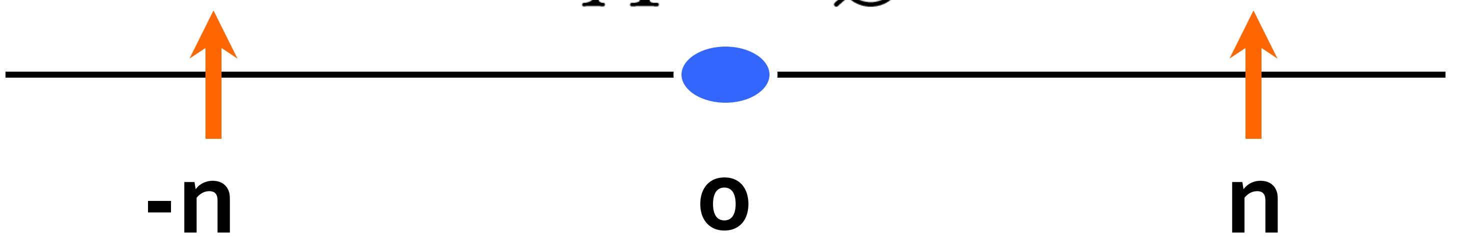
$$A = \varnothing$$



$$+\mathbf{v}\mathbf{v}^T \qquad \mathbf{v} \in \{v_e\}$$

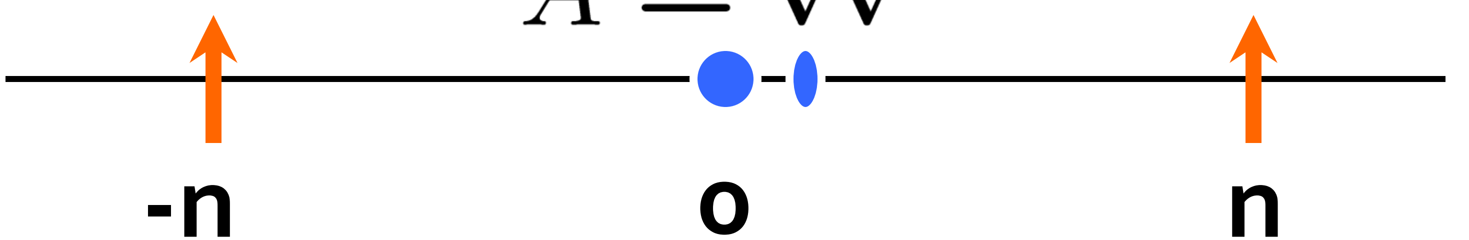
Step 1

$$A = \varnothing$$



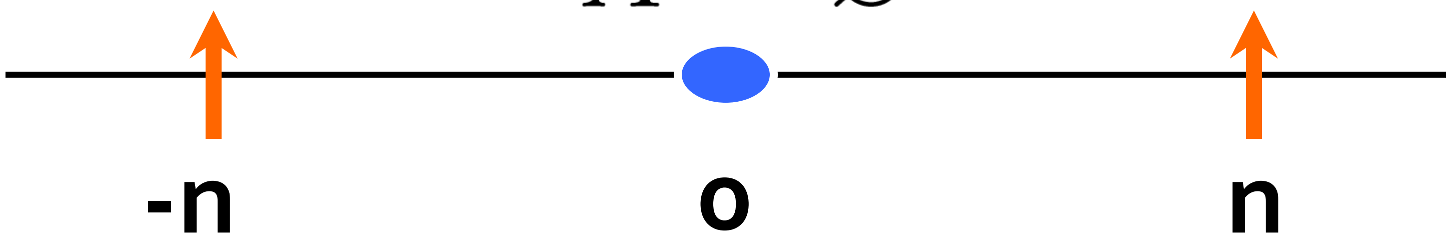
$$+\mathbf{v}\mathbf{v}^T \qquad \mathbf{v} \in \{v_e\}$$

$$A = \mathbf{v}\mathbf{v}^T$$



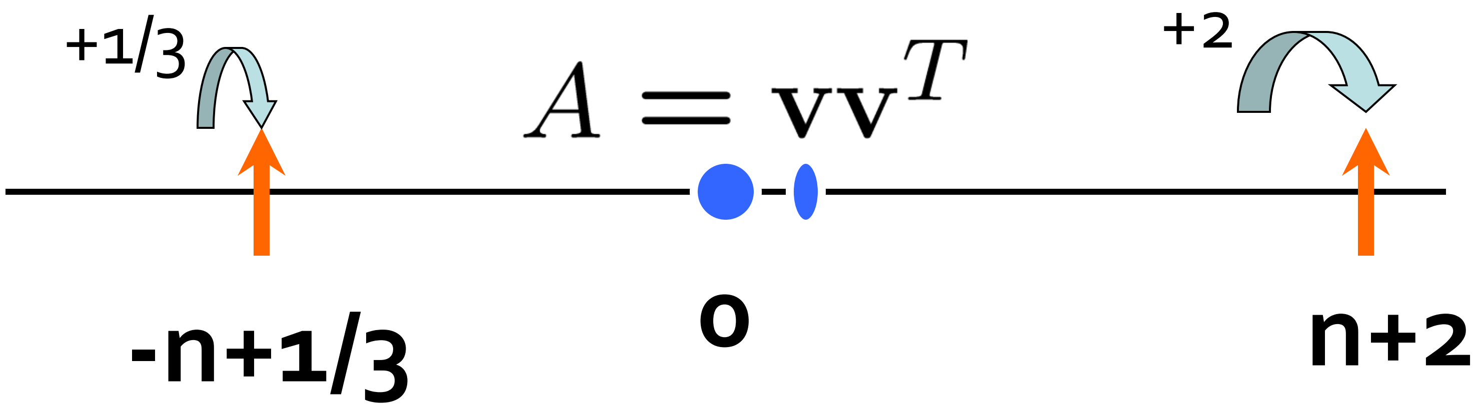
Step 1

$$A = \varnothing$$

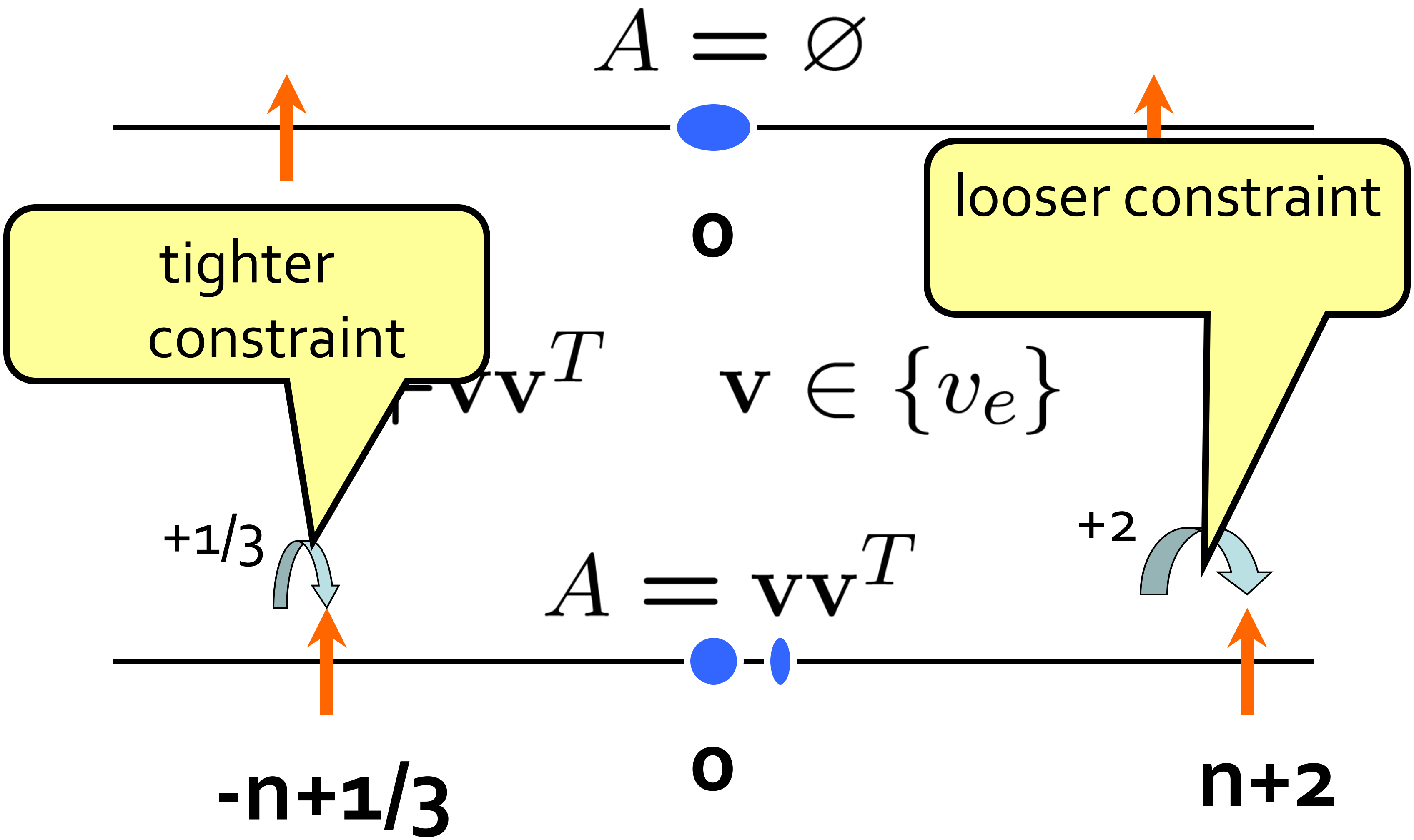


$$+\mathbf{v}\mathbf{v}^T \qquad \mathbf{v} \in \{v_e\}$$

$$A = \mathbf{v}\mathbf{v}^T$$



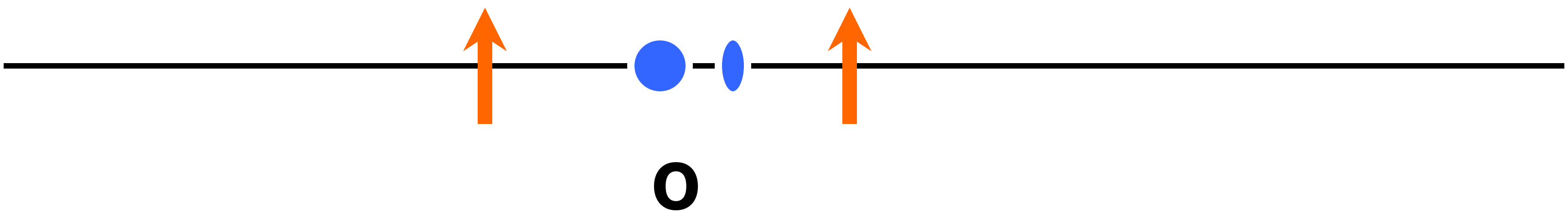
Step 1

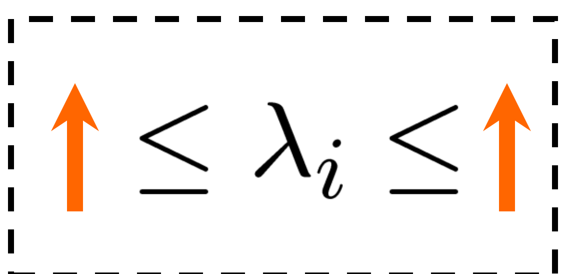


$A = \varnothing$
0
tighter constraint
looser constraint
$\mathbf{v}\mathbf{v}^T$
$\mathbf{v} \in \{v_e\}$
+1/3
+2
$A = \mathbf{v}\mathbf{v}^T$
-n+1/3
0
n+2

Step i+1

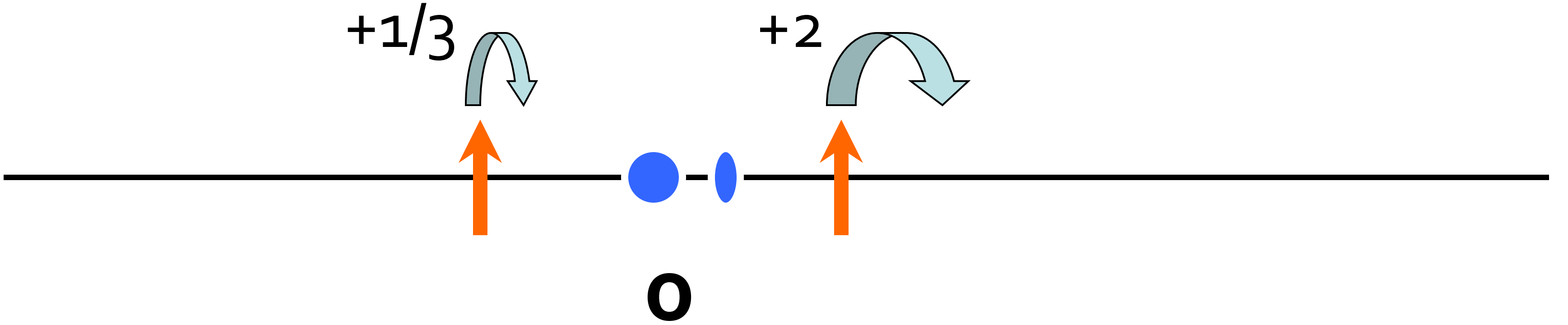
$A^{(i)}$





Step i+1

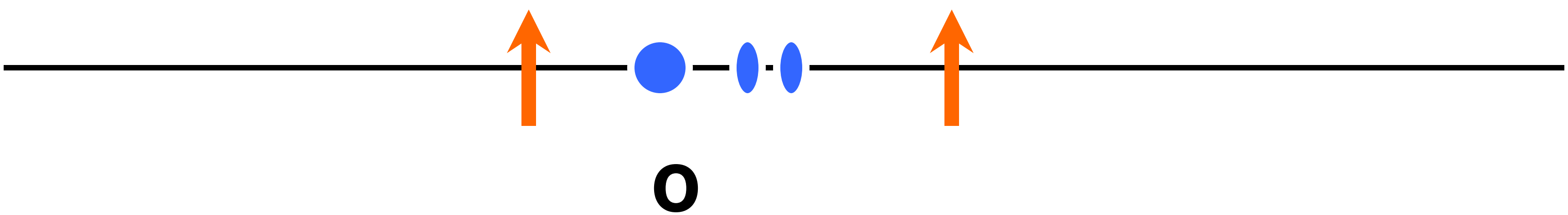
$A^{(i)}$

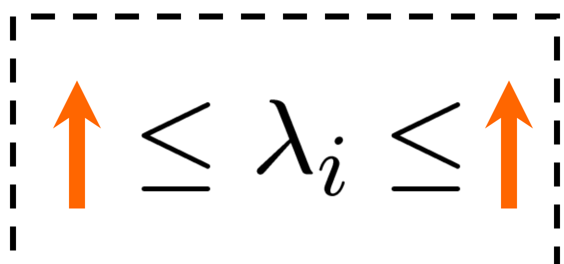


$+\mathbf{v}\mathbf{v}^T$

Step i+1

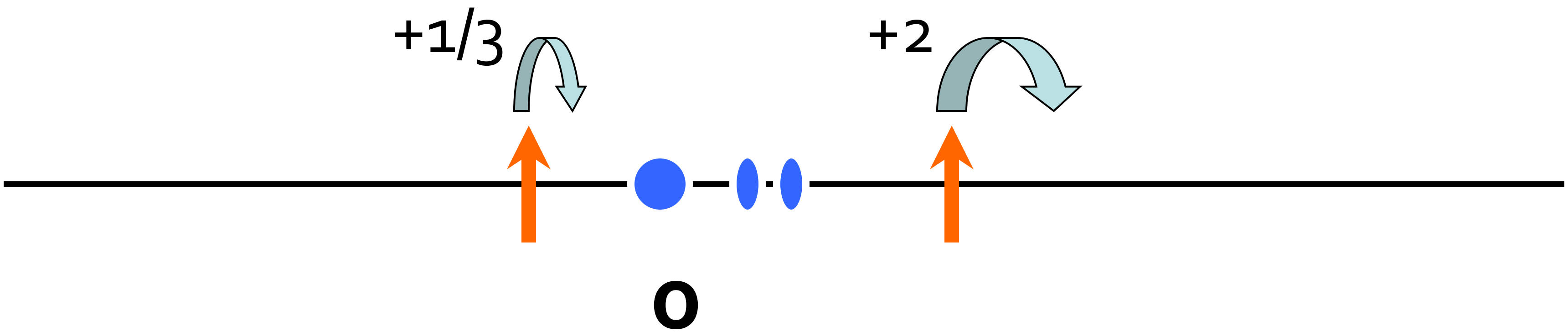
$$A^{(i)}, A^{(i+1)}$$





Step i+1

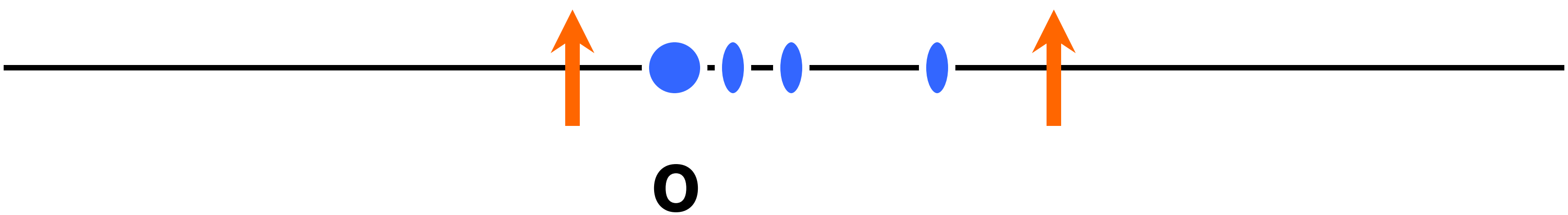
$A^{(i)}, A^{(i+1)}$

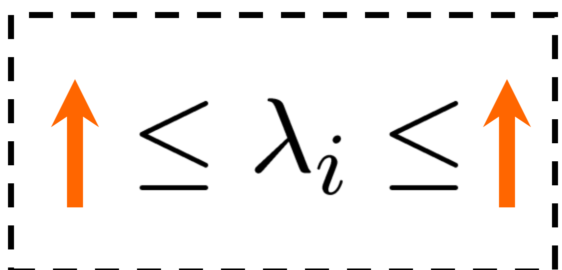


$+\mathbf{v}\mathbf{v}^T$

Step i+1

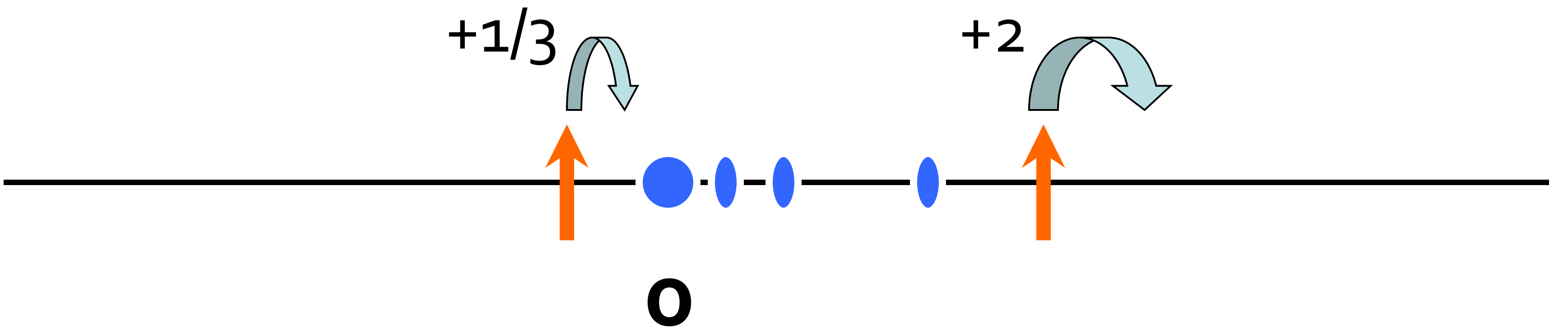
$$A^{(i)}, A^{(i+1)}, A^{(i+2)}$$





Step i+1

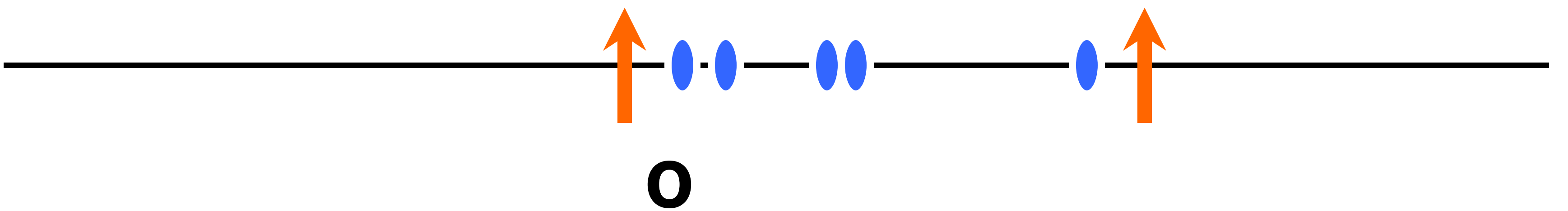
$$A^{(i)}, A^{(i+1)}, A^{(i+2)}$$

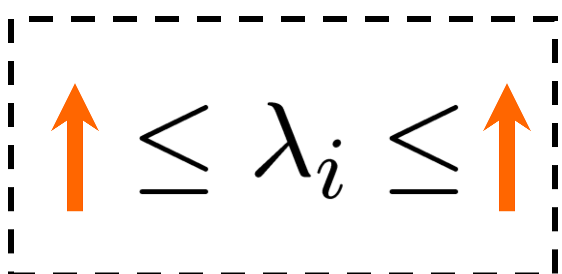


$$+\mathbf{v}\mathbf{v}^T$$

Step i+1

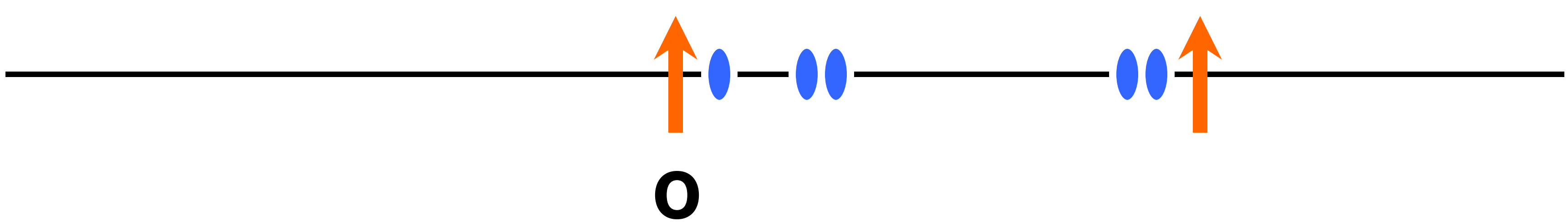
$$A^{(i)}, A^{(i+1)}, A^{(i+2)}, A^{(i+3)}$$

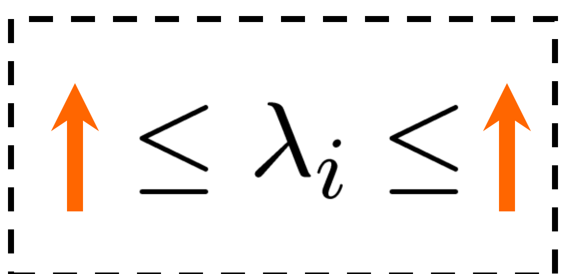




Step i+1

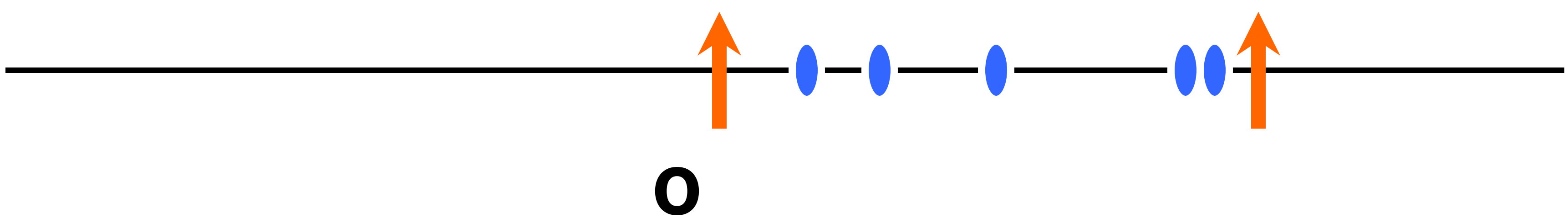
$$A^{(i)}, A^{(i+1)}, A^{(i+2)}, A^{(i+3)}, \ldots$$

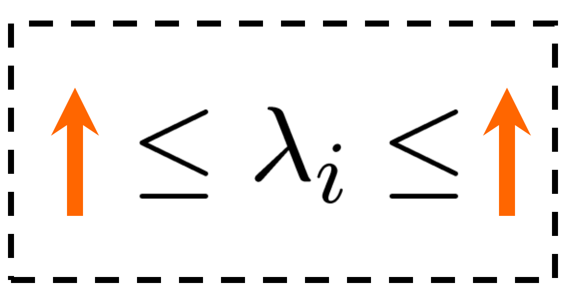




Step i+1

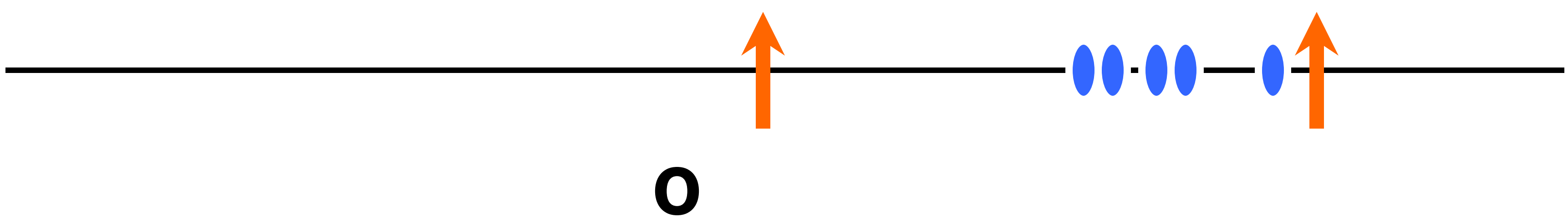
$$A^{(i)}, A^{(i+1)}, A^{(i+2)}, A^{(i+3)}, \ldots$$

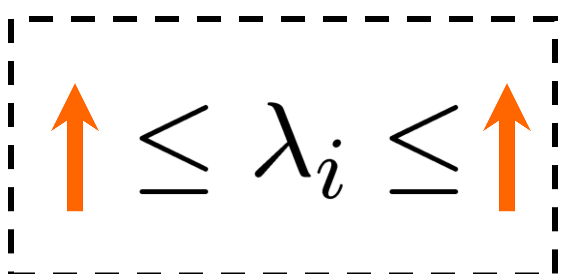




Step i+1

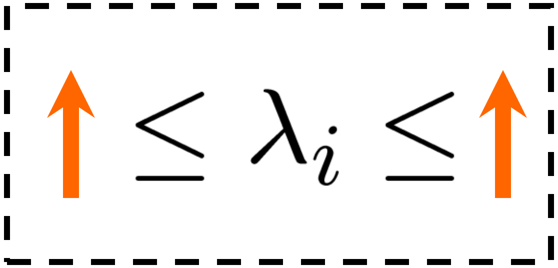
$$A^{(i)}, A^{(i+1)}, A^{(i+2)}, A^{(i+3)}, \ldots$$





Step 6n

$$A^{(i)}, A^{(i+1)}, A^{(i+2)}, A^{(i+3)}, \ldots, A^{(6n)}$$



Step 6n

$$A^{(i)}, A^{(i+1)}, A^{(i+2)}, A^{(i+3)}, \ldots, A^{(6n)}$$

13-approximation with 6n vectors.

$\square$

Problem

need to show that an appropriate

$$v_e v_e^T$$

always exists.

Problem

need to show that an appropriate

$$v_e v_e^T$$

always exists.

$$\uparrow \leq \lambda_i \leq \uparrow$$

is not strong enough to do the induction.

Problem

need to show that an appropriate

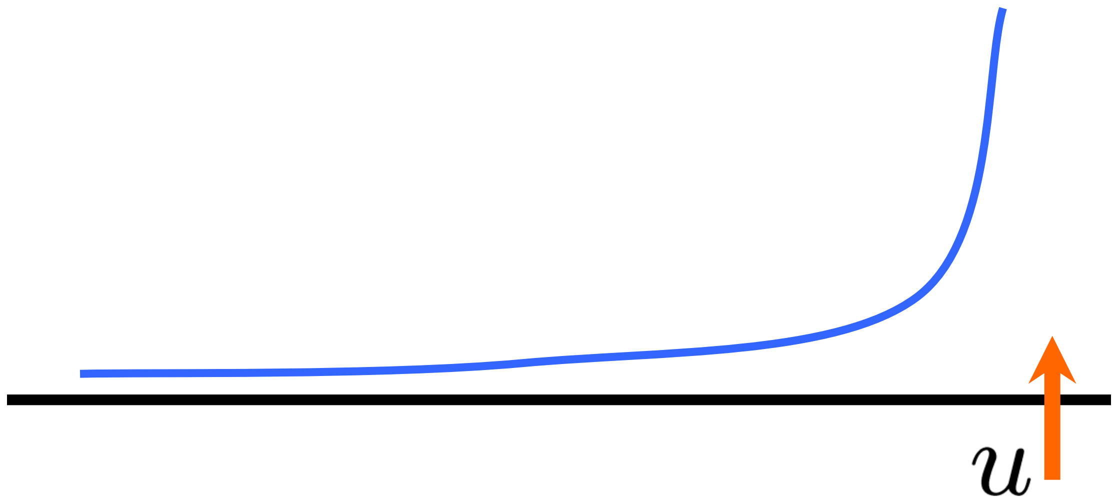
need a better way to measure quality of eigenvalues.

$$\uparrow \leq \lambda_i \leq \uparrow$$

is not strong enough to do the induction.

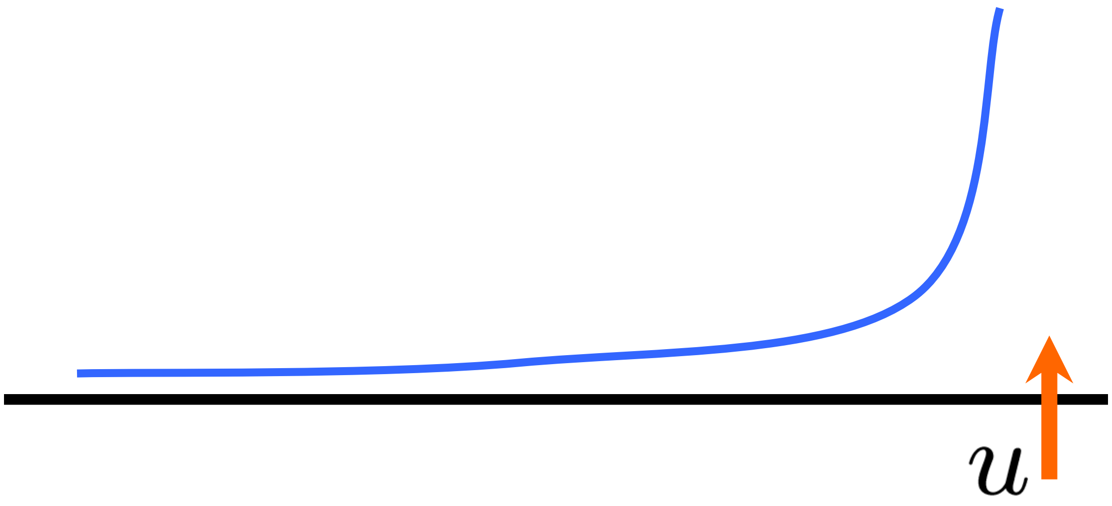
The Upper Barrier

$$\Phi^u(A) = \mathrm{Tr}(uI - A)^{-1} = \sum_i \frac{1}{u-\lambda_i}$$



The Upper Barrier

$$\Phi^u(A) = \mathrm{Tr}(uI - A)^{-1} = \sum_i \frac{1}{u - \lambda_i}$$



$$\Phi^u(A) \leq 1 \Rightarrow \lambda_{\max}(A) \ll u$$

The Upper Barrier

$$\Phi^u(A) = \mathrm{Tr}(uI - A)^{-1} = \sum_i \frac{1}{u-\lambda_i}$$

No λ_i within dist. **1**

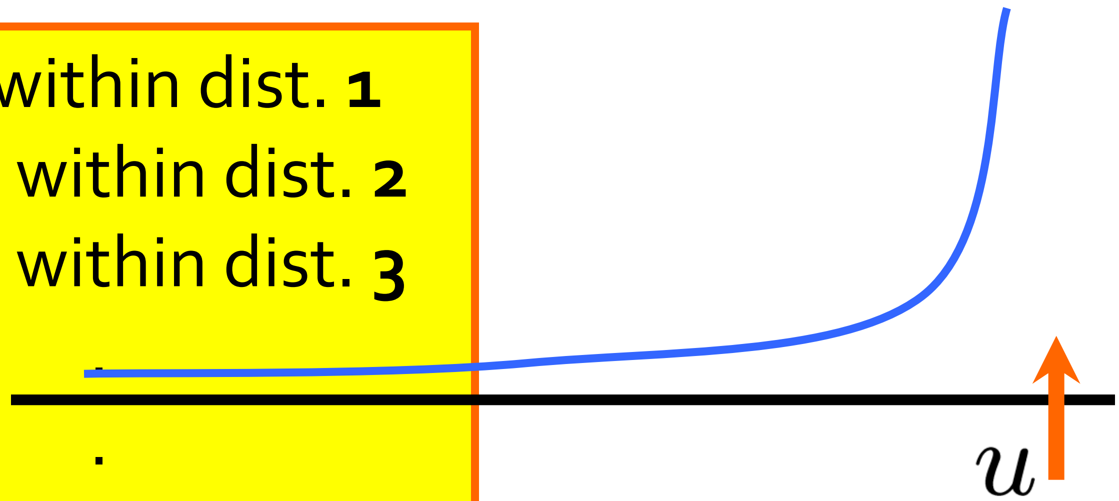
No **2** λ_i within dist. **2**

No **3** λ_i within dist. **3**

.

.

No **k** λ_i within dist. **k**



$$\Phi^u(A) \leq 1 \Rightarrow \lambda_{\mathrm{max}}(A) \ll u$$

‘Total repulsion’ in physical model

…per Barrier

$$\Phi^u(A) = \mathrm{Tr}(uI - A)^{-1} = \sum_i \frac{1}{u-\lambda_i}$$

No λ_i within dist. **1**

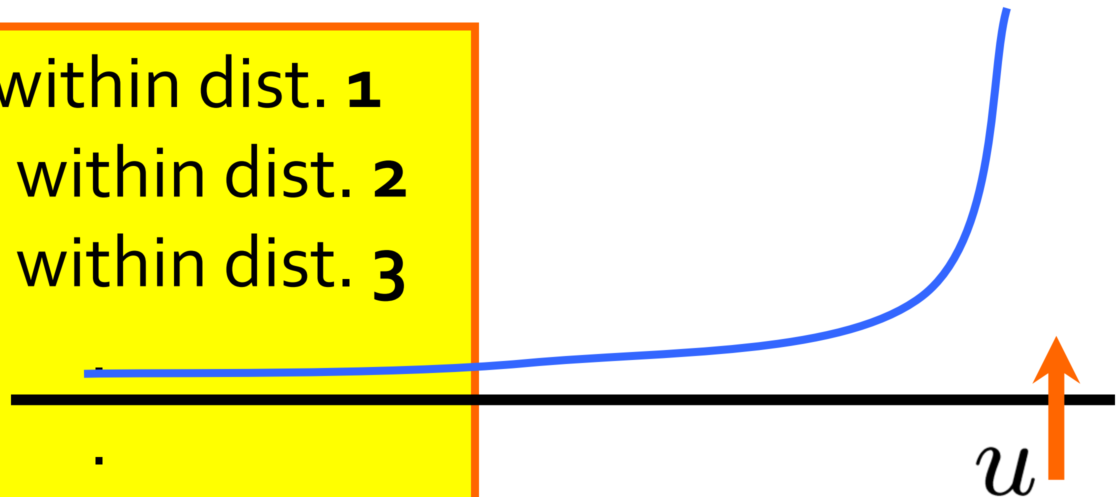
No **2** λ_i within dist. **2**

No **3** λ_i within dist. **3**

.

.

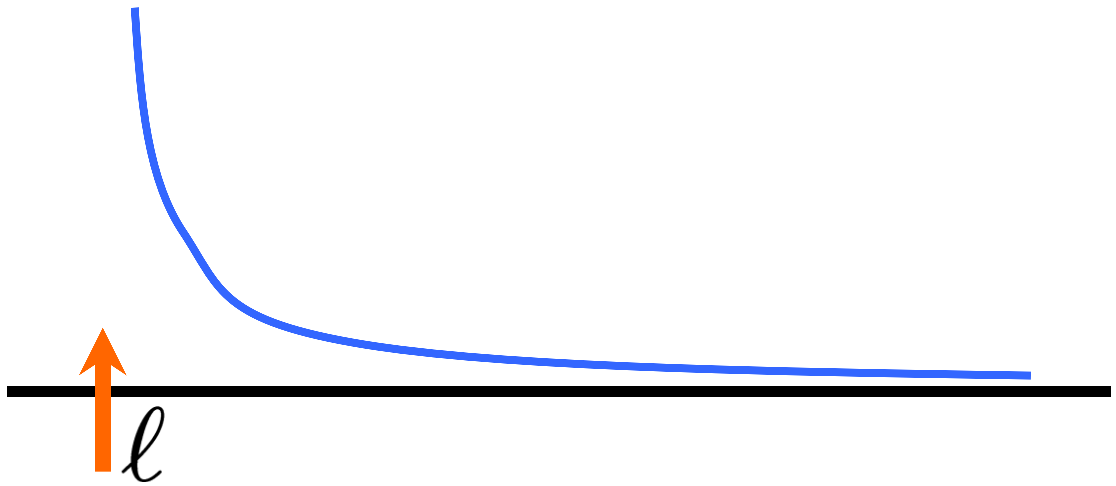
No **k** λ_i within dist. **k**



$$\Phi^u(A) \leq 1 \Rightarrow \lambda_{\max}(A) \ll u$$

The Lower Barrier

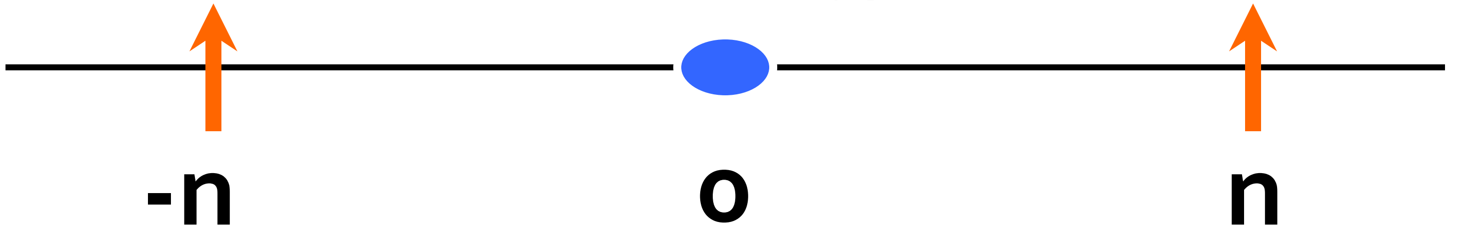
$$\Phi_{\ell}(A) = \mathrm{Tr}(A - \ell I)^{-1} = \sum_i \frac{1}{\lambda_i - \ell}$$



$$\Phi_{\ell}(A) \leq 1 \Rightarrow \lambda_{\mathrm{min}}(A) \gg \ell$$

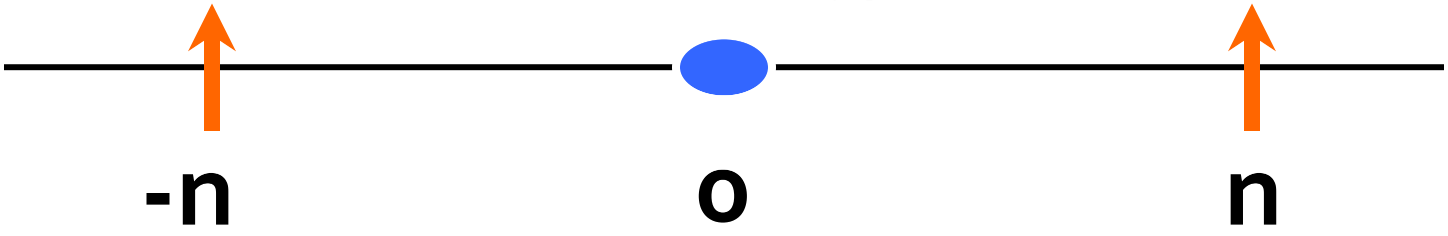
The Beginning

$$A = \varnothing$$



The Beginning

$$A = \varnothing$$

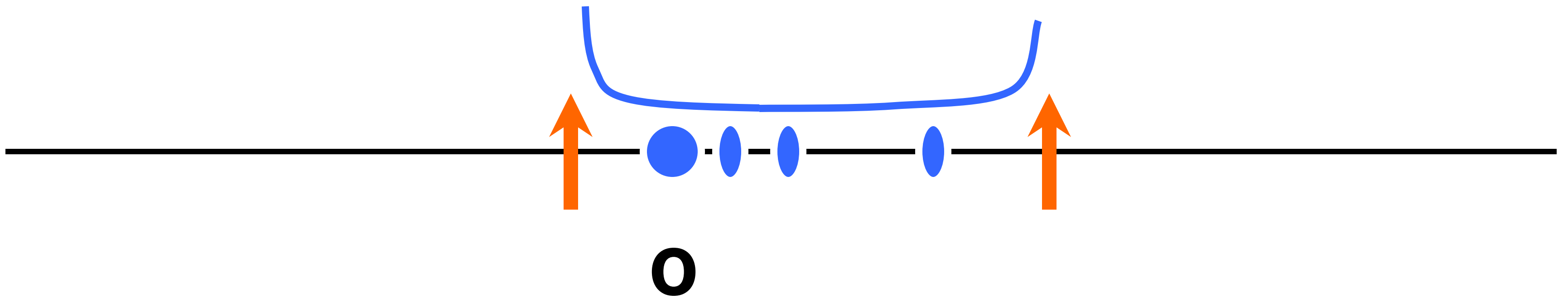


$$\Phi^n(\varnothing) = \mathrm{Tr}(nI)^{-1} = 1$$

$$\Phi_{-n}(\varnothing) = \mathrm{Tr}(nI)^{-1} = 1.$$

Step i+1

$$A^{(i)}, A^{(i+1)}, A^{(i+2)}$$

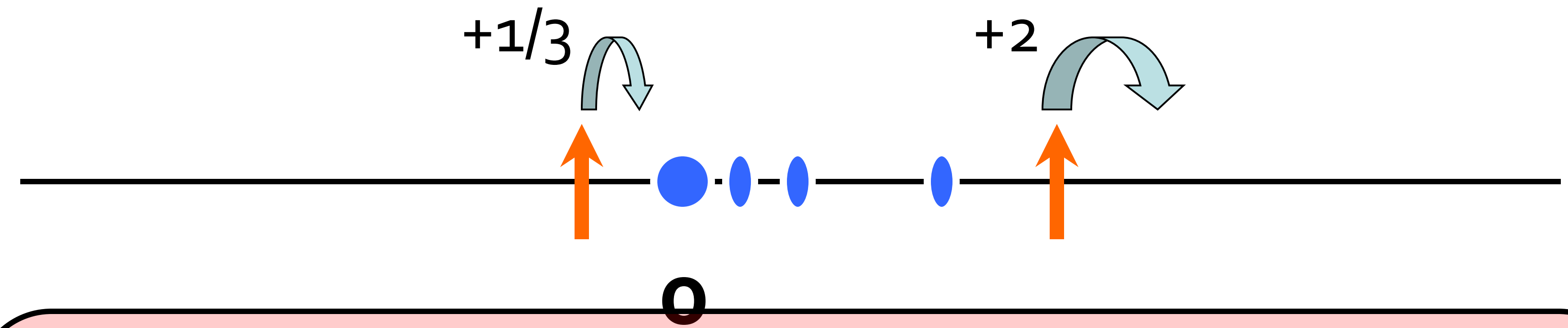


$$\Phi^{u}(A) \leq 1$$

$$\Phi_{\ell}(A) \leq 1.$$

Step i+1

$$A^{(i)}, A^{(i+1)}, A^{(i+2)}$$



Lemma.

can always choose $+s\mathbf{v}\mathbf{v}^T$

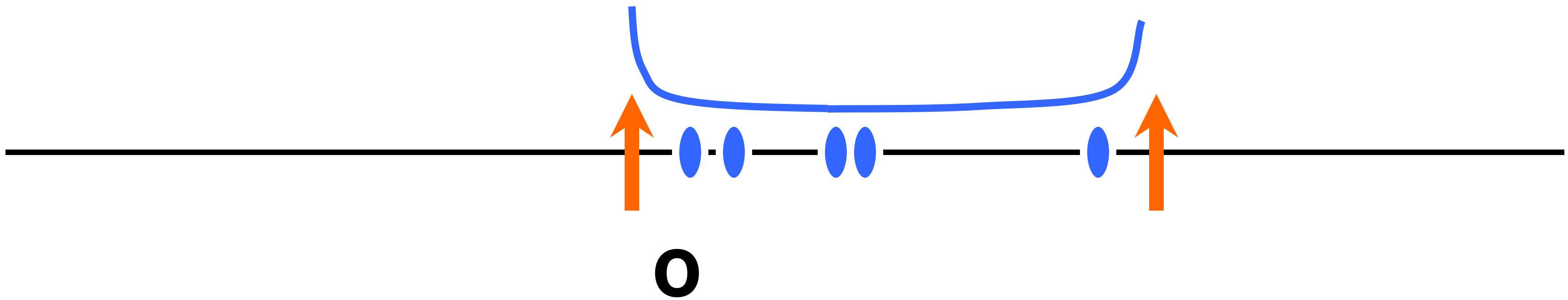
so that potentials do not increase

$$\Phi^u(A) \leq 1$$

$$\Phi_\ell(A) \leq 1.$$

Step i+1

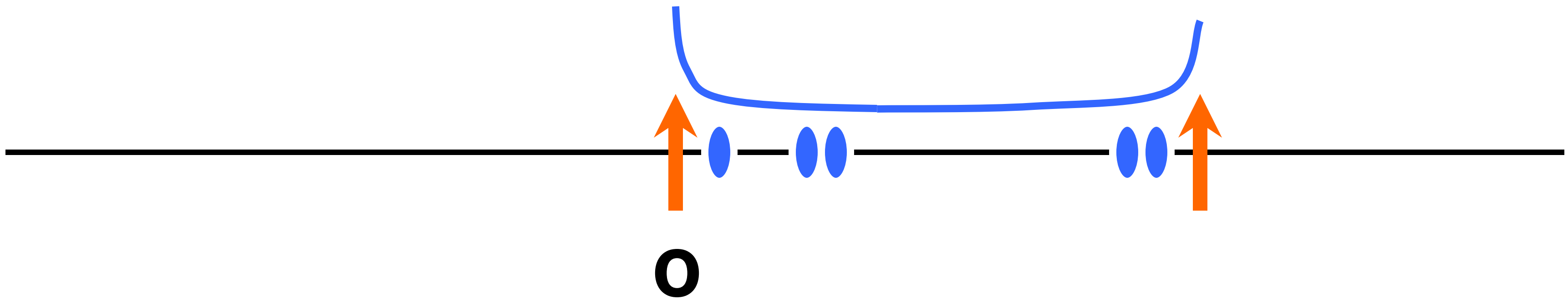
$$A^{(i)}, A^{(i+1)}, A^{(i+2)}, A^{(i+3)}$$



$$\Phi^u(A) \leq 1$$
$$\Phi_\ell(A) \leq 1.$$

Step i+1

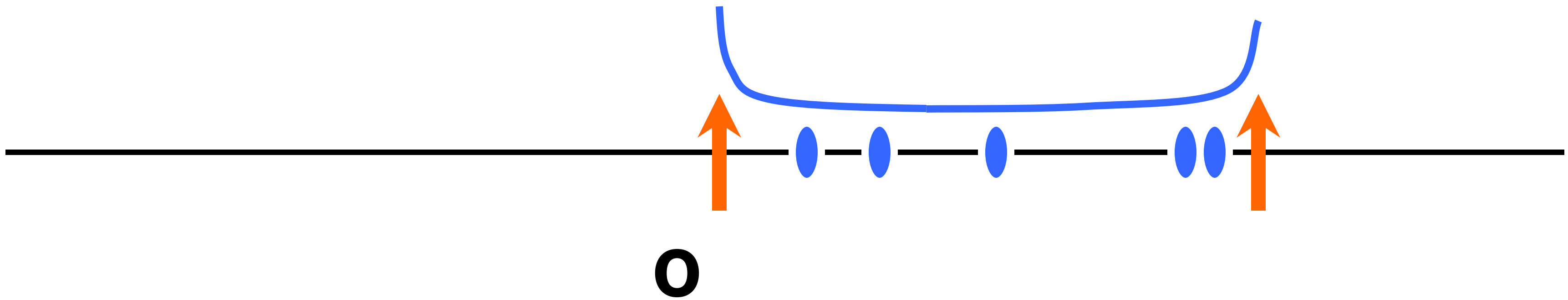
$$A^{(i)}, A^{(i+1)}, A^{(i+2)}, A^{(i+3)}, \ldots$$



$$\Phi^u(A) \leq 1$$
$$\Phi_\ell(A) \leq 1.$$

Step i+1

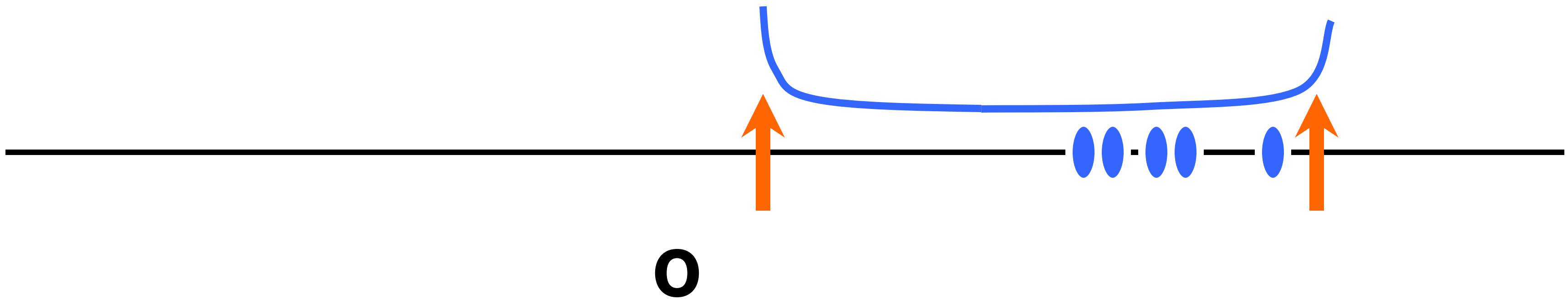
$$A^{(i)}, A^{(i+1)}, A^{(i+2)}, A^{(i+3)}, \dots$$



$$\Phi^u(A) \leq 1$$
$$\Phi_\ell(A) \leq 1.$$

Step i+1

$$A^{(i)}, A^{(i+1)}, A^{(i+2)}, A^{(i+3)}, \ldots$$



$$\Phi^u(A) \leq 1$$

$$\Phi_\ell(A) \leq 1.$$

Step 6n

$$A^{(i)}, A^{(i+1)}, A^{(i+2)}, A^{(i+3)}, \ldots, A^{(6n)}$$

$$\Phi^u(A) \leq 1$$
$$\Phi_\ell(A) \leq 1.$$

Step 6n

$$A^{(i)}, A^{(i+1)}, A^{(i+2)}, A^{(i+3)}, \ldots, A^{(6n)}$$

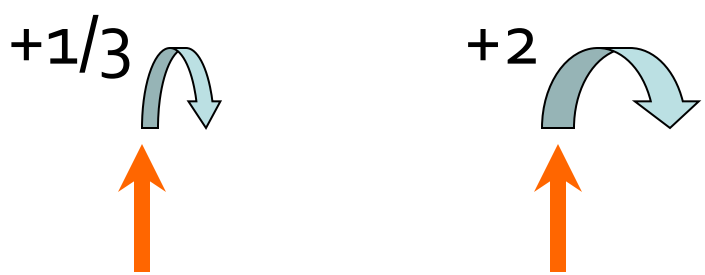
13-approximation with 6n vectors.

□

Goal

Lemma.
can always choose $+s\mathbf{v}\mathbf{v}^T$ so that *both* potentials do not increase.

$$\Phi^u(A) \le 1$$
$$\Phi_\ell(A) \le 1.$$



$$+s\mathbf{v}\mathbf{v}^T$$

The Right Question

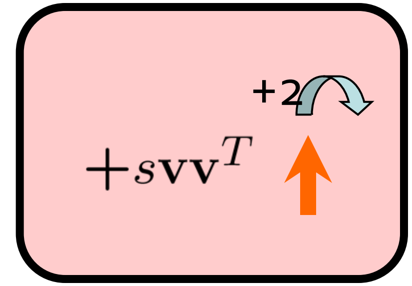
"Which vector should we add?"

The Right Question

~~“Which vector should we add?”~~

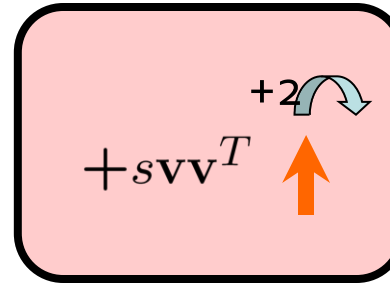
“Given a vector, how much of it can we add?”

Upper Barrier Update



Add $s\mathbf{v}\mathbf{v}^T$ & ***set*** $u' \leftarrow u + 2.$

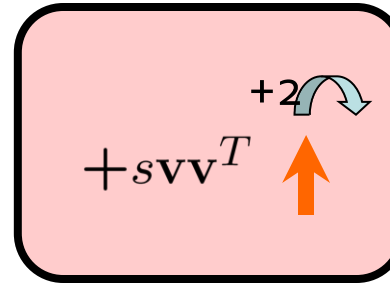
Upper Barrier Update



Add $s\mathbf{v}\mathbf{v}^T$ & ***set*** $u' \leftarrow u + 2.$

$$\Phi^{u'}(A + s\mathbf{v}\mathbf{v}^T)$$
$$= \mathrm{Tr}(u'I - A - s\mathbf{v}\mathbf{v}^T)^{-1}$$

Upper Barrier Update



Add $s\mathbf{v}\mathbf{v}^T$ & ***set*** $u' \leftarrow u + 2.$

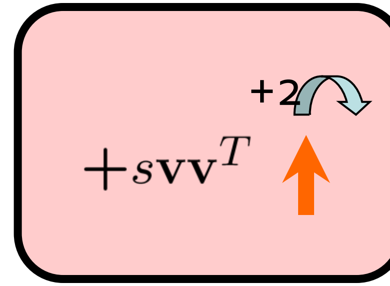
$$\Phi^{u'}(A + s\mathbf{v}\mathbf{v}^T)$$

$$= \mathrm{Tr}(u'I - A - s\mathbf{v}\mathbf{v}^T)^{-1}$$

$$\mathrm{Tr}(A + vv^T)^{-1} = \mathrm{Tr}A^{-1} - \frac{v^T A^{-2} v}{1 + v^T A^{-1} v}$$

Sherman-Morrisson

Upper Barrier Update



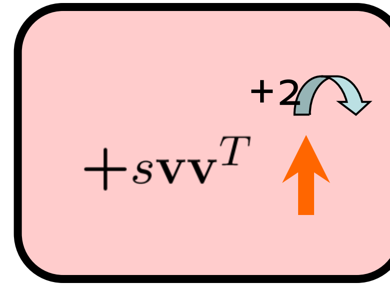
Add $s\mathbf{v}\mathbf{v}^T$ & ***set*** $u' \leftarrow u + 2.$

$$\Phi^{u'}(A + s\mathbf{v}\mathbf{v}^T)$$

$$= \mathrm{Tr}(u'I - A - s\mathbf{v}\mathbf{v}^T)^{-1}$$

$$= \Phi^{u'}(A) + \frac{\mathbf{v}^T(u'I - A)^{-2}\mathbf{v}}{1/s - \mathbf{v}^T(u'I - A)^{-1}\mathbf{v}}$$

Upper Barrier Update



Add $s\mathbf{v}\mathbf{v}^T$ & ***set*** $u' \leftarrow u + 2.$

$$\Phi^{u'}(A + s\mathbf{v}\mathbf{v}^T)$$

$$= \mathrm{Tr}(u'I - A - s\mathbf{v}\mathbf{v}^T)^{-1}$$

$$= \Phi^{u'}(A) + \frac{\mathbf{v}^T(u'I - A)^{-2}\mathbf{v}}{1/s - \mathbf{v}^T(u'I - A)^{-1}\mathbf{v}}$$

want $\leq \Phi^u(A).$

How much of $\mathbf{v}\mathbf{v}^T$ can we add?

Rearranging:

$$\Phi^{u'}(A + s\mathbf{v}\mathbf{v}^T) \leq \Phi^{u}(A)$$

$$\Longleftrightarrow$$

$$\frac{1}{s} \geq \mathbf{v}^T \left(\frac{(u'I - A)^{-2}}{\Phi^{u}(A) - \Phi^{u'}(A)} + (u'I - A)^{-1} \right) \mathbf{v}$$

How much of $\mathbf{v}\mathbf{v}^T$ can we add?

Rearranging:

$$\Phi^{u'}(A + s\mathbf{v}\mathbf{v}^T) \leq \Phi^{u}(A)$$

$$\Longleftrightarrow$$

$$\frac{1}{s} \geq \mathbf{v}^T \left(\frac{(u'I - A)^{-2}}{\Phi^{u}(A) - \Phi^{u'}(A)} + (u'I - A)^{-1} \right) \mathbf{v}$$

$$\boxed{\frac{1}{s} \geq U_A \bullet \mathbf{v}\mathbf{v}^T}$$

The Lower Barrier

Similarly:

$$\Phi_{\ell'}(A + s\mathbf{v}\mathbf{v}^T) \leq \Phi_{\ell}(A)$$

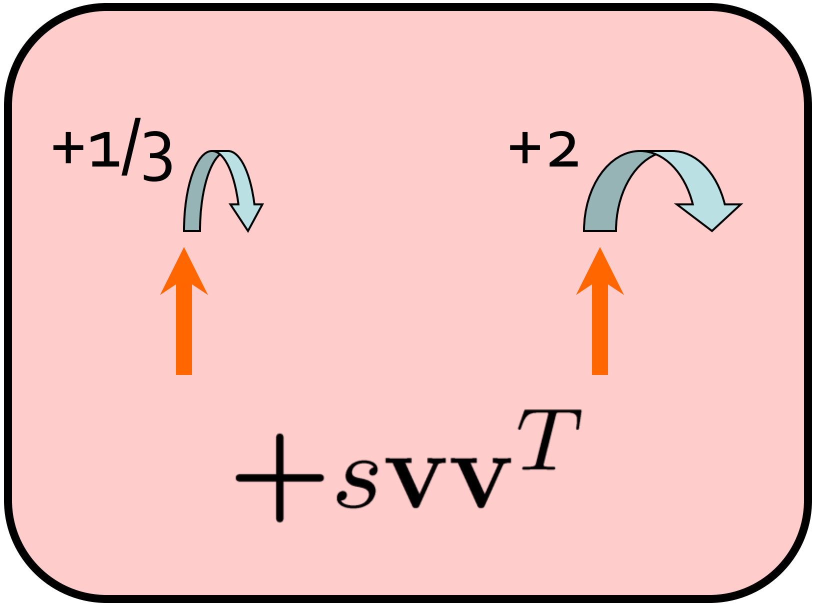
$$\Longleftrightarrow$$

$$\frac{1}{s} \leq \mathbf{v}^T \left(\frac{(A - \ell' I)^{-2}}{\Phi_{\ell'}(A) - \Phi_{\ell}(A)} - (A - \ell' I)^{-1} \right) \mathbf{v}$$

$$\boxed{\frac{1}{s} \leq L_A \bullet \mathbf{v}\mathbf{v}^T}$$

Goal

Show that we can always add some vector while respecting *both* barriers.



Both Barriers

There is always a vector with

$$U_A \bullet \mathbf{v}\mathbf{v}^T \leq L_A \bullet \mathbf{v}\mathbf{v}^T$$

can add

Both Barriers

must add

There is always a vector with

$$U_A \bullet \mathbf{v}\mathbf{v}^T \leq L_A \bullet \mathbf{v}\mathbf{v}^T$$

can add

Both Barriers

must add

There is always a vector with

$$U_A \bullet \mathbf{v}\mathbf{v}^T \leq L_A \bullet \mathbf{v}\mathbf{v}^T$$

Then, can squeeze scaling factor in between:

$$U_A \bullet \mathbf{v}\mathbf{v}^T \leq \frac{1}{s} \leq L_A \bullet \mathbf{v}\mathbf{v}^T$$

Taking Averages

$$\exists \mathbf{v}, U_A \bullet \mathbf{v}\mathbf{v}^T \leq L_A \bullet \mathbf{v}\mathbf{v}^T$$

$$\begin{aligned}\sum_{\mathbf{v} \in \{v_e\}} U_A \bullet \mathbf{v}\mathbf{v}^{\mathbf{T}} &= U_A \bullet \left(\sum_e v_e v_e^T \right) \\ &= U_A \bullet I \\ &= \mathsf{Tr}(U_A).\end{aligned}$$

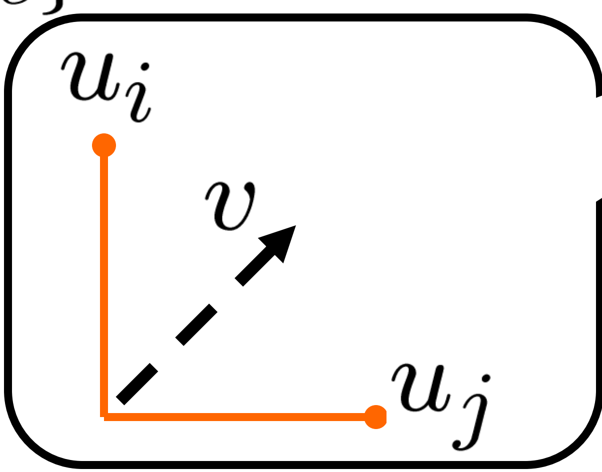
Taking Averages

$$\exists \mathbf{v}, U_A \bullet \mathbf{v}\mathbf{v}^T \leq L_A \bullet \mathbf{v}\mathbf{v}^T$$

$$\sum_{\mathbf{v} \in \{v_e\}} U_A \bullet \mathbf{v}\mathbf{v}^{\mathbf{T}} = U_A \bullet \left(\sum_e v_e v_e^T \right)$$

$$U_A \bullet I$$

$$= \mathrm{Tr}(U_A).$$



Bounding ***Tr(U_A)***

$$\frac{\operatorname{Tr}(u'I - A)^{-2}}{\Phi^{u}(A) - \Phi^{u'}(A)} + \operatorname{Tr}(u'I - A)^{-1}$$

Bounding *Tr(U_A)*

$$\frac{\mathrm{Tr}(u'I - A)^{-2}}{\Phi^{u}(A) - \Phi^{u'}(A)} + \boxed{\Phi^{u'}(A)}$$

Bounding ***Tr(U$_A$)***

$$\frac{\mathrm{Tr}(u'I - A)^{-2}}{\Phi^{u}(A) - \Phi^{u'}(A)} + \boxed{} \leq \Phi^{u}(A)$$

Bounding $Tr(U_A)$

$$\frac{\mathrm{Tr}(u'I - A)^{-2}}{\Phi^{u}(A) - \Phi^{u'}(A)} + \boxed{\leq 1}$$

induction

Bounding ***Tr(U_A)***

$$\frac{-\frac{\partial}{\partial u'}\Phi^{u'}(A)}{\Phi^{u}(A) - \Phi^{u'}(A)} + \leq 1$$

induction

(Recall $\Phi^{u}(A) = \mathrm{Tr}(uI - A)^{-1}$.)

Bounding $Tr(U_A)$

$$\frac{-\frac{\partial}{\partial u'}\Phi^{u'}(A)}{\geq \delta_u\left(-\frac{\partial}{\partial u'}\Phi^{u'}(A)\right)} + \leq 1$$

convexity

induction

$$\text{(Recall } \Phi^u(A) = \mathrm{Tr}(uI - A)^{-1}.\text{)}$$

Bounding *Tr(U_A)*

$$\frac{-\frac{\partial}{\partial u'}\Phi^{u'}(A)}{\geq \delta_u(-\frac{\partial}{\partial u'}\Phi^{u'}(A))} + \leq 1$$

convexity

induction

$$\mathrm{Tr}(U_A) \leq \frac{1}{\delta_u} + 1$$

Taking Averages

$$\exists \mathbf{v}, U_A \bullet \mathbf{v}\mathbf{v}^T \leq L_A \bullet \mathbf{v}\mathbf{v}^T$$

$$\sum_{\mathbf{v} \in \{v_e\}} U_A \bullet \mathbf{v}\mathbf{v}^{\mathbf{T}} \leq \frac{1}{\delta_u} + 1.$$

Taking Averages

$$\exists \mathbf{v}, U_A \bullet \mathbf{v}\mathbf{v}^T \leq L_A \bullet \mathbf{v}\mathbf{v}^T$$

$$\sum_{\mathbf{v} \in \{v_e\}} U_A \bullet \mathbf{v}\mathbf{v}^{\mathbf{T}} \leq \frac{1}{\delta_u} + 1.$$

$$\sum_{\mathbf{v} \in \{v_e\}} L_A \bullet \mathbf{v}\mathbf{v}^{\mathbf{T}} \geq \frac{1}{\delta_\ell} - 1.$$

Taking Averages

$$\exists \mathbf{v}, U_A \bullet \mathbf{v}\mathbf{v}^T \leq L_A \bullet \mathbf{v}\mathbf{v}^T$$

$$\sum_{\mathbf{v} \in \{v_e\}} U_A \bullet \mathbf{v}\mathbf{v}^{\mathbf{T}} \leq \frac{1}{2} + 1.$$

=3/2

$$\sum_{\mathbf{v} \in \{v_e\}} L_A \bullet \mathbf{v}\mathbf{v}^{\mathbf{T}} \geq \frac{1}{\delta_\ell} - 1.$$

Taking Averages

$$\exists \mathbf{v}, U_A \bullet \mathbf{v}\mathbf{v}^T \leq L_A \bullet \mathbf{v}\mathbf{v}^T$$

$$\sum_{\mathbf{v} \in \{v_e\}} U_A \bullet \mathbf{v}\mathbf{v}^{\mathrm{T}} \leq \frac{1}{2} + 1.$$

=3/2

$$\sum_{\mathbf{v} \in \{v_e\}} L_A \bullet \mathbf{v}\mathbf{v}^{\mathrm{T}} \geq \frac{1}{1/3} - 1.$$

2

Taking Averages

$$\exists \mathbf{v}, U_A \bullet \mathbf{v}\mathbf{v}^T \leq L_A \bullet \mathbf{v}\mathbf{v}^T$$ √

$$\sum_{\mathbf{v} \in \{v_e\}} U_A \bullet \mathbf{v}\mathbf{v}^{\mathbf{T}} \leq \frac{1}{2} + 1.$$

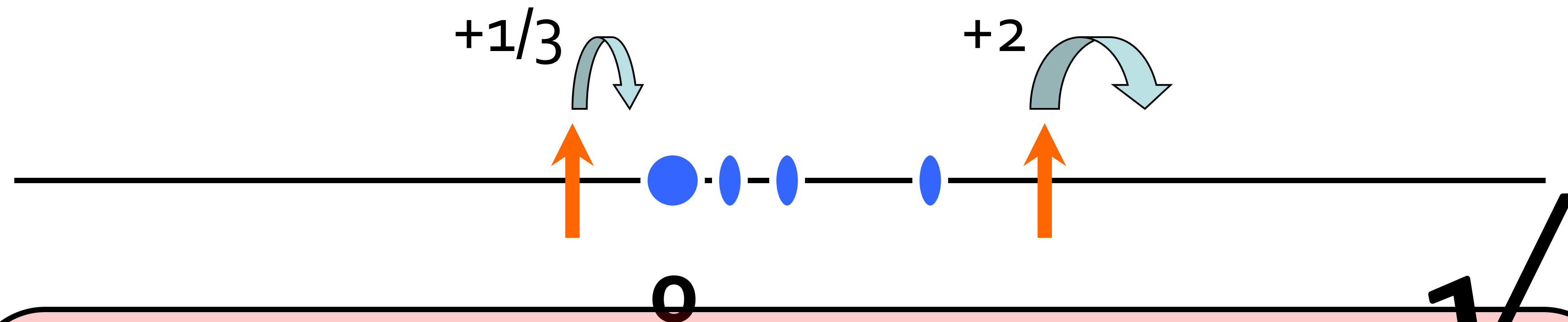
=3/2

$$\sum_{\mathbf{v} \in \{v_e\}} L_A \bullet \mathbf{v}\mathbf{v}^{\mathbf{T}} \geq \frac{1}{1/3} - 1.$$

2

Step i+1

$$A^{(i)}, A^{(i+1)}, A^{(i+2)}$$



Lemma.

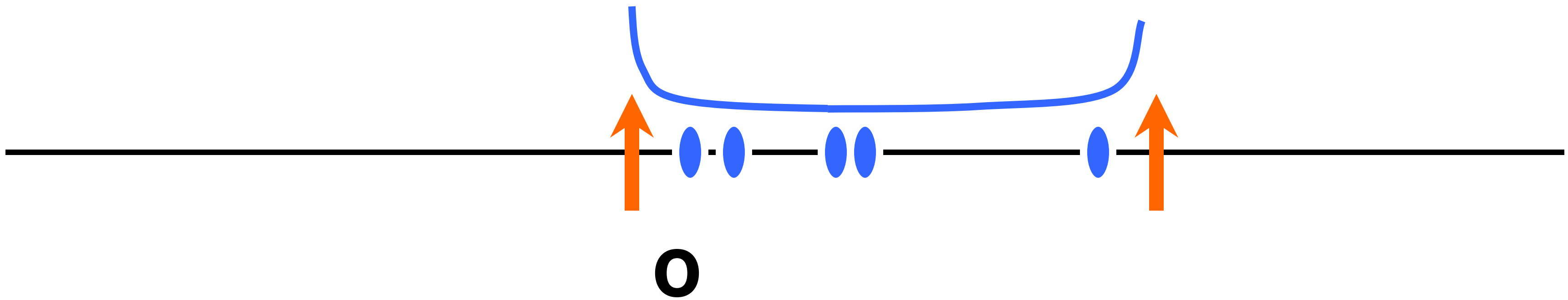
can always choose $+s\mathbf{v}\mathbf{v}^T$ so that potentials do not increase.

$$\Phi^u(A) \leq 1$$

$$\Phi_\ell(A) \leq 1.$$

Step i+1

$$A^{(i)}, A^{(i+1)}, A^{(i+2)}, A^{(i+3)}$$

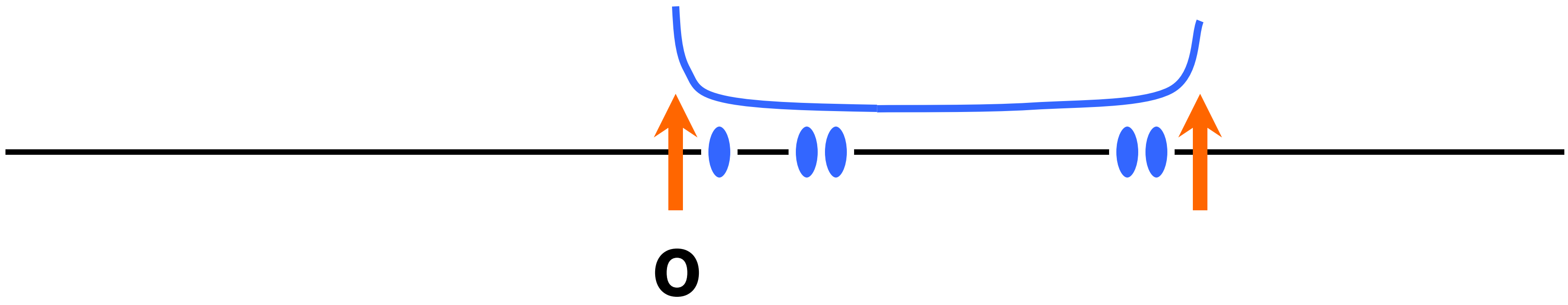


$$\Phi^u(A) \leq 1$$

$$\Phi_\ell(A) \leq 1.$$

Step i+1

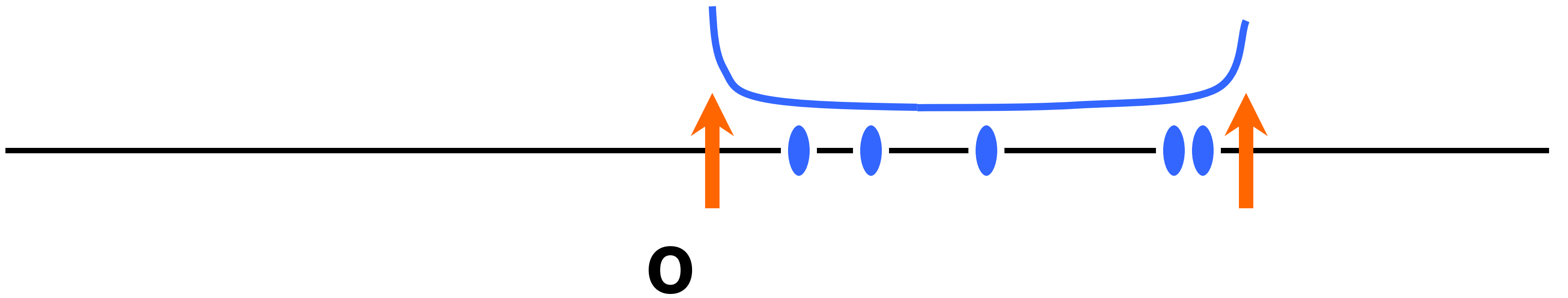
$$A^{(i)}, A^{(i+1)}, A^{(i+2)}, A^{(i+3)}, \ldots$$



$$\Phi^u(A) \leq 1$$
$$\Phi_\ell(A) \leq 1.$$

Step i+1

$$A^{(i)}, A^{(i+1)}, A^{(i+2)}, A^{(i+3)}, \ldots$$

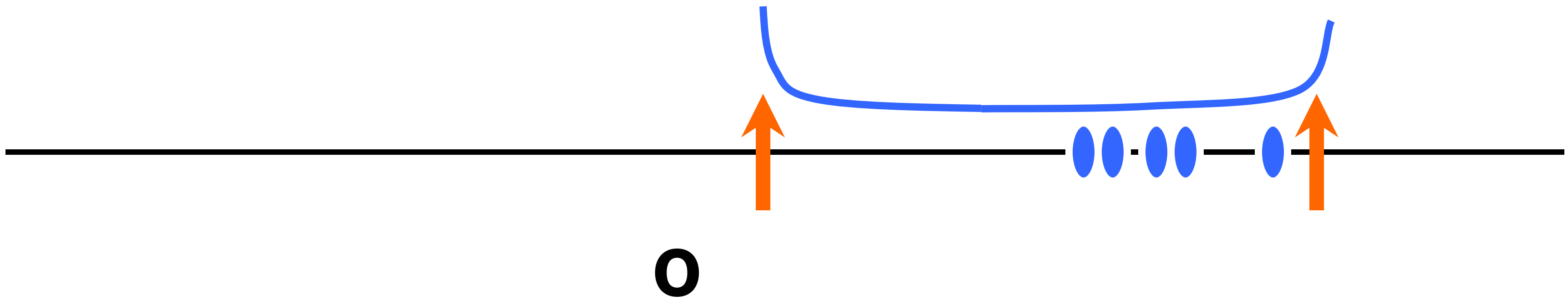


$$\Phi^u(A) \leq 1$$

$$\Phi_\ell(A) \leq 1.$$

Step i+1

$$A^{(i)}, A^{(i+1)}, A^{(i+2)}, A^{(i+3)}, \ldots$$



$$\Phi^u(A) \leq 1$$

$$\Phi_\ell(A) \leq 1.$$

Step 6n

$$A^{(i)}, A^{(i+1)}, A^{(i+2)}, A^{(i+3)}, \ldots, A^{(6n)}$$

$$\Phi^u(A) \leq 1$$
$$\Phi_\ell(A) \leq 1.$$

Step 6n

$$A^{(i)}, A^{(i+1)}, A^{(i+2)}, A^{(i+3)}, \ldots, A^{(6n)}$$

13-approximation with 6n vectors.

□

Main Sparsification Theorem ✓

If

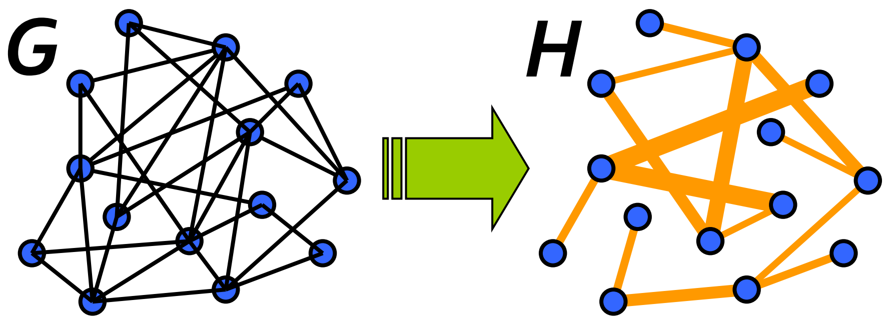
$$\sum_e v_e v_e^T = I_n$$

then there are scalars $s_e \geq 0$ with

$$1 \leq \lambda(\sum_e s_e v_e v_e^T) \leq 13$$

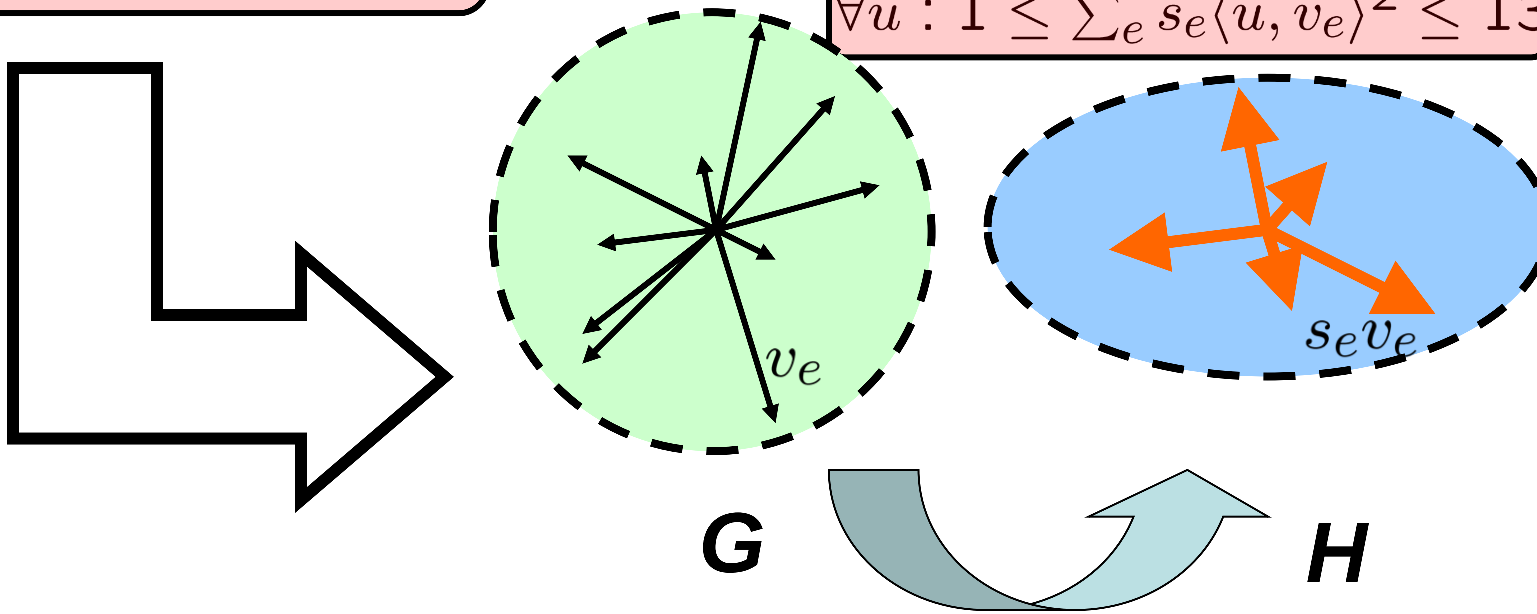
and $|\{s_e \neq 0\}| \leq 6n$.

Sparsification of Graphs



$$1 \leq \frac{x^T L_H x}{x^T L_G x} \leq 13 \quad \forall x \in \mathbb{R}^n$$

$$\forall u : 1 \leq \sum_e s_e \langle u, v_e \rangle^2 \leq 13$$



Twice-Ramanujan

Fixing dn steps and tightening parameters gives

$$\frac{d+1+2\sqrt{d}}{d+1-2\sqrt{d}}.$$

(zeros of Laguerre polynomials).

□

Major Themes

- Electrical model of **interlacing** is useful
- Can use barrier potential to **iteratively** construct matrices with desired spectra
- Analysis of progress is **greedy / local**
- Requires **fractional weights** on vectors

Sparsification of PSD Matrices

Theorem. If

$$V = \sum_i v_i v_i^T$$

then there are scalars $s_i \geq 0$ for which

$$V \preceq \sum_i s_i v_i v_i^T \preceq (1+\epsilon) V$$

and at most n/ϵ^2 are nonzero.

To put this in context…

Given: $V = \sum_{i \leq m} v_i v_i^T$

Spectral Theorem: If rank(V)=n then

$$V = \sum_{i \leq n} \lambda_i u_i u_i^T$$

n terms!

for eigenvectors $\boldsymbol{u_i}$.

To put this in context…

Given: $V = \sum_{i \le m} v_i v_i^T$

Spectral Theorem: If rank(V)=n then

$$V = \sum_{i \le n} \lambda_i u_i u_i^T$$

n terms

***u**$_i$* need not be 'meaningful' directions…

(e.g., ***v**$_i$* = edges of graph)

To put this in context…

Given: $V = \sum_{i \leq m} v_i v_i^T$

Spectral Theorem: If rank(V)=n then

$$V = \sum_{i \leq n} \lambda_i u_i u_i^T$$

n terms

This Theorem. Can find scalars s_i so that:

$$V \sim_\epsilon \sum_i s_i v_i v_i^T$$

n/ϵ^2 terms!

Open Questions

- The Ramanujan bound
- Unweighted sparsifiers for $\mathbf{K_n}$
- A faster algorithm
- Directed graphs? (must be weaker notion)
- The Kadison-Singer Conjecture

Open Questions

- The Ramanujan bound
- Unweighted sparsifiers for $\mathbf{K_n}$
- A faster algorithm
- Directed graphs? (must be weaker notion)
- The Kadison-Singer Conjecture

Sums of Outer Products

$$A = \sum_{i \leq m} v_i v_i^T \qquad v_i \in \mathbf{R}^n, m \gg n$$

Sums of Outer Products

$$A = \sum_{i \leq m} v_i v_i^T \qquad v_i \in \mathbf{R}^n, m \gg n$$

- elementary / "meaningful" directions

Sums of Outer Products

$$A = \sum_{i \leq m} v_i v_i^T \qquad v_i \in \mathbf{R}^n, m \gg n$$

- **Q**: Can we write A as a *sparse* sum?

Sums of Outer Products

$$A = \sum_{i \le m} v_i v_i^T \qquad v_i \in \mathbf{R}^n, m \gg n$$

- **Q**: Can we write A as a *sparse* sum?
- **A**: Yes, Spectral Theorem:

$$A = \sum_{i \le n} \lambda_i u_i u_i^T \text{ for eigvecs } u_i$$

Sums of Outer Products

$$A = \sum_{i \leq m} v_i v_i^T \qquad v_i \in \mathbf{R}^n, m \gg n$$

- **Q**: Can we write A as a *sparse* sum?
- **A**: Yes, Spectral Theorem:

$$A = \sum_{i \leq n} \lambda_i u_i u_i^T \text{ for eigvecs } u_i$$

- **Good**: only ***n*** terms = optimal!
- **Bad**: $\boldsymbol{u_i}$ hard to interpret in terms of $\boldsymbol{v_i}$

Example: Graphs

Undirected graph G(V,E)

Laplacian Matrix

$$L_G = \sum_{ij \in E} (e_i - e_j)(e_i - e_j)^T$$

Example: Graphs

Undirected graph G(V,E)

Laplacian Matrix

$$L_G = \sum_{ij \in E} (e_i - e_j)(e_i - e_j)^T$$

Interested in `sparse approximation': write $\boldsymbol{L_G}$ as sum of a *small number* of edges.

Example: Graphs

Undirected graph G(V,E)

Laplacian Matrix

$$L_G = \sum_{ij \in E} (e_i - e_j)(e_i - e_j)^T$$

Interested in `sparse approximation': write $\boldsymbol{L_G}$ as sum of a *small number* of edges.

Spectral Thm: $\boldsymbol{L_G = \sum_{i=1:n} \lambda_i u_i u_i^T}$

Sparse, but $\boldsymbol{u_i}$ do not correspond to edges… : (

Spectral Sparsification [BSS'09]

Theorem. Given $A = \sum_{i \leq m} v_i v_i^T$

there are nonnegative weights $s_i \geq 0$ s.t.

$$A \sim \tilde{A} = \sum_i s_i v_i v_i^T$$

and at most **1.1n** of the s_i are nonzero.

Spectral S … cation [BSS'09]

• same v_i

Theorem. Given $A = \sum_{i \le m} v_i v_i^T$

there are nonnegative weights $s_i \geq 0$ s.t.

$$A \sim \tilde{A} = \sum_i s_i v_i v_i^T$$

and at most **1.1n** of the s_i are nonzero.

• cf. n terms for $\lambda_i u_i u_i^T$

Back to Graphs

Undirected graph G(V,E)

Laplacian Matrix

$$L_G = \sum_{ij \in E} (e_i - e_j)(e_i - e_j)^T$$

Back to Graphs

Undirected graph G(V,E)

Laplacian Matrix

$$L_G = \sum_{ij \in E}(e_i - e_j$$

• $\leq$ 1.1n **edges** !

Apply **Theorem**:

$$\tilde{L_G} = \sum_{ij \in E} s_i(e_i - e_j)(e_i - e_j)^T$$

$$L \sim \tilde{L}$$