The basic idea is to generalize $S\Lambda_d$. Again we introduce variables $Y_I$. For all subsets $I \subseteq [n]$, $|I| \leq t$.

But now we will argue that the matrix

\[
\begin{pmatrix}
I & \mu \\
& \Omega
\end{pmatrix}
\]

\[|I|, |J| \leq \frac{d}{2}, \quad |I \cup J| \leq d\]

\[Y_{I \cup J}\]

is PSD.

- Since PSD matrices are convex, can still solve (via SDP) to arbitrary accuracy in time $n^{O(d)}$.

- Duality to proof system (SOS) follows by extended Farkas lemma and relationship between SOS polynomials and PSD matrices.

- PSD condition implies the $S\Lambda$ inequalities (so degree-$d$ SOS tightening are at least as good as degree-$d$ $S\Lambda_d$).
All vectors over Reals

Vectors \( u, v \) are column vectors, so

\[
\begin{align*}
\mathbf{u}^T \mathbf{v} &= \langle \mathbf{u}, \mathbf{v} \rangle = \text{a scalar} \\
\mathbf{u} \mathbf{v}^T &= \mathbf{u} \otimes \mathbf{v} = \text{tensor product} \\
(u \otimes v)_{ij} &= u_i \cdot v_j
\end{align*}
\]

Fact \( x^T A x = (x x^T) \cdot A \)

<table>
<thead>
<tr>
<th>dot product as vectors</th>
</tr>
</thead>
<tbody>
<tr>
<td>outer product</td>
</tr>
</tbody>
</table>

Fact \( \text{Tr} (AB) = \text{Tr}(BA) \)

Fact \( A, B \text{ symmetric} \implies \text{Tr}(AB) = \mathbf{A} \cdot \mathbf{B} \)

\( \text{Tr} = \text{sum 9 diagonal entries} \)
\[ \|v\|_p = \left( \sum_{i=1}^n |v_i|^p \right)^{\frac{1}{p}} \quad p \geq 1 : \text{Euclidean norm} \]

\[ \|v\|_\infty = \max_{i \in [n]} |v_i| \]

A, B matrices, entries \( A_{ij} \), \( B_{ij} \)

\( A \otimes B \), the Kronecker product is a matrix

with entries \((A \otimes B)_{ij,ij} = A_{ij} B_{ij}\)

\( A^{\otimes k} = A \otimes A \otimes \cdots \otimes A \)

A, B square matrices

\[ \text{Det}(AB) = \text{Det}(A) \cdot \text{Det}(B) \]

A diagonal matrix or (upper or lower) triangular. Then

\[ \text{Det}(A) = \text{product of diagonal entries} \]
Let $A$ be an $n \times n$ matrix. Then the sum of eigenvalues $\theta A = \text{trace}(A)$ and product of eigenvalues $= \text{Det}(A)$.
**PSD matrices**

Let $A$ be an $n \times n$ symmetric matrix on Reals.

$A$ is **positive-semidefinite** (PSD) if $\forall v \in \mathbb{R}^n$

$$v^T A v \geq 0 \quad \left( v^T A v = \sum_{i,j} A_{ij} v_i v_j \right)$$

$A$ is **positive definite** if it is positive semi-definite and non-singular.

**Notation**

- $A \succeq 0$ means $A$ is PSD
- $A > 0$ means $A$ is positive definite

**Fact**

If $M_1, \ldots, M_k$ are PSD, so is any nonnegative combination

$$\alpha_1 M_1 + \alpha_2 M_2 + \ldots + \alpha_k M_k \quad \alpha_i \geq 0$$

(so $\text{PSD}_n$ - all $n$-by-$n$ PSD matrices forms a cone.)
Matrix Decompositions

Let \( A > 0 \) (A n-by-n)

Then \( A \) has the following decompositions

1. \( A = U^T U \) Cholesky decomp

   - \( U \) is unique
   - Witnesses \( A > 0 \) since \( x^T A x = x^T U U^T x = |Ux|^2 > 0 \)

2. \( A = L D L^T \) LDL decompositions

   - \( D \) diagonal and nonneg
   - \( L \) lower triangular
   - \( L, D \) are unique

LDL holds for any symmetric matrix, but \( A > 0 \) \( \Rightarrow \) \( D \) has positive entries
\[ A = \mathbf{Q} \Lambda \mathbf{Q}^T \]

**SVD**  
**spectral decomp**  
**eigen decomp**  
*(not unique)*

\( \mathbf{Q} \) is orthogonal matrix

columns of \( \mathbf{Q} \) are unitary eigenvectors of \( A \)

ie. \( \mathbf{Q} \mathbf{Q}^T = \mathbf{I} \)

\( \mathbf{\Lambda} \): diagonal matrix corresponding to eigenvalues, all are positive

A symmetric \( \Rightarrow \) eigenvectors with different eigenvalues are orthogonal:

\( x \) has eigenvalue \( \lambda_x \neq 0 \)

\( y \) " " \( \lambda_y \neq 0 \)

Then \( x^T A y = \lambda_x x^T y \)

\( y^T A x = \lambda_y y^T x = \lambda_y x^T y \)

so \( \lambda_x x^T y = \lambda_y x^T y \Rightarrow \lambda_x = \lambda_y \)

A symmetric + real \( \Rightarrow \) eigenvalues are real

say \( x \) has eigenvalue \( \lambda \). Then

\( (Ax)^T A x = x^T A^T A x = x^T \mathbf{\Lambda} A x = \mathbf{\Lambda} (1 \times 1)^2 \)

is real + positive so \( \lambda \) is real
**LDL Decomposition**

show by induction that $A = L^TDL$

using $(\text{for } x > 0)$:

$$
\begin{bmatrix}
\alpha & v^T \\
v & C
\end{bmatrix} =
\begin{bmatrix}
1 & 0 \\
v & 1
\end{bmatrix}
\begin{bmatrix}
\frac{\alpha}{v} & 0 \\
0 & C - \frac{vv^T}{\alpha}
\end{bmatrix}
\begin{bmatrix}
1 & v^T\
0 & 1
\end{bmatrix}
$$

$v^T$: length $n-1$ vector

$C, I$: $(n-1) \times (n-1)$ matrix

Need to show: $B = C - \frac{vv^T}{\alpha}$ is positive definite.

Fix $u \in \mathbb{R}^{n-1}$, $u$ nonzero

Fix $x^T = \begin{bmatrix} -\frac{v^Tu}{\alpha}, u^T \end{bmatrix}$. scalar

Then $u^T Bu = x^T Ax > 0$ so positive definite
From LDL, get cholesky decomp. by:

\[ A = L^T D L = (L^T D^\frac{1}{2} L) \]

since \( D \) is diagonal

\[ = (LD^\frac{1}{2})^T (LD^\frac{1}{2}) \]

\[ = U^T U \]

If \( A \) is PSD, all decompositions still work. But then the entries of \( U + D \) can be zero.

Also \( U, L, D \) not unique

and \( U \) can have zero entries

\[ [\text{For LDL if } A \text{ is PSD, } \alpha \geq 0. \]

If \( \alpha = 0 \) then \( v \) is also 0 since otherwise there would be a \( y \) st \( y^T A y < 0 \)
Lemma: A symmetric $n \times n$ matrix on $\mathbb{R}$

$A$ is PSD iff $A \cdot B \geq 0$ for all PSD matrices $B$

Proof: $A, B$ PSD $\Rightarrow$ can write $A$ as

$Q^T \Lambda Q$ (spectral decom)

Then

$A \cdot B = \text{Tr}(AB) = \text{Tr}(Q^T \Lambda Q B)$

$(\text{Tr}(M) = \text{sum of diag entries})$

$= \text{Tr}(\Lambda Q B Q)$

$\Lambda$ diagonal, non-neg

$\Rightarrow \text{Tr}(\Lambda B')$

$B' = Q B Q$ is PSD

$= \sum \lambda_i b'_{ii}$

$\geq 0$

$\Leftarrow$: say $\exists y^* \text{ s.t. } y^T A y < 0$

Fix $B = y y^T$. Then $A \cdot B = y^T A y$

$B$ is PSD by construction

$+ A \cdot B < 0$.

Lemma: $A \succ 0$ then $A \cdot B > 0 \forall B \succ 0$ unless $B = 0$

$A$ pos def.
Lemma (another characterization)

A n-by-n square, symmetric matrix over \( \mathbb{R} \)

A is positive definite iff determinant of all upper left submatrices are positive.

Thm A n-by-n symmetric. TFAE:

1. A is psd (\( A \succeq 0 \))
2. All eigenvalues are non-neg
3. Determinant of all upper left submatrices is non-negative
4. The quadratic polynomial \( P_A(x) = \Sigma A_{ij} x_i x_j \)
   is a sos's,
   i.e. \( \exists \) linear forms \( l_i, \ldots, l_m \) s.t.
   \[ P_A = \Sigma_i (l_i)^2 \]
5. \( A = U^T U \) (cholesky)
6. There are correlated rv's \( X_1, \ldots, X_m \) s.t. \( \forall i,j \) \( x_i x_j = A_{ij} \) and \( X_i \) distributed like \( N(0, A_{ii}) \)
**SOS - some history**

1. Late 1800's Minkowski/Hilbert asked:
   
   Can every nonneg multivar poly over $\mathbb{R}$ be written as a SOS's $(p_1^2 + p_2^2 + \ldots + p_k^2)$

   Motzkin [1960's] - NO!

   - $1 + x^4 y^2 + x^2 y^4 - 3 x^2 y^2$ is nonnegative.

     \[
     \frac{1 + x^4 y^2 + x^2 y^4 - 3 x^2 y^2}{3} \geq (1 \cdot x^2 y^2, x^2 y^4)^T
     \]

     so nonneg.

     Can also show it cannot be written as SOS.

2. Hilbert's 17th Problem (1900 address):

   Can every nonneg poly over $\mathbb{R}$ be written as sum-of-squares of Rational functions?

   Artin (1927) - yes!
More on Motzkin’s polynomial

\[ 1 + x^4y^2 + x^2y^4 - 3x^2y^2 \text{ nonnegativity} \]

**Arithmetic Mean - Geometric Mean (AM-GM) Ineq:**

\[ \text{AM of } x_1, \ldots, x_n : \frac{x_1 + x_2 + \cdots + x_n}{n} \]

\[ \text{GM of } x_1, \ldots, x_n : (x_1 \cdot x_2 \cdots \cdot x_n)^\frac{1}{n} \]

\[ \text{AM-GM} : \frac{x_1 + \cdots + x_n}{n} \geq (x_1 \cdot x_2 \cdots \cdot x_n)^\frac{1}{n} \]

\[
\begin{align*}
\text{2 dimensions:} & \quad \text{perim. of } x_1 \text{-by-} x_2 \text{ rect is } 2x_1 + 2x_2 \\
\text{perim of square with the same area } x_1 x_2 \text{ is } 4 \sqrt{x_1 x_2} \\
\text{says square has smallest perimeter of all rectangles of same area}
\end{align*}
\]

Lots of proofs of AM-GM ineq
Jansen's Ineq.

Let $f$ be concave function. Then

$$f \left( \frac{x_1 + \ldots + x_n}{n} \right) \geq \frac{1}{n} \sum_i f(x_i)$$

$f$ on mean $\bar{x} = \text{mean} \{ f(x_1) \ldots f(x_n) \} \overline{x_i}$

Proof of AM-GM from Jansen's Ineq:

$$f = \log$$

$$\log ( \frac{\sum x_i}{n} ) \geq \frac{1}{n} \sum_i \log x_i$$

$$= \frac{1}{n} \sum_i \log x_i^x$$

$$= \log ( \prod x_i^x )$$

Back to Nonneg of Motzkin's Poly:

$$1 + x^4 y^2 + y^4 x^2 - 3x^2 y^2$$

apply AM-GM to $1, x^4 y^2, y^4 x^2$

$$\frac{-3x^2 y^2 + 1 + x^4 y^2 + y^4 x^2}{3} \geq -x^2 y^2 + \left( \frac{x^4 y^6}{3} \right)^{1/3}$$

$$= (3^3 - 1)x^2 y^2 \geq 0$$
Motzkin Poly cont'd

To see $M(x,y) = x^2 y^4 + x^4 y^2 + 1 - 3x^2 y^2$

is non-neg, here's a better way.

$(x^2 + y^2 + 1)$ is positive

and we have

$$(x^2 + y^2 + 1) M(x,y) =$$

$$(x^2 y - y)^2 + (xy^2 - x)^2 +$$

$$(x^2 y^2 - 1)^2 + \frac{3}{4} (xy^3 - x^3 y)$$

$$+ \frac{3}{4} (xy^3 + x^3 y - 2xy)^2$$

so $M(x,y) = \frac{p_1^2 + p_2^2 + p_3^2 + p_4^2 + p_5^2}{Q_1^2 + Q_2^2 + 1}$

ie. it is a sum-of-squares (always nonneg)

divided by something positive

so $M(x,y)$ is nonneg!
so this poly is nonneg.
and can be written as \( \frac{\text{SOS}}{\text{sos}} \)

But cannot be written as \( \text{SOS}! \)

Luckily

In some special cases, \( \text{SOS} \) is equivalent to non-negativity.

1. Univariate case \( n=1 \)
2. Degree \( 2 \) \( d=2 \) \quad \text{Hilbert 1888}
3. \( n=2 \) and \( d=4 \)

* 4. Functions over hypercube
   \( f: [0,1]^n \to \mathbb{R} \)
Univariate case

\[ f(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_{2d} x^{2d} \]

\[ = [1 \ x \ x^2 \ \ldots \ x^d] \begin{bmatrix} q_{00} & q_{01} & \ldots & q_{0d} \\ q_{01} & q_{11} & \ldots & q_{1d} \\ \vdots & \vdots & \ddots & \vdots \\ q_{0d} & q_{1d} & \ldots & q_{dd} \end{bmatrix} \begin{bmatrix} 1 \\ x \\ x^2 \\ \vdots \\ x^d \end{bmatrix} \]

\[ = \sum_{i=0}^{d} \left( \sum_{j+k=i} q_{jk} \right) x^i \]

In univ case, SOS condition is equivalent to nonnegativity !
SOS over Boolean Hypercube

**Lemma** Every nonnegative function $f : [0,1]^n \to \mathbb{R}$ has a degree $2n$ SOS certificate.

**Proof**: Let $g : [0,1]^n \to \mathbb{R}$ be the multilinear (unique) function that agrees with $f$ on the hypercube. Then $f = g^2$ over $[0,1]^n$ and has degree $\leq 2n$.

We'll soon see a better, more constructive algorithm to find an SOS certificate (but still exp time in worst case).
\textbf{sos certificates via PSD matrices}

**Lemma**

Let $f(x_1, x_2, \ldots, x_n) \in \mathbb{R}[x_1, x_2, \ldots, x_n]$ be a degree $2d$ poly

Let $z$ be a vector of all degree $\leq 2d$ monomials.

Then $f$ is sos iff $f(x) = z^T Q z$, $Q \succeq 0$

\textbf{Proof}

Factorize $Q = U^T U$.

Then $f(x) = z^T U^T U z = \|Uz\|^2$

\[= \sum_i (Uz)_i^2 \]

The terms in decomposition are given by $g_i = (Uz)_i$

The # $g$ squares = rank of $Q$

\[\text{i.e. } (Uz)^T = [y_1, \ldots, y_r] \quad r \text{ pd's}\]

\[(Uz)^T (Uz) = [y_1, \ldots, y_r] \begin{bmatrix} y_1^2 \\ \vdots \\ y_r^2 \end{bmatrix} = y_1^2 + \cdots + y_r^2\]
So \( f = (Uz)^T \cdot (Uz) = y_1^2 + \ldots + y_r^2 \)

the \( i \)th row of \( U \) are coeffs \( y_i \) of \( Y_i \) since highest terms in \( y_i \) cannot cancel degrees of \( y_i \)'s are at most \( d \)

\[ \begin{align*}
\Rightarrow \quad f &= y_1^2 + y_2^2 + \ldots + y_r^2 \\
\text{create matrix} \ Q \ \text{s.t.} \ \\
Q_{ij} &= y_i \cdot y_j
\end{align*} \]

\[ \begin{pmatrix}
y_1^2 & \ldots & y_r^2 \\
\text{coeff} \ y_i \cdot \text{coeff} \ y_j
\end{pmatrix} \]

\( y_i \)s over \( Z \). Can think \( \mathbb{F}_2^d : \mathbb{F}_2 \), where nonzero \( Z_i \)s correspond to degree \( \leq d \) monomials in \( x \) vars

But \( f: \mathbb{R}[x_1, \ldots, x_n]_{2d} \to \mathbb{R} \) nonneg is \( \text{not equivalent to} \ f \ \text{being SOS} \)
Example: Let \( p = 2x^4 + 2x^3y - x^2y^2 + 5y^4 \)

If \( p \) a sos, since \( p \) is homogeneous, the poly's in the SOS's representation have degree 2

[so terms are \( x^2, y^2, xy \)]

\[
\begin{pmatrix} x_1 & x_2 & x_3 \\
 x & y & xy \end{pmatrix}
\begin{bmatrix} q_{11} & q_{12} & q_{13} \\
 q_{21} & q_{22} & q_{23} \\
 q_{31} & q_{32} & q_{33} \end{bmatrix}
\begin{pmatrix} x^2 \\
 y^2 \\
 xy \end{pmatrix}
= \sum_{i,j} q_{ij} x_i x_j
\]

\[
= q_{11} x_1 x_1 + q_{22} x_2 x_2 + q_{33} x_3 x_3 +
2 q_{12} x_1 x_2 + 2 q_{23} x_2 x_3 + 2 q_{13} x_1 x_3
\]

\[
= q_{11} x^4 + q_{22} y^4 + q_{33} x^2 y^2 + 2 q_{12} x^2 y^2 
+ 2 q_{23} x y^3 + 2 q_{13} x y^3
\]

\[
= q_{11} x^4 + q_{22} y^4 + (q_{33} + 2 q_{12}) x^2 y^2 + 2 q_{23} x y^3 
+ 2 q_{13} x y^3
\]
The existence of a PSD $Q$ is equivalent to feasibility of an SDP in standard primal form:

$$Q \succeq 0 \text{ s.t.} \begin{align*}
q_{11} &= 2 \\
q_{22} &= 0 \\
q_{13} &= 2 \\
q_{23} &= 0 \\
q_{33} + 2q_{32} &= -1
\end{align*}$$
Matching coefficients of $p$

$$p = 2x^4 + 2x^3y - x^2y^2 + 5y^4 + 0xy^3$$

gives $q_{11}, q_{13}, q_{22}, 2q_{23}$

$$-1 = q_{23} + 2q_{11}$$

Solving SDP gives

$$Q = \begin{pmatrix} 2 & -3 & 1 \\ -3 & 5 & 0 \\ 1 & 0 & 5 \end{pmatrix}$$

$\text{Rank 2}$

so $p$ is the sum of 2 squares

Cholesky factorization:

$$Q = \frac{1}{\sqrt{2}} \begin{pmatrix} 2 & -3 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 2 & -3 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix}$$

$$U$$

$$W = \frac{1}{\sqrt{2}} \begin{pmatrix} 2 & -3 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 2 & -3 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix}$$

$$xy = \begin{pmatrix} 2x^2 - 3y^2 + xy \\ y^2 + 3xy \end{pmatrix}$$

so $p = y_1^2 + y_2^2$
Now let's consider $f: \{0,1\}^n \to \mathbb{R}$. Here, $f$ non-neg is equivalent to $f$ having an SOS certificate.

But now SOS certificate could have much larger degree than $f$.

Let $f: \{0,1\}^n \to \mathbb{R}$. Let $Z$ be vector of all multilinear monomials on $x_1, \ldots, x_n$ ($|Z| = 2^n$).

**Thm** $f: \{0,1\}^n \to \mathbb{R}$ is non-neg. iff $f$ has an SOS certificate iff $\exists Q$ s.t. $Z^TQZ = f(x)$, $Q \succeq 0$.

**Pf** (1) We already saw $f$ non-neg iff $f$ has an SOS certificate.

(2) Same as previous pf but now can't restrict to monomials of degree $\leq \deg(f)$.
Example

Here's an example of a function \( f : \{0,1\}^n \rightarrow \mathbb{Z} \) that is non-Neg, has low degree (3) but requires SOS certificates of degree \( 2(n) \).

Start with random \texttt{UNSAT} 3-CNF

\[
\text{random mod 2 eqns, or Tseitin on degree-3 expander}
\]

\( f = \bigwedge \bigwedge \bigwedge \bigwedge \) over \( x_1 \ldots x_n \)

Convert each clause \( C \rightarrow P_C \):

\[
C = (x_1 \lor \neg x_2 \lor x_3) \rightarrow P_C = (1-x_1)(x_2)(1-x_3)
\]

\( \forall \alpha \) falsifying \( C \rightarrow P_C (\alpha) = 1 \)

\( \forall \alpha \) satisfying \( C \rightarrow P_C (\alpha) = 0 \)

Let \( P_f = \sum_{C \in \text{ef}} P_C - 1 \)

Thm (Grigoriev) \( P_f \) requires SOS degree \( 2(n) \) [we will do this soon]

and Non-Negativity if \( f \) is \texttt{UNSAT}
To find SOS certificate, find a feasible solution to:

\[ P_f = \sum Q_{ij} z_i z_j \]

\[ Q \geq 0 \]

This is a bunch of equations equating coefficient's \( Q_j P_f \) to poly's (as in example)

Variables are \( q_{ij} \)

where \( i,j \) correspond to \( i,j \in [n] \)
So already we can see that sum-of-squares forms the basis of a proof system (for nonnegativity of a single function over hypercube.)

Let \( f: \{0,1\}^n \to \mathbb{R} \)

Then \( f \preceq 0 \iff -f-1 \succeq 0 \)

is unsolvable over \( \{0,1\}^n \) iff

\[ \exists \text{ sos } g \text{ s.t. } g = f. \]

\[ \iff \text{Let } g \text{ be a sos, } g = f. \text{ Then } g \succeq 0, \quad -f-1 \succeq 0 \]

\[ \iff g - f - 1 = -1 \succeq 0 \quad \# \]

\[ \iff \text{ If } f \preceq 0 \text{ unsolvable over } \{0,1\}^n \]

then \( f \succeq 0 \), so by completeness \( \exists \text{ sos } g \text{ s.t. } g = f \)

\[ \text{[degree of } g \text{ is } \leq 2n] \]
More general proof system [more than one poly]

[Krivine, Steenge] 60-70's

Positivstellensatz (generalizes Nullstellensatz)

Take a set of poly inequalities asserting all are nonnegative

\[ p_i \geq 0, \ldots, p_m \geq 0, \ i = 0 \]

If unsolvable (an IR) then they imply -1

\[ \sum P_i \cdot Q_i = -1, \quad \text{where} \quad Q_i's \text{ are SOS's} \]

This is an SOS proof