Expander and Derandomization
Many derandomization results are based on the assumption that certain random/hard objects exist.

Some unconditional derandomization can be achieved using explicit constructions of pseduorandom objects.
Synopsis

1. Basic Linear Algebra
2. Random Walk
3. Expander Graph
4. Explicit Construction of Expander Graph
5. Reingold’s Theorem
Basic Linear Algebra
Three Views

All boldface lower case letters denotes column vectors.

Matrix = Linear transformation : $\mathbb{Q}^n \to \mathbb{Q}^m$

1. $f(u + v) = f(u) + f(v)$, $f(cu) = cf(u)$
2. the matrix $M_f$ corresponding to $f$ has $f(e_j)$ as the $j$-th column

Interpretation of $v = Au$

1. Dynamic view: $u$ is transformed to $v$, movement in one basis
2. Static view: $u$ in the column basis is the same as $v$ in the standard basis, movement of basis

Equation, Geometry (row picture), Algebra (column picture)

- Linear equation, hyperplane, linear combination
Suppose $M$ is a matrix, $\mathbf{c}_1, \ldots, \mathbf{c}_n$ are column vectors, and $\mathbf{r}_1, \ldots, \mathbf{r}_n$ are row vectors.

\begin{align*}
M(\mathbf{c}_1, \ldots, \mathbf{c}_n) &= (Mc_1, \ldots, Mc_n) \\ 
\begin{pmatrix}
\mathbf{r}_1 \\
\mathbf{r}_2 \\
\vdots \\
\mathbf{r}_n
\end{pmatrix} 
&= \mathbf{c}_1 \mathbf{r}_1 + \mathbf{c}_2 \mathbf{r}_2 + \ldots + \mathbf{c}_n \mathbf{r}_n
\end{align*}
Inner Product, Projection, Orthogonality

1. Inner product $\mathbf{u}^\dagger \mathbf{v}$ measures the degree of colinearity of $\mathbf{u}$ and $\mathbf{v}$
   - $\frac{\mathbf{u}}{\|\mathbf{u}\|}$ is the normalization of $\mathbf{u}$
   - $\mathbf{u}$ and $\mathbf{v}$ are orthogonal if $\mathbf{u}^\dagger \mathbf{v} = 0$
   - $\frac{\mathbf{u}^\dagger \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|}$ is the projection of $\mathbf{v}$ onto $\mathbf{u}$, where $\|\mathbf{u}\| = \sqrt{\mathbf{u}^\dagger \mathbf{u}}$ is the length of $\mathbf{u}$
   - projection matrix $P = \frac{\mathbf{u} \mathbf{u}^\dagger}{\|\mathbf{u}\| \|\mathbf{u}\|} = \frac{\mathbf{u}}{\|\mathbf{u}\|} \cdot \frac{\mathbf{u}^\dagger}{\|\mathbf{u}\|}$
   - suppose $\mathbf{u}_1, \ldots, \mathbf{u}_m$ are linearly independent. the projection of $\mathbf{v}$ onto the subspace spanned by $\mathbf{u}_1, \ldots, \mathbf{u}_m$ is $P\mathbf{v}$, where the projection matrix $P$ is $A(A^\dagger A)^{-1}A^\dagger$.
     - if $\mathbf{u}_1, \ldots, \mathbf{u}_m$ are orthonormal, $P = \mathbf{u}_1 \mathbf{u}_1^\dagger + \ldots + \mathbf{u}_m \mathbf{u}_m^\dagger = I_m$.

2. Basis, orthonormal basis, orthogonal matrix
3. $Q^{-1} = Q^\dagger$ for every orthogonal matrix $Q$
   - Gram-Schmidt orthogonalization, $A = QR$

Cauchy-Schwartz Inequality. $\cos \theta = \frac{\mathbf{u}^\dagger \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|} \leq 1.$
We look for fixpoints of a linear transformation $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$.

$$A\mathbf{v} = \lambda \mathbf{v}.$$  

If there are $n$ linear independent fixpoints $\mathbf{v}_1, \ldots, \mathbf{v}_n$, then every $\mathbf{v} \in \mathbb{R}^n$ is some linear combination $c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n$. By linearity,

$$A\mathbf{v} = c_1A\mathbf{v}_1 + \ldots + c_nA\mathbf{v}_n = c_1\lambda_1\mathbf{v}_1 + \ldots + c_n\lambda_n\mathbf{v}_n.$$  

If we think of $\mathbf{v}_1, \ldots, \mathbf{v}_n$ as a basis, the effect of the transform $A$ is to stretch the coordinates in the directions of the axes.
Eigenvalue, Eigenvector, Eigenmatrix

If $A - \lambda I$ is singular, an eigenvector $x$ satisfies $x \neq 0$, $Ax = \lambda x$; and $\lambda$ is the eigenvalue.

1. $S = [x_1, \ldots, x_n]$ is the eigenmatrix. By definition $AS = S\Lambda$.
2. If $\lambda_1, \ldots, \lambda_n$ are different, $x_1, \ldots, x_n$ are linearly independent.
3. If $x_1, \ldots, x_n$ are linearly independent, $A = S\Lambda S^{-1}$.

Suppose $c_1x_1 + \ldots + c_nx_n = 0$. Then $c_1\lambda_1x_1 + \ldots + c_n\lambda_nx_n = 0$. It follows that $c_1(\lambda_1 - \lambda_n)x_1 + \ldots + c_{n-1}(\lambda_{n-1} - \lambda_n)x_{n-1} = 0$. By induction we eventually get $c_1(\lambda_1 - \lambda_2)\ldots(\lambda_1 - \lambda_n)x_1 = 0$. Thus $c_1 = 0$. Similarly $c_2 = \ldots = c_n = 0$.

We shall write the spectrum $\lambda_1, \lambda_2, \ldots, \lambda_n$ such that $|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_n|$. $\rho(A) = |\lambda_1|$ is called spectral radius.
Similarity Transformation

Similarity Transformation = Change of Basis

1. \( A \) is similar to \( B \) if \( A = MBM^{-1} \) for some invertible \( M \).
2. \( \mathbf{v} \) is an eigenvector of \( A \) iff \( M^{-1}\mathbf{v} \) is an eigenvector of \( B \).

\[ \]

\( A \) and \( B \) describe the same transformation using different bases.

1. The basis of \( B \) consists of the column vectors of \( M \).
2. A vector \( \mathbf{x} \) in the basis of \( A \) is transformed into the vector \( M^{-1}\mathbf{x} \) in the basis of \( B \), that is \( \mathbf{x} = M(M^{-1}\mathbf{x}) \).
3. \( B \) then transforms \( M^{-1}\mathbf{x} \) into some \( \mathbf{y} \) in the basis of \( B \).
4. In the basis of \( A \) the vector \( A\mathbf{x} \) is \( M\mathbf{y} \).

\[ \]

Fact. Similar matrices have the same eigenvalues.
Diagonalization transformation is a special case of similarity transformation. In diagonalization $Q$ provides an orthogonal basis.

**Question.** Is every matrix similar to a diagonal matrix?

**Schur’s Lemma.** For each matrix $A$ there is a unitary matrix $U$ such that $T = U^{-1}AU$ is triangular. The eigenvalues of $A$ appear in the diagonal of $T$. 
What are the matrices that are similar to diagonal matrices?

A matrix $N$ is normal if $NN^\dagger = N^\dagger N$.

**Theorem.** A matrix $N$ is normal iff $T = U^{-1}NU$ is diagonal iff $N$ has a complete set of orthonormal eigenvectors.

**Proof.**
If $N$ is normal, $T$ is normal. It follows from $T^\dagger = T$ that $T$ is diagonal. If $T$ is diagonal, it is the eigenvalue matrix of $N$, and $NU = UT$ says that the column vectors of $U$ are precisely the eigenvectors.
## Hermitian Matrix and Symmetric Matrix

<table>
<thead>
<tr>
<th></th>
<th>real matrix</th>
<th>complex matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>length</td>
<td>$|x| = \sqrt{\sum_{i \in [n]} x_i^2}$</td>
<td>$|x| = \sqrt{\sum_{i \in [n]}</td>
</tr>
<tr>
<td>conjugate transpose</td>
<td>$A^\dagger$</td>
<td>$A^\dagger$</td>
</tr>
<tr>
<td>inner product</td>
<td>$x^\dagger y = \sum_{i \in [n]} x_i y_i$</td>
<td>$x^\dagger y = \sum_{i \in [n]} \overline{x_i} y_i$</td>
</tr>
<tr>
<td>orthogonality</td>
<td>$x^\dagger y = 0$</td>
<td>$x^\dagger y = 0$</td>
</tr>
<tr>
<td>symmetric/Hermitian</td>
<td>$A^\dagger = A$</td>
<td>$A^\dagger = A$</td>
</tr>
<tr>
<td>diagonalization</td>
<td>$A = Q\Lambda Q^\dagger$</td>
<td>$A = U\Lambda U^\dagger$</td>
</tr>
<tr>
<td>orthogonal/unitary</td>
<td>$Q^\dagger Q = I$</td>
<td>$U^\dagger U = I$</td>
</tr>
</tbody>
</table>

**Fact.** If $A^\dagger = A$, then $x^\dagger A x = (x^\dagger A x)^\dagger$ is real for all complex $x$.

**Fact.** If $A^\dagger = A$, the eigenvalues are real since $v^\dagger A v = \lambda v^\dagger v = \lambda \|v\|^2$.

**Fact.** If $A^\dagger = A$, the eigenvectors of different eigenvalues are orthogonal.

**Fact.** $\|Ux\|^2 = \|x\|^2$ and $\|Qx\|^2 = \|x\|^2$. 
Spectral Theorem

**Theorem.** Every Hermitian matrix $A$ can be diagonalized by a unitary matrix $U$. Every symmetric matrix $A$ can be diagonalized by an orthogonal matrix $Q$.

\[
U^\dagger AU = \Lambda, \\
Q^\dagger AQ = \Lambda.
\]

The eigenvalues are in $\Lambda$; the orthonormal eigenvectors are in $Q$ respectively $U$.

---

**Corollary.** Every Hermitian matrix $A$ has a spectral decomposition.

\[
A = U\Lambda U^\dagger = \sum_{i \in [n]} \lambda_i u_i u_i^\dagger.
\]

Notice that $I = UU^\dagger = \sum_{i \in [n]} u_i u_i^\dagger$. 

---


Positive Definite Matrix

Symmetric matrixes with positive eigenvalues are at the center of many applications.

A symmetric matrix $A$ is positive definite if $x^\top A x > 0$ for all $x \neq 0$.

**Theorem.** Suppose $A$ is symmetric. The following are equivalent.

1. $x^\top A x > 0$ for all $x \neq 0$.
2. $\lambda_i > 0$ for all the eigenvalues $\lambda_i$.
3. $A = R^\top R$ for some matrix $R$ with independent columns.

If we replace $>$ by $\geq$, we get positive semidefinite matrices.
Singular Value Decomposition

Consider an \( m \times n \) matrix \( A \). Both \( AA^\dagger \) and \( A^\dagger A \) are symmetric.

1. \( AA^\dagger \) is positive semidefinite since \( x^\dagger AA^\dagger x = \|A^\dagger x\|^2 \geq 0 \).

2. \( AA^\dagger = U \Sigma' U^\dagger \), where \( U \) consists of the orthonormal eigenvectors \( u_1, \ldots, u_m \) and \( \Sigma' \) is the diagonal matrix made up from the eigenvalues \( \sigma_1^2 \geq \ldots \geq \sigma_r^2 \).

3. \( A^\dagger A = V \Sigma'' V^\dagger \).

4. \( AA^\dagger u_i = \sigma_i^2 u_i \) implies that \( (\sigma_i^2, A^\dagger u_i) \) is an eigenpair for \( A^\dagger A \). So \( v_i = \frac{A^\dagger u_i}{\|A^\dagger u_i\|} \).

5. \( u_i^\dagger AA^\dagger u_i = u_i^\dagger \sigma_i^2 u_i = \sigma_i^2 \). So \( \|A^\dagger u_i\| = \sigma_i \).

6. \( Av_i = A \frac{A^\dagger u_i}{\|A^\dagger u_i\|} = \frac{\sigma_i^2 u_i}{\sigma_i} = \sigma_i u_i \).

Hence \( AV = U \Sigma \), or \( A = U \Sigma V^\dagger \). Notice that \( \Sigma \) an \( m \times n \) matrix.
Singular Value Decomposition

We call

1. \( \sigma_1, \ldots, \sigma_r \) the singular values of \( A \), and
2. \( U \Sigma V^\dagger \) the singular value decomposition, or SVD, of \( A \).

Lemma. If \( A \) is normal, then \( \sigma_i = |\lambda_i| \) for all \( i \in [n] \).

Proof.
Since \( A \) is normal, \( A = U \Lambda U^\dagger \) by diagonalization. Now \( A^\dagger A = AA^\dagger = U \Lambda^2 U^\dagger \). So the spectrum of \( A^\dagger A/AA^\dagger \) is \( \lambda_1^2, \ldots, \lambda_n^2 \). 

\[ \square \]
Rayleigh Quotient

Suppose $A$ is an $n \times n$ Hermitian matrix, $(\lambda_1, v_1), \ldots, (\lambda_n, v_n)$ are the eigenpairs.

The Rayleigh quotient of $A$ and nonzero $x$ is defined as follows:

$$R(A, x) = \frac{x^\dagger A x}{x^\dagger x} = \frac{\sum_{i \in [n]} \lambda_i \|v_i^\dagger x\|^2}{\sum_{i \in [n]} \|v_i^\dagger x\|^2}.$$  \hspace{1cm} (3)

It is clear from (3) that

- if $\lambda_1 \geq \ldots \geq \lambda_n$, then $\lambda_i = \max_{x \perp v_1, \ldots, x \perp v_{i-1}} R(A, x)$, and
- if $|\lambda_1| \geq \ldots \geq |\lambda_n|$, then $|\lambda_i| = \max_{x \perp v_1, \ldots, x \perp v_{i-1}} |R(A, x)|$. 

One can use Rayleigh quotient to derive lower bound for $\lambda_i$. 

Vector Norm

The norm of a vector is a measure of its magnitude/size/length.

A norm on $\mathbb{F}^n$ is a function $\|\cdot\| : \mathbb{F}^n \rightarrow \mathbb{R}^{\geq 0}$ satisfying the following:

1. $\|v\| = 0$ iff $v = \mathbf{0}$.
2. $\|a v\| = |a| \cdot \|v\|$.
3. $\|v + w\| \leq \|v\| + \|w\|$.

A vector space with a norm is called a normed vector space.

1. $L^1$-norm. $\|v\|_1 = |v_1| + \ldots + |v_n|$.
2. $L^2$-norm. $\|v\|_2 = \sqrt{|v_1|^2 + \ldots + |v_n|^2} = \sqrt{v^\dagger v}$.
3. $L^p$-norm. $\|v\|_p = \sqrt[p]{|v_1|^p + \ldots + |v_n|^p}$.
4. $L^\infty$-norm. $\|v\|_\infty = \max\{|v_1|, \ldots, |v_n|\}$. 
Matrix Norm

We define matrix norm in compatible with vector norm. Suppose $\mathbb{F}^n$ is a normed vector space over field $\mathbb{F}$.

An induced matrix norm is a function $\|\cdot\| : \mathbb{F}^{n \times n} \rightarrow \mathbb{R}^+$ satisfying the following properties.

1. $\|A\| = 0$ iff $A = 0$.
2. $\|aA\| = |a| \cdot \|A\|$.
3. $\|A + B\| \leq \|A\| + \|B\|$.
4. $\|AB\| \leq \|A\| \cdot \|B\|$.
**Matrix Norm**

A matrix norm measures the amplifying power of a matrix. Define

\[ \|A\| = \max_{v \neq 0} \frac{\|Av\|}{\|v\|}. \]

It satisfies (1-4). Additionally \( \|Ax\| \leq \|A\| \cdot \|x\| \) for all \( x \).

**Lemma.** \( \rho(A) \leq \|A\|. \)
Spectral Norm

\[ \|A\|_2 \] is called the spectral norm of \( A \).

\[
\frac{1}{\sqrt{n}} \|A\|_1 \leq \|A\|_2 \leq \sqrt{n} \|A\|_1.
\]

**Lemma.** \( \|A\|_2 = \sigma_1 \).

**Corollary.** If \( A \) is a normal matrix, then \( \|A\|_2 = |\lambda_1| \).

Let \( A^\dagger A = V \Sigma V^\dagger \), let \( V = (v_1, \ldots, v_n) \), and let \( x = a_1 v_1 + \ldots + a_n v_n \). Then

\[
\|Ax\|_2^2 = x^\dagger (A^\dagger A x) = x^\dagger (\sum_{i \in [n]} \sigma_i^2 a_i v_i) \leq \sigma_1^2 \|x\|_2^2.
\]

The equality holds when \( x = v_1 \). Therefore \( \|A\|_2 = \sigma_1 \).
MIT Open Course

https://ocw.mit.edu/courses/mathematics/18-06-linear-algebra-spring-2010/video-lectures/
Random Walk
Graphs are the prime objects of study in combinatorics.

The matrix representation of graphs lends itself to an algebraic treatment to these combinatorial objects. It is especially effective in the treatment of regular graph.
Our digraph admit both self-loops and parallel edges. An undirected edge is seen as two directed edges in opposite directions.

In this lecture whenever we say graph, we mean undirected graph.
The reachability matrix $M$ of a digraph $G$ is defined by $M_{i,j} = 1$ if there is an edge from vertex $j$ to vertex $i$; $M_{i,j} = 0$ otherwise.

The random walk matrix $A$ of a $d$-regular digraph $G$ is $\frac{1}{d} M$.

Let $p$ be a probability distribution over the vertices of $G$ and $A$ is the random walk matrix of $G$. Then $A^k p$ is the distribution after $k$-step random walk.
Random Walk Matrix

Consider the following periodic digraph with $dn$ vertices.

- The vertices are arranged in $n$ layers, each consisting of $d$ vertices. There is an edge from every vertex in the $i$-th layer to every vertex in the $j$-th layer, where $j = i + 1 \mod n$.

Does $A^k p$ converge to a stationary state?

Bipartite digraph, $n = 2$. Undirected bipartite graphs are special bipartite digraph.
In spectral graph theory graph properties are characterized by graph spectrums.

Suppose $G$ is a $d$-regular graph and $A$ is the random walk matrix of $G$.

1. $1$ is an eigenvalue of $A$ and its associated eigenvector is the stationary distribution vector $\mathbf{1} = (\frac{1}{n}, \ldots, \frac{1}{n})^\dagger$. In other words $A\mathbf{1} = \mathbf{1}$.

2. All eigenvalues have absolute values $\leq 1$.

3. $G$ is disconnected if and only if $1$ is an eigenvalue of multiplicity $\geq 2$.

4. If $G$ is connected, $G$ is bipartite if and only if $-1$ is an eigenvalue of $A$.

In 2 and 3($\iff)$ and 4($\iff$), consider the entry with the largest absolute value.
Rate of Convergence

For a regular graph $G$ with random walk matrix $A$, we define

$$\lambda_G \overset{\text{def}}{=} \max_{p} \frac{\|Ap - 1\|_2}{\|p - 1\|_2} = \max_{v \perp 1} \frac{\|Av\|_2}{\|v\|_2} = \max_{v \perp 1, \|v\|_2 = 1} \|Av\|_2,$$

where $p$ is over all probability distribution vectors.

The two definitions are equivalent.

1. $(p - 1) \perp 1$ and $Ap - 1 = A(p - 1)$.
2. For each $v \perp 1$, $p = \alpha v + 1$ is a probability distribution for a sufficiently small $\alpha$.

By definition $\|Av\|_2 \leq \lambda_G \|v\|_2$ for all $v$ such that $v \perp 1$. 
**Lemma.** \( \lambda_G = |\lambda_2| \).

Let \( \mathbf{v}_2, \ldots, \mathbf{v}_n \) be the eigenvectors corresponding to \( \lambda_2, \ldots, \lambda_n \).

Given \( \mathbf{x} \perp \mathbf{1} \), let \( \mathbf{x} = c_2 \mathbf{v}_2 + \ldots + c_n \mathbf{v}_n \). Then

\[
\| \mathbf{A} \mathbf{x} \|_2^2 = \| \lambda_2 c_2 \mathbf{v}_2 + \ldots + \lambda_n c_n \mathbf{v}_n \|_2^2
= \lambda_2^2 c_2^2 \| \mathbf{v}_2 \|_2^2 + \ldots + \lambda_n^2 c_n^2 \| \mathbf{v}_n \|_2^2
\leq \lambda_2^2 (c_2^2 \| \mathbf{v}_2 \|_2^2 + \ldots + c_n^2 \| \mathbf{v}_n \|_2^2)
= \lambda_2^2 \| \mathbf{x} \|_2^2.
\]

So \( \lambda_G^2 \leq \lambda_2^2 \). The equality holds since \( \| \mathbf{A} \mathbf{v}_2 \|_2^2 = \lambda_2^2 \| \mathbf{v}_2 \|_2^2 \).
The spectral gap $\gamma_G$ of a graph $G$ is defined by

$$\gamma_G = 1 - \lambda_G.$$ 

A graph $G$ has spectral expansion $\gamma$, where $\gamma \in (0, 1)$, if $\gamma_G \geq \gamma$.

In an expander the spectral expansion provides a bound on the expansion ratio.
Lemma. Let $G$ be an $n$-vertex regular graph and $p$ a probability distribution over the vertices of $G$. Then

$$\|A^\ell p - 1\|_2 \leq \lambda_G^\ell \|p - 1\|_2 < \lambda_G^\ell.$$ 

The first inequality holds because

$$\frac{\|A^\ell p - 1\|_2}{\|p - 1\|_2} = \frac{\|A^\ell p - 1\|_2}{\|A^{\ell-1} p - 1\|_2} \cdots \frac{\|Ap - 1\|_2}{\|p - 1\|_2} \leq \lambda_G^\ell.$$ 

The second inequality holds because

$$\|p - 1\|_2^2 = \|p\|_2^2 + \|1\|_2^2 - 2\langle p, 1 \rangle \leq 1 + \frac{1}{n} - 2\frac{1}{n} < 1.$$ 

In terms of random walk, $\lambda_G$ bounds the speed of mixing time. [if $G$ is bipartite, $\lambda_G = 1$.]
Lemma. If $G$ is an $n$-vertex $d$-regular graph with self-loops at each vertex, $\gamma_G \geq \frac{1}{12n^2}$.

Let $u$ be the unit vector such that $u \perp 1$ and $\lambda_G = \|Au\|_2$, and let $v = Au$.

- If we can prove $1 - \|v\|_2^2 \geq \frac{1}{6n^2}$, we will get $\lambda_G = \|v\|_2 \leq 1 - \frac{1}{12n^2}$, hence the lemma.
- It’s easy to show $1 - \|v\|_2^2 = \|u\|_2^2 - \|v\|_2^2 = \|u\|_2^2 - 2\langle Au, v \rangle + \|v\|_2^2 = \sum_{i,j} A_{i,j}(u_i - v_j)^2$.

Now $u_i - u_j \geq \frac{1}{\sqrt{n}}$ for some $i,j \in [n]$. Let $i \to i_1 \to \ldots \to i_k \to j$ be minimal from $i$ to $j$. Then

$$1/\sqrt{n} \leq u_i - u_j \leq |u_i - v_i| + |v_i - u_{i_1}| + |u_{i_1} - v_{i_1}| + \ldots + |v_{i_k} - u_j| \tag{4}$$

$$\leq \sqrt{(u_i - v_i)^2 + (v_i - u_{i_1})^2 + \ldots + (v_{i_k} - u_j)^2} \cdot \sqrt{2D}, \tag{5}$$

where $D$ is the diameter of $G$. Notice that there are $k$ edges and $k$ self-loops in (4). Thus

$$\sum_{i,j} A_{i,j}(u_i - v_j)^2 \geq 1/(dn(2D + 1)) \geq 1/(6n^2) \tag{6}$$

by (5) and $A_{h,h}, A_{h,h+1} \geq 1/d$ and $D \leq 3n/(d + 1)$. [see next slide.]
If two nodes, say $u$ and $v$, in a shortest path between two nodes are of distance 3, then the neighbors of $u$ plus $u$ are disjoint from the neighbors of $v$ plus $v$. It follows that

$$(d + 1) \cdot \frac{D}{3} \leq n.$$
**Corollary.** Let $G$ be a $d$-degree $n$-vertex graph with self-loop on every vertex. Let $s, t$ be connected. Let $\ell > 24n^2 \log n$ and let $X_\ell$ denote the vertex distribution after $\ell$ step random walk from $s$. Then $\Pr[X_\ell = t] > \frac{1}{2n}$.

Graphs with self-loops are not bipartite. According to the Lemmas,

$$\|A_\ell e_s - 1\|_2 < \left(1 - \frac{1}{12n^2}\right)^{24n^2 \log n} < \frac{1}{n^2}.$$  

It follows that $(A_\ell e_s)(i) - \frac{1}{n} > -\frac{1}{n^2}$.

If the walk is repeated for $2n^2$ times, the error probability is reduced to below $\frac{1}{2n}$. 

---

*Computational Complexity, by Fu Yuxi*
Theorem. $\text{UPATH (Undirected Connectivity)}$ is in $\text{RL}$.

An undirected graph can be turned into a non-bipartite regular graph by introducing enough self-loops.
Can the random algorithm for \textsc{UPath} be derandomized? Recall that

\[ L \subseteq RL \subseteq NL. \]
Expander Graph
Expander graphs, defined by Pinsker in 1973, are sparse and well connected. They behave approximately like complete graphs.

▶ Sparsity should be understood in an asymptotic sense.

Well-connectedness can be characterized in a number of manners.

1. Algebraically, expanders are graphs whose second largest eigenvalue is bounded away from 1 by a constant.

2. Combinatorially, expanders are highly connected. Every set of vertices of an expander has a large boundary geometrically.

3. Probabilistically, expanders are graphs in which a random walk converges to the stationary distribution quickly.
Algebraic Property

Intuitively the faster random walk converges, the better the graph is connected. According to Lemma, the smaller $\lambda_G$ is, the faster random walk converges to 1.

Suppose $d \in \mathbb{N}$ and $\lambda \in (0, 1)$ are constants.

A $d$-regular graph $G$ with $n$ vertices is an $(n, d, \lambda)$-graph if $\lambda_G \leq \lambda$.

- It follows from a result on page 28 that a $(n, d, \lambda)$-graph is connected.

$\{G_n\}_{n \in \mathbb{N}}$ is a $(d, \lambda)$-expander graph family if $G_n$ is an $(n, d, \lambda)$-graph for all $n \in \mathbb{N}$. 
Probabilistic Property

In an expander random walk converges to the uniform distribution in logarithmic steps.

$$\|A^{\log \frac{1}{\lambda}(n)} p - 1\|_2 < \lambda^{\log \frac{1}{\lambda}(n)} = \frac{1}{n}. \quad (6)$$

In other words, the mixing time of an expander is logarithmic.

It follows from the inequality in (6) that for every $i \in [n]$,

$$\left( A^{\log \frac{1}{\lambda}(n)} p \right)(i) > 0.$$ 

**Fact.** The diameter of an $n$-vertex expander graph is $\Theta(\log n)$. 

---

*Computational Complexity, by Fu Yuxi*  Expander and Derandomization 42 / 91
Combinatorial Property

Suppose $G = (V, E)$ is an $n$-vertex $d$-regular graph.

- Let $\overline{S}$ stand for $V \setminus S$ for $S \subseteq V$.
- Let $E(S, T)$ be the set of edges $i \rightarrow j$ with $i \in S$ and $j \in T$.
- Let $\partial S = E(S, \overline{S})$ for $|S| \leq \frac{n}{2}$.

The expansion constant $h_G$ of $G$ is defined as follows:

$$h_G = \min_{|S| \leq \frac{n}{2}} \frac{|\partial S|}{|S|}.$$

Constant $\rho > 0$. An $n$-vertex $d$-regular graph $G$ is an $(n, d, \rho)$-edge expander if $\frac{h_G}{d} \geq \rho$.

- There are $d|S|$ edges emitting from the nodes of $S$. 

Existence of Expander

**Theorem.** Let $\epsilon > 0$. There exists $d = d(\epsilon)$ and $N \in \mathbb{N}$ such that for every $n > N$ there exists an $(n, d, \frac{1}{2} - \epsilon)$ edge expander.
Theorem. Let $G = (V, E)$ be a finite, connected, $d$-regular graph. Then

$$\frac{\gamma_G}{2} \leq \frac{h_G}{d} \leq \sqrt{2\gamma_G}.$$
Let $S$ be such that $|S| \leq \frac{n}{2}$ and $\frac{\partial(S)}{|S|} = h_G$. Define $x \perp 1$ by $x_i = \begin{cases} |\bar{S}|, & i \in S, \\ -|S|, & i \in \bar{S}. \end{cases}$

\[
\|x\|^2_2 = n|S||\bar{S}|,
\]
\[
x^\dagger A x = (|\bar{S}|1_S - |S|1_{\bar{S}})^\dagger A (|\bar{S}|1_S - |S|1_{\bar{S}})
\]
\[
= \frac{1}{d} \left( |\bar{S}|^2 |E(S, S)| + |S|^2 |E(\bar{S}, \bar{S})| - 2|S||\bar{S}||E(S, \bar{S})| \right)
\]
\[
= \frac{1}{d} \left( dn|S||\bar{S}| - n^2|E(S, \bar{S})| \right),
\]

where $= is due to $d|S| = |E(S, \bar{S})| + |E(S, S)|$ and $d|\bar{S}| = |E(\bar{S}, S)| + |E(\bar{S}, \bar{S})|.$

The Rayleigh quotient $R(A, x)$ provides a lower bound for $\lambda_G.$

\[
\lambda_G \geq \frac{x^\dagger A x}{\|x\|_2^2} = \frac{1}{d} \frac{dn|S||\bar{S}| - n^2|E(S, \bar{S})|}{n|S||\bar{S}|} = 1 - \frac{1}{d} \frac{n}{|\bar{S}|} \frac{|\partial(S)|}{|S|} \geq 1 - \frac{2h_G}{d}.
\]
Let \( \mathbf{u} \perp \mathbf{1} \) be such that \( A\mathbf{u} = \lambda_2 \mathbf{u} \). Write \( \mathbf{u} = \mathbf{v} + \mathbf{w} \), where \( \mathbf{v} \) respectively \( \mathbf{w} \) is defined from \( \mathbf{u} \) by replacing the negative respectively positive components by 0.

Wlog, assume that the number of positive components of \( \mathbf{v} \) is \( \leq \frac{n}{2} \).

Wlog, let the coordinates of \( \mathbf{v} \) satisfy \( v_1 \geq v_2 \geq \ldots \geq v_n \). Then

\[
\sum_{i,j} A_{i,j} |v_i^2 - v_j^2| = 2 \sum_{i<j} \sum_{k=i}^{j-1} (v_k^2 - v_{k+1}^2) = 2 \sum_{i=1}^{n/2} \sum_{j=i+1}^{n/2} \sum_{k=i}^{j-1} (v_k^2 - v_{k+1}^2)
\]

\[
= 2 \sum_{k=1}^{n/2} |\partial[k]|(v_k^2 - v_{k+1}^2) \geq 2 \sum_{k=1}^{n/2} h_G k (v_k^2 - v_{k+1}^2) = \frac{2h_G}{d} \|\mathbf{v}\|^2.
\]

The equality \( = \) is valid because \( v_k = 0 \) for all \( k > n/2 \).
\[
\frac{h_G}{d} \leq \sqrt{2\gamma_G}
\]

\[
\langle Av, v \rangle \geq \langle Av, v \rangle + \langle Aw, v \rangle = \lambda_2 \|v\|^2_2 \text{ because } Au = \lambda_2 u, \langle w, v \rangle = 0 \text{ and } \langle Aw, v \rangle \leq 0.
\]

\[
1 - \lambda_G \geq 1 - \frac{\langle Av, v \rangle}{\|v\|^2_2} = \frac{\|v\|^2_2 - \langle Av, v \rangle}{\|v\|^2_2} = \frac{\sum_{i,j} A_{i,j}(v_i - v_j)^2}{2\|v\|^2_2}. \tag{8}
\]

Using Cauchy-Schwartz Inequality,

\[
\sum_{i,j} A_{i,j}(v_i - v_j)^2 \cdot \sum_{i,j} A_{i,j}(v_i + v_j)^2 \geq \left( \sum_{i,j} A_{i,j} |v_i^2 - v_j^2| \right)^2. \tag{9}
\]

Now \( \langle Av, v \rangle \leq \lambda_1 \|v\|^2_2 = \|v\|^2_2 \). Therefore

\[
2\|v\|^2_2 \cdot \sum_{i,j} A_{i,j}(v_i + v_j)^2 \leq 2\|v\|^2_2 \cdot (2\|v\|^2_2 + 2\langle Av, v \rangle) \leq 8\|v\|^4_2. \tag{10}
\]

(7)+(8)+(9)+(10) implies \( \sqrt{8(1 - \lambda_G)} \geq \frac{2h_G}{d} \).
Combinatorial definition and algebraic definition are equivalent.

1. The inequality $\frac{1-\lambda_G}{2} \leq \frac{h_G}{d}$ implies that if $G$ is an $(n, d, \lambda)$-expander graph, then it is an $(n, d, \frac{1-\lambda}{2})$ edge expander.

2. The inequality $\frac{h_G}{d} \leq \sqrt{2(1-\lambda_G)}$ implies that if $G$ is an $(n, d, \rho)$ edge expander, then it is an $(n, d, 1-\frac{\rho^2}{2})$-expander graph.
Convergence in Entropy

Rényi 2-Entropy:

\[ H_2(p) = \log \left( \frac{1}{\|p\|_2^2} \right). \]

**Fact.** If \( A \) is the random walk matrix of an \((n, d, \lambda)\)-expander, then \( H_2(Ap) \geq H_2(p) \). The equality holds if and only if \( p \) is uniform.

**Proof.**

Let \( p = \mathbf{1} + w \). Then \( w \perp \mathbf{1} \) and \( \langle Aw, \mathbf{1} \rangle = w^\dagger A^\dagger \mathbf{1} = w^\dagger A = w^\dagger \mathbf{1} = 0 \). Therefore

\[
\|Ap\|_2^2 = \|\mathbf{1}\|_2^2 + \|Aw\|_2^2 \leq \|\mathbf{1}\|_2^2 + \lambda \|w\|_2^2 \leq \|\mathbf{1}\|_2^2 + \|w\|_2^2 = \|p\|_2^2.
\]

The equality holds when \( p = \mathbf{1} \).

Random walks increase randomness.
The smaller the spectral gap, or the larger the spectral expansion, the more expander graphs behave like random graphs. This is what the next lemma says.
Lemma. Let $G = (V, E)$ be an $(n, d, \lambda)$-expander graph. Let $S, T \subseteq V$. Then

$$\left| |E(S, T)| - \frac{d}{n} |S||T| \right| \leq \lambda d \sqrt{|S||T|}. \quad (11)$$

Notice that (11) implies

$$\left| \frac{|E(S, T)|}{dn} - \frac{|S|}{n} \cdot \frac{|T|}{n} \right| \leq \lambda. \quad (12)$$

The edge density \approx the product of the vertex densities. This property is called mixing.

Proof of Expander Mixing Lemma

Let \([v_1, \ldots, v_n]\) be the eigenmatrix of \(G\) and set \(v_1 = (\frac{1}{\sqrt{n}}, \ldots, \frac{1}{\sqrt{n}})^\dagger\).

Let \(1_S = \sum_i \alpha_i v_i\) and \(1_T = \sum_j \beta_j v_j\) be the characteristic vectors of \(S, T\) respectively.

\[
|E(S, T)| = (1_S)^\dagger (dA) 1_T = \left( \sum_i \alpha_i v_i \right)^\dagger (dA) \left( \sum_j \beta_j v_j \right) = \sum_i d\lambda_i \alpha_i \beta_i.
\]

Since \(\alpha_1 = (1_S)^\dagger v_1 = \frac{|S|}{\sqrt{n}}\) and \(\beta_1 = (1_T)^\dagger v_1 = \frac{|T|}{\sqrt{n}}\), by Cauchy-Schwartz Inequality,

\[
\left| |E(S, T)| - \frac{d}{n} |S||T| \right| = \sum_{i=2}^n d\lambda_i \alpha_i \beta_i \leq d\lambda \sum_{i=2}^n \alpha_i \beta_i \leq d\lambda \|\alpha\|_2 \|\beta\|_2.
\]

Finally observe that \(\|\alpha\|_2 \|\beta\|_2 = \|1_S\|_2 \|1_T\|_2 = \sqrt{|S||T|}\).
Suppose $A(x, r)$ is a random algorithm with error probability 1/3. The algorithm uses $r(n)$ random bits on input $x$ with $|x| = n$.

1. Reduce the error probability exponentially by repeating the algorithm $t(n)$ times.
2. Altogether $r(n)t(n)$ random bits are used.

The goal is to achieve the same error reduction rate using far fewer random bits, in fact $r(n) + O(t(n))$ random bits.

The key observation is that a $t$-step random walk in an expander graph looks like $t$ vertices sampled uniformly and independently.

- Confer the inequality (12).
$K_n$ is perfect from the viewpoint of random walk.

- No matter what distribution it starts with, random walk reaches the uniform distribution in one step.

Let $J_n = [1, \ldots, 1]$ be the random walk matrix of $K_n$ with self-loop.
Lemma. Suppose $G$ is an $(n, d, \lambda)$-expander and $A$ is its random walk matrix. Then $A = \gamma J_n + \lambda E$ for some $E$ such that $\|E\| \leq 1$.

We may think of a random walk on an expander as a convex combination of two random walks of different type:

- with probability $\gamma$ it walks randomly on a complete graph, and
- with probability $\lambda$ it walks randomly according to an error matrix that does not amplify the distance to the uniform distribution.
Decomposition for Random Walk on Expander

We need to prove that \( \|Ev\|_2 \leq \|v\|_2 \) for all \( v \), where \( E \) is defined by

\[
E = \frac{1}{\lambda} (A - (1 - \lambda)J_n).
\]

The following proof methodology should now be familiar.

- Let \( \alpha = \sum_{i\in[n]} v_i \). Then \( v = \alpha 1 + w \) with \( w \perp 1 \).
- \( A1 = 1 \) and \( J_n 1 = 1 \). Consequently \( E(\alpha 1) = \alpha 1 \).
- \( J_n w = 0 \) and \( Aw \perp \alpha 1 \). Hence \( Ew = \frac{1}{\lambda} Aw \).
- \( \|Aw\|_2 \leq \lambda \|w\|_2 \).

Thus \( \|Ev\|_2^2 = \|\alpha 1 + \frac{1}{\lambda} Aw\|_2^2 = \|\alpha 1\|_2^2 + \frac{1}{\lambda} \|Aw\|_2^2 \leq \|\alpha 1\|_2^2 + \|w\|_2^2 = \|v\|_2^2 \). Done.
Expander Random Walk Theorem

**Theorem.** Let $G$ be an $(n, d, \lambda)$ expander graph, and let $B \subseteq [n]$ satisfy $|B| \leq \beta n$ for some $\beta \in (0, 1)$. Let $X_1$ be a random variable denoting the uniform distribution on $[n]$ and let $X_k$ be a random variable denoting a $k - 1$ step random walk from $X_1$. Then

$$\Pr \left[ \bigwedge_{i \in [k]} X_i \in B \right] \leq \left( \gamma \sqrt{\beta} + \lambda \right)^{k-1}.$$
Expander Random Walk Theorem

Let $B_i$ stand for $X_i \in B$. We need to bound the following.

$$
\Pr \left[ \bigwedge_{i \in [k]} X_i \in B \right] = \Pr[B_1 \ldots B_k] = \Pr[B_1] \cdot \Pr[B_2|B_1] \ldots \Pr[B_k|B_1 \ldots B_{k-1}].
$$

(13)

By seeing $B$ as a diagonal matrix, we define the distribution vector $p_i$ by

$$
p_i = \frac{BA}{\Pr[B_i|B_1 \ldots B_{i-1}]} \cdot \ldots \cdot \frac{BA}{\Pr[B_2|B_1]} \cdot \frac{B1}{\Pr[B_1]},
$$

where $\frac{BA}{\Pr[B_2|B_1]} \cdot \frac{B1}{\Pr[B_1]}$ for example is the normalization of $BA \cdot \frac{B1}{\Pr[B_1]}$. So the probability in (13) is bounded by $\|(BA)^{k-1}B1\|_1$. We will prove

$$
\|(BA)^{k-1}B1\|_2 \leq \frac{1}{\sqrt{n}} \left((1 - \lambda)\sqrt{\beta} + \lambda\right)^{k-1}.
$$
Expander Random Walk Theorem

Using Lemma,

\[ \|BA\| = \|B((1 - \lambda)J_n + \lambda E)\| \leq (1 - \lambda)\|BJ_n\| + \lambda\|BE\| = (1 - \lambda)\sqrt{\beta} + \lambda\|BE\| \leq (1 - \lambda)\sqrt{\beta} + \lambda. \]

Therefore

\[ \|(BA)^{k-1}B1\|_2 \leq \|BA\|_2^{k-1}\cdot\|B1\|_2 \leq \left(\frac{\sqrt{\beta}}{\sqrt{n}}\right)((1 - \lambda)\sqrt{\beta} + \lambda)^{k-1} \leq \frac{1}{\sqrt{n}} \left((1 - \lambda)\sqrt{\beta} + \lambda\right)^{k-1}. \]

Suppose \(\|v\|_2 = 1\) and \(\alpha = \sum_{i \in [n]} v_i\). Then \(v = \alpha 1 + w\) and \(w \perp 1\). Now

- \(\|BJ_n v\|_2 = \|BJ_n \alpha 1\|_2 = \alpha \|B1\|_2 \leq \sqrt{n}\|B1\|_2 = \sqrt{n}\cdot\frac{\sqrt{\beta}}{\sqrt{n}} = \sqrt{\beta}\), and consequently
- \(\|BJ_n\| = \max\{\|BJ_n v\|_2 : \|v\|_2 = 1\} = \sqrt{\beta}\).

The equality holds when \(v = (\frac{1}{\sqrt{n}}, \ldots, \frac{1}{\sqrt{n}})^\dagger\).
Suppose $A(x, r)$ is a random algorithm with error probability $\beta$.

Let $B$ be the set of $r$'s for which $A$ errs on $x$.

Choose an explicit $(2^{\lceil r(|x|) \rceil}, d, \lambda)$-graph $G = (V, E)$ with $V = \{0, 1\}^{\lceil r(|x|) \rceil}$.

Algorithm $A'$.

1. Pick $v_0 \in_R V$.
2. Generate a random walk $v_0, \ldots, v_t$.
3. Output $\bigvee_{i=0}^{t} A(x, v_i)$.

By the Theorem, the error probability of $A'$ is no more than $(\gamma \sqrt{\beta} + \lambda)^{t-1}$. 
Error Reduction for **BPP**

**Algorithm $A''$.**

1. Pick $v_0 \in \mathbb{R}V$.
2. Generate a random walk $v_0, \ldots, v_t$.
3. Output $\text{Maj}\{A(x, v_i)\}_{i \in [t]}$.

Fix a set of indices $K \subseteq \{0, 1, \ldots, t\}$ such that $|K| \geq \frac{t+1}{2}$.

\[
\Pr[\forall i \in K. v_i \in B] \leq \left(\gamma \sqrt{\beta} + \lambda\right)^{|K|-1} \leq \left(\gamma \sqrt{\beta} + \lambda\right)^{\frac{t-1}{2}} \leq \left(\frac{1}{4}\right)^{t-1},
\]

assuming $\gamma \sqrt{\beta} + \lambda \leq 1/16$. Applying union bound on the choices of $K$,

\[
\Pr[A'' \text{ fails}] \leq 2^t \left(\frac{1}{4}\right)^{t-1} = O(2^{-t}).
\]
Explicit Construction of Expander Graph
Explicit Construction

If random strings are of log size, explicit expander family is good enough.

- An expander family $\{G_n\}_{n \in \mathbb{N}}$ is explicit if there is a P-time algorithm that outputs the random walk matrix of $G_n$ whenever the input is $1^n$. [$\text{poly}(n)$]

If random strings are of polynomial size, strongly explicit expander family is necessary.

- An expander family $\{G_n\}_{n \in \mathbb{N}}$ is strongly explicit if there is a P-time algorithm that on input $\langle n, v, i \rangle$ outputs the index of the $i$-th neighbor of $v$ in $G_n$. [$\text{polylog}(n)$]
We will look at several graph product operations. We then show how to use these operations to construct explicit expander graphs.

Path Product

Suppose $G, G'$ are $n$-vertex graphs with degree $d$ respectively $d'$. Let $A, A'$ be their random walk matrices.

The path product $G'G$ is defined by the random walk matrix $A'A$.

- $G'G$ is $n$-vertex $dd'$-degree.

Lemma. $\lambda_{G'G} \leq \lambda_{G'}\lambda_G$.

Proof. $\lambda_{G'G} = \max_{\mathbf{v} \perp 1} \frac{\|A'A\mathbf{v}\|_2}{\|\mathbf{v}\|_2} = \max_{\mathbf{v} \perp 1} \|A'\mathbf{v}\|_2 \cdot \|\mathbf{v}\|_2 \leq \max_{\mathbf{v} \perp 1} \|A'\mathbf{v}\|_2 \cdot \max_{\mathbf{v} \perp 1} \|A\mathbf{v}\|_2 \leq \lambda_{G'}\lambda_G$ using the fact that $A\mathbf{v} \perp 1$ whenever $\mathbf{v} \perp 1$.

Lemma. $\lambda_{G^k} = (\lambda_G)^k$.

Proof. $(\lambda_G)^k$ is the second largest eigenvalue of $G^k$. 

Computational Complexity, by Fu Yuxi

Expander and Derandomization
Tensor Product

Suppose $G$ is an $n$-vertex $d$-degree graph and $G'$ is an $n'$-vertex $d'$-degree graph. The random walk matrix of the tensor product $G \otimes G'$ is $nn'$-vertex $dd'$-degree.

$$A \otimes A' = \begin{pmatrix} a_{11}A' & a_{12}A' & \cdots & a_{1n}A' \\ a_{21}A' & a_{22}A' & \cdots & a_{2n}A' \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}A' & a_{n2}A' & \cdots & a_{nn}A' \end{pmatrix}.$$  

$(u, u') \rightarrow (v, v')$ in $G \otimes G'$ iff $u \rightarrow v$ in $G$ and $u' \rightarrow v'$ in $G'$. 
Lemma. $\lambda_{G \otimes G'} = \max\{\lambda_G, \lambda_{G'}\}$.

If $(\lambda, \mathbf{v})$ is an eigenpair of $A$ and $(\lambda', \mathbf{v}')$ is an eigenpair of $A'$, then $(\lambda \lambda', \mathbf{v} \otimes \mathbf{v}')$ is an eigenpair of $A \otimes A'$. 
Let $A$ be the random walk matrix of an $n$-vertex regular graph $G$ of degree $D$. The rotation matrix $\hat{A}$ is an $(nD) \times (nD)$ adjacent matrix such that $\hat{A}_{(v,j),(u,i)} = 1$ if $v$ is the $i$-th neighbor of $u$, and $u$ is the $j$-th neighbor of $v$. 
The zig-zag product $G \boxtimes H$ is the $d^2$-degree graph defined by $(I_n \otimes B) \hat{A} (I_n \otimes B)$.

- $G$ is an $n$-vertex regular graph of degree $D$, and $A$ is the random walk matrix of $G$.
- $H$ is a $D$-vertex regular graph of degree $d$, and $B$ is the random walk matrix of $H$.

$(v, m)$ is the $(i, j)$-th neighbor of $(u, l)$: $l'$ is the $i$-th neighbor of $l$ in $H$; $v$ is the $l'$-th neighbor of $u$ and $u$ is the $m'$-th neighbor of $v$; $m$ is the $j$-th neighbor of $m'$ in $H$. 
Lemma. \((I_n \otimes J_D) \hat{A}(I_n \otimes J_D) = A \otimes J_D.\)

\[
\left( (I_n \otimes J_D) \hat{A}(I_n \otimes J_D) \right)_{(v, m), (u, l)} = \frac{1}{D} \cdot 1 \cdot \frac{1}{D} = \frac{1}{D} \cdot \frac{1}{D} = (A \otimes J_D)_{(v, m), (u, l)}.\]
Claim. If $\|C\|_2 \leq 1$ then $\lambda_C \leq 1$.

Proof.

$\lambda_C = \max_{v \perp 1} \frac{\|Cv\|_2}{\|v\|_2} \leq \max_{v \perp 1} \frac{\|C\|_2 \|v\|_2}{\|v\|_2} \leq \|C\|_2 \leq 1.$

Claim. $\lambda_{A+B} \leq \lambda_A + \lambda_B$ for symmetric matrices $A, B$.

Proof.

$\lambda_{A+B} = \max_{v \perp 1} \frac{\|(A+B)v\|_2}{\|v\|_2} \leq \max_{v \perp 1} \frac{\|Av\|_2 + \|Bv\|_2}{\|v\|_2} \leq \lambda_A + \lambda_B.$
Zig-Zag Product

**Lemma.** $\lambda_{G \boxtimes H} \leq \lambda_G + 2\lambda_H$ and $\gamma_{G \boxtimes H} \geq \gamma_G \gamma_H^2$.

Let $A$, $B$ and $M$ be the random walk matrices of $G$, $H$ and $G \boxtimes H$ respectively.

- $\hat{A}$ is the $(nD) \times (nD)$ rotation matrix of $G$.
- $B = (1 - \lambda_H)J_D + \lambda_H E$ for some $E$ with $\|E\|_2 \leq 1$. This is the Lemma.

Now

$$M = (I_n \otimes B)\hat{A}(I_n \otimes B) = ((1 - \lambda_H)I_n \otimes J_D + \lambda_H I_n \otimes E) \hat{A} ((1 - \lambda_H)I_n \otimes J_D + \lambda_H I_n \otimes E)$$

$$= (1 - \lambda_H)^2(I_n \otimes J_D)\hat{A}(I_n \otimes J_D) + \ldots = (1 - \lambda_H)^2(A \otimes J_D) + \ldots,$$

where $= \text{ is due to Lemma.}$ Using Lemma and the Claims, one gets

$$\lambda_M \leq (1 - \lambda_H)^2 \lambda_{A \otimes J_D} + 1 - (1 - \lambda_H)^2 \leq \max\{\lambda_G, \lambda_{J_D}\} + 2\lambda_H = \lambda_G + 2\lambda_H.$$

For the inequality $\gamma_{G \boxtimes H} \geq \gamma_G \gamma_H^2$, consider $1 - \leq$. 

---

*Computational Complexity, by Fu Yuxi*
Comment on Zig-Zag Product

1. Typically \( d \ll D \).

2. A \( t \)-step random walk uses \( O(t \log d) \) rather than \( O(t \log D) \) random bits.

3. The last lemma is useful when both \( \lambda_G \) and \( \lambda_H \) are small. If not, a different upper bound can be derived. Both upper bounds are discussed in the following paper.

One idea is to use path product and zig-zag product to produce an expander family.

<table>
<thead>
<tr>
<th>Product</th>
<th>Size</th>
<th>Degree</th>
<th>Expansion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Path Product</td>
<td>-</td>
<td>↑</td>
<td>↑</td>
</tr>
<tr>
<td>Tensor Product</td>
<td>↑</td>
<td>↑</td>
<td>↓</td>
</tr>
<tr>
<td>Zigzag Product</td>
<td>↑</td>
<td>↓</td>
<td>↓</td>
</tr>
</tbody>
</table>
The crucial point of the zig-zag construction is that we can use a constant graph to build a constant degree graph family.

Let $H$ be a $(D^4, D, 1/8)$-graph constructed by brutal force. Define

\[ G_1 = H^2, \]
\[ G_{k+1} = G_k^2 \odot H. \]

**Fact.** $G_k$ is a $(D^{4k}, D^2, 1/2)$-graph.

**Proof.**
The base case is clear from Lemma, and the induction step is taken care of by the previous lemma. \qed
Expander Construction I

The time to access to a neighbor of a node is given by the following inductive equation

\[
\text{time}(G_k) = 2 \cdot \text{time}(G_{k-1}) + \text{time}(H) \\
= 2^{k-1} \cdot \text{time}(H^2) + (2^{k-2} + \ldots + 2 + 1) \cdot \text{time}(H) \\
= 2^{O(k)} \\
= \text{poly}(|G_k|).
\]

The time to compute a neighbor is a polynomial of the graph size. We conclude that the expander family is explicit, but not strongly explicit.

We will see that \( \text{space}(G_k) \) is \( \text{polylog}(|G_k|) \).
Replacement Product

The replacement product $G \Box H$ is the $2d$-degree graph defined by $\frac{1}{2}\hat{A} + \frac{1}{2}(I_n \otimes B)$.

- $G$ is an $n$-vertex regular graph of degree $D$, and $A$ is the random walk matrix of $G$.
  $H$ is a $D$-vertex regular graph of degree $d$, and $B$ is the random walk matrix of $H$.

If $\hat{G}(u, l) = (v, m)$, place $d$ parallel edges from the $l$-th vertex of $H_u$ to the $m$-th vertex of $H_v$. 
Lemma. $\lambda_{G \oplus H} \leq 1 - \frac{(1 - \lambda_G)(1 - \lambda_H)^2}{24}$ and $\gamma_{G \oplus H} \geq \frac{1}{24} \gamma_G \gamma_H^2$.

Let $A, B$ be the random walk matrices of $G, H$ respectively.

\[
(A \oplus B)^3 = \left(\frac{1}{2} \hat{A} + \frac{1}{2} (I_n \otimes B)\right)^3 = \left(\frac{1}{2} \hat{A} + \frac{1}{2} (I_n \otimes (\lambda_H E + \gamma_H J_D))\right)^3
\]

\[
= \frac{1}{8} \left(\hat{A} + \lambda_H (I_n \otimes E) + \gamma_H (I_n \otimes J_D)\right)^3 = \frac{1}{8} \left(\hat{A}^3 + \ldots + \gamma_H^2 (I_n \otimes J_D) \hat{A} (I_n \otimes J_D)\right)
\]

\[
= \frac{1}{8} \left(\hat{A}^3 + \ldots + \gamma_H^2 (A \otimes J_D)\right),
\]

where the last equality is due to Lemma. Applying Lemma and the Claims, we get

\[
(\lambda_{A \oplus B})^3 = \lambda_{(A \oplus B)^3} \leq 1 - \frac{\gamma_H^2}{8} + \frac{\gamma_H^2}{8} \lambda_{A \otimes J_D} \leq 1 - \frac{\gamma_H^2}{8} + \frac{\gamma_H^2}{8} \lambda_G = 1 - \frac{\gamma_H^2}{8} \gamma_G.
\]

We have proved that $(\lambda_{G \oplus H})^3 \leq 1 - \frac{\gamma_G \gamma_H^2}{8} \leq \left(1 - \frac{\gamma_G \gamma_H^2}{24}\right)^3$. Hence $\gamma_{G \oplus H} \geq \frac{1}{24} \gamma_G \gamma_H^2$. 
Theorem. There exists a strongly explicit \((4, \lambda)\)-expander family for some \(\lambda < 1\).

As a first step we prove that we can efficiently construct a family \(\{G_k\}_{k \in \omega}\) of graphs where each \(G_k\) has \((2d)^{100k}\) vertices.

1. Let \(H\) be a \(((2d)^{100}, d, 0.01)\)-expander graph, \(G_1\) a \(((2d)^{100}, 2d, 0.5)\)-expander graph, and \(G_2\) a \(((2d)^{100 \times 2}, 2d, 0.5)\)-expander graph, all found by brutal force.

2. For \(k > 2\) define

\[
G_k = \left( G_{\lfloor \frac{k-1}{2} \rfloor} \otimes G_{\lceil \frac{k-1}{2} \rceil} \right)^{50} \bar{\otimes} H.
\]

\(G_k\) is a \(((2d)^{100k}, 2d, 0.98)\)-expander graph.
Fact. $G_k$ is a $((2d)^{100k}, 2d, 0.98)$-expander graph.

1. Let $n_k$ be the number of vertices of $G_k$.

$$n_k = n_{\left\lfloor \frac{k-1}{2} \right\rfloor} n_{\left\lceil \frac{k-1}{2} \right\rceil} (2d)^{100} = (2d)^{100\left\lfloor \frac{k-1}{2} \right\rfloor}(2d)^{100\left\lceil \frac{k-1}{2} \right\rceil}(2d)^{100} = (2d)^{100k}.$$

2. $G_{\left\lfloor k-1 \right\rfloor}, G_{\left\lceil k-1 \right\rceil}$ degree $2d \Rightarrow G_{\left\lfloor k-1 \right\rfloor} \otimes G_{\left\lceil k-1 \right\rceil}$ degree $(2d)^2 \Rightarrow (G_{\left\lfloor k-1 \right\rfloor} \otimes G_{\left\lceil k-1 \right\rceil})^50$ degree $(2d)^{100} \Rightarrow G_k$ degree $2d$.

3. $\lambda_{G_{\left\lfloor k-1 \right\rfloor}}, \lambda_{G_{\left\lceil k-1 \right\rceil}} \leq 0.98 \Rightarrow \lambda_{G_{\left\lfloor k-1 \right\rfloor} \otimes G_{\left\lceil k-1 \right\rceil}} \leq 0.98 \Rightarrow \lambda(G_{\left\lfloor k-1 \right\rfloor} \otimes G_{\left\lceil k-1 \right\rceil})^{50} \leq 0.5 \Rightarrow \lambda_{G_k} \leq 1 - 0.5(0.99)^2/24 < 0.98.
There is a \( \text{poly}(k) \)-time algorithm that upon receiving a label \( v \) of a vertex in \( G_k \) and an index \( j \) in \([2d]\) finds the \( j \)-th neighbor of \( v \).

\[
|\langle n, v, i \rangle| = \text{polylog}(n).
\]

\[
\text{time}(G_{k+1}) = 50 \cdot \text{time}(G_{\lfloor k/2 \rfloor}) + 50 \cdot \text{time}(G_{\lceil k/2 \rceil}) + \text{time}(H)
\]

\[
= 2^{O(\log k)} \cdot \text{time}(G_2) + (2^{O(\log k - 1)} + \ldots + 2^{O(1)} + O(1)) \cdot \text{time}(H)
\]

\[
= 2^{O(\log k)}
\]

\[
= \text{poly}(k).
\]

The time to compute a neighbor is \( \text{polylog}(n) \). The expander family is strongly explicit.
Expander Construction II

Suppose \((2d)^{100k} < i < (2d)^{100(k+1)}\). Let \((2d)^{100(k+1)} = xi + r\).

- Divide the \((2d)^{100(k+1)}\) vertices into \(i\) classes among which \(r\) classes being of size \(x + 1\) and \(i - r\) classes being of size \(x\).
- Contract every class into a mega-vertex.
- Add \(2d\) self-loops to each of the \(i - r\) mega-vertices.

This is a \((i, 2d(x + 1), (2d)0.01/(x + 1))\) edge expander.

We get a \(((2d)^{101}, 0.01/(2d)^{99})\) edge expander family.
Reingold’s Theorem
Theorem. $\text{UPATH} \in L$.

The Idea

Connectivity Algorithm for $d$-degree expander graph is easy.

- The diameter of an expander graph is of length $O(\log(n))$.
- An exhaustive search can be carried out in logspace.

Reingold’s idea is twofold.

1. Transforming conceptually the input graph $G$ to a graph $G'$ so that a connected component in $G$ turns to an expander in $G'$ and unconnected vertices in $G$ remain unconnected in $G'$.

2. Finding a neighbor of a given vertex in the imaginary $G'$ can be done in logspace.
The Algorithm

Fix the \((D^4, D, 1/8)\)-graph \(H\), and apply Construction I.

1. Let \(s_0 = s\) and \(t_0 = t\).
2. Convert the input graph \(G\) to a \(D^2\)-degree graph \(G_0\) on the fly.
   2.1 Add self-loops to increase degree.
   2.2 Replace a large degree vertex by a cycle to decrease degree.
3. \(G_k = G_{k-1}^2 \oplus H\) is constructed on the fly. Let \(s_k\) be a node in the “cloud” corresponding to \(s_{k-1}\) and \(t_k\) be a node in the “cloud” corresponding to \(t_{k-1}\).
4. Apply Connectivity Algorithm to the expander \(G_{10 \log n}\).

Correctness.

- \(G_k\) is a \((D^2, 1/2)\)-graph.
- \(s_k\) and \(t_k\) are connected if and only if \(s_{k-1}\) and \(t_{k-1}\) are connected.
The Data Structure

The algorithm imagines a tree structure for $G_{10\log n}$, and exhausts all paths starting from $s$ by carrying out depth first traversal of the imaginary tree repeatedly.

If $r$ is the $(i, j)$-th neighbor of $s$ and $t$ is the $(i', j')$-th neighbor of $r$, then $(i, j)(i', j')$ is stored in the record for $s G_{10\log n} t$. The algorithm only stores the current vertex for backtracking.
The Complexity

One need to investigate the space complexity of accessing a neighbor of a vertex in $G_k$.

1. $G_0$ can be constructed in logspace.

2. To visit a neighbor of a node in $G_k$ using the rotation matrix of $G_k$, it makes use of the rotation matrix of $G_{k-1}$ twice. The key point is the following.

\[
\begin{align*}
\text{space}(G_{k-1}^2) &= \text{space}(G_{k-1}) + O(1), \\
\text{space}(G_{k-1}^2 \odot H) &= \text{space}(G_{k-1}^2) + O(1).
\end{align*}
\]

The size of the additional space depends only on $D$, which is a constant.

3. The depth first tree traversal keeps a stack of depth bounded by $10 \log n$. 

Lewis and Papadimitriou introduced $\text{SL}$ as the class of problems solvable in logspace by an NTM that satisfies the following.

1. If the answer is 'yes,' one or more computation paths accept.
2. If the answer is 'no,' all paths reject.
3. If the machine can make a transition from configuration $C$ to configuration $D,$ then it can also goes from $D$ to $C.$

**Theorem.** $\text{UPATH}$ is $\text{SL}$-complete.

**Corollary.** $\text{UPATH}$ is $\text{L}$-complete.

**Proof.**
Reingold Theorem implies that $\text{L} = \text{SL}.$
The problem “$\text{RL} = \text{L}$” is open. The best we know is $\text{RL} \subseteq \text{L}^{3/2}$. 