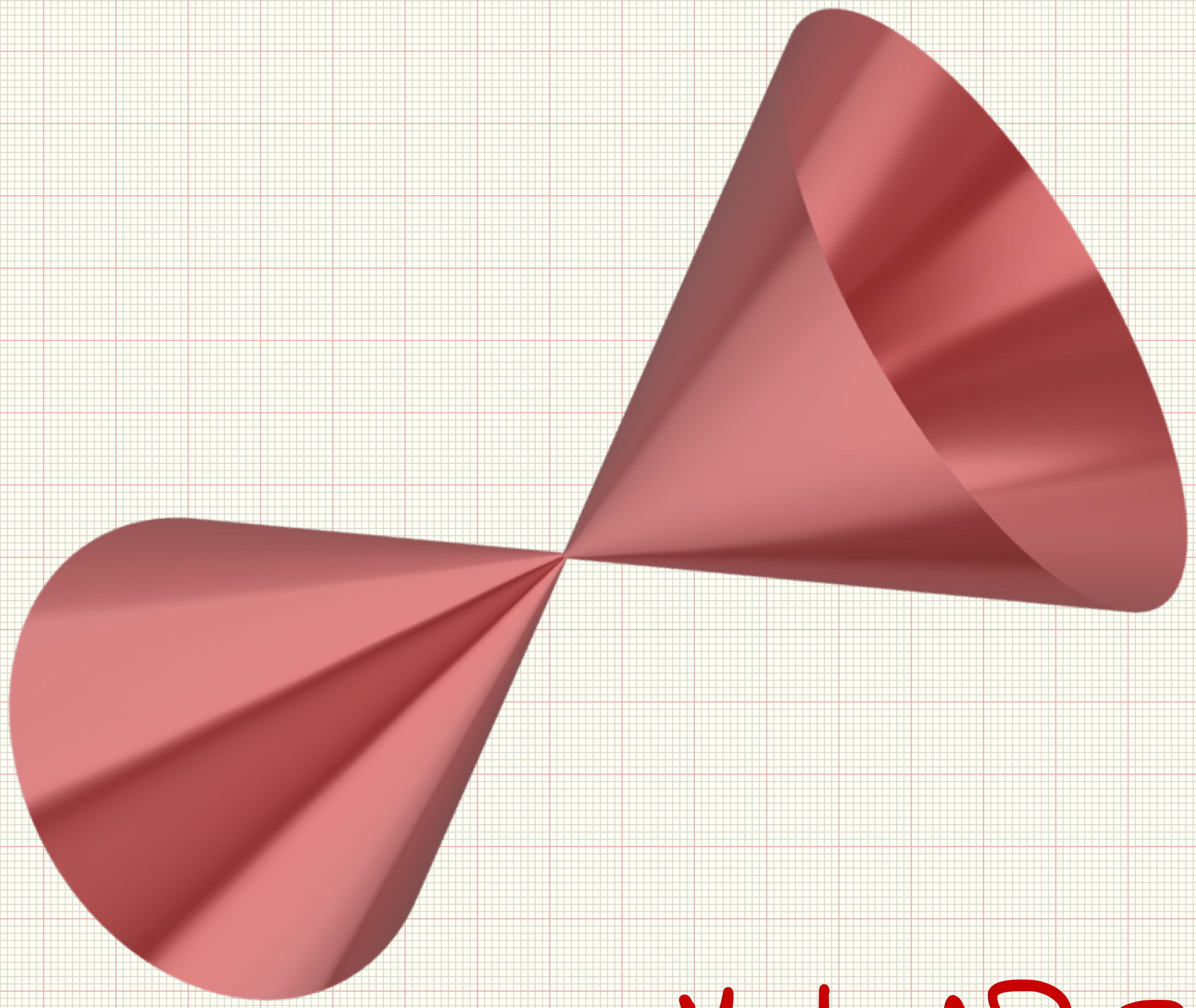


Polar Duality in Three Liftings

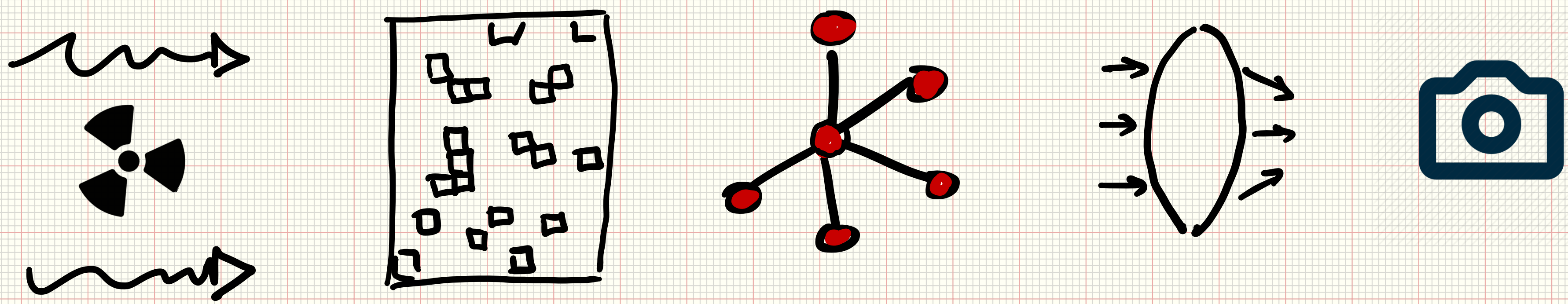
Michael P. Friedlander
University of British Columbia

Banff, January 29, 2018

1^{ST} LIFT:

QUADRATIC TO LINEAR

X-RAY CRYSTALLOGRAPHY



x-rays   mask   molecule   lens   detector

x — measurement #1 → $|\langle a_1, x\rangle|^2 =: b_1$

x — measurement #2 → $|\langle a_2, x\rangle|^2 =: b_2$

$\vdots$

x — measurement #m → $|\langle a_m, x\rangle|^2 =: b_m$

- vecs a_i encode mask config.
- magnitude-only measurements

PHASE RETRIEVAL PROBLEM

recover x from measurements (a_i, b_i)

Convex Relaxation

Formulations of phase retrieval (PR).

(quadratic) Find vector $x \in \mathbb{C}^n$ that solves $|\langle a_i, x \rangle|^2 = b_i$

$\downarrow$ LIFT

(Linear) Find matrix $X = xx^*$ that solves $\langle a_i a_i^*, X \rangle = b_i$

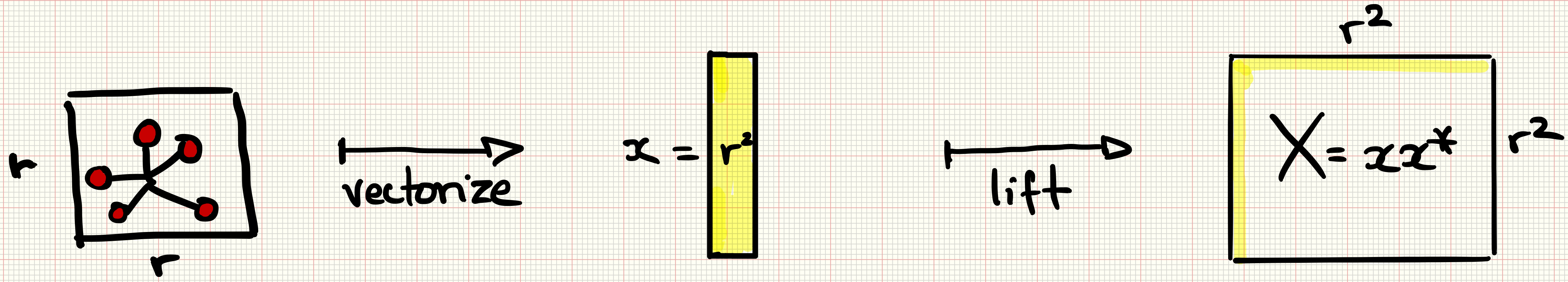
$\downarrow$ RELAX to semidefinite program (SDP)

$$\underset{X \in \mathcal{H}^{n \times n}}{\text{minimize}} \quad \sum_{i=1}^{n} \lambda_i(X) \quad \text{subj to} \quad \langle a_i a_i^*, X \rangle = b_i, \quad X \succcurlyeq 0$$

GOOD. Convex relaxation is exact w.h.p. if $m = O(n)$

[Candès, Strohmer, Voroninski '13]

BAD. Resulting problem is unmanageably large



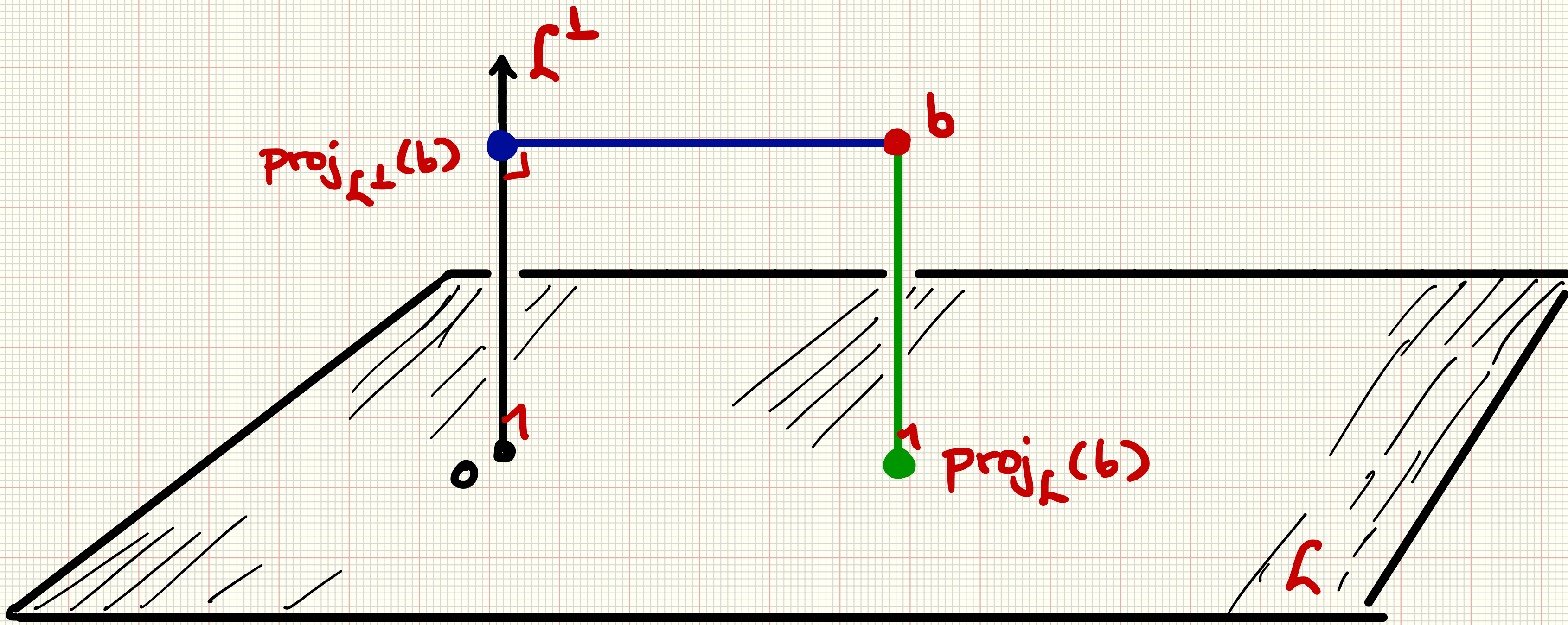
WORSE. SDP solvers typically require $O(r^4)$ storage

$O(r^6)$ work/itn

2ND LIFT:

POLARITY & CONES

Orthogonal Subspaces



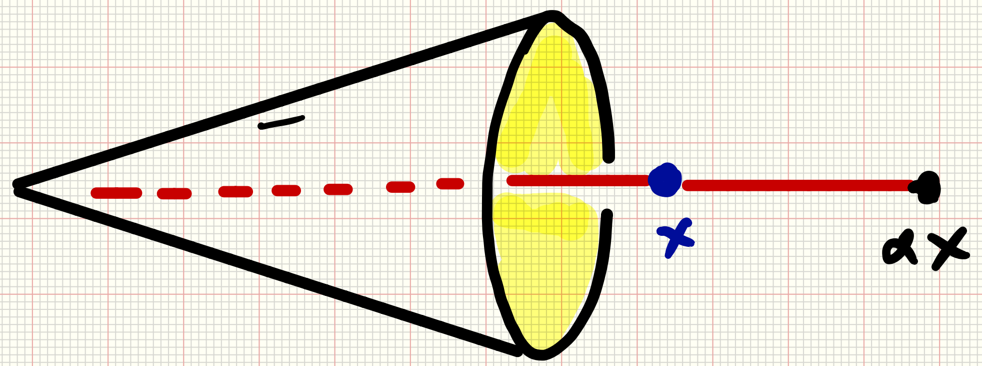
$$\mathrm{proj}_{\mathcal{L}}(b) + \mathrm{proj}_{\mathcal{L}^\perp}(b) = b$$

$$\mathrm{dist}_{\mathcal{L}}^2(b) + \mathrm{dist}_{\mathcal{L}^\perp}^2(b) = \|b\|_2^2$$

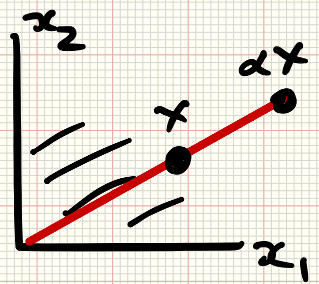
$$\mathrm{dist}_{\mathcal{C}}^2(b) = \inf_{z \in \mathcal{C}} \|z-b\|_2^2$$

Convex Cones

A set $K \subseteq \mathbb{R}^n$ is a <u>cone</u> if $x \in K \implies \alpha x \in K \quad \forall \alpha > 0$



Examples:



$\mathbb{R}^n_+ = \{$ real n-vectors w/ non-neg entries $\}$

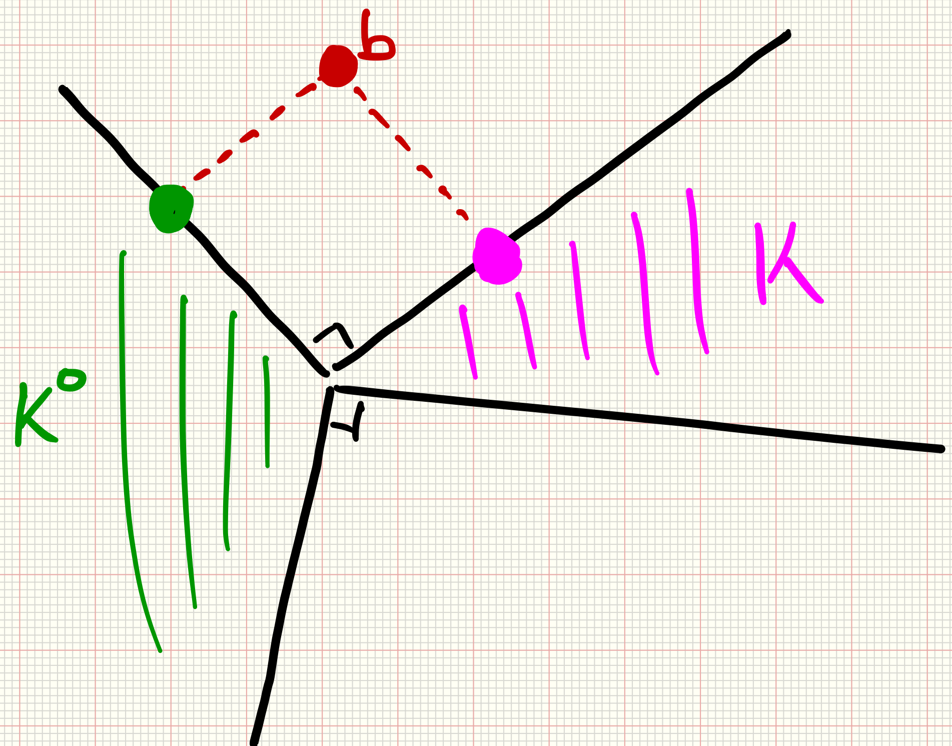
$\mathbb{S}^n_+ = \{$ Hermitian matrices w/ non-neg e-vals $\}$

$\mathbb{P}^n_+ = \{$ non-negative polynomials (even degree) $\}$

Polarity Generalizes Orthogonality

Cone $\quad K = \{ \text{convex cone} \}$ $\qquad K =$ 

Polar $\quad K^\circ = \{ z \mid \langle x, z \rangle \le 0 \quad \forall\, x \in K \}$ $\qquad K^\circ =$ 

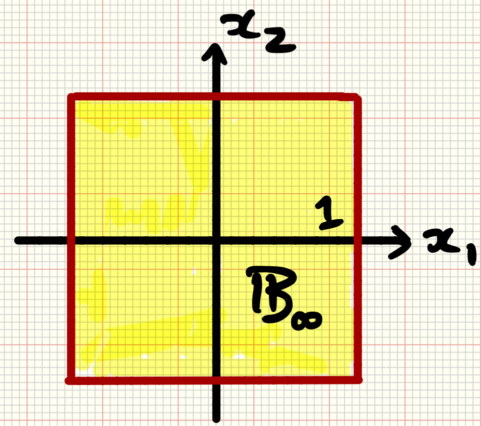


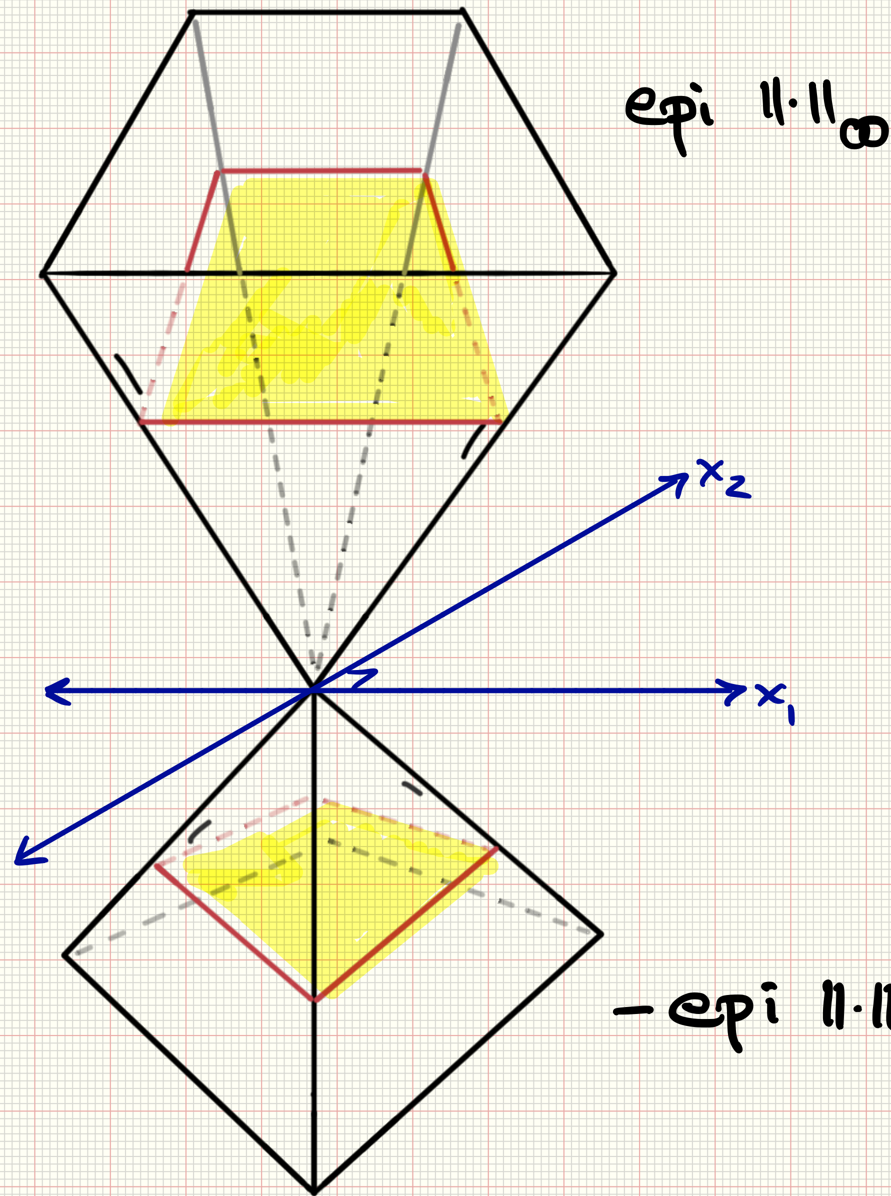
$$\mathrm{proj}_K(b) + \mathrm{proj}_{K^\circ}(b) = b$$

$$\mathrm{dist}_K^2(b) + \mathrm{dist}_{K^\circ}^2(b) = \|b\|^2$$

Polarity of Norms [Example w/ 1- and ∞-norms]

$$\|x\|_\infty = \max_j |x_j|$$

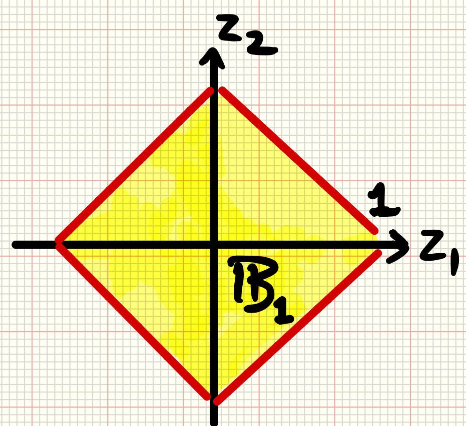




polar or "dual" norms

$$\mathbb{B}_\infty = (\mathbb{B}_1)^\circ$$

$$\|z\|_1 = \sum_j |z_j|$$

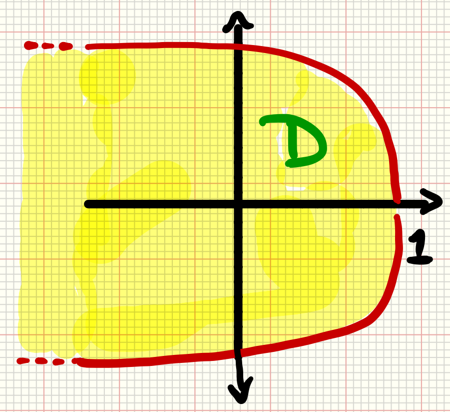


Hölder Inequality: $\langle x, z \rangle \leq \|x\|_p \cdot \|z\|_q$ w/ $\frac{1}{p} + \frac{1}{q} = 1$

POLARITY OF "NORM-LIKE" FUNCTIONS

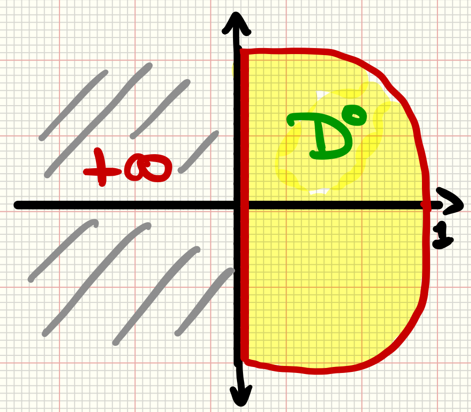
Take any cvx set D. Defines a gauge function.

$$\phi(x) := \inf\{\lambda > 0 \mid x \in \lambda D\}$$



Polar gauge functions

$$\phi^\circ(z) := \inf\{\lambda > 0 \mid z \in \lambda D^\circ\}$$



Generalized Hölder Inequality:

$$\langle x, z\rangle \le \phi(x)\cdot\phi^\circ(z)$$

Example: Linear Programming

$$\begin{array}{ll} \min & \langle c, x \rangle \\ \text{st} & Ax = b \\ & x \geq 0 \end{array} \iff \begin{array}{ll} \min & \phi(x) \\ \text{st} & Ax = b \end{array}$$

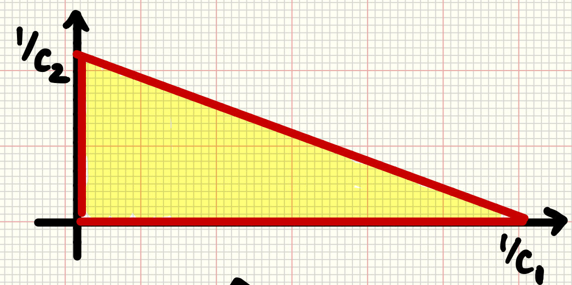
$$\phi(x) := \langle c, x \rangle + \delta_{\geq 0}(x)$$

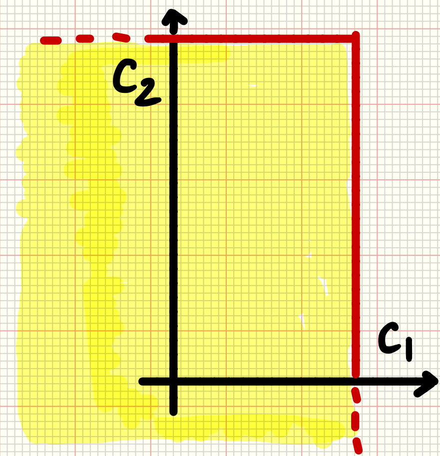
$$\delta_{\geq 0}(x) := \begin{cases} 0 & \text{if } x \geq 0 \\ +\infty & \text{else} \end{cases}$$

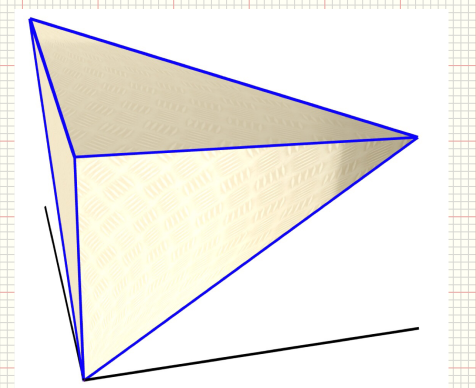
[For simplicity, assume vector $c > 0$. WLOG req's $c \geq 0$]

$$\phi(x) := \langle c, x \rangle + \delta_{\geq 0}(x)$$

$$\phi^{o}(z) = \max \left\{ 0, z_1/c_1, \dots, z_n/c_n \right\}$$







GAUGE OPTIMIZATION

Gauge Optimization

For any gauge function $\phi : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ and cvx set C:

primal $\quad \underset{x}{\text{minimize}} \ \phi(x) \quad \text{st} \quad x \in C$

polar dual $\quad \underset{z}{\text{minimize}} \ \phi^\circ(z) \quad \text{st} \quad z \in C'$

where C' is the <u>anti</u> <u>polar</u> of C:

$$C' = \{ z \mid \langle x, z \rangle \geq 1 \ \text{ for all } x \in C \}$$

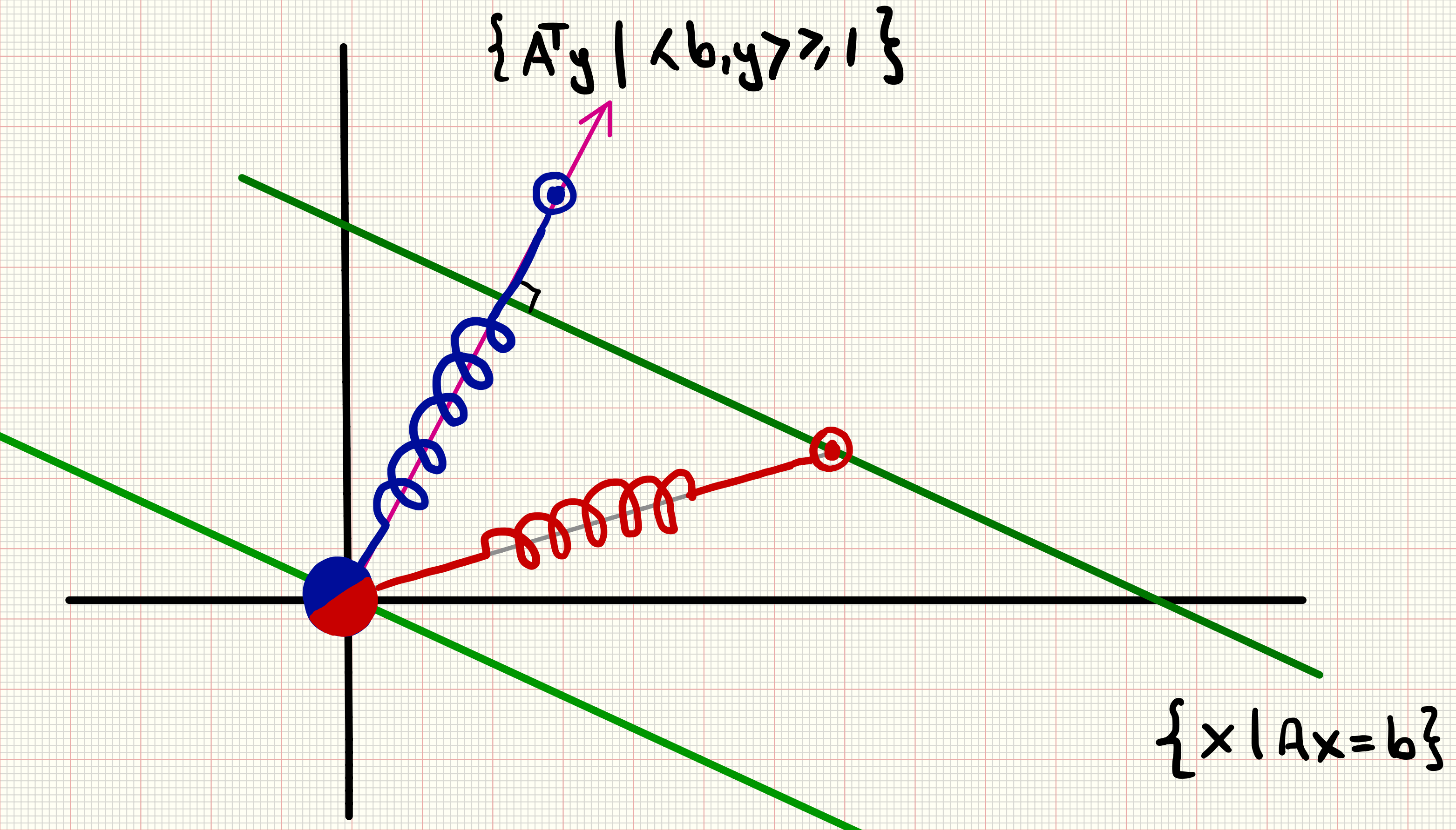
weak duality: $1 \leq \langle x, z \rangle \leq \phi(x) \cdot \phi^\circ(z)$ for <u>all</u> $x \in C, z \in C'$

strong duality: $1 = \langle x, z \rangle = \phi(x) \cdot \phi^\circ(z) \iff x, z$ primal/dual optimal

Canonical Example

min-norm $\quad \min_x \ \|x\| \quad$ subj to $\quad Ax = b \quad (A = \boxed{})$

polar dual $\quad \min_y \ \|A^T y\|_* \quad$ subj to $\quad \langle b, y \rangle \geq 1$



x, y optimal

$\Updownarrow$

$1 = \|x\| \cdot \|A^T y\|_*$

Spectral Examples

<u>Nuclear Norm</u> $\quad \min \ \|X\|_1 = \sum \sigma_i(X) \quad \text{st} \quad \mathcal{A}X = b$

polar dual $\quad \min \ \|\mathcal{A}^* y\|_\infty = \sigma_{max}(\mathcal{A}^* y) \quad \text{st} \quad \langle b, y \rangle = 1$

<u>Semi Def Prog</u> $\quad \min\limits_{X} \ \text{trace}(CX) \quad \text{st} \quad \mathcal{A}X = b, \ X \succeq 0 \quad (C \succ 0)$

polar dual $\quad \min \ \lambda_1(\mathcal{A}^* y, C) \quad \text{st} \quad \langle b, y \rangle = 1$

[Arbitrary convex constraints also possible]

NUMERICS:

BACK TO PHASE RETRIEVAL

Algorithm Summary

For Phase Retrieval, obtain rank-1 solution $X = xx^*$ via

$$\min_{X \in \mathcal{H}^{n\times n}} \operatorname{trace}(X) \quad \text{subj to} \quad \mathcal{A}X = b,\ X \succeq 0$$

Approach:

① apply first-order method to polar dual $y_k \to y_*$:

$$\min_{y \in \mathbb{R}^n} \lambda_1(\mathcal{A}^* y) \quad \text{st} \quad \langle b, y \rangle = 1$$

② each iteration uses Lanczos for rightmost eigenpairs

$$\mathcal{A}^* y_k = \lambda_1 U_1 U_1^* + U_{>1} \Lambda_{>1} U_{>1}^*$$

③ primal solutions shares simultaneously ordered e-val decomp:

$$X^* = U_1 S U_1^* \qquad S \succeq 0 \text{ via small SDP}$$

Hölder: $\langle x, \mathcal{A}^* y \rangle \leq \phi(x) \cdot \phi^\circ(\mathcal{A}^* y)$

1-D synthetic signal: $n = 128$

	GAUGE		TFOCS		
L	nDFT	xErr	nDFT	xErr	%
12	18,330	$2{\cdot}10^{-6}$	2,341,800	$4{\cdot}10^{-3}$	100
11	19,256	$2{\cdot}10^{-6}$	2,427,546	$4{\cdot}10^{-3}$	100
10	19,045	$1{\cdot}10^{-6}$	2,857,650	$5{\cdot}10^{-3}$	100
9	21,933	$2{\cdot}10^{-6}$	11,752,236	$7{\cdot}10^{-3}$	89
8	23,144	$2{\cdot}10^{-6}$	10,975,744	$1{\cdot}10^{-2}$	22
7	25,781	$2{\cdot}10^{-6}$	6,853,245	$2{\cdot}10^{-2}$	0
6	34,689	$3{\cdot}10^{-6}$	2,664,126	$6{\cdot}10^{-2}$	0

L = # of illuminations ≥ 4 (information theoretic bnd)

Gauge: First-order method on polar dual problem

TFOCS: First-order method for SDP. (Becker, Bobin, Candes '11)

2-D image: 1800×1350 (7.3 MB)

	GAUGE		WFLOW	
L	nDFT	xErr	nDFT	xErr
15	5,700	$2 \cdot 10^{-6}$	8,100	$4 \cdot 10^{-6}$
10	12,280	$9 \cdot 10^{-7}$	12,340	$2 \cdot 10^{-5}$

Lifted image $X = xx^*$

$\approx$ 50 TB

WFLOW $\equiv$

Non-convex gradient descent

(Candès, Li, Soltanolkotabi '15)

3rd LIFT:

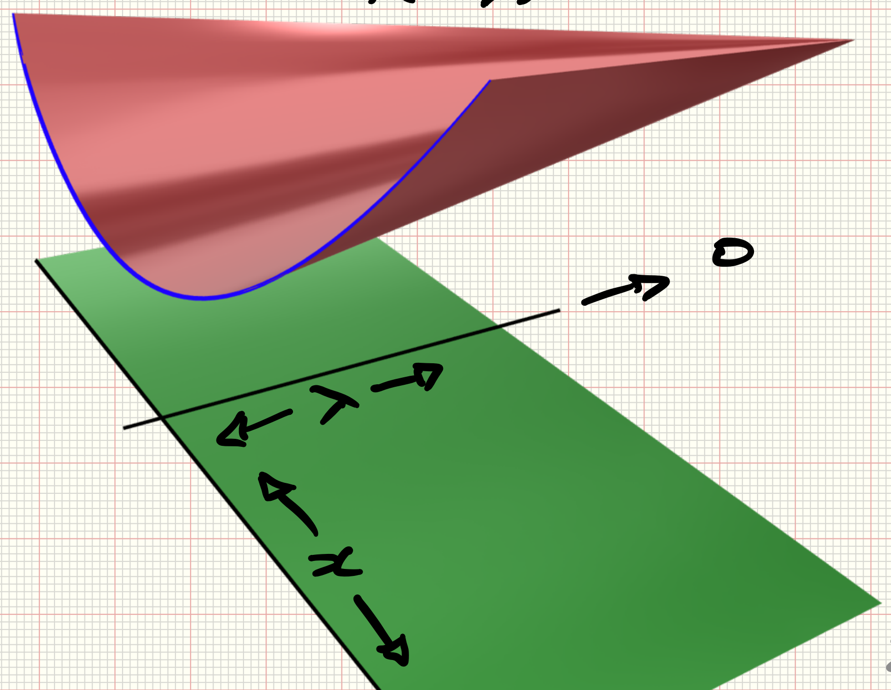
PERSPECTIVE

Extension to all Convex Optimization

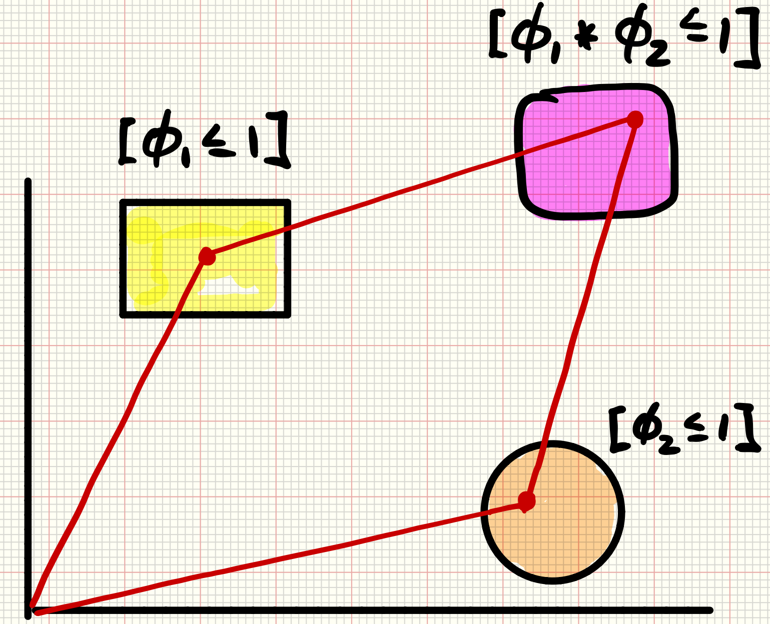
Perspective of any convex function $f: \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$:

$$f^{\pi}(x, \lambda) = \lambda f(x/\lambda) \quad (\lambda > 0) \quad \Leftarrow \text{"norm-like" in } (x, \alpha)$$

$$\boxed{\min_{x \in C} f(x) \iff \min_{\substack{x \in C \\ \lambda = 1}} f^{\pi}(x, \lambda)}$$



Lots To Do!



Polar Convolution:

smooth approximations by "mixing" level sets

Polar primal-dual algo's:

specialized algo's for polar dual pairs.

Software:

- modeling language for building gauge problems (cf. CVX)
- general-purpose solvers

REFERENCES

(0) R. Freund, "Dual gauge programs, with applications to quadratic programming and the minimum-norm problem", Math Prog, 1987

(1) Friedlander, Macêdo, Pong, "Gauge optimization and duality", SIOPT, 2014

(2) Macêdo, "Gauge duality & low-rank spectral optimization", UBC PhD Thesis, 2015

(3) Friedlander, Macêdo, "Low-rank spectral optimization", SIMAX, 2016

(4) Aravkin, Burke, F., "Foundations of gauge & perspetive duality", arXiv, 2017

(5) Friedlander and Pong, "Polar convolution", in preparation