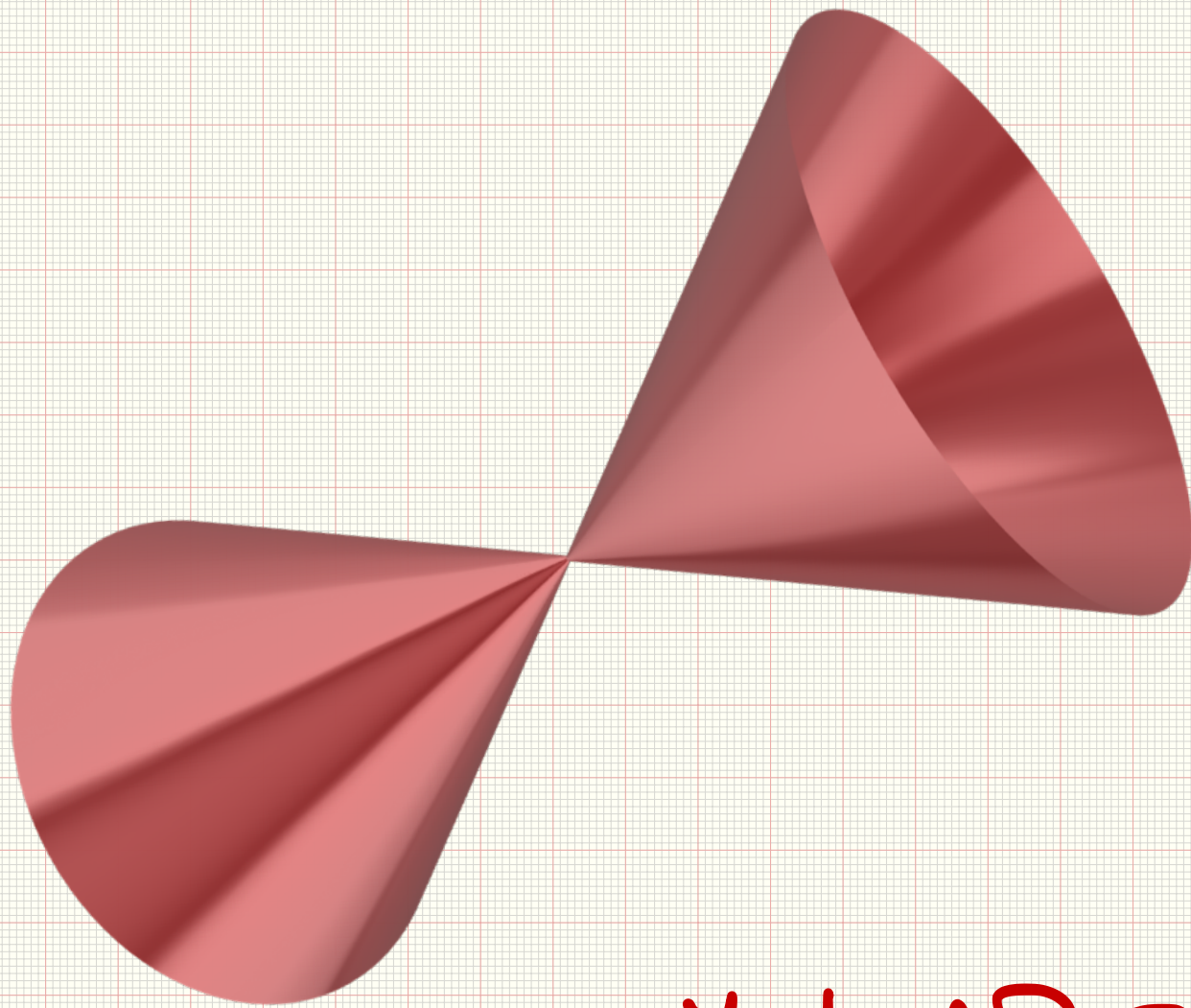




POLAR DUALITY IN THREE LIFTINGS

Michael P. Friedlander

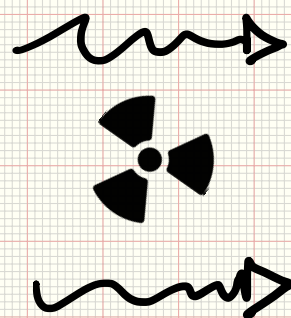
University of British Columbia

NOVEMBER 28, 2017

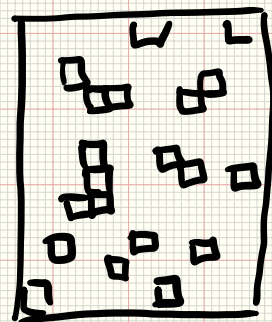
1ST LIFT:

QUADRATIC TO LINEAR

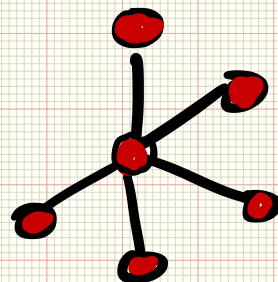
X-RAY CRYSTALLOGRAPHY



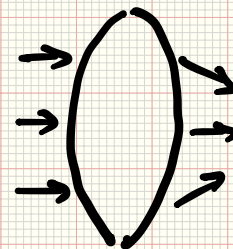
x-rays



mask



molecule



lens



detector

X

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

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

$\vdots$

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

- 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 $|\langle a_i, x \rangle|^2 = b_i$

↓ LIFT

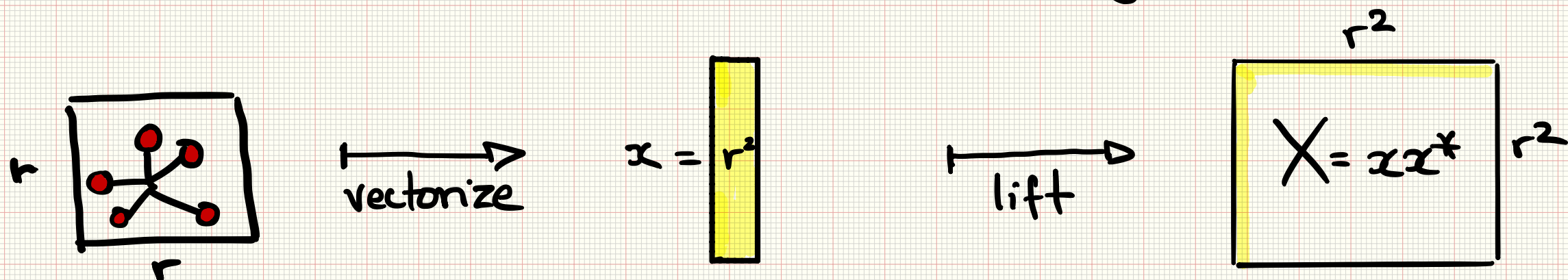
(Linear~) Find matrix $X = xx^*$ that solves $\langle a_i a_i^*, X \rangle = b_i$

↓ RELAX to semidefinite program (SDP)

minimize $\sum_{i=1}^m \lambda_i(X)$ subj to $\langle a_i a_i^*, X \rangle = b_i$, $X \succeq 0$
 $X \in \mathbb{H}^{n \times n}$

GOOD. Convex relaxation is **exact** w.h.p. if $m = O(n)$
[Candès, Strohmer, Vershyninski '13]

BAD. Resulting problem is unmanageably large

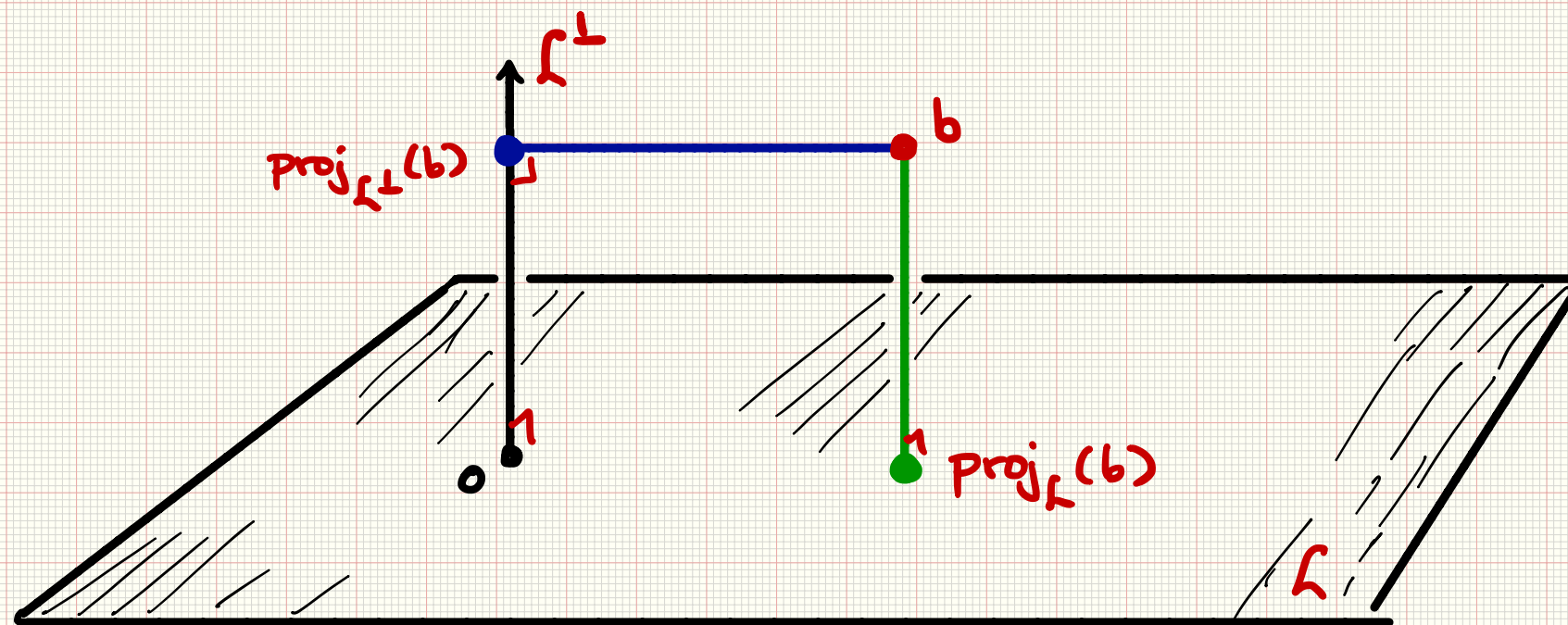


WORSE. SDP solvers typically require $O(r^4)$ storage
 $O(r^6)$ work/itn

2ND LIFT:

POLARITY ∇ CONES

ORTHOGONAL SUBSPACES



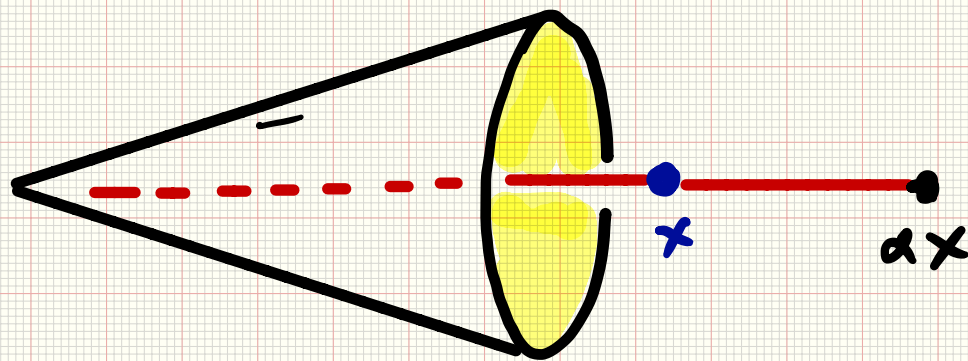
$$\text{proj}_L(b) + \text{proj}_{L^\perp}(b) = b$$

$$\underline{\text{dist}_L^2(b)} + \underline{\text{dist}_{L^\perp}^2(b)} = \underline{\|b\|_2^2}$$

$$\text{dist}_C^2(b) = \inf_{z \in C} \|z - b\|_2^2$$

CONVEX CONES

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

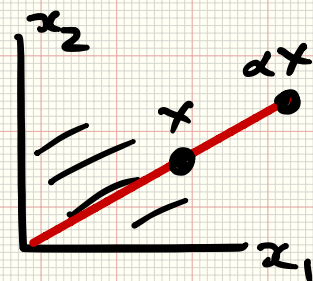


EXAMPLES:

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

$$\mathcal{S}_+^n = \{ \text{Hermitian matrices w/ non-neg e-vals} \}$$

$$\mathcal{P}_+^n = \{ \text{non-negative polynomials (even degree)} \}$$



POLARITY GENERALIZES ORTHOGONALITY

CONE

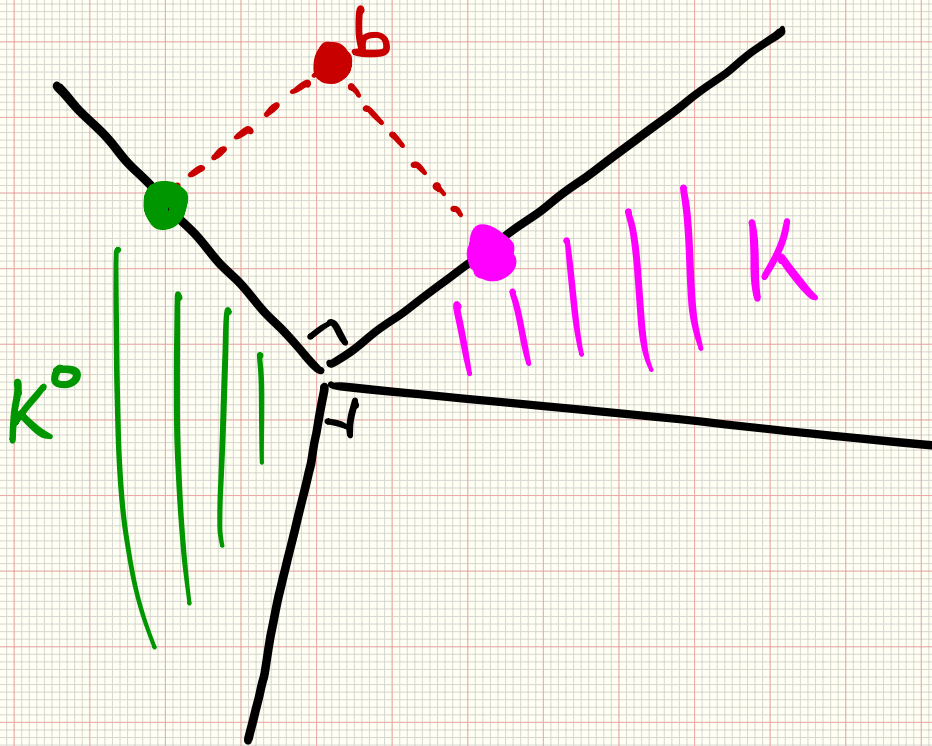
$$K = \{\text{convex cone}\}$$

$$K = \triangle$$

POLAR

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

$$K^\circ = \triangleright$$



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

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

POLARITY OF NORMS

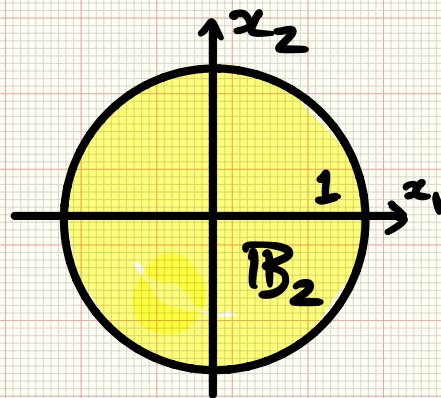
Extend definition of polarity to norms. Begin w/ 2-norm.

$$\| \cdot \|_2 = \sqrt{\sum (\cdot)^2}$$

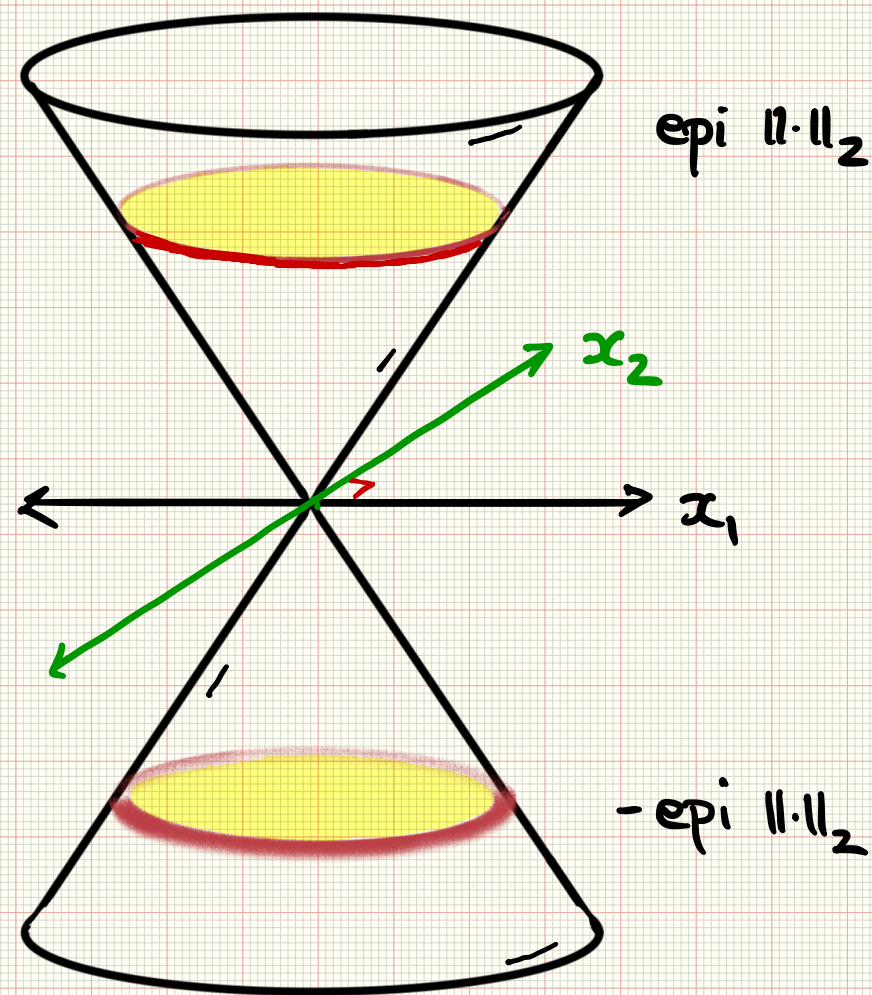
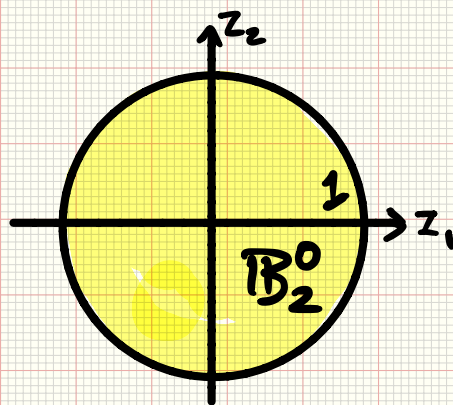
↑
self polar
(self dual)

↓

$$(\| \cdot \|_2)^{\circ} = \| \cdot \|_2$$



$$B_2 = B_2^{\circ}$$



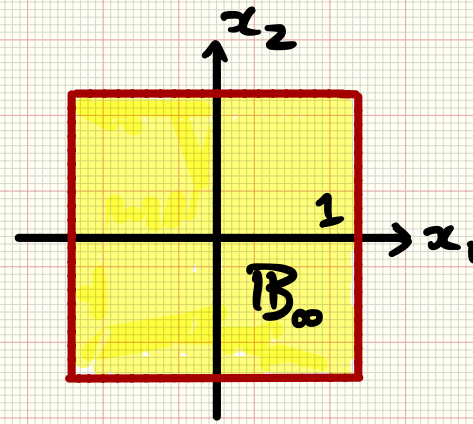
Cauchy-Schwarz Inequality: $\langle x, z \rangle \leq \|x\|_2 \cdot \|z\|_2$

SECOND EXAMPLE: 1- AND ∞ -NORMS

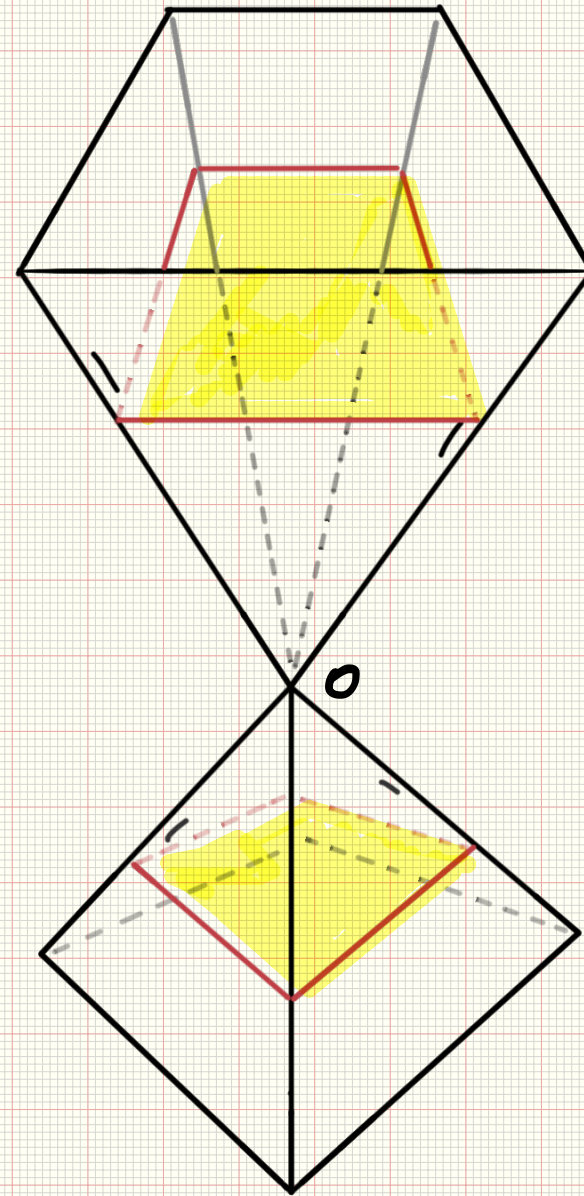
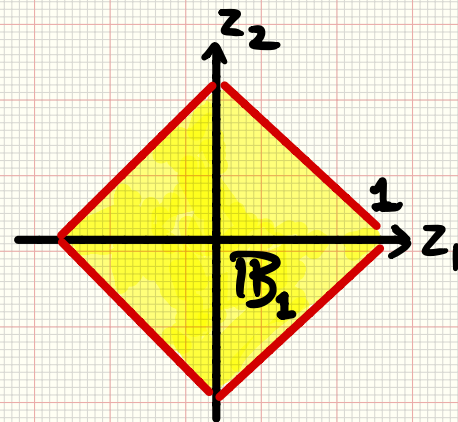
$$\|x\|_{\infty} = \max_j |x_j|$$

↑
polar or "dual"
norms

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



$$B_{\infty} = (B_1)^{\circ}$$



$\text{epi } \|\cdot\|_{\infty}$

$-\text{epi } \|\cdot\|_1$

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

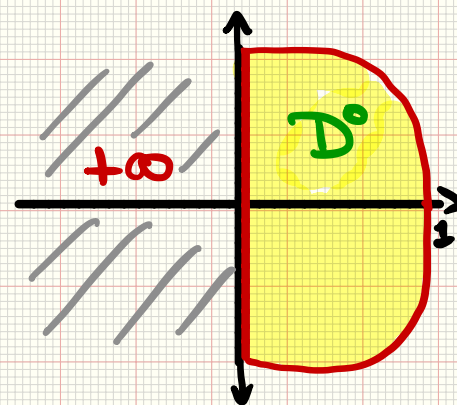
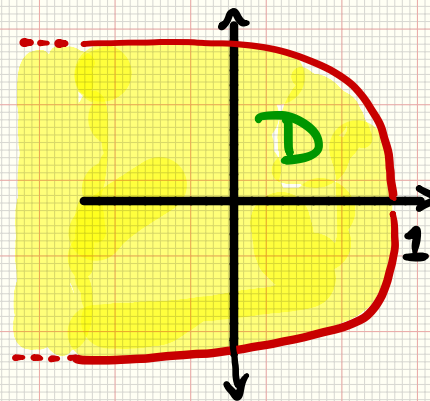
POLARITY OF "NORM-LIKE" FUNCTIONS

Take any cvx set D . Defines a gauge function.

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

Polar gauge functions

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

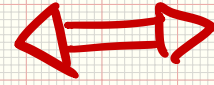


Generalized Hölder inequality:

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

EXAMPLE: LINEAR PROGRAMMING

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



$$\begin{array}{ll} \min & \phi(x) \\ \text{st} & Ax = b \end{array}$$

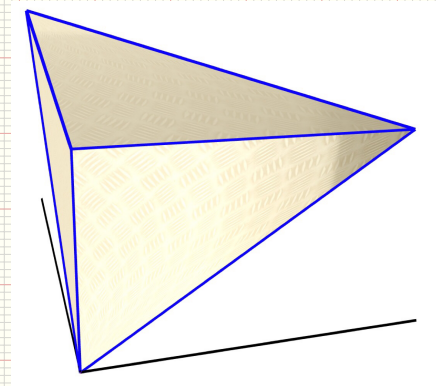
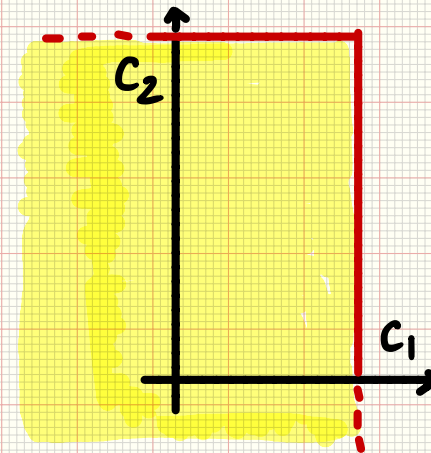
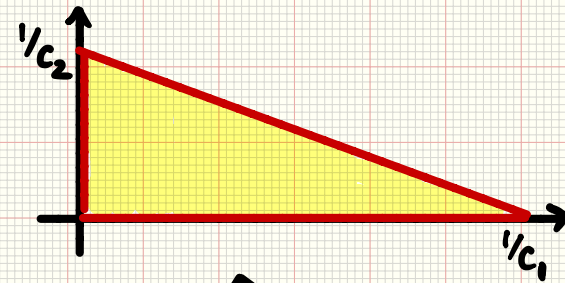
$$\phi(x) := \langle c, x \rangle + \delta_{\geq 0}(x)$$

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

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

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

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



GAUGE

OPTIMIZATION

GAUGE OPTIMIZATION

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

primal

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

polar dual

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

where C' is the anti polar of C :

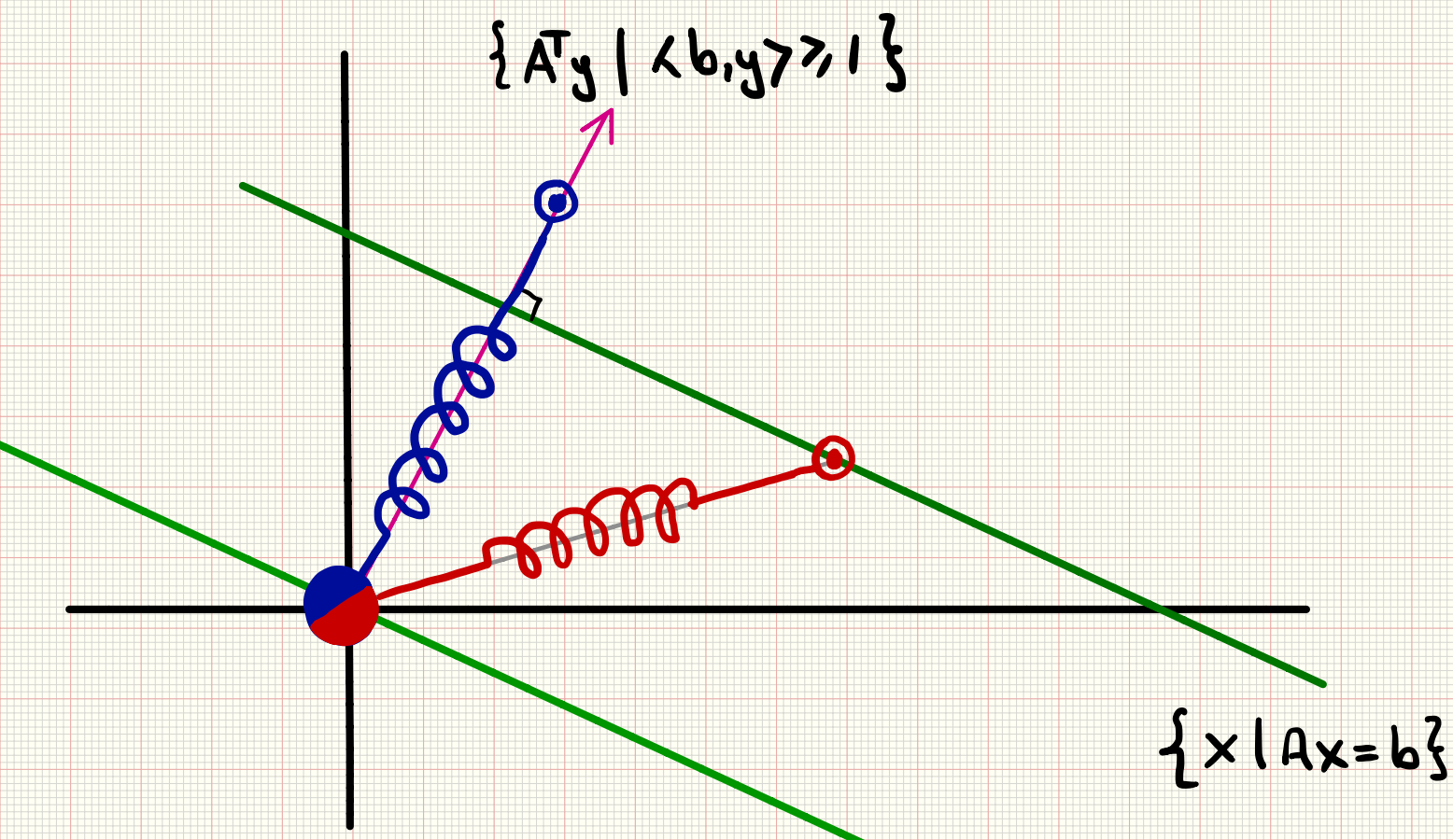
$$C' = \{z \mid \langle x, z \rangle \geq 1 \text{ for all } x \in C\}$$

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

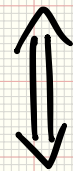
strong duality: $1 = \langle x, z \rangle = \phi(x) \cdot \phi^\circ(z) \iff x, z$ primal/dual
optimal

CANONICAL EXAMPLE

$$\begin{array}{ll}
 \text{min-norm} & \min_x \|x\| \quad \text{subj to } Ax = b \quad (A = \boxed{}) \\
 \text{polar dual} & \min_y \|A^T y\|_* \quad \text{subj to } \langle b, y \rangle \geq 1
 \end{array}$$



x, y optimal



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

SPECTRAL EXAMPLES

Nuclear Norm

$$\min \|X\|_1 = \sum \sigma_i(X) \quad \text{st} \quad AX = b$$

polar dual

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

SemiDef Prog

$$\min_X \text{trace}(CX) \quad \text{st} \quad AX = b, \quad X \succeq 0 \quad (C \succeq 0)$$

polar dual

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

[Arbitrary convex constraints also possible]

NUMERICS:
BACK TO PHASE RETRIEVAL

ALGORITHM SUMMARY

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

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

Approach:

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

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

② each iteration uses Lanczos for rightmost eigenpairs

$$A^*y_k = \lambda_1 U_1 U_1^* + U_2 \Lambda_2 U_2^*$$

③ primal solution shares simultaneously ordered e-val decomp:

$$X^* = U_1 S U_1^*$$

$S \succeq 0$ via small SDP

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

1-D synthetic signal: $n = 128$

L	GAUGE		TFOCS		%
	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

$L = \# \text{ of illuminations} \geq 4$ (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)

L	GAUGE		WFLOW	
	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 $X = xx^*$

≈ 50 TB

WFLOW $\equiv$

Non-convex gradient descent
(Candès, Li, Soltanolkotabi '15)

3rd LIFT :

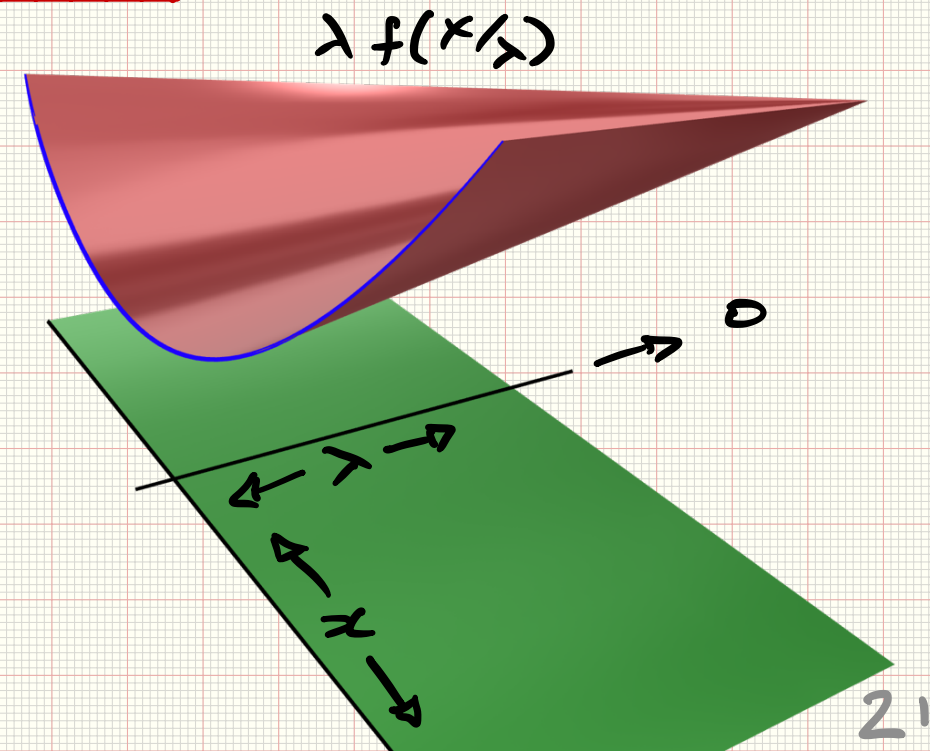
PERSPECTIVE

EXTENSION TO ALL CONVEX OPTIMIZATION

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

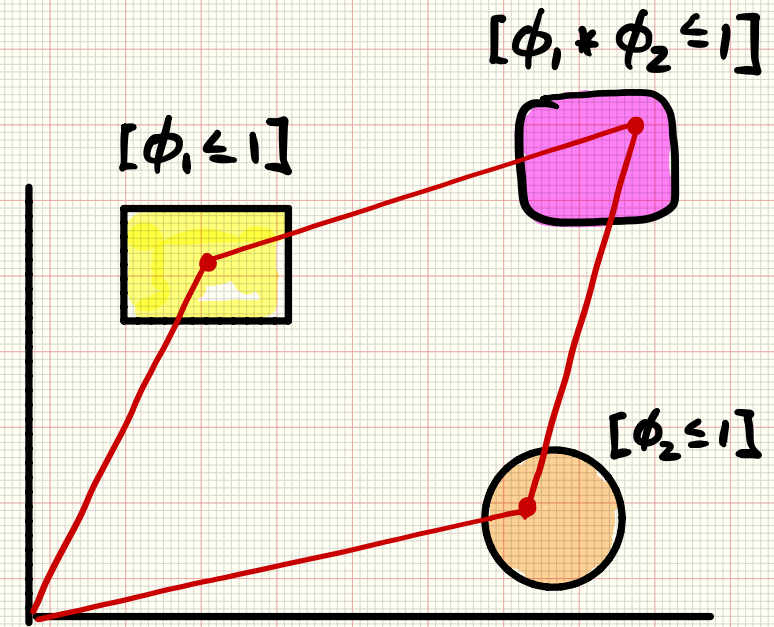
$$f^\pi(x, \lambda) = \lambda f(x/\lambda) \quad (\lambda > 0) \quad \Leftarrow \text{"norm-like" in } (x, \lambda)$$

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



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", SIOP, 2014
- (2) Macêdo, "Gauge duality $\hat{=}$ low-rank spectral optimization", UBC PhD Thesis, 2015
- (3) Friedlander, Macêdo, "Low-rank spectral optimization", SIMAX, 2016
- (4) Aravkin, Burke, F., "Foundations of gauge $\hat{=}$ perspective duality", arXiv, 2017
- (5) Friedlander and Pong, "Polar convolution", in preparation