Projection methods in geodesic metric spaces

Ian Searston

Computer Assisted Research Mathematics & Applications
University of Newcastle
http://carma.newcastle.edu.au/

CARMA Retreat
17 August MMXIII

Abstract

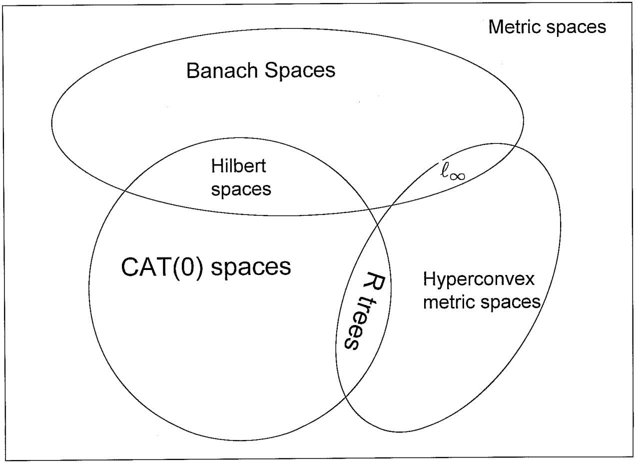
The feasibility problem associated with nonempty closed convex sets A and B is to find some $x \in A \cap B$. Projection algorithms in general aim to compute such a point.

These algorithms play key roles in optimization and have many applications outside mathematics - for example in medical imaging.

Until recently convergence results were only available in the setting of linear spaces (more particularly, Hilbert spaces) and where the two sets are closed and convex.

The extension into geodesic metric spaces allows their use in spaces where there is no natural linear structure, which is the case for instance in tree spaces, state spaces, phylogenomics and configuration spaces for robotic movements.

Diagram


Metric spaces
Banach Spaces
Hilbert
spaces
ℓ_∞
CAT(0) spaces
R trees
Hyperconvex
metric spaces

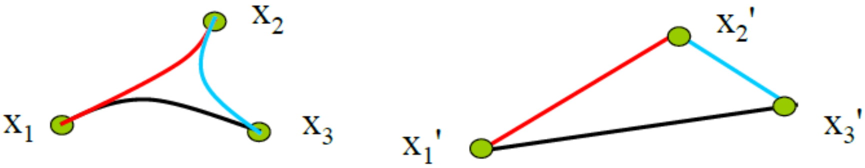
convex sets

Definition

A subset C of a geodesic metric space is *convex* if whenever x, y are in C every metric segment from x to y also lies in C.

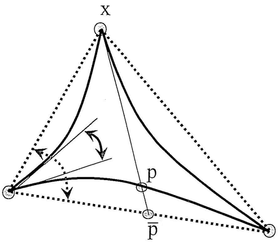
CAT(κ) spaces

A geodesic metric space X is a $CAT(\kappa)$-**space** if every geodesic triangle in triangle in X satisfies the $CAT(\kappa)$-condition (inequality) relative to its comparison triangle in the (comparison) space $\mathbb{M}^2_\kappa$.



comparison triangle in $\mathbb{E}^2$

CAT(κ) condition



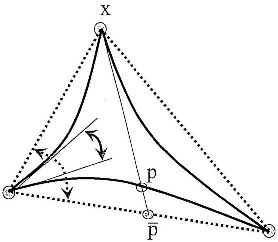
CAT(0)-condition

Here the **comparison spaces** are:

$$\mathbb{M}^2_\kappa = \begin{cases} \mathbb{S}^2_\kappa, & \text{if } \kappa > 0; \\ \mathbb{E}^2, & \text{if } \kappa = 0; \\ \mathbb{H}^2_\kappa, & \text{if } \kappa < 0. \end{cases}$$

Where:

CAT(κ) condition



CAT(0)-condition

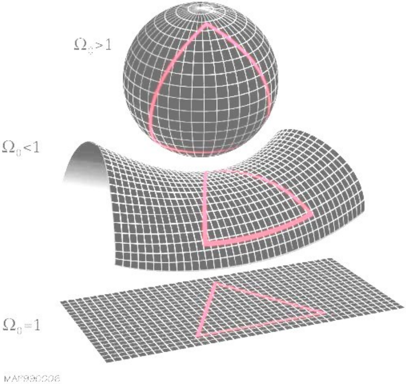
Here the **comparison spaces** are:

$$\mathbb{M}^2_{\kappa} = \begin{cases} \mathbb{S}^2_{\kappa}, & \text{if } \kappa > 0; \\ \mathbb{E}^2, & \text{if } \kappa = 0; \\ \mathbb{H}^2_{\kappa}, & \text{if } \kappa < 0. \end{cases}$$

Where:

Comparison spaces

$\mathbb{S}^2_\kappa$ is the 2-sphere of radius $\frac{1}{\sqrt{\kappa}}$,
$\mathbb{E}^2$ is two dimensional Euclidean space, and
$\mathbb{H}^2_\kappa$ is the hyperbolic two manifold of constant negative curvature κ,

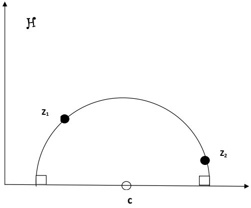


Poincareé's upper half-plane model for $\mathbb{H}^2_\kappa$

We identify $\mathbb{H}^2_\kappa$ with the Poincaré upper half-plane $\{z \in \mathbb{C} : \Im z > 0\}$ equipped with the metric $\frac{1}{\sqrt{-\kappa}} d_P$ where

$$d_P(z_1, z_2) \;=\; \int_{z_1}^{z_2} \frac{|dz|}{\Im z} \;=\; \cosh^{-1}\left(1 + \frac{|z_1 - z_2|^2}{2\Im z_1 \Im z_2}\right).$$

In which case, geodesics are semicircles with centres on the extended real axis.



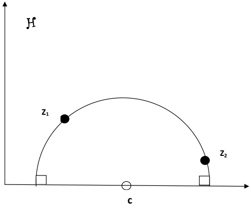
The geodesic through the points z_1 and z_2 in $\mathbb{H}^2_{-1}$

Poincareé's upper half-plane model for $\mathbb{H}^2_\kappa$

We identify $\mathbb{H}^2_\kappa$ with the Poincaré upper half-plane $\{z \in \mathbb{C} : \Im z > 0\}$ equipped with the metric $\frac{1}{\sqrt{-\kappa}} d_P$ where

$$d_P(z_1, z_2) \;=\; \int_{z_1}^{z_2} \frac{|dz|}{\Im z} \;=\; \cosh^{-1}\left(1 + \frac{|z_1 - z_2|^2}{2\Im z_1 \Im z_2}\right).$$

In which case, geodesics are semicircles with centres on the extended real axis.



The geodesic through the points $\mathbf{z_1}$ and $\mathbf{z_2}$ in $\mathbb{H}^2_{-1}$

$CAT(\kappa)$ spaces

CAT(κ) spaces with $\kappa \leq 0$ display much of the geometry inherent in Euclidean space with geodesics playing the role of lines.

Spaces with *curvature bounded below*; that is, spaces X for which $\inf\{\kappa : X \text{ is a } CAT(\kappa) \text{ space}\} > -\infty$, have non-bifurcating geodesics and are important to us because if geodesics are extendable their extensions are unique.

We will be especially interested in spaces which are $CAT(0)$. A class which includes all $CAT(\kappa)$-spaces with $\kappa < 0$.

$CAT(\kappa)$ spaces

CAT(κ) spaces with $\kappa \leq 0$ display much of the geometry inherent in Euclidean space with geodesics playing the role of lines.

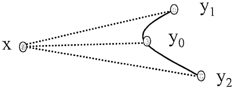
Spaces with *curvature bounded below*; that is, spaces X for which $\inf\{\kappa : X \text{ is a } CAT(\kappa) \text{ space}\} > -\infty$, have non-bifurcating geodesics and are important to us because if geodesics are extendable their extensions are unique.

We will be especially interested in spaces which are $CAT(0)$. A class which includes all $CAT(\kappa)$-spaces with $\kappa < 0$.

Equivalents for $CAT(0)$ spaces

X satisfies the CN-inequality of *B*ruhat and Titts:. That is for any three points $x, y_1, y_2 \in X$, and noting that y_0 is the metric midpoint of y_1 and y_2 if $d(y_1, y_2) = d(y_1, y_0) + d(y_0, y_2)$ and $d(y_1, y_0) = d(y_0, y_2)$ then

$$d(x, y_0)^2 \leq \frac{1}{2}d(x, y_1)^2 + \frac{1}{2}d(x, y_2)^2 - \frac{1}{4}d(y_1, y_2)^2.$$



cf the parallelogram law.

Nearest Point Projections

Proposition

Let X be a $CAT(0)$ space and C be a convex subset which is complete in the induced metric. Then,

(1) *for every $x \in X$ there exists a unique point $P_c(x) \in X$ such that $d(x, P_C(x)) = d(x, C) := \inf_{y \in C} d(x, y)$;*

(2) *if y belongs to the geodesic segment $[x, P_C(x)]$ we have $P_C(y) = P_C(X)$;*

(3) *for any $x \in X \backslash C$ and $y \in C \backslash P_C(x)$ we have $\angle_{P_C(x)}(x, y) \geq \frac{\pi}{2}$*

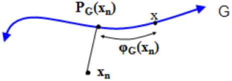
(4) *P_C is a firmly nonexpansive (in the sense of Bruck†) retraction onto C.*

$*\dagger$ $T : X \to X$ is firmly nonexpansive if $d(Tx, Ty) \leq d(tx \oplus (1-t)Tx, ty \oplus (1-t)Ty)$ all $t \in [0, 1]$.

A reformulation of weak-convergence in CAT(0) spaces

For $x \in X$ and G, any geodesic through x, we define the function $\phi_G : X \to \mathbb{R}$ by

$$\phi_G(x_n) := d(x, P_G(x_n)).$$

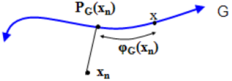


In a CAT(0)-space a sequence (x_n) is weakly convergent to x if and only if $\lim_{n\to\infty} \phi_G(x_n) = 0$, for every geodesic G containing x

A reformulation of weak-convergence in CAT(0) spaces

For $x \in X$ and G, any geodesic through x, we define the function $\phi_G : X \to \mathbb{R}$ by

$$\phi_G(x_n) := d(x, P_G(x_n)).$$



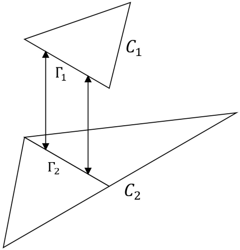
In a CAT(0)-space a sequence (x_n) is weakly convergent to x if and only if $\lim_{n\to\infty} \phi_G(x_n) = 0$, for every geodesic G containing x

Reshetnyak Gluings

Theorem

(Reshetnyak's Gluing Theorem.) Let $\{(X_i, d_i)\}, i = 1, 2$ *be two complete spaces of curvature* $\leq k$. *Suppose that there are convex sets* $C_i \in X_i$ *and an isometry* $f : C_1 \to C_2$. *Attach these spaces together along the isometry* f.
Then the resulting space (X, d) *is a space of curvature* $\leq k$.

We illustrate with an example.



Projection Algorithms in CAT(0) spaces

The convex feasibility problem associated with the nonempty closed convex sets A, B is to

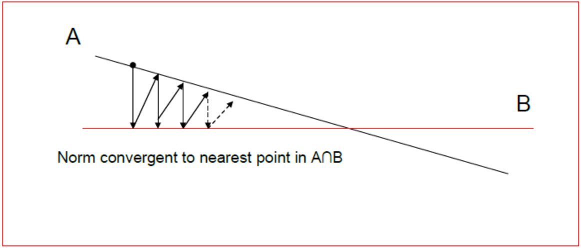
$$\text{“find some } x \in A \cap B\text{”}.$$

Projection algorithms in general aim to compute such a point. We consider two such algorithms in the context of CAT(0) spaces.

This allows us to treat feasability problems where the sets are metrically, but not necessarily algebraically, convex. For example star shaped sets in E^2.

Alternating Projection method in CAT(0) spaces

The method of **alternating projection** into convex sets (sometimes known as "project, project") emerged from initial work by John von Neumann (1903 – 1957) who, in the $1930s$, proved that when A and B were closed affine manifolds of a Hilbert space the iterative scheme $x_{n+1} = P_B P_A x_n$ converged in norm for any initial starting point x_0 to $P_{A \cap B} x_0$. .



Alternating Projection method in CAT(0) spaces

In 1965, weak convergence was established by L. M. Bregman when $A, B \in H$ are closed convex sets in a Hilbert space with $A \cap B \neq \emptyset$. Examples show that norm convergence need not occur.

The Hilbert space proof can be adapted to obtain an analogous result in CAT(0) spaces [Bacak, S & Sims].

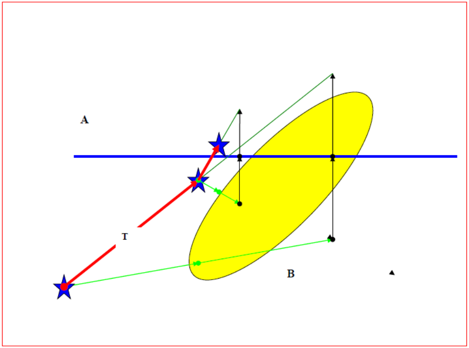
Theorem

Let X be a complete CAT(0) space and $A, B \subset X$ convex, closed subsets such that $A \cap B \neq \emptyset$. Let $x_0 \in X$ be a starting point and (x_n) be the sequence generated by alternating projections. Then (x_n) weakly converges to a point $x \in A \cap B$.
Strong convergence pertains when A and B satisfy certain "regularity" conditions and various estimates on the rate of convergence are possible.

Douglas-Rachford method

Starting with any initial point x_0, the Douglas-Rachford algorithm is the iterative scheme

$$x_{n+1} := T(x_n) \text{ where, } T = \frac{1}{2}(R_A R_B + I).$$



Douglas-Rachford method II

Provided A and B are convex and have a non empty intersection the Douglas-Rachford algorithm was shown to converge weakly to a point x with $P_B x \in A \cap B$, by P-L. Lions and B. Mercier in 1979.

Pierre-Louis Lions **Bertrand Mercier**

Douglas-Rachford method in CAT(0) spaces

Impediments to extending Douglas-Rachford into CAT(0) spaces.:
How to define reflection?
How to show convergence?
To discuss reflections in $CAT(0)$ spaces we require geodesics to be extendable. We also require that the extension is unique which happens if and only if the curvature is bounded below.

Douglas-Rachford method in CAT(0) spaces

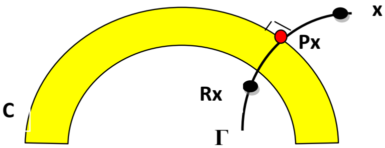
Impediments to extending Douglas-Rachford into CAT(0) spaces.:
How to define reflection?
How to show convergence?
To discuss reflections in $CAT(0)$ spaces we require geodesics to be extendable. We also require that the extension is unique which happens if and only if the curvature is bounded below.

Reflections in $CAT(0)$ spaces

With the above conditions we can define the reflection of a point x in a closed convex subset C of X, a $CAT(0)$ space, to be a point $R_C(x)$ on a geodesic which is an extension of the segment $[x, P_C(x)]$ such that

$$d(R_C(x), P_C(x)) = d(x, P_C(x)),$$

where $P_C x$ is the projection of x onto the set C.



Non-expansivity of Reflections in $CAT(0)$ spaces

It is well known that reflections in Hilbert space are non-expansive; this follows since the closest point projection is firmly nonexpansive, something which is also true in an appropriate sense in CAT(0) spaces.

constant curvature spaces

Using the appropriate "Law of cosines" Fernández-Leon – Nicolae [2012] proved the following.

Proposition

For $k \in R$ *and* $n \in N$. *Suppose* C *is a nonempty closed and convex subset of* M_k^n *and* $x, y \in M_k^n$ *such that* $dist(x, C), dist(y, C) < D_k/2$. *Then,*

$$d(R_C x, R_C y) \leq d(x, y).$$

Using this they go on to establish weak convergence of Douglas-Rachford in the classical spaces M_k^n of constant curvature.

However, in general reflections in CAT(0) spaces need not be nonexpansive.

constant curvature spaces

Using the appropriate "Law of cosines" Fernández-Leon – Nicolae [2012] proved the following.

Proposition

For $k \in R$ *and* $n \in N$*. Suppose* C *is a nonempty closed and convex subset of* M_k^n *and* $x, y \in M_k^n$ *such that* $dist(x, C), dist(y, C) < D_k/2$*. Then,*

$$d(R_C x, R_C y) \leq d(x, y).$$

Using this they go on to establish weak convergence of Douglas-Rachford in the classical spaces M_k^n of constant curvature.

However, in general reflections in CAT(0) spaces need not be nonexpansive.

non-constant curvature

To conclude we construct and investigate a special instance of a CAT(0) space of non-constant curvature.
We begin with the CAT(0) space Φ consisting of the geodesic (convex subset) $|z| = 1$ in the Poincar/'e upper half-plane which we may identify with

$$\Phi = \left(\left(-\frac{\pi}{2}, \frac{\pi}{2}\right), d_P\right)$$

where d_P is the restriction of the "Poincaré metric" given by

$$d_P(\phi_1, \phi_2) \;=\; \int_{\phi_1}^{\phi_2} \frac{d\phi}{\cos(\phi)} \;=\; [\ln(\sec(\phi) + \tan(\phi))]_{\phi_1}^{\phi_2}.$$

Since the function $\ln(\sec(\phi) + \tan(\phi))$ occurs frequently in what follows we will denote it by $H(\phi)$

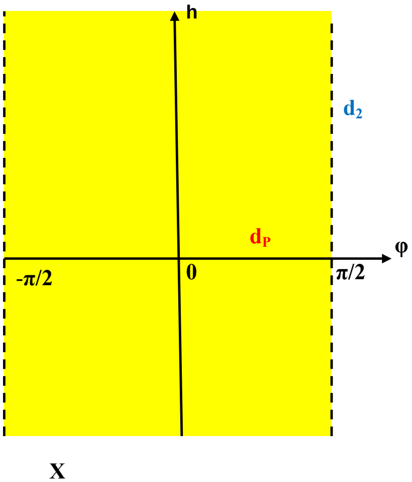
$X := \Phi \otimes_2 \mathbb{E}^1$

$X := \Phi \otimes_2 \mathbb{E}^1$ ($X_+ := \Phi \otimes_2 |\mathbb{E}^1|$); the ℓ_2^2-direct product of Φ with $\mathbb{E}^1$ – 1-dimensional, Euclidean space ($|\mathbb{E}^1|$ – the positive cone in 1-dimensional, Euclidean space), and metric given by,

$$\begin{aligned} d_X((\phi_1, h_1), (\phi_2, h_2)) &= \sqrt{(d_P(\phi_1, \phi_2))^2 + (h_1 - h_2)^2} \\ &= \sqrt{(H(\phi_2) - H(\phi_1)^2 + (h_1 - h_2)^2}, \end{aligned}$$

for $h_1, h_2 \in \mathbb{R}$ (> 0) and $-\pi/2 < \phi_1, \phi_2 < \pi/2$.

X



NOTE: X and X_+ are constant curvature (flat; curvature 0) CAT(0) spaces.

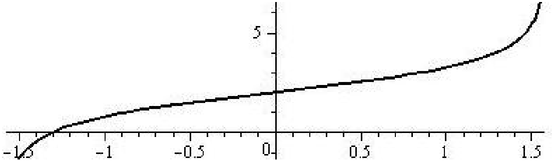
Geodesics in X

The unique **geodesic** Γ in X (or X_+) passing through two distinct points $P_1 :\ (\phi_1,\, h_1)$ and $P_2 : (\phi_2,\, h_2)$ is:
the‘vertical’ line (half line) $\{(\phi_1,\, h) : h \in \mathbb{R}(h > 0)\}$ if $\phi_1 = \phi_2$,
otherwise, using the Euler-Lagrange equation to minimize the length of a curve from P_1 to P_2, we find Γ has equation,

$$h(\phi) = AH(\phi) + B,$$

where the constants A and B are uniquely determined from the condition that $P_1,\ P_2 \in \Gamma$, in particular $A = \dfrac{h_1 - h_2}{H(\phi_1) - H(\phi_2)}$.

A geodesic in X



A geodesic in $X := \Phi \otimes_2 \mathbb{E}^1$

Poincaré revisited

In the upper half plane model of the hyperbolic space H^2_{-1} let $Y = \{z : \Im z > 0, |z| \leq 1\}$ equipped with the metric d_P inherited from H^2_{-1}.

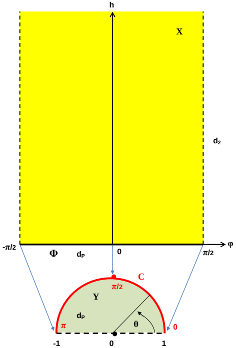
Y is a closed, convex subset of H^2_{-1} and hence a $CAT(0)$ space of constant curvature -1.

Let C be the geodesic in Y given by $C = \{e^{i\theta} : 0 < \theta < \pi\}$. Then, C is also a closed, convex subset of Y and under the mapping $\phi \mapsto \ e^{i(\frac{\pi}{2}-\phi)}$, Φ **is isometric to** C.

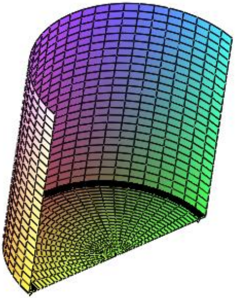
The space Z

Z is obtained by gluing X_+ to Y under the identification of Φ with C which by Reshetnyak's gluing theorem is a CAT(0) space of non-constant curvature, bounded below by -1.

Geodesics in Z are uniquely extendable, and so **reflection in closed convex sets of Z is well defined**.



3-d model of Z



A model for Z as a submanifold of $\mathbb{E}^3$

An upper-half plane model for Z

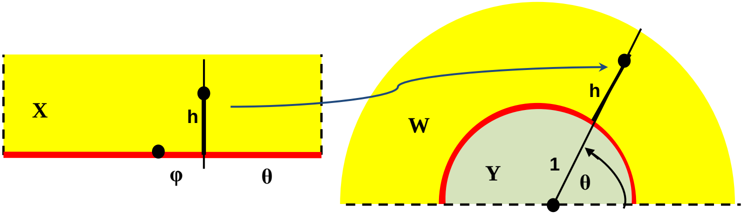
We model Y in the upper half-plane as above and identify points in X_+ with points in $W := \{\rho e^{i\theta} : \rho \geq 1, 0 < \theta < \pi\}$ under the mapping

$$(\phi, h) \mapsto (1+h)e^{i(\frac{\pi}{2}-\phi)}.$$

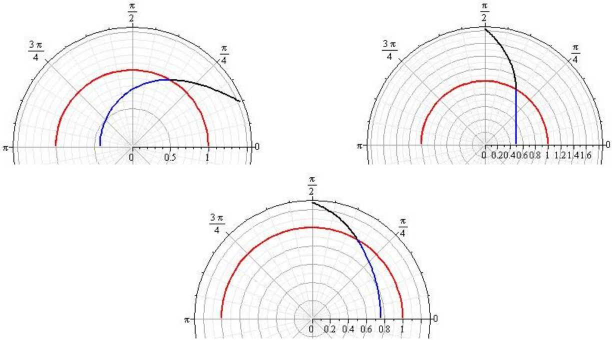
This naturally identifies Θ with C and is an isometry when W is equipped with the metric,

$$d_W(\rho_1 e^{i\theta_1}, \rho_2 e^{i\theta_2} = \sqrt{(R(\theta_2) - R(\theta_1))^2 \ + \ (\rho_2 - \rho_1)^2},$$

where $R(\theta) := H(\frac{\pi}{2} - \theta) = \ln(\operatorname{cosec}(\theta) + \cot(\theta))$.

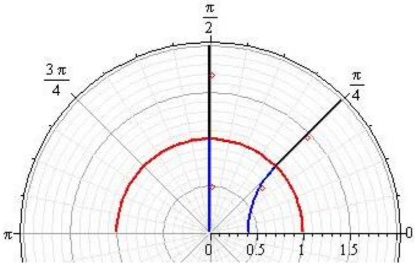


Some geodesics in Z



Some geodesics in the upper half-plane model of Z

View of reflections in C



reflection in C

Reflection need not be nonexpansiive

In general $R_C|_Y$ is nonexpansive, but $R_C|_W = (R_C|_Y)^{-1}$ need not be.

For instance, as illustrated in the previous slide, the points $P_1 = i/2$ and $P_2 = 0.5439 + 0.4925i$ in Y
have
$Q_1 := R_C(P_1) = 1.6931i$ and $Q_2 := R_C(P_2) = 1.453e^{\pi/4}$
and

$$d_W(Q_1, Q_2) = 0.9135 \ < \ d_Y(P_1, P_2) = 1.0476$$

and so,

$$d_Z(R_C(Q_1), R_C(Q_2)) = d_Y(P_1, P_2) > d_Z(Q_1, Q_2)\,\textbf{!}$$

Reflection need not be nonexpansiive

In general $R_C|_Y$ is nonexpansive, but $R_C|_W = (R_C|_Y)^{-1}$ need not be.

For instance, as illustrated in the previous slide, the points
$P_1 = i/2$ and $P_2 = 0.5439 + 0.4925i$ in Y
have
$Q_1 := R_C(P_1) = 1.6931i$ and $Q_2 := R_C(P_2) = 1.453e^{\pi/4}$
and

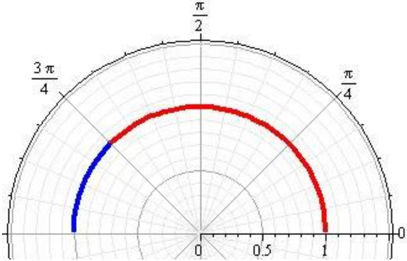
$$d_W(Q_1, Q_2) = 0.9135 \; < \; d_Y(P_1, P_2) = 1.0476$$

and so,

$$d_Z(R_C(Q_1), R_C(Q_2)) = d_Y(P_1, P_2) > d_Z(Q_1, Q_2)\,\textbf{!}$$

Douglas-Rachford in Z II

We take as our two convex sets in Z the closed half-rays $A = \{e^{i\theta} : 3\pi/4 \leq \theta \leq \pi\}$ and $B = \{e^{i\theta} : 0 \leq \theta \leq 3\pi/4\}$ of C, so $A \cap B = \{e^{3\pi i/4}\} = \{-0.7071, 0.7071)\}$.



Douglas-Rachford in Z III

The following table shows three iteration of Douglas-Rachford, starting from $x_1 = (0.5439, 0.4925)$; points z in Y are specified by $(\Re z, \Im z)$ and those in X_+ by (θ, h).

	$n = 1$	$n = 2$	$n = 3$
x_n	$(0.5439, 0.4925)$	$(-0.607669, 0.625647)$	$(-0.624135, 0.613204)$
$P_B x_n$	$(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$	$(-0.690260, 0.723561)$	$(-0.707009, 0.707205)$
$R_B x_n$	$(\frac{\pi}{4}, 0.4530)$	$(-0.761849, 0.190098)$	$(-0.785260, 0.190012)$
$P_A R_B x_n$	$(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$	$(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$	$(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$
$R_A R_B x_n$	$(-0.895023, 0.122895)$	$(-0.639941, 0.600583)$	$(-0.624326, 0.613055)$
x_{n+1}	$(-0.607669, 0.625647)$	$(-0.624135, 0.613204)$	$(-0.624231, 0.613130)$

The iterates appear to be rapidly stabilizing with $P_B x_n$ converging to the feasible point.

Douglas-Rachford in Z IV

The iterates from the alternative starting point $y_1 = (0, 0.5)$ behave similarly.

	$n = 1$	$n = 2$	$n = 3$
y_n	$(0, 0.5)$	$(-0.509291, 0.485280)$	$(-0.533052, 0.474962)$
$P_B y_n$	$(0, 1)$	$(-0.681383, 0.731927)$	$(-0.706154, 0.708058)$
$R_B y_n$	$(0, 0.6931)$	$(-0.749650, 0.492872)$	$(-0.784052, 0.495575)$
$P_A R_B y_n$	$(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$	$(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$	$(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})$
$R_A R_B y_n$	$(-0.744747, 0.231161)$	$(-0.555758, 0.463751)$	$(-0.534775, 0.474034)$
y_{n+1}	$(-0.509291, 0.485280)$	$(-0.533052, 0.474962)$	$(-0.533914, 0.474499)$

These two tables also show that in these instances the iterated map $T := \frac{1}{2}(I + R_A R_B)$ is nonexpansive, even though some intermediary steps are not.

Douglas-Rachford in Z V

Specifically;

$$
\begin{aligned}
d_Y(Tx_1, Ty_1) = 0.3098 \leq d_Y(x_1, y_1) = 1.0476 \\
d_Y(Tx_2, Ty_2) = 0.3056 \leq d_Y(x_2, y_2) = 0.3098 \\
d_Y(Tx_3, Ty_3) = 0.3056 \leq d_Y(x_3, y_3) = 0.3056
\end{aligned}
$$

While $d(R_A R_B x_1, R_A R_B y_1) = 1.0500 > d(x_1, y_1) = 1.0476$ **!**

Thus, while reflections in CAT(0) spaces of non-constant curvature need not be nonexpansive, it appears that the averaging process in Douglas-Rachford iteration may compensate for this.
This seems deserving of further investigation.

Douglas-Rachford in Z V

Specifically;

$$\begin{aligned} d_Y(Tx_1, Ty_1) &= 0.3098 \leq d_Y(x_1, y_1) = 1.0476 \\ d_Y(Tx_2, Ty_2) &= 0.3056 \leq d_Y(x_2, y_2) = 0.3098 \\ d_Y(Tx_3, Ty_3) &= 0.3056 \leq d_Y(x_3, y_3) = 0.3056 \end{aligned}$$

While $d(R_A R_B x_1, R_A R_B y_1) = 1.0500 > d(x_1, y_1) = 1.0476$ **!**

Thus, while reflections in CAT(0) spaces of non-constant curvature need not be nonexpansive, it appears that the averaging process in Douglas-Rachford iteration may compensate for this.
This seems deserving of further investigation.