Piecewise linear operators

Rima Alaifari, Daniel Freeman*, Dorsa Ghoreishi,
Mitchell Taylor, and Pedro Tradacete

Department of Mathematics and Statistics, Saint Louis University, USA

ICMAT: Workshop on Banach spaces and Banach lattices

21 May, 2024

Supported by NSF grant DMS-2154931

Philosophy of linear analysis of non-linear objects

1. There are many powerful tools for working with Banach spaces and Banach lattices. :)

2. Not everything is nice and linear. :(

3. Given some non-linear object X (such as a manifold or metric space), we can often construct a map $T : X \to Z$ which embeds X into a Banach lattice Z. This then allows us to apply our linear tools to study X!

Outline

1. Discuss a simple but very useful linear method for the reconstruction of vectors in a Hilbert space H. This involves inverting a linear operator.
2. Introduce a situation where we instead need to reconstruct elements of a non-linear quotient $H/\sim$ of H.
3. This now involves inverting a non-linear operator, and we will discuss a new way of thinking about such operators.
4. Go off on a tangent about non-linear quotients of Banach spaces.

Frames for Hilbert spaces

A collection of vectors $(x_j)_{j\in J} \subseteq H$ is a *frame* of a Hilbert space H if there are universal constants $0 < A \leq B < \infty$ such that

$$A\|x\|^2 \leq \sum_{j\in J} |\langle x, x_j\rangle|^2 \leq B\|x\|^2 \qquad \text{for all } x \in H.$$

The *analysis operator* of $(x_j)_{j\in J}$ is the map $\Theta : H \to \ell_2(J)$ given by

$$\Theta(x) = (\langle x, x_j\rangle)_{j\in J} \qquad \text{for all } x \in H.$$

Note: $(x_j)_{j\in J}$ is a frame of $H \Leftrightarrow \Theta$ is an embedding of H into $\ell_2(J)$.

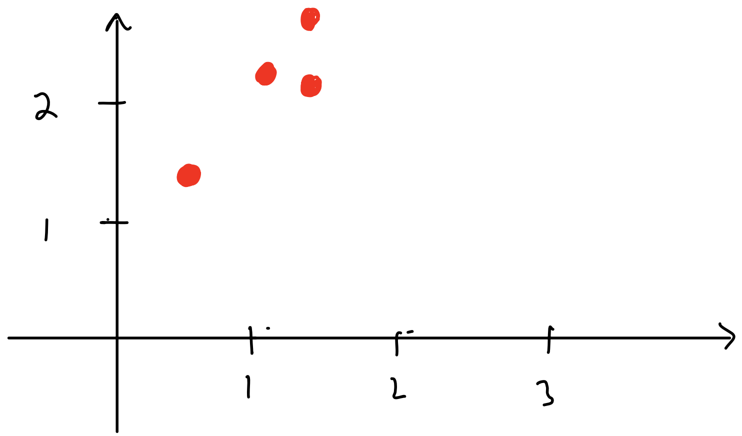
Vector reconstruction using a frame:

1. x is some object in a Hilbert space H that we are interested in.
2. The frame coefficients $\Theta(x) = (\langle x, x_j\rangle)_{j\in J}$ are a collection of rank-1 linear measurements of x.
3. We can stably reconstruct x from these measurements by applying the linear operator $(\Theta^*\Theta)^{-1}\Theta^*$ giving $x = (\Theta^*\Theta)^{-1}\Theta^*(\Theta x)$.

Equivalence classes

Let G be a finite group of isometries on H. We are interested in the vector $x \in H$, but we consider x to be equivalent to all the vectors $\{gx\}_{g \in G}$.

Example 1: x is a point cloud of data in $\mathbb{R}^n$.



We can express the point cloud as any of the following $4x2$ matrices:

$$\begin{bmatrix} .5 & 1.4 \\ 1.1 & 2.3 \\ 1.3 & 2.1 \\ 1.3 & 2.5 \end{bmatrix}, \begin{bmatrix} 1.1 & 2.3 \\ 1.3 & 2.1 \\ .5 & 1.4 \\ 1.3 & 2.5 \end{bmatrix}, \begin{bmatrix} 1.1 & 2.3 \\ 1.3 & 2.1 \\ .5 & 1.4 \\ 1.3 & 2.5 \end{bmatrix}, \dots$$

Thus, if M_{4x2} is the space of $4x2$ matrices and G is the group of all permutations of rows then a point cloud $x \in M_{4x2}$ is equivalent to gx for every $g \in G$.

Example 2: Phase retrieval

Let $(x_j)_{j \in J}$ be a frame of H with analysis operator $\Theta : H \to \ell_2(J)$.

The goal of phase retrieval is to recover a vector $x \in H$ from the magnitude of the frame coefficients,

$$|\Theta(x)| = (|\langle x, x_j \rangle|)_{j \in \Omega}.$$

Note: We cannot distinguish between x and λx for any scalar with $|\lambda| = 1$ as $|\Theta(x)| = |\Theta(\lambda x)|$.

Thus, we consider the equivalence relation $x \sim \lambda x$ for all scalars $|\lambda| = 1$. We say that $(x_j)_{j \in J}$ does phase retrieval if the map $|\Theta| : H/\sim \to \ell_2(J)$ is one-to-one.

That is, doing phase retrieval means that we can recovery any $[x]_\sim \in X/\sim$ from the magnitude of the frame coefficients $|\Theta(x)|$.

Nonlinear quotients of Hilbert spaces

Let G be a finite group of isometries on H. We consider the equivalence relation $x \sim gx$ for all $g \in G$.
The quotient space $H/\sim$ is the space of equivalence classes $H/\sim = \{[x]_\sim : x \in H\}$.
The quotient metric is defined as

$$d([x]_\sim, [y]_\sim) = \min_{g \in G} \|x - gy\|$$

To do analysis and reconstruct elements of $H/\sim$, we want to embed it back into a Hilbert space. That is, we need a map $T_G : H/\sim \to \ell_2(J)$.

For this to be stable in applications, we need T_G to be bi-Lipschitz.
That is, there exists $A, B > 0$ so that

$$A \min_{g \in G} \|x - gy\| \leq \|Tx - Ty\| \leq B \min_{g \in G} \|x - gy\| \qquad \text{for all } x, y \in H.$$

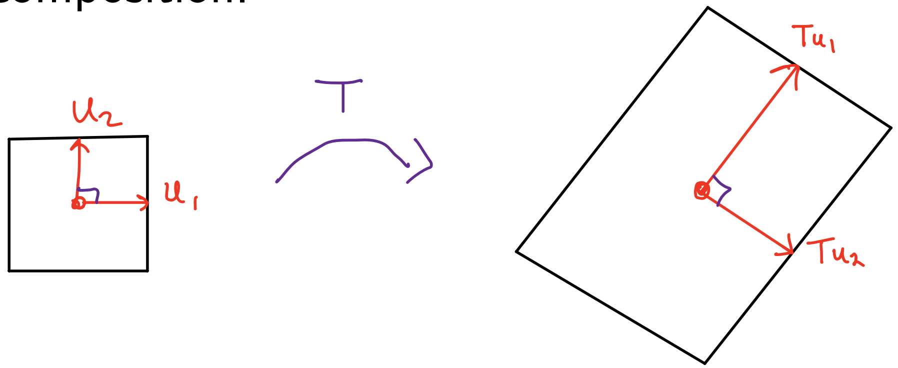
We can measure properties of the embedding $T_G : H/\sim \to \ell_2(J)$ by working with the lifting $T : H \to \ell_2(J)$ where $T(x) = T_G([x]_\sim)$.

How can we think about $T : H \to \ell_2(J)$?
What are some nice properties which we can exploit?
Note: $T : H \to \ell_2(J)$ will never be linear! (except for the trivial case $G = \{Id_H\}$.)

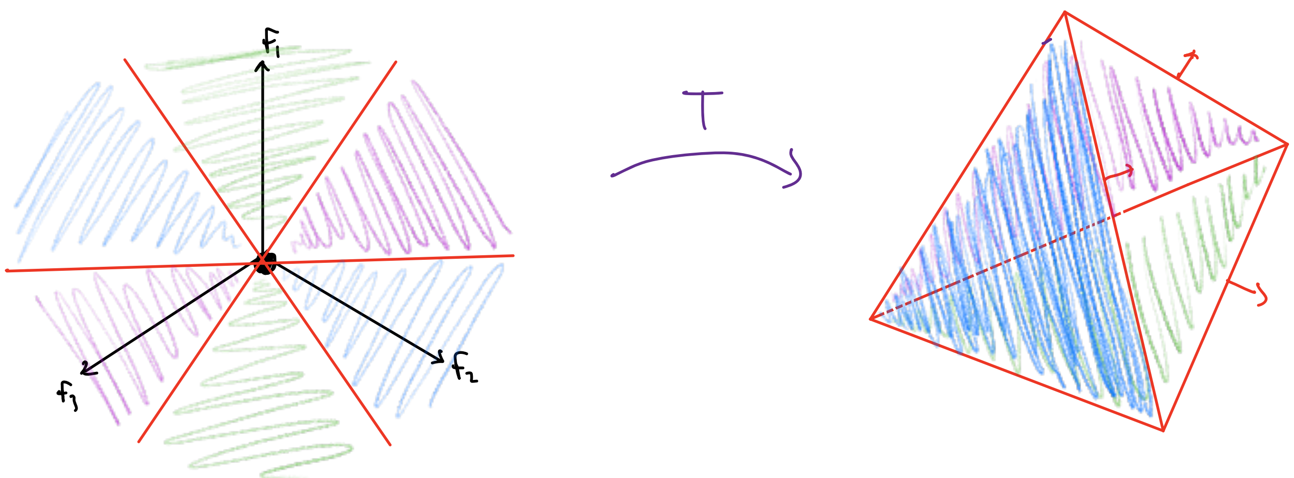
Visualizing linear maps vs Phase retrieval

Case 1: $T : \mathbb{R}^2 \to \mathbb{R}^3$ is linear.

We can visualize T by considering an ortho-normal basis u_1, u_2 corresponding to the singular value decomposition.



Case 2: $T : \mathbb{R}^2 \to \mathbb{R}^3$ is given by $T(x) = (|\langle x, f_1\rangle|, |\langle x, f_2\rangle|, |\langle x, f_3\rangle|)$ where $f_1 = (0, 1)$, $f_2 = (\sqrt{3}/2, -1/2)$, $f_3 = (-\sqrt{3}/2, -1/2)$.



Piecewise-linear operators

We say that a function $T : \mathbb{R}^d \to \mathbb{R}^n$ is a piecewise linear operator with linear decomposition $(X_j, T_j)_{j=1}^N$ if for each $1 \leq j \leq N$ we have that $X_j \subseteq \mathbb{R}^d$ is a finite intersection of closed half spaces and $T : \mathbb{R}^d \to \mathbb{R}^n$ is a linear operator such that

1. $\mathbb{R}^d = \bigcup_{j=1}^m X_j$,
2. $T|_{X_j} = T_j|_{X_j} \qquad \text{for all } 1 \leq j \leq N.$

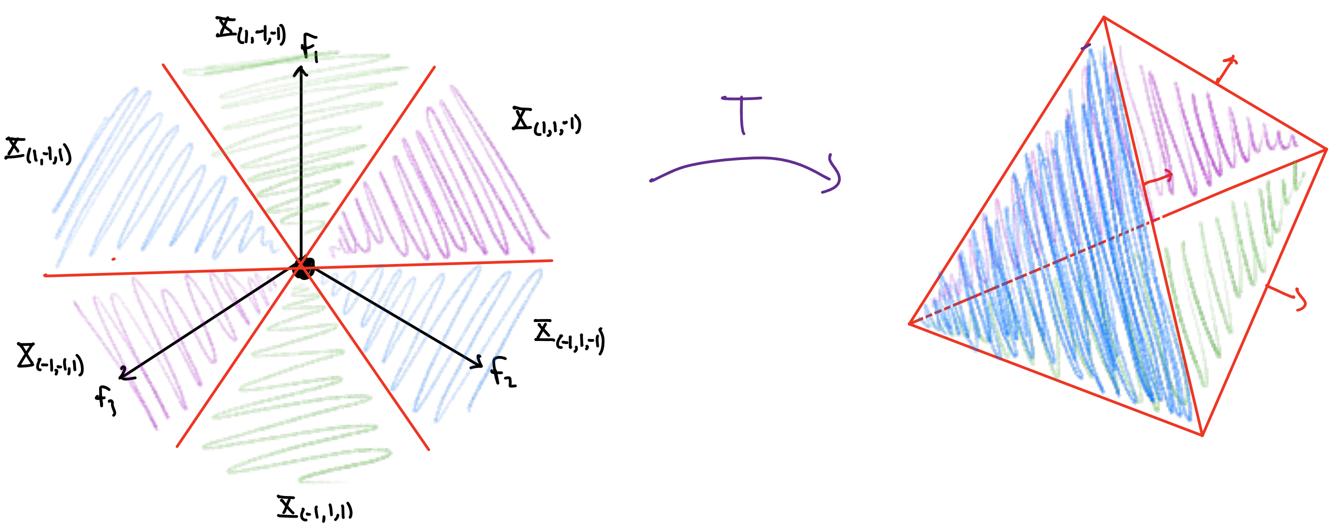
Let $(f_j)_{j=1}^N$ be a frame of $\mathbb{R}^d$.
$T : \mathbb{R}^d \to \mathbb{R}^N$, $T(x) = (|\langle x, f_j\rangle|)_{j=1}^N$.

For each choice of signs $\varepsilon_j = \pm 1$ for $1 \leq j \leq N$,

$$X_{(\varepsilon_j)_{j=1}^N} = \{x \in \mathbb{R}^d : \varepsilon_j \langle x, f_j\rangle \geq 0.\}$$

$T_{(\varepsilon_j)_{j=1}^N} : \mathbb{R}^d \to \mathbb{R}^N$ is the linear operator $T_{(\varepsilon_j)_{j=1}^N}(x) = \big(\langle x, \varepsilon_j f_j\rangle\big)_{j=1}^N$ for all $x \in \mathbb{R}^d$.

Case: $T : \mathbb{R}^2 \to \mathbb{R}^3$ is given by $T(x) = (|\langle x, f_1\rangle|, |\langle x, f_2\rangle|, |\langle x, f_3\rangle|)$
where $f_1 = (0, 1)$, $f_2 = (\sqrt{3}/2, -1/2)$, $f_3 = (-\sqrt{3}/2, -1/2)$.



Piecewise linear operators for max filtering

G is a finite group of isometries on $\mathbb{R}^d$. $(f_j)_{j=1}^N$ is a frame of $\mathbb{R}^d$. The map $T : \mathbb{R}^d \to \mathbb{R}^N$ is given by

$$T(x) = \big(\max_{g \in G} \langle x, gf_j \rangle \big)_{j=1}^N \qquad \text{for all } x \in \mathbb{R}^d.$$

(Real phase retrieval is max filtering for the case $G = \{Id, -Id\}$.)

For each choice of $(g_j)_{j=1}^N \in G^N$,

$$X_{(g_j)_{j=1}^N} = \{x \in \mathbb{R}^d : \langle x, g_j f_j \rangle = \max_{g \in G} \langle x, gf_j \rangle.\}$$

$T_{(g_j)_{j=1}^N} : \mathbb{R}^d \to \mathbb{R}^N$ is the linear operator $T_{(g_j)_{j=1}^N}(x) = \big(\langle x, g_j f_j \rangle \big)_{j=1}^N \qquad \text{for all } x \in \mathbb{R}^d.$

Summary

1. G is a finite group of isometries on $\mathbb{R}^d$.
2. $x \sim gx$ for all $g \in G$ gives an equivalence relation on $\mathbb{R}^d$.
3. We want a bi-Lipschitz embedding $T_G : \mathbb{R}^d / \sim \to \mathbb{R}^N$ which lifts to a piecewise linear operator $T : \mathbb{R}^d \to \mathbb{R}^N$.

It is well known for the case of phase retrieval that if $T_G : \mathbb{R}^d / \sim \to \mathbb{R}^N$ is one-to-one then T_G is bi-Lipschitz.

Question (J. Cahill, J. Iverson, D. Mixon, and D. Packer (2022))

Let G be a finite group of isometries on $\mathbb{R}^d$, and let $T_G : \mathbb{R}^d / \sim \to \mathbb{R}^N$ be the max filtering map. Does T_G being one-to-one imply that T_G is bi-Lipschitz?

Theorem (R. Balan and E. Tsoukanis (2023))

Let G be a finite group of isometries on $\mathbb{R}^d$, and let $T_G : \mathbb{R}^d / \sim \to \mathbb{R}^N$ be co-orbit embedding corresponding to some ordering (this is a generalization of max filtering). If T_G is one-to-one then T_G is bi-Lipschitz.

Theorem (R.Alaifari, F., D. Ghoreishi, M. Taylor, and P. Tradacete (2024))

Let G be a finite group of isometries on $\mathbb{R}^d$, and let $T_G : \mathbb{R}^d / \sim \to \mathbb{R}^N$ be a map which lifts to a piece-wise linear operator. If T_G is one-to-one then T_G is bi-Lipschitz.

Upper Lipschitz bound

Theorem (R.Alaifari, F., D. Ghoreishi, M. Taylor, and P. Tradacete (2024))

Let G be a finite group of isometries on $\mathbb{R}^d$ and let $T_G : \mathbb{R}^d/\!\sim \to \mathbb{R}^N$ be a one-to-one function which lifts to a piece-wise linear operator $T : \mathbb{R}^d \to \mathbb{R}^N$. Let $(X_j, T_j)_{j=1}^m$ be a minimal linear decomposition of T. Let $\beta = \max_{1\leq j\leq m} \|T_j\|$. Then,

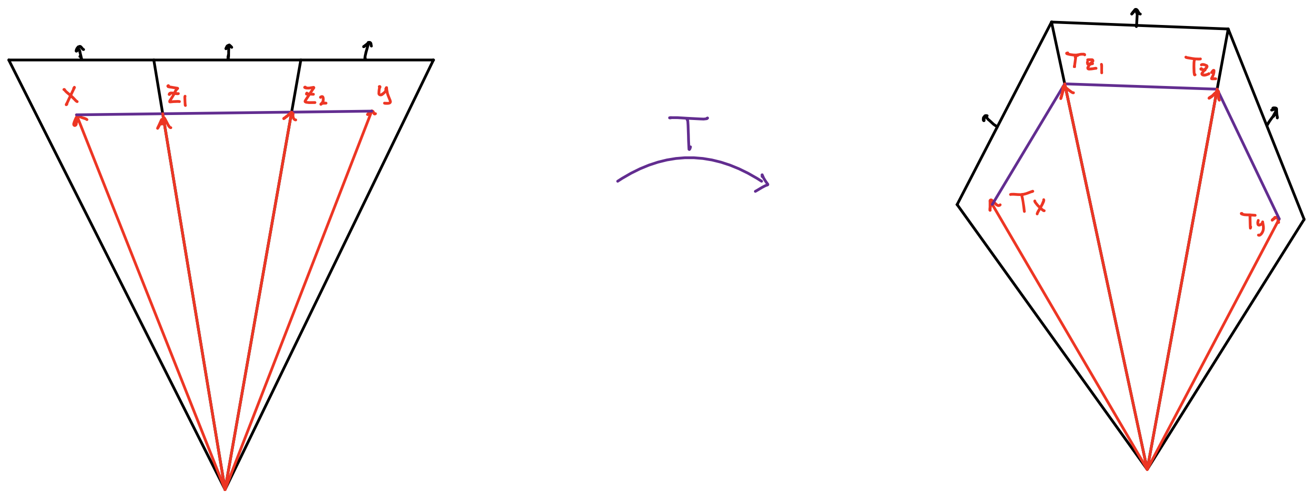
$$\|Tx - Ty\| \leq \beta \min_{g\in G} \|x - gy\| \text{ for all } x, y \in \mathbb{R}^d.$$

Furthermore, there exists $x, y \in \mathbb{R}^d$ with $x \not\sim y$ so that

$$\|Tx - Ty\| = \beta \min_{g\in G} \|x - gy\|.$$

Sketch of proof for upper Lipschitz bound

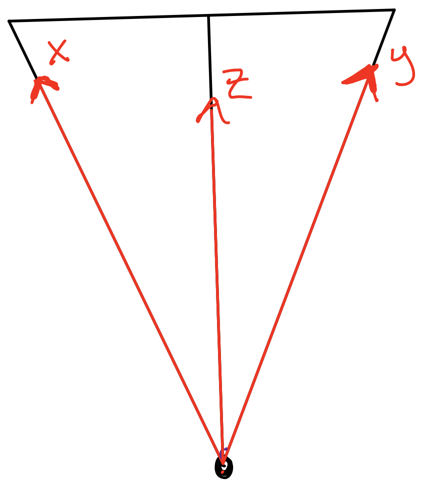
Without loss of generality, $\|x - y\| = \min_{g \in G} \|x - gy\|$.

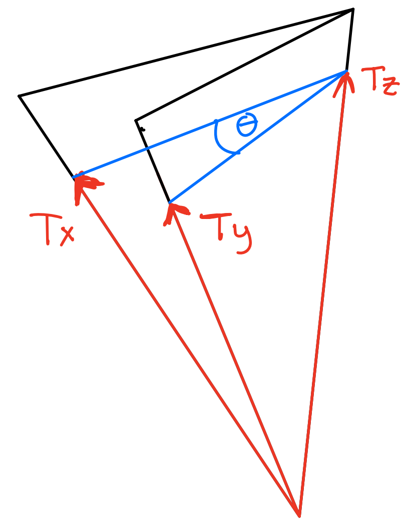


$$
\begin{aligned}
\|Tx - Ty\| &= \|Tx - Tz_1 + Tz_1 - Tz_2 + Tz_2 - Ty\| \\
&\le \|Tx - Tz_1\| + \|Tz_1 - Tz_2\| + \|Tz_2 - Ty\| \\
&= \|T_1 x - T_1 z_1\| + \|T_2 z_1 - T_2 z_2\| + \|T_3 z_2 - T_3 y\| \\
&\le \|T_1\| \|x - z_1\| + \|T_2\| \|z_1 - z_2\| + \|T_3\| \|z_2 - y\| \\
&\le \beta \Big(\|x - z_1\| + \|z_1 - z_2\| + \|z_2 - y\| \Big) \\
&= \beta \|x - y\|
\end{aligned}
$$

Idea of proof for lower Lipschitz bound

Without loss of generality, $\|x - y\| = \min_{g \in G} \|x - gy\|$.





$$
\begin{aligned}
\|Tx - Ty\|^2 &= \|Tx - Tz\|^2 + \|Ty - Tz\|^2 - 2\|Tx - Tz\|\|Ty - Tz\|\cos(\theta) \\
&\geq \|Tx - Tz\|^2 + \|Ty - Tz\|^2 - \Big(\|Tx - Tz\|^2 + \|Ty - Tz\|^2\Big)\cos(\theta) \\
&= (1 - \cos(\theta))\Big(\|Tx - Tz\|^2 + \|Ty - Tz\|^2\Big) \\
&\geq (1/2)(1 - \cos(\theta))\Big(\|Tx - Tz\| + \|Ty - Tz\|\Big)^2 \\
&\geq (1/2)(1 - \cos(\theta))\Big(\alpha\|x - z\| + \alpha\|y - z\|\Big)^2 \\
&\geq (1/2)(1 - \cos(\theta))\alpha^2\|x - y\|^2
\end{aligned}
$$

Principle angles

Given subspaces $X, Y \subseteq \mathbb{R}^d$, the first principle angle θ between X and Y is defined as the largest angle $\theta \in [0, \pi/2]$ such that the angle between any vector $x \in X$ and $y \in Y$ is at least θ. That is,

$$\langle x, y\rangle \leq \|x\|\|y\| \cos(\theta) \qquad \text{for all } x \in X, y \in Y. \tag{1}$$

Higher order principle angles can be defined inductively by choosing x and y to give equality in (1) then finding the principle angle between $X \cap x^{\perp}$ and $Y \cap y^{\perp}$.

If $X \cap Y \neq \{0\}$ then the first principle angle between X and Y is 0. In which case, we are interested in the first non-zero principle angle between X and Y.

Principle angles between cones

In order to prove the lower Lipschitz bound, we need to bound an angle between two cones X and Y which are each the intersection of finitely many closed half-spaces.

The first principle angle between X and Y is the largest angle $\theta \in [0, \pi]$ such that the angle between any vector $x \in X$ and $y \in Y$ is at least θ. That is,

$$\langle x, y \rangle \leq \|x\| \|y\| \cos(\theta) \qquad \text{for all } x \in X, y \in Y. \tag{2}$$

However, we cannot define higher order principle angles in the same way we do for subspaces because we cannot orthogonalize!

We can define the first non-zero principle angle in a different way though.

Theorem (R.Alaifari, F., D. Ghoreishi, M. Taylor, and P. Tradacete (2024))

Let $X, Y \subseteq \mathbb{R}^d$ be finite intersections of closed half-spaces such that $X \not\subseteq Y$ and $Y \not\subseteq X$. Then there is an angle $\theta \in (0, \pi]$ so that for all $x \in X$ and $y \in Y$, there exists $z \in X \cap Y$ so that the angle between $x - z$ and $y - z$ is at least θ. That is,

$$\langle x - z, y - z \rangle \leq \|x - z\| \|y - z\| \cos(\theta).$$

We define the first non-zero principle angle between X and Y to be the largest $\theta \in (0, \pi]$ which satisfies (2).