Nonlinear analysis in geodesic metric spaces

Ian J Searston BSc, MMaths
The University of Newcastle
Callaghan, NSW 2308, Australia

A thesis submitted for the degree of

Doctor of Philosophy (Mathematics)

July, 2014
This thesis contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. I give consent to the final version of my thesis being made available worldwide when deposited in the University’s Digital Repository, subject to the provisions of the Copyright Act 1968.

Signature: .................................................. Date: ..........................
Acknowledgements

To have reached this stage in my mathematical journey I owe much to my principal supervisor, Associate Professor Brailey Sims, and my co-supervisor, Professor George Willis. Over this period Associate Professor Sims has been a tower of strength. He has been patient and encouraging at all times, his scholarship has been inspiring and has displayed a continuous enthusiasm for mathematics and the teaching thereof. Professor Willis has been patient and always showed a keen interest in my research.

I also wish to acknowledge the help of Laureate Professor Jonathan Borwein, Director of the Priority Research Centre for Computer-Assisted Mathematics and its Applications (CARMA). He has offered advice, suggested further avenues to explore and presented many thought provoking talks. As well, CARMA has provided some financial assistance to attend conferences. My thanks also go to two other members of staff, Miss Rebecca Smith and Mr Matthew Skerritt, for their advice and assistance with technical issues.

I also wish to acknowledge financial assistance for conferences from the School of Mathematical and Physical Sciences as well as the Faculty of Science and Information Technology.

Finally my thanks to the office staff who, although generally pushed for time, have always been most helpful and efficient.
Contents

1 Overview 1
  1.1 General ideas ................................. 1
  1.2 Summary of following chapters ............... 3
  1.3 Historical notes ................................ 5

2 Metric spaces and geodesic spaces 7
  2.1 Introduction .................................. 7
  2.2 Metric spaces .................................. 9
    2.2.1 Metrics and metric spaces ............... 9
    2.2.2 Linear spaces, normed linear spaces and Banach spaces ... 11
    2.2.3 The dual space and weak convergence ....... 14
    2.2.4 Inner product spaces and Hilbert spaces ..... 16
  2.3 Geodesic Metric Spaces ......................... 20
    2.3.1 Arcs, curves and geodesics in metric spaces ..... 20
    2.3.2 Convexity in geodesic spaces ............... 22
    2.3.3 Angles in geodesic metric spaces ............ 28
    2.3.4 Hyperbolic metric spaces .................... 29
    2.3.5 Extending Geodesics ........................ 30
  2.4 Conclusion .................................... 31
  2.5 Historical Notes ................................ 31

3 CAT(0) spaces 35
  3.1 Introduction .................................... 35
  3.2 CAT(κ) spaces .................................. 36
## CONTENTS

3.3 Polarization in CAT(0) spaces ............................................... 42
3.4 Building new CAT(0) spaces from old ones ............................. 45
  3.4.1 Reshetnyak’s Theorem ................................................. 45
3.5 Nearest Point Projections in CAT(0) spaces ............................. 47
3.6 Analogues of Weak Convergence in CAT(0) spaces .................... 50
  3.6.1 Asymptotic centres ................................................... 50
  3.6.2 $\Delta$-convergence .................................................. 51
  3.6.3 $\phi$ - convergence .................................................. 53
3.7 Conclusion ........................................................................... 55
3.8 Historical Notes .................................................................. 56

4 Convex Analysis in CAT(0) spaces .......................... 59
  4.1 Introduction ........................................................................ 59
  4.2 Convex sets in CAT(0) spaces ............................................ 60
    4.2.1 Hyperplanes and half-spaces in CAT(0) spaces .............. 60
    4.2.2 A Separation Theorem in CAT(0) spaces ..................... 61
    4.2.3 Hyperplane characterization of weak convergence ........ 63
  4.3 Convex functions in CAT(0) spaces ................................... 64
    4.3.1 Background ............................................................. 64
    4.3.2 Convex functions in CAT(0) spaces ............................ 64
    4.3.3 Tangent cones to points of a geodesic metric space ...... 69
    4.3.4 Differentiability of Convex Functions .......................... 70
    4.3.5 Subdifferentials ....................................................... 72
  4.4 Conclusion .......................................................................... 74
  4.5 Historical Notes ............................................................... 74

5 Fixed Point Theory in CAT(0) spaces .................. 77
  5.1 Introduction ........................................................................ 77
  5.2 Fixed point theory in CAT(0) spaces ................................ 77
    5.2.1 Approximate fixed point property ............................... 79
  5.3 Averaged maps .................................................................. 82
    5.3.1 Asymptotic regularity ............................................... 83
    5.3.2 Firmly nonexpansive maps ............................... 86
## CONTENTS

5.4 Fixed point Property in product spaces ........................................ 88
5.5 Locally nonexpansive mapping ....................................................... 89
5.6 Asymptotically nonexpansive mapping ........................................ 91
5.7 Conclusion ................................................................. 92
5.8 Historical Notes .......................................................... 92

6 Projection and Reflection Algorithms in CAT(0) spaces ..................... 97
6.1 Introduction ............................................................... 97
6.2 Alternating Projections in CAT(0) spaces ...................................... 102
  6.2.1 Preliminaries ......................................................... 103
  6.2.2 Convergence Results ................................................. 106
  6.2.3 Applications of results .............................................. 107
6.3 Averaged projection method in CAT(0) spaces ............................... 110
6.4 Reflections in CAT(0) spaces ................................................ 112
  6.4.1 Geodesics and reflections. ........................................... 112
  6.4.2 Non-expansivity of Reflections in CAT(0) spaces ............... 113
6.5 Conclusions ............................................................... 118
6.6 Historical Notes .......................................................... 118

7 A Prototype CAT(0) space of non-constant curvature ....................... 123
7.1 Introduction ............................................................... 123
7.2 The Poincaré upper-half plane ................................................ 123
7.3 The spaces Φ and X ......................................................... 126
  7.3.1 Definitions ............................................................ 126
  7.3.2 Geodesics in X ......................................................... 127
  7.3.3 The length of geodesic segments in X (X+) .......................... 129
  7.3.4 Mid-points of geodesic segments in X (X+) ....................... 129
  7.3.5 Projections onto Geodesics in X ................................... 130
  7.3.6 Reflections in Geodesics in X ..................................... 131
7.4 The spaces Y and Y − X+ .................................................. 132
  7.4.1 Geodesic extensions between Y and X+ ............................ 134
  7.4.2 An upper-half plane model for Y − X+ ............................ 137
  7.4.3 Geodesics in the upper half-plane Y − W model for Y − X+ 138
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.4.4</td>
<td>Using the $Y - W$ model</td>
<td>139</td>
</tr>
<tr>
<td>7.4.5</td>
<td>Reflections in $C$ need not be nonexpansive</td>
<td>142</td>
</tr>
<tr>
<td>7.5</td>
<td>The Douglas-Rachford algorithm</td>
<td>143</td>
</tr>
<tr>
<td>7.6</td>
<td>Conclusions</td>
<td>147</td>
</tr>
<tr>
<td>8</td>
<td>Conclusions</td>
<td>149</td>
</tr>
<tr>
<td>9</td>
<td>Appendix</td>
<td>153</td>
</tr>
</tbody>
</table>

**BIBLIOGRAPHY** .................................................. 169
# List of Figures

1.1 Comparison Spaces ........................................... 2  
1.2 É. Cartan, A. Alexandrov and V. Toponogov  ........... 5  
1.3 Mikhail Gromov .................................................. 6  

2.1 A metric tree.................................................. 23  
2.2 Menger (or metric) convexity ............................... 23  
2.3 Menger convexity gives midpoint convexity ............... 24  
2.4 Convexity of distance function in Busemann convex space . 27  
2.5 The Alexandrov angle ......................................... 29  
2.6 The 'thin triangle' condition ................................. 30  
2.7 Jacques Hadamard, Maurice Fréchet and Felix Hausdorff . 32  
2.8 David Hilbert and Stefan Banach ......................... 33  
2.9 Karl Menger and Herbert Busemann ....................... 34  

3.1 Comparison triangles ........................................... 37  
3.2 Relationship of CAT(0) spaces to other metric spaces ... 38  
3.3 The Alexandrov angle ......................................... 39  
3.4 The CN-inequality of Bruhat and Tits ..................... 40  
3.5 The 4-point condition ........................................ 41  
3.6 Pre-gluing diagram ............................................ 46  
3.7 Example of Reshetnyak’s Theorem ........................... 46  
3.8 An enlargement of $\triangle \text{abc}$ in Figure 3.7 ........... 47  
3.9 A specialization of Sosov’s notion of $\phi$–convergence ... 54  
3.10 Michael Edelstein and Zdzislaw Opial  .................... 56  
3.11 François Bruhat, Jacques Tits and Yurii Reshetnyak ....... 57  

vi
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Potential lack of convexity of $H^+_γ(x_0)$ in the absence of property N</td>
<td>62</td>
</tr>
<tr>
<td>4.2</td>
<td>A separation theorem</td>
<td>62</td>
</tr>
<tr>
<td>4.3</td>
<td>Hyperplane characterization of weak convergence</td>
<td>63</td>
</tr>
<tr>
<td>4.4</td>
<td>Infimum at a unique point</td>
<td>68</td>
</tr>
<tr>
<td>4.5</td>
<td>Hermann Minkowski and Victor Klee</td>
<td>75</td>
</tr>
<tr>
<td>4.6</td>
<td>J. L. Jensen and O. Hölder</td>
<td>76</td>
</tr>
<tr>
<td>5.1</td>
<td>Mark Krasnosel’skii</td>
<td>93</td>
</tr>
<tr>
<td>5.2</td>
<td>Felix Browder and W. Art Kirk</td>
<td>95</td>
</tr>
<tr>
<td>6.1</td>
<td>Projection Algorithms in $E^2$</td>
<td>98</td>
</tr>
<tr>
<td>6.2</td>
<td>Reflection of $x$ in a convex set in Hilbert space</td>
<td>100</td>
</tr>
<tr>
<td>6.3</td>
<td>The Douglas-Rachford algorithm in a Hilbert space</td>
<td>101</td>
</tr>
<tr>
<td>6.4</td>
<td>Extendable geodesics</td>
<td>112</td>
</tr>
<tr>
<td>6.5</td>
<td>Bifurcation of a geodesic</td>
<td>113</td>
</tr>
<tr>
<td>6.6</td>
<td>Reflection in a closed convex set $C$ of a CAT(0) space</td>
<td>113</td>
</tr>
<tr>
<td>6.7</td>
<td>Reflection in R-trees</td>
<td>114</td>
</tr>
<tr>
<td>6.8</td>
<td>John von Neumann</td>
<td>119</td>
</tr>
<tr>
<td>6.9</td>
<td>Jim Douglas Jr. and Henry H. Rachford Jr.</td>
<td>120</td>
</tr>
<tr>
<td>6.10</td>
<td>Pierre-Louis Lions and Bertrand Mercier</td>
<td>121</td>
</tr>
<tr>
<td>7.1</td>
<td>This figure shows the geodesic through the points $z_1$ and $z_2$</td>
<td>125</td>
</tr>
<tr>
<td>7.2</td>
<td>Space $X := \Phi \otimes_2 E^1$</td>
<td>127</td>
</tr>
<tr>
<td>7.3</td>
<td>A geodesic in $X := \Phi \otimes_2 E^1$</td>
<td>128</td>
</tr>
<tr>
<td>7.4</td>
<td>Two geodesics intersecting orthogonally in $X$</td>
<td>131</td>
</tr>
<tr>
<td>7.5</td>
<td>The gluing of $X_+$ to $Y$.</td>
<td>133</td>
</tr>
<tr>
<td>7.6</td>
<td>A model for $Y - X_+$ as a submanifold of $\mathbb{E}^3$</td>
<td>134</td>
</tr>
<tr>
<td>7.7</td>
<td>From $X_+$ to $Y$</td>
<td>135</td>
</tr>
<tr>
<td>7.8</td>
<td>Calculation to allow for change in metric</td>
<td>136</td>
</tr>
<tr>
<td>7.9</td>
<td>An Upper-half plane model for $Y - X_+$</td>
<td>138</td>
</tr>
<tr>
<td>7.10</td>
<td>Geodesic in upper half-plane $Y - W$ model</td>
<td>139</td>
</tr>
<tr>
<td>7.11</td>
<td>Geodesic in upper half-plane $Y - W$ model</td>
<td>139</td>
</tr>
<tr>
<td>7.12</td>
<td>Geodesic in upper half plane $Y - W$ model</td>
<td>140</td>
</tr>
</tbody>
</table>
Abstract

This thesis deals with nonlinear analysis in geodesic metric spaces, particularly in CAT(0) spaces. A major aim is to investigate the convex feasibility problem associated with the nonempty closed convex sets $A$ and $B$. This problem is normally investigated in a Hilbert space setting, but in this work we have placed it into a CAT(0) setting.

In chapter 2 we begin with metric spaces and move into geodesic metric spaces in order to obtain the structure which we will need in later chapters. In chapter 3 we introduce CAT(0) spaces and investigate their properties, especially those that we will need in chapters 5, 6 and 7. We also introduce new work - “Polarization in CAT(0) spaces”. More new work occurs in Chapter 4 where we develop hyperplanes and half-spaces in a CAT(0) space setting and then prove a separation theorem in CAT(0) spaces. Chapter 5 covers fixed point theory in CAT(0) spaces. We use an appropriate notion of weak sequential convergence (introduced in chapter 3) to develop a fixed point theory for nonexpansive type mapping.

In chapter 6 we prove that the project project algorithm works in CAT(0) spaces. This formed part of my research and the results presented here were published jointly by Miroslav Bacak, Ian Searston and Brailey Sims in the Journal of Mathematical Analysis and Applications ([6]). In chapter 7 we set up a prototype for a CAT(0) space of non-constant curvature, develop its geometry and then investigate the Douglas-Rachford algorithm in this space. This work was also a significant part of my research.
1

Overview

1.1 General ideas

Metric spaces in which every pair of points can be joined by an arc isometric to a compact interval of the real line are known as geodesic metric spaces. If every triangle in a geodesic metric space satisfies the “CAT(κ) inequality” they are known as CAT(κ) geodesic metric spaces, or more simply CAT(κ) spaces. These are of great interest because not only does this inequality capture the concept of non-positive curvature very well, but spaces satisfying this condition have much of the geometry inherent in Euclidean space (a CAT(0) space). It is in CAT(0) spaces that much of our interest lies.

The pioneering work of A. D. Alexandrov involved several equivalent definitions of what it means for a metric space to have curvature bounded above by a real number κ. Mikhail Gromov gave prominence to one of Alexandrov’s definitions when he termed it the CAT(κ) inequality.

Alexandrov pointed out that the curvature bounds on a space can be defined by comparing triangles in that space to triangles in the comparison spaces $\text{M}^2_\kappa$ where,
1.1 General ideas

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

Here:

\[ S^2_\kappa \] is classical two dimensional spherical space of constant positive curvature \( \kappa \); namely, the two sphere,

\[ \{ x = (x_1, x_2, x_3) \in \mathbb{R}^3 : \|x\| := \sqrt{\langle x, x \rangle} = 1 \} \]

where \( \langle x, y \rangle := x_1y_1 + x_2y_2 + x_3y_3 \), equipped with the metric, \( d(x, y) := \frac{1}{\sqrt{\kappa}} \cos^{-1}(\langle x, y \rangle) \),

\( E^2 \) is two dimensional Euclidean space, and

\( H^2_\kappa \) is the hyperbolic two manifold of constant negative curvature \( \kappa \); namely,

\[ \{ x = (x_1, x_2, x_3) \in \mathbb{R}^3 : \langle x, x \rangle = -1, x_3 > 0 \} \]

where here \( \langle x, y \rangle := x_1y_1 + x_2y_2 - x_3y_3 \). The distance between \( x \) and \( y \) in \( H^2_\kappa \) is \( d(x, y) := \frac{1}{\sqrt{-\kappa}} \cosh^{-1}(-\langle x, y \rangle) \).

Figure 1.1: Comparison Spaces
We note that $\mathbb{H}_k^2$ may be identified isometrically with $\{z \in \mathbb{C} : |z| < 1\}$, the Poincaré disc, or the Poincaré upper half plane, $\{z \in \mathbb{C} : \Im z > 0\}$, the model we will most frequently use.

In recent years, CAT(-1) and CAT(0) spaces have come to play an important role both in the study of groups from a geometrical view point and in the proofs of certain rigidity theorems in geometry.

### 1.2 Summary of following chapters

In this section we give a summary of the following chapters. Each chapter concludes with historical and other notes to add further perspectives to the development of the theory.

Since CAT(0) spaces are geodesic metric spaces we begin chapter two with a summary of some preliminaries that will be used in the rest of our work. We begin with metrics and metric spaces because they are the spaces on which the rest is built. We include convergence and completeness because of their importance in later work. We then deal with linear and normed linear spaces and importantly Hilbert spaces and weak convergence therein.

The second part of the chapter deals with geodesic metric spaces. We begin with definitions of arcs, curves and geodesics and then deal with convexity, both Menger and Busemann, including a proof of Menger’s theorem. We conclude with geodesic extensions.

In chapter three we begin with a definition of CAT($\kappa$) spaces and then move to a special class, CAT(0) spaces, which are of most interest to us. We look at properties of these spaces and include a section on polarization in CAT(0) spaces in which we introduce a polarization identity for CAT(0) spaces and new work on angles. Next we look at “building new spaces from old” including product spaces and using Reshetnyak’s theorem. Importantly for us the nearest point projection algorithm works as expected in CAT(0) spaces. We next examine projections onto
closed convex sets in CAT(0) spaces which importantly for us prove to be nonexpansive. We conclude this chapter by examining analogues of weak convergence in CAT(0) spaces. Many Hilbert space arguments involving weak-compactness can be replaced by asymptotic centre arguments or the notion of $\triangle$-convergence or $\phi$-convergence (which we take as the analogue of weak sequential convergence).

Chapter four begins our development of convex analysis in CAT(0) spaces. We introduce the concept of hyperplanes and half-spaces in CAT(0) and then develop a separation theorem (theorem 4.2.6) followed by a hyperplane characterization of weak convergence in CAT(0) spaces. We move then to convex functions, tangent cones to points of a geodesic metric space, the differentiability of convex functions and finally subdifferentials.

In chapter five we examine fixed point theory in CAT(0) spaces and find that the notion of weak sequential convergence proves to be sufficient to develop a rich fixed point theory for nonexpansive type mapping which parallels that in Hilbert space.

Chapters six and seven contain some of our major work in CAT(0) spaces. In chapter six we prove that the projection algorithms commonly known as the “Project Project Algorithm” and the “Project Project Average Algorithm” in Hilbert space can be extended to CAT(0) spaces. We also present an application of the alternating projection method to convex optimization in CAT(0) spaces. For the remainder of the chapter we consider the reflection algorithm in CAT(0) spaces and find that while it is nonexpansive in spaces of constant curvature this is not generally true for spaces of non-constant curvature. In chapter seven we develop a prototype CAT(0) space of non-constant curvature and use it both to show the failure of reflections to be nonexpansive and to investigate the Douglas-Rachford algorithm in such a space. We develop formulae for the geodesics in this space and use Maple to show that the Douglas-Rachford algorithm still appears to work in our prototype space, despite the lack of nonexpansivity.
1.3 Historical notes

The acronym CAT is derived from the names Élie Joseph Cartan (1869 - 1951), Aleksandr Danilovich Alexandrov (1912 - 1979) and Victor Andreevich Toponogov (1930 - 2004) in recognition of their pioneering work in the area.

Cartan was an influential French mathematician whose work achieved a synthesis between the areas of differential equations and geometry, Lie algebras and continuous groups. Alexandrov gave several equivalent definitions of what it means for a metric space to have a curvature bounded above by a real number $\kappa$. The work of Toponogov was influenced by that of Alexandrov. He was also noted for contributions to differential geometry and Riemannian geometry “in the large”.

Mikhael Gromov (b 1943) was born in Russia. In 1974 he left Russia and became a professor at the State University of New York. In 1981 he moved to the Université de Paris and the following year became a permanent professor at the Institut des Hautes Études Scientifiques in Bures-sur-Yvette, France. He has explained the main features of the global geometry of manifolds of non-positive curvature, essentially by basing his account on the CAT(0) inequality.

He has won numerous prizes and honours, including the AMS Steele prize in 1997, the Swiss Balzan Prize in 1999, and the Kyoto Prize for Basic Sciences in 2002.
He was awarded the Abel Prize for mathematics in 2009 “for his revolutionary contributions to geometry”.

For further historical notes we refer to pages 1 to 9 in [82].
2

Metric spaces and geodesic spaces

2.1 Introduction

The theory of spaces of non-positive curvature is considered to have its foundation in an 1898 paper by Jacques Hadamard [52] on surfaces of non-positive curvature. He was one of the first mathematicians to strongly emphasise the importance of topological methods in the study of these spaces. A theory of “metric spaces with non-positive curvature”, that is a theory that does not make any differentiability assumption and whose methods use the distance function alone, without the local coordinates provided by an embedding in Euclidean space or by another Riemannian metric structure, was developed several decades after the above paper. Hadamard indicated in [51] how to extend some of the ideas he proved for surfaces to higher dimensions. The development of these ideas in the general setting of Riemannian manifolds of non-positive curvature was largely carried out by Élie Cartan, see in particular [29].

A few years after Maurice Fréchet (a protégé of Hadamard) introduced the axioms for metric spaces, Karl Menger initiated a theory of geodesics in these spaces. A geodesic in a metric space is a continuous arc whose length is equal to the distance
between its endpoints. He introduced new methods that did not make any use of local coordinates or of differentials, but only of equalities involving the distance function and of the triangle inequality. He also based work on his definition of betweenness, that is a point \(z\) in a metric space is said to lie between two distinct points \(x\) and \(y\) if \(z\) is distinct from \(x\) and \(y\) and

\[
d(x, y) = d(x, z) + d(z, y).
\]

Menger also introduced the notion of “comparison configurations” which is at the basis of the various definitions of curvature that make sense in general metric spaces and which we have already mentioned in the previous chapter and will subsequently elaborate further.

Finally we mention the works of Herbert Busemann and A. D. Alexandrov (and collaborators). They started their works in the 1940s with no real interaction between the two. The importance of their theories continue to grow especially with the attention given to non-positive curvature by Mikhael Gromov in the 1970s. Busemann precisely defined non-positive curvature by a convexity property of the distance function. He showed, using this definition, that most of the important properties of non-positively curved Riemannian manifolds are valid in a setting which is much wider than that of Riemannian geometry. A non-positively curved space in the sense of Busemann is sometimes referred to as a “locally convex metric space” or “local Busemann space” and the terminology “non-positively curved space” in this sense is due to Busemann. As mentioned in the previous chapter Alexandrov based his definition on the notion of angle. Alexandrov introduced the notion of angle in a metric space as a generalization of the notion of angle in a surface.

We conclude this section by noting that a metric space which is non-positively curved in the sense of Alexandrov is also non-positively curved in the sense of Busemann, but that the converse is not true. For example any finite-dimensional normed vector space whose unit ball is strictly convex is non-positively curved in the sense of Busemann, but if the norm of such a space is not associated to an
inner product, then this space fails to be non-positively curved in the sense of Alexandrov.

2.2 Metric spaces

2.2.1 Metrics and metric spaces

In many branches of mathematics - in geometry as well as analysis - we need to have available a notion of distance which is applicable to the elements of abstract sets. A metric space is a non-empty set equipped with a concept of distance which is suitable for the treatment of convergent sequences in the set and continuous functions defined on the set. Our general references are [61], [87] and [90].

Definition 2.2.1. Let $X$ be a non-empty set with $x, y, z \in X$. A metric on $X$ is a real function $d$ of ordered pairs of elements of $X$ which satisfies the following three conditions:

1. $d(x, y) \geq 0$ and $d(x, y) = 0 \iff x = y$;
2. $d(x, y) = d(y, x)$ (symmetry);
3. $d(x, y) \leq d(x, z) + d(z, y)$ (the triangle inequality).

The function $d$ assigns to each pair $(x, y)$ of elements of $X$ a non-negative real number $d(x, y)$. We call $d(x, y)$ the distance between the $x$ and $y$.

Definition 2.2.2. A metric space consists of two objects; a non-empty set $X$ and a metric $d$ on $X$.

The elements of $X$ are called the points of the metric space $(X, d)$. When it can be done without confusion we denote the metric space $(X, d)$ simply by $X$. It almost always happens that different metrics can be defined on the same non-empty set and so distinct metrics make distinct metric spaces.

From the definition of a metric space we have the concept of the distance from one point to another. We can now define the distance from a point to a set and the
2.2 Metric spaces

diameter of a set.

**Definition 2.2.3.** Let $X$ be a metric space with metric $d$, $A \subset X$ and $x \in X$. Then the distance from $x$ to $A$ is defined by

$$d(x, A) = \inf \{d(x, a) : a \in A\};$$

and the diameter of the set $A$ is defined by

$$d(A) = \sup \{d(a_1, a_2) : a_1, a_2 \in A\}.$$ 

It is possible to generalise the Euclidean notion of a circle (or sphere) into any metric space.

**Definition 2.2.4.** Let $(X, d)$ be a metric space.

(a) The closed ball of radius $r$ and centre $x$ is

$$B_r[x] = \{y \in X : d(y, x) \leq r\},$$

(b) The open ball of radius $r$ and centre $x$ is

$$B_r(x) = \{y \in X : d(y, x) < r\},$$

(c) The sphere of radius $r$ and centre $x$ is

$$S_r(x) = \{y \in X : d(y, x) = r\}.$$ 

We next define **convergence** and **completeness**. Let $X$ be a metric space with metric $d$, and let $(x_n) = (x_1, x_2, \ldots, x_n, \ldots)$ be a sequence of points in $X$.

**Definition 2.2.5.** $(x_n)$ is **convergent** if there exists $x$ such that

for each $\epsilon > 0$, there exists a positive integer $n_0$ such that

$$n \geq n_0 \Rightarrow d(x_n, x) < \epsilon; \text{ or equivalently,}$$

10
for each open ball $B_r(x)$ centred on $x$, there exists a positive integer $n_0$ such that $x_n$ is in $B_r(x)$ for all $n \geq n_0$.

When $(x_n)$ converges to $x$ we often write $x_n \to x$ and we call $x$ the limit of $(x_n)$; $x = \lim_{n} (x_n)$.

A sequence $(x_n)$ is called a **Cauchy sequence** if for each $\epsilon > 0$ there exists a positive integer $n_0$ such that $m, n \geq n_0 \Rightarrow d(x_m, x_n) < \epsilon$. It follows from the above definitions that all convergent sequences are Cauchy sequences, but we can see that the converse is not true.

**Definition 2.2.6.** A **complete metric space** is a metric space in which every Cauchy sequence is convergent.

### 2.2.2 Linear spaces, normed linear spaces and Banach spaces

We begin with a nonempty set $L$ on which a binary operation $x + y$ is defined and satisfies:

1. $x + y = y + x$;
2. $x + (y + z) = (x + y) + z$;
3. there exists in $L$ a unique element, denoted by 0, which we call the zero element, or the origin, such that $x + 0 = x$ for all $x$;
4. for each element in $x$ in $L$ there corresponds a unique element in $L$, which we denote as $-x$ and call the negative of $x$, such that $x + (-x) = 0$.

We refer to the system of real numbers as **scalars**. We make the assumption that each scalar $\alpha$ and each element $x$ in $L$ can be combined by a process called scalar multiplication to yield an element of $L$ denoted by $\alpha x$ in such a way that

1. $\alpha(x + y) = \alpha x + \alpha y$;
2. $(\alpha \beta)x = \alpha x + \beta x$;
3. $(\alpha \beta)x = \alpha(\beta x)$;
2.2 Metric spaces

(8) \(1 \cdot x = x\).

The algebraic system of \(L\) equipped with these operations and axioms is a (real) linear or vector space.

**Definition 2.2.7.** A **normed linear space** is a linear space on which there is defined a **norm**, that is a function which assigns to each element \(x\) in the space a real number \(\|x\|\) in such a manner that

\[
(9) \quad \|x\| \geq 0 \text{ and } \|x\| = 0 \iff x = 0;
\]
\[
(10) \quad \|x + y\| \leq \|x\| + \|y\|;
\]
\[
(11) \quad \|\alpha x\| = |\alpha|\|x\|.
\]

In general terms a normed linear space is simply a linear space in which there is available a satisfactory notion of the distance from an arbitrary element to the origin. By writing \(-x = (-1)x\) and using (11) we obtain \(\|-x\| = |-1||x| = \|x\|\).

We also note that a normed linear space is a metric space with respect to the induced metric defined by

\[
d(x, y) = \|x - y\|.
\]

Of particular importance when \((X, \|\cdot\|)\) is a normed linear space is the unit sphere of \(X\), \(S_1(0)\), which we sometimes denote as \(S(X)\). Thus

\[
S(X) = \{x \in X : \|x\| = 1\}.
\]

Also the (closed) unit ball of a normed linear space is

\[
B[X] = \{x \in X : \|x\| \leq 1\}(=B[1][0]).
\]

Once the unit sphere is known all other spheres and balls in the normed linear space \((X, \|\cdot\|)\) are essentially determined. Indeed, all other spheres are translates of dilates of the unit sphere. Thus if we write \(rS(X)\) for the dilate \(\{ru : u \in S(X)\}\) and \(x + rS(X)\) for the translate \(\{x + w : w \in rS(X)\}\) we have \(S_r(x) = rS(X) + x = \{y \in X : \|y - x\| = r\}\). Finally we note that \(B_r[x] = B_r(x) \cup S_r(x)\).
At this stage we define what we mean by a neighbourhood of a point.

**Definition 2.2.8.** A subset $A$ of $X$ is said to be a **neighbourhood** of a point $x \in X$ if and only if there is a positive real number $r$ such that $B_r(x) \subseteq A$.

Since $x \in B_r(x)$ for $r > 0$, if $A$ is a neighbourhood of $x$ then $x \in A$. Intuitively if $A$ is a neighbourhood of $x$ then $A$ contains all points of $X$ that are ‘sufficiently near to $x$’. It is easy to see that a set is open if and only if it is a neighbourhood of each of its points.

An important concept in the study of normed linear spaces is that of a **convex set**. We start with the line segment, $[x, y]$, between the two points $x$ and $y$ of $(X, \| \cdot \|)$ which consists of all those points on $L(x, y) = \{ z \in X : z = \lambda x + (1 - \lambda)y, \lambda \in \mathbb{R} \}$ which correspond to values of $\lambda$ between 0 and 1. Consequently we have

$$[x, y] = \{ z \in X : z = \lambda x + (1 - \lambda)y \text{ where } 0 \leq \lambda \leq 1 \}.$$

**Definition 2.2.9.** A subset $C$ of the normed linear space $(X, \| \cdot \|)$ is said to be convex if whenever $x$ and $y$ are two points of $C$ the line segment between $x$ and $y$ lies entirely in $C$.

Thus $C$ is convex if and only if $\lambda x + (1 - \lambda)y \in C$ whenever $x, y \in C$ and $\lambda \in [0, 1]$. The simplest examples of convex sets are the open and closed balls.

We now define the Minkowski gauge

**Definition 2.2.10.** For $K$, a bounded convex neighbourhood of 0 in a normed linear space $X$ we define its Minkowski gauge by

$$p_K(x) = \inf \{ \xi : \xi > 0 \text{ and } \frac{1}{\xi} x \in K \}$$

for all $x \in X$.

Since $K$ is a neighbourhood of 0, there exists $\alpha > 0$ such that $\{ x \in X : \|x\| \leq \alpha \} \subseteq K$. Therefore if $x \in X$ and $x \neq 0$, then the point $\alpha \|x\|^{-1} x \in K$. This shows that the set $\{ \xi : \xi > 0 \text{ and } \xi^{-1} x \in K \}$ is nonempty for each $x \in X$. The closed unit ball of a normed linear space $X$ is a convex neighbourhood of 0 and we can
see that its Minkowski gauge is just the norm on $X$.

We finish this subsection with the following definition.

**Definition 2.2.11.** A Banach space is a normed linear space which is complete with respect to the induced metric. Thus $(X, \| \cdot \|)$ is a Banach space if $\|x_n - x_m\| \to 0$ implies there exists $x \in X$ with $\|x_n - x\| \to 0$.

The next subsection deals with the dual space of a Banach space (or a normed space) and the concept of weak convergence, the importance of which will be seen in later sections.

### 2.2.3 The dual space and weak convergence

If $X$ is a Banach space, then a mapping $f : X \to \mathbb{R}$ is called a linear functional if for each $x, y \in X$ and $\alpha, \beta \in \mathbb{R}$ we have

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y).$$

The space of all continuous linear functionals on $X$ is denoted by $X^*$. Since $\lim_{n \to \infty} x_n = x$ if and only if $\lim_{n \to \infty} (x_n - x) = 0$, a linear functional is continuous on the whole space if and only if it is continuous at 0. From this a linear functional $f$ is continuous if and only if it is bounded in the sense that there exists a constant $c$ such that for each $x \in X$

$$|f(x)| \leq c\|x\|.$$

We also have the converse true. The following two propositions are important in this work and their proofs may be found in [61].

**Proposition 2.2.12.** A continuous linear functional $f$ on a Banach (or normed) space $X$ is always bounded.

In view of Proposition 2.2.12, for each $f \in X^*$ there is a number $\|f\|_* \in \mathbb{R}^+$ such
that
\[ \|f\|_* = \sup \{|f(x)| : x \in S_X \}, \]
where, as per usual notation we have \( S_X = \{ x \in X : \|x\| = 1 \} \). The space \((X^*, \| \cdot \|_*)\) is called the dual (or conjugate) space of \( X \). The next proposition shows that \((X^*, \| \cdot \|_*)\) is a normed space which is also complete.

**Proposition 2.2.13.** If \( X \) is a normed space, then its dual space \((X^*, \| \cdot \|_*)\) is always a Banach space.

We note that Proposition 2.2.13 offers an easy way to see that certain normed spaces are complete. For example, once it is known that \( l_1 \) is complete and that \((l_1)^* = l_\infty\), then it follows that \( l_\infty \) is complete. It seems appropriate at this stage that we should also state the Hahn-Banach Theorem. The proof is readily available.

**Theorem 2.2.14.** (Hahn-Banach) Let \( X \) be a Banach space and let \( H \) be a linear subspace of \( X \). Then given any continuous linear functional \( f \in H^* \) there is a continuous linear functional \( \tilde{f} \in X^* \) such that

(i) \( f(x) = \tilde{f}(x) \) for each \( x \in H \); and

(ii) \( \|\tilde{f}\|_* = \|f\|_* \).

We should note that the norm \( \|f\|_* \) used in the above statement is relative to the dual space \( H^* \) and we also have the following corollary.

**Corollary 2.2.15.** Let \( X \) be a normed space and let \( x \in X, x \neq 0 \). Then there exists \( f_x \in X^* \) such that \( \|f_x\|_* = 1 \) and for which \( f_x(x) = \|x\| \).

In a normed space we now introduce a notion of convergence of sequences weaker than that relative to the induced norm.
Definition 2.2.16. We say that the sequence \((x_n)\) converges \textbf{weakly} to \(x \in X\), and we write 
\[ x_n \rightharpoonup x \]
if, for all \(f \in X^*\), we have 
\[ f(x_n) \to f(x), \quad \text{as } n \to \infty. \]

Weak convergence of a sequence corresponds to convergence in the \textbf{weak topology} \(\sigma(X, X^*)\) which has as a sub-base all sets of the form \(H(f, \alpha) = \{x \in X : f(x) < \alpha\}\) where \(f \in X^*\) and \(\alpha \in \mathbb{R}\).

### 2.2.4 Inner product spaces and Hilbert spaces

Just as metrics are induced by the richer structure of a norm, a norm sometimes results because of other structure carried by the space. In particular this is so when the space has an \textit{inner-product} defined on it.

**Definition 2.2.17.** An \textit{inner-product} for the vector space \(X\) is a mapping from ordered pairs of elements of \(X\) into the real field;
\[ X \times X \to \mathbb{R} : (x, y) \mapsto \langle x, y \rangle, \]
which satisfies:

1. \(\langle x, x \rangle > 0\) for all \(x \in X\) and \(x \neq 0\); (positivity)
2. \(\langle x, y \rangle = \langle y, x \rangle\) for all \(x, y \in X\); (symmetry)
3. \(\langle \alpha x, y \rangle = \alpha \langle y, x \rangle\) for all \(x, y \in X\) and \(\alpha \in \mathbb{R}\); (homogeneity)
4. \(\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle\) for all \(x, y, z \in X\). (additivity)

A vector space \(X\) together with an inner-product \(\langle \cdot, \cdot \rangle\) will be referred to as an \textit{inner-product space}.
Example 2.2.18. An inner product on $l_2$, the space of all square summable sequences, may be defined by

$$\langle x, y \rangle = \sum_{n=1}^{\infty} x_n y_n.$$  

The importance of an inner-product space is that the formula

$$\|x\| = \sqrt{\langle x, x \rangle}$$

defines a norm on $X$. We note that if an inner product space $(X, \| \cdot \|)$ is complete relative to its norm metric, then $X$ is called a **Hilbert space** and the space $l_2$ is an example of a Hilbert space. The axioms in 2.2.7 can easily be established as follows:

1. $\langle x, x \rangle$ is greater than 0 if $x \neq 0$ and equals 0 if $x = 0$, so for all $x$, $\langle x, x \rangle \geq 0$ and so $\|x\| \geq 0$. Also $\|x\| = 0 \iff \langle x, x \rangle = 0 \iff x = 0$;

2. $\|\alpha x\| = \sqrt{\langle \alpha x, \alpha x \rangle} = |\alpha|\|x\|$;

3. The proof of $\|x + y\| \leq \|x\| + \|y\|$ is in part (5) of the following theorem.

**Theorem 2.2.19.** In any inner-product space the following are true;

1. $\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2$ (parallelogram law);

2. $\langle x, y \rangle = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2)$ (polarization identity);

3. If $\langle x, y \rangle = 0$ then $\|x + y\|^2 = \|x\|^2 + \|y\|^2$ (Pythagorean identity);

4. $|\langle x, y \rangle| \leq \|x\||y||$ (Cauchy-Schwarz inequality);

5. $\|x + y\| \leq \|x\| + \|y\|$ (triangle inequality);

**Proof.** An outline of these proofs follows;
(1) \[
\|x + y\|^2 = \langle x + y, x + y \rangle \\
= \langle x, x \rangle + 2\langle x, y \rangle + \langle y, y \rangle \\
= \|x\|^2 + 2\langle x, y \rangle + \|y\|^2
\]
\[
\|x - y\|^2 = \langle x - y, x - y \rangle \\
= \|x\|^2 - 2\langle x, y \rangle + \|y\|^2.
\]

By adding these two identities we obtain the parallelogram law.

(2) Subtraction of the two identities in (1) followed by division of both sides by 4 gives the polarization identity.

(3) This part follows immediately from the first three lines of the proof in part (1) with \(\langle x, y \rangle = 0\).

(4) \[
\langle x, y \rangle = \frac{\|x + y\|^2}{4} - \frac{\|x - y\|^2}{4} \quad \text{(polarisation identity)}
\]
\[
= \frac{\|x\|^2}{2} + \frac{\|y\|^2}{2} - \frac{\|x - y\|^2}{2} \quad \text{(parallelogram law)}
\]
\[
\leq \frac{1}{2} \left( (\|x\|^2 + \|y\|^2) - (\|x\| - \|y\|)^2 \right) \quad \text{(using } \|x\| - \|y\| \leq \|x - y\|) \]
\[
= \|x\|\|y\| \quad \text{(obtained by replacing } x \text{ by } -x) \]
so \[|\langle x, y \rangle| \leq \|x\|\|y\|.
\]

We note that equality occurs if and only if \(\|x\| = \|y - x\| + \|y\|\).

(5) We now use the result of (4),
\[
\|x + y\|^2 = \langle x + y, x + y \rangle \\
= \langle x, x \rangle + 2\langle x, y \rangle + \langle y, y \rangle \\
\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \\
= (\|x\| + \|y\|)^2
\]
so by taking the square root of both sides we obtain the triangle inequality, which is otherwise known as Minkowski’s inequality in this context.

Norms arising from an inner-product have a special property; that is they satisfy the parallelogram law. The Euclidean norm in \( \mathbb{R}^n \) which arises from an inner-product satisfies the parallelogram law, which is not satisfied by all norm functions. For example in \( l^2_1 \) the norm \( \|x\|_1 = |x_1| + |x_2| \) does not satisfy the parallelogram rule. This shows that \( \| \cdot \|_1 \) does not arise from any inner-product according to the formula \( \sqrt{\langle x, x \rangle} \). Thus the parallelogram law characterizes inner-product spaces (the Jordan-von Neumann characterization).

**Theorem 2.2.20.** The inner-product \( \langle x, y \rangle \) is a continuous function on \( X \times X \).

**Proof.** Proof follows using the Cauchy-Schwarz inequality.

The importance of a Hilbert space lies in the close relation of its dual to the space itself. We note that in an inner-product space \( X \), properties (3) and (4) in Definition 2.2.17 imply that for any \( y \in X \) the inner-product generates a linear functional \( f_y \) on \( X \) defined by \( f_y(x) = \langle x, y \rangle \) for all \( x \in X \). Now the Cauchy-Schwarz inequality gives us that

\[
|f_y(x)| = |\langle x, y \rangle| \leq \|x\| \|y\| \quad \text{for all} \quad x \in X
\]

so \( f_y \) is continuous. But also we have \( \|f_y\| \leq \|y\| \). However

\[
f_y\left(\frac{y}{\|y\|}\right) = \frac{\langle y, \frac{y}{\|y\|}, y \rangle}{\|y\|} = \|y\| \\
\text{so} \quad \|f_y\| = \|y\|.
\]

The inner product then produces continuous linear functionals on \( X \), but we need to know the form of all the continuous linear functionals on \( X \). For a Hilbert space we have the Riesz [Fréchet] Representation Theorem.
2.3 Geodesic Metric Spaces

**Theorem 2.2.21.** For any continuous linear functional $f$ on a Hilbert space $H$ there exists a unique $y \in H$ such that

$$f(x) = \langle x, y \rangle \quad \text{for all} \quad x \in H.$$  

*Proof.* For a proof we refer to [45] page 39. \qed

For a real Hilbert space the mapping $y \mapsto f_y$ of $H$ into $H^*$ is, by the Riesz Representation Theorem, linear (real) and is an isometric isomorphism.

The main geometric concept missing in an abstract space such as a Banach space is that of an angle between two vectors. This is remedied in an inner-product space. From the Cauchy-Schwarz inequality, with non-zero vectors $x, y$, there exists an angle $\theta$, $0 \leq \theta \leq \pi$, such that

$$\cos \theta = \frac{\langle x, y \rangle}{\|x\|\|y\|}.$$  

We refer to $\theta$ as the angle between the vectors $x$ and $y$, which is consistent with the cosine rule in $\mathbb{E}^2$: $\|x - y\|^2 = \|x\|^2 + \|y\|^2 - 2\|x\|\|y\|\cos(\theta)$.

2.3 Geodesic Metric Spaces

2.3.1 Arcs, curves and geodesics in metric spaces

Let $(X, d)$ be a metric space.

**Definition 2.3.1.** A (simple) arc is a continuous 1-1 map from a connected interval $I \subset \mathbb{R}$ into $X$. We denote by $\Gamma_\gamma$ the image of $I$ under $\gamma$.

We note that points of $\Gamma_\gamma$ can be parameterized by $t \in I$ so that $\gamma$ induces a total order on $\Gamma_\gamma$; that is, $\gamma(t) \prec \gamma(t')$ if $t < t'$.

For an arc $\gamma$ with $x = \gamma(t_x)$ and $y = \gamma(t_y)$ where $t_x < t_y$ the arc length along $\gamma$
from \(x\) to \(y\) is given by

\[
l(\gamma_{xy}) = \sup \sum_{i=1}^{n} d(t_{i-1}, t_{i})
\]

where the supremum is taken over all possible partitions; \(t_x = t_0 < t_1 < \cdots t_n = t_y\).

The length of \(\gamma_{xy}\) is either a non-negative number or it is infinite. The arc \(\gamma\) is said to be rectifiable if \(l(\gamma_{xy}) < \infty\) for all \(x < y\) in \(\Gamma_{\gamma}\).

**Remark 2.3.2.** If \((X,d)\) is a metric space and for each pair \(x,y\) \(\in X\) there exists a rectifiable arc from \(x\) to \(y\) then \(d^*(x,y) = \inf l(\gamma_{xy})\) defines a metric on \(X\), where the infimum is taken over all possible rectifiable arcs \(\gamma_{xy}\) from \(x\) to \(y\) and \(d^*(x,y) \geq d(x,y)\).

**Definition 2.3.5.** A geodesic from \(x \in X\) to \(y \in X\) is an arc \(\gamma\) from a closed interval \([0,d(x,y)]\) \(\subset \mathbb{R}\) to \(X\) such that \(\gamma(0) = x, \gamma(d(x,y)) = y\) and \(d(\gamma(t), \gamma(t')) = |t - t'|\) for all \(t, t' \in [0, t]\). Equivalently \(\gamma\) is an isometry from \([0,d(x,y)]\) to \(X\).

**Remark 2.3.4.**

1. \(\gamma\) is a geodesic if and only if for any points \(x, y, z \in \Gamma_{\gamma}\) with \(x < z < y\) we have \(d(x,y) = d(x,z) + d(z,y)\).

2. In particular if the infimum in remark 2.3.2 is achieved then \(\gamma_{xy}\) is a geodesic in \((X,d^*)\) between \(x\) and \(y\).

3. For a geodesic \(\gamma_{xy}\) from \(x\) to \(y\), the point \(z_t = \gamma_{xy}(td(x,y))\) is such that \(d(x,z_t) = td(x,y)\) and \(d(x,z_t) + d(z_t,y) = d(x,y)\).

**Definition 2.3.6.** A geodesic line is an isometry \(\gamma : \mathbb{R} \to X\) such that \(d(\gamma(t_1), \gamma(t_2)) = |t_1 - t_2|\) for all \(t_1, t_2 \in \mathbb{R}\).

**Definition 2.3.7.** The metric space \((X,d)\) is called a geodesic metric space, (or, more briefly, a geodesic space) if for every pair of points \(x,y \in X\) there is a geodesic from \(x\) to \(y\).

**Remark 2.3.8.** If \((X,d)\) is a geodesic metric space then \(d^*\) as defined in Remark 2.3.2 is equal to \(d\). Conversely, if the infimum in Remark 2.3.2 is always achieved then \((X,d^*)\) is a geodesic metric space.
Definition 2.3.9. We say \((X,d)\) is uniquely geodesic if for all \(x,y \in X\) there is exactly one geodesic from \(x\) to \(y\).

Remark 2.3.10. In a uniquely geodesic metric space for \(x,y \in X\)

1. we denote the point \(z_t\) of Remark 2.3.4 by \((1-t)x \oplus ty\); and
2. denote \(\Gamma_{xy}\) by \([x,y]\) and refer to it as the geodesic segment from \(x\) to \(y\).
   That is \([x,y] = \{(1-t)x \oplus ty : t \in [0,1]\}\).

Definition 2.3.11. A geodesic triangle \(\triangle\) in a geodesic space \(X\) consists of three points \(x,y,z \in X\) and three geodesic arcs \(\gamma_{xy}, \gamma_{yz}, \gamma_{xz}\).

Definition 2.3.12. In a geodesic space \((X,d), C \subseteq X\) is said to be a convex subset of \((X,d)\) if \((C,d|_C)\) is a convex metric space; that is any geodesic joining any two of its points lies entirely in \(C\).

We finish this subsection with some examples of uniquely geodesic metric spaces;

Example 2.3.13. (i) Euclidean and Hyperbolic spaces. For every \(n \geq 0\), the Euclidean space \(\mathbb{E}^n\) and the hyperbolic space \(\mathbb{H}^n\) are uniquely geodesic spaces;

(ii) Normed vector spaces. Normed vector spaces are examples of geodesic spaces, but not all of them are uniquely geodesic. Any normed vector space whose norm is strictly convex (in particular an inner product space) is an example of an uniquely geodesic space;

(iii) \(\mathbb{R}\)-trees. An \(\mathbb{R}\)-tree is a metric space \(X\) such that:

1. there is a unique geodesic segment, \([x,y]\), joining each pair of points \(x,y \in X\);
2. if \([x,z] \cap [z,y] = \{z\}\), then \([x,z] \cup [z,y] = [x,y]\).

2.3.2 Convexity in geodesic spaces

We look at two different notions of convexity; those of Karl Menger and Herbert Busemann. Menger approached convexity from the idea of “betweenness” while Busemann used the idea of “distance”. It is with Menger’s approach that we begin.
2.3 Geodesic Metric Spaces

In $(X,d)$ the metric segment between $x, y \in X$ is

$$\text{seg}[x,y] = \{z : d(x,z) + d(z,y) = d(x,y)\}.$$

Note that we always have $x, y \in \text{seg}[x,y]$ and we refer to these as the end points of the segment. A point $z \in [x,y]$ which is distinct from $x$ and $y$ is said to be between $x$ and $y$. The following definition uses Menger’s notion of “betweenness”.

**Definition 2.3.14.** $(X,d)$ is Menger (or metrically) convex if for all $x, y \in X$, with $x \neq y$ there exists a point $z$ between them. That is, there exists $z \in X$ with $z \neq x, z \neq y$ and $d(x,z) + d(z,y) = d(x,y)$.

Figure 2.2 illustrates this definition.

We note that the relation of betweenness is symmetric in the sense that if $y$ lies between $x$ and $z$, then $y$ lies between $z$ and $x$ and also transitive. We now show that in complete metric spaces, metric convexity is equivalent to being a geodesic metric space.

We establish Menger’s theorem using arguments in the spirit of those found in
his original proof. Alternative proofs may be found in Blumenthal [18] (due to Aronszajn) and Goebel-Kirk [46] (using Caristi’s theorem).

**Theorem 2.3.15** (Menger). A complete space $X$ is a geodesic metric space if and only if it is metrically convex.

**Proof.** ($\Rightarrow$) For $x, y \in X$ let $\gamma : [0, d(x, y)] \to \mathbb{R}$ be a geodesic joining $x$ and $y$, then for any $t \in (0, 1)$, $z_t = \gamma(td(x, y))$ satisfies the Menger convexity condition for $x$ and $y$.

($\Leftarrow$) **STEP 1.** (Menger convexity $\Rightarrow$ midpoint convexity)
Given two distinct points $x, y \in X$, let $x_1 = x$, $y_1 = y$ and let $z_1$ be a point between $x$ and $y$ (the existence of which is guaranteed by Menger convexity). Now, suppose we have points $x_n, y_n$ and $z_n$ ($z_n$ between $x_n$ and $y_n$). Then, if $d_n := \frac{d(x, z_n)}{d(x, y)} = \frac{1}{2}$ we are done,
else for $d_n < \frac{1}{2}$ let $x_{n+1} = z_n$ and $y_{n+1} = y_n$
or for $d_n > \frac{1}{2}$ let $x_{n+1} = x_n$ and $y_{n+1} = z_n$ and then let $z_{n+1}$ be a point between $x_{n+1}$ and $y_{n+1}$, and continue.

![Figure 2.3: Menger convexity gives midpoint convexity](image)

If this procedure has not terminated at a midpoint of $x$ and $y$ ($d_n = \frac{1}{2}$, some $n$), then we have that $l_n := d(x, x_n)$ is an increasing sequence bounded above by $\frac{1}{2}d(x, y)$ and so convergent to some $l_0$, further $d(x_n, x_m) = |l_n - l_m| \to 0$. So $(x_n)$ is a Cauchy sequence convergent to $x_w$ say. Similarly $y_n \to y_w$. 


If either \( d(x, x_w) \) or \( d(x, y_w) \) is \( \frac{1}{2}d(x, y) \) we are done, otherwise proceed transfinitely to determine \( x_{w+1}, y_{w+1} \) and a \( z_{w+1} \) between them. Continue in this way.

Since at each step either \( x_\lambda \) or \( y_\lambda \) is distinct from all of its predecessors the procedure must stop before the first ordinal of cardinality greater than the continuum is reached, yielding a midpoint of \( x \) and \( y \).

**STEP 2.** (contruction of a metric segment)

For the next part of this proof we follow the method in [46]. Let \( x_0, x_1 \in X \), with \( x_0 \neq x_1 \). From Step 1, there exists \( x_{\frac{1}{2}} \in X \) such that \( d(x_0, x_{\frac{1}{2}}) = \frac{1}{2}d(x_0, x_1) \). That is \( x_{\frac{1}{2}} \) is a ‘midpoint’ of the pair \( (x_0, x_1) \). Let \( d = d(x_0, x_1) \) and define the mapping \( \gamma \) by taking

\[
\gamma(0) = x_0, \quad \gamma(d/2) = x_{\frac{1}{2}}, \quad \gamma(d) = x_1.
\]

Again by Step 1 there exist points \( x_{\frac{1}{4}}, x_{\frac{3}{4}} \) which are respective midpoints of the pairs \( (x_0, x_{\frac{1}{2}}) \) and \( (x_{\frac{1}{2}}, x_1) \). We now set

\[
\gamma(d/4) = x_{\frac{1}{4}} \text{ and } \gamma(3d/4) = x_{\frac{3}{4}}.
\]

By using the transitivity of betweenness, we conclude that \( \gamma \) is an isometry on the set \( \{0, d/4, d/2, 3d/4, d\} \). By induction it is possible to obtain points \( \{x_{p/2^n}\} \), \( 1 \leq p \leq 2^n - 1 \) \( (n = 1, 2, \ldots) \), in \( X \) such that the mapping \( \gamma : pd/2^n \to xd/2^n \) is an isometry. Since \( \{pd/2^n\} \) is a dense subset of \([0, d]\) and since \( X \) is complete, \( \gamma \) extends to the entire interval \([0, d]\) and thus we obtain a metric segment in \( X \) joining \( x_0 \) and \( x_1 \). This completes the proof.

\[\square\]

**Corollary 2.3.16.** If \( C \) is a complete subset of the metric space \((X, d)\), then \( C \) is convex if and only if \((C, d|_C)\) is metrically convex (that is Menger convex).

In a 1948 paper, [27], Busemann introduced the notion of non-positive curvature in metric spaces and he showed that many of the basic results in convexity theory have the flavour of non-positive curvature. This notion is now known as **Busemann convexity** and spaces with this property are called **Busemann spaces**.
Definition 2.3.17. A geodesic metric space \((X,d)\) is called Busemann convex if given any pair of geodesics \(\gamma_1 : [0, l_1] \rightarrow X\) and \(\gamma_2 : [0, l_2] \rightarrow X\) with \(\gamma_1(0) = \gamma_2(0)\) we have \(d(\gamma_1(t_1), \gamma_2(t_2)) \leq td(\gamma_1(l_1), \gamma_2(l_2))\) for all \(t \in [0, 1]\).

Proposition 2.3.18. In a Busemann convex metric space \(X\) any two points are joined by a unique geodesic segment.

Proof. Let us assume \(\gamma_1 : [0, l] \rightarrow X\) and \(\gamma_2 : [0, l] \rightarrow X\) are two geodesics joining the two points. Then by definition 2.3.17 we have

\[
d(\gamma_1(t), \gamma_2(t)) \leq td(\gamma_1(l), \gamma_2(l)) = 0.
\]

Hence \(\gamma_1 = \gamma_2\) so the space \(X\) is uniquely geodesic. \(\square\)

Remark 2.3.19. It follows that, in a Busemann convex space \(X\), for any geodesic \(\gamma : [0, l] \rightarrow X\) and \(t \in [0, 1]\) we can write \(\gamma(t) = (1 - t)\gamma(0) \oplus t\gamma(l)\).

Proposition 2.3.20. In a Busemann convex space the distance between matching points on any two geodesic segments is a convex function. That is, if \(\gamma_i : [0, l_i] \rightarrow X\) for \(i = 1, 2\) are any two geodesics, then for \(0 \leq t \leq 1\) we have

\[
d((1-t)\gamma_1(0) \oplus t\gamma_1(l_1), (1-t)\gamma_2(0) \oplus t\gamma_2(l_2)) \leq (1-t)d(\gamma_1(0), \gamma_2(0)) + td(\gamma_1(l_1), \gamma_2(l_2)).
\]

Proof. Let \(\gamma_3 : [0, l_3] \rightarrow X\) be a geodesic such that \(\gamma_3(0) = \gamma_2(0)\) and \(\gamma_3(l_3) = \gamma_1(l_1)\). From Definition 2.3.17 we have \(d(\gamma_1(t_1), \gamma_3(t_3)) \leq (1 - t)d(\gamma_1(0), \gamma_3(0))\) and \(d(\gamma_2(t_2), \gamma_3(t_3)) \leq td(\gamma_2(l_2), \gamma_3(l_3))\). Hence

\[
d(\gamma_1(t_1), \gamma_2(t_2)) \leq d(\gamma_1(t_1), \gamma_3(t_3)) + d(\gamma_2(t_2), \gamma_3(t_3))
\]

\[
\leq (1 - t)d(\gamma_1(0), \gamma_3(0)) + td(\gamma_2(l_2), \gamma_3(l_3))
\]

\[
\leq (1 - t)d(\gamma_1(0), \gamma_2(0)) + td(\gamma_1(l_1), \gamma_2(l_2)).
\]

Rewriting using Remark 2.3.19 gives the required form. \(\square\)

This proposition is illustrated by Figure 2.4.
Proposition 2.3.21. A complete geodesic metric space $X$ is Busemann convex if and only if for points $x, y, z \in X$

$$d\left(\frac{x \oplus y}{2}, \frac{x \oplus z}{2}\right) \leq \frac{1}{2}d(y, z),$$

that is, $X$ is "midpoint" Busemann convex.

Proof. ($\Rightarrow$) This is simply the case of taking $\gamma_1$ to be the geodesic from $x$ to $y$, $\gamma_2$ to be the geodesic from $x$ to $z$ and $t = \frac{1}{2}$ in the definition of Busemann convex.

($\Leftarrow$) Let $x = \gamma_1(0) = \gamma_2(0)$, $y = \gamma_1(l_1)$, $z = \gamma_2(l_2)$. then

$$d\left(\frac{x \oplus y}{2}, \frac{x \oplus z}{2}\right) \leq \frac{1}{2}d(y, z)$$

translates to $d(\gamma_1(\frac{1}{2}l_1), \gamma_2(\frac{1}{2}l_2)) \leq \frac{1}{2}d(\gamma_1(l_1), \gamma_2(l_2))$. From this we can obtain

$$d(\gamma_1(tl_1), \gamma_2(tl_2)) \leq td(\gamma_1(l_1), \gamma_2(l_2))$$

for every dyadic $t = k/2^p$, with $k = 1, \ldots 2^p - 1$ and $p \in \mathbb{N}$. A density argument extends the last inequality to all $t \in [0, l]$. 

Remark 2.3.22. Menger convexity concerns convexity of the space (the space is a convex geodesic metric space), while Busemann convexity is concerned with convexity properties of the distance function.
2.3 Geodesic Metric Spaces

2.3.3 Angles in geodesic metric spaces

In [1] and [2] Alexandrov gives a definition of an angle between geodesics issuing from a common point in a metric space. In this subsection we will look at this notion of angle, but first we need to define a comparison triangle so that we can compare the geometry of a metric space to that of the Euclidean space.

**Definition 2.3.23.** For a (geodesic) metric space $X$ and any three points $x, y, z \in X$, the comparison triangle in $\mathbb{E}^2$ is a triangle $\bar{\Delta} = \bar{\Delta}(x, y, z)$ with vertices such that $d_2(\bar{x}, \bar{y}) = d(x, y)$, $d_2(\bar{y}, \bar{z}) = d(y, z)$ and $d_2(\bar{z}, \bar{x}) = d(z, x)$. For a point $p \in [x, y]$ the comparison point in $\bar{\Delta}$ is the point $\bar{p} \in [\bar{x}, \bar{y}]$ with $d_2(\bar{x}, \bar{p}) = d(x, p)$. We define comparison points in $[y, x]$ and $[z, x]$ in a similar way.

We note that the interior angle of the comparison triangle $\bar{\Delta}(x, y, z)$ at $\bar{x}$ is called the comparison angle to that subtended at $x$ by $y$ and $z$ and we denote it by $\bar{\angle}_x(y, z)$. We note that $\bar{\angle}_x(y, z)$ is the unique angle $\{\theta : 0 \leq \theta \leq \pi\}$ such that

$$d(x, y)^2 = d(x, z)^2 + d(z, y)^2 + 2d(x, z)d(z, y)\cos(\theta).$$

The following definition is what Alexandrov called the upper angle between geodesics in a metric space.

**Definition 2.3.24.** Let $\gamma_1 : [0, a_1] \to X$ and $\gamma_2 : [0, a_2] \to X$ be two geodesics with $\gamma_1(0) = \gamma_2(0)$. Given $t_1 \in (0, a_1]$ and $t_2 \in (0, a_2]$ we consider the comparison triangle $\bar{\Delta}(\gamma_1(t_1), \gamma_1(t_1), \gamma_2(t_2))$ and the comparison angle $\bar{\angle}_{\gamma_1(0)}(\gamma_1(t_1), \gamma_2(t_2))$ given by

$$\cos(\bar{\angle}_{\gamma_1(0)}(\gamma_1(t_1), \gamma_2(t_2))) = \frac{1}{2t_1t_2}(t_1^2 + t_2^2 - d(\gamma_1(t_1), \gamma_2(t_2))^2).$$

The Alexandrov angle (or upper angle) between the geodesics $\gamma_1$ and $\gamma_2$ at $\gamma_1(0)$ is the number $\angle(\gamma_1, \gamma_2) \in [0, \pi]$ given by

$$\angle(\gamma_1, \gamma_2) = \limsup_{t, t' \to 0} \bar{\angle}_{\gamma_1(0)}(\gamma_1(t_1), \gamma_2(t_2)).$$

The angle between two geodesic segments which have a common endpoint is defined
2.3 Geodesic Metric Spaces

Figure 2.5: The Alexandrov angle

to be the angle between the unique geodesics which issue from this point and whose images are the given segments. For $p, x, y \in X$, a uniquely geodesic space, with $p \neq x$ and $p \neq y$, we denote the angle between the geodesic segments $[p, x]$ and $[p, y]$ by $\angle_p(x, y)$.

We note that in $\mathbb{E}^2$, the Alexandrov angle is equal to the usual Euclidean angle. Finally we note that in a metric tree, the angle between two geodesic segments which have a common endpoint is either 0 or $\pi$.

2.3.4 Hyperbolic metric spaces

The class of geodesic metric spaces called hyperbolic spaces was introduced by Mikhael Gromov and are sometimes called Gromov-hyperbolic spaces to distinguish them from classical hyperbolic spaces.

**Definition 2.3.25.** A geodesic metric space is hyperbolic (in the sense of Gromov) if it satisfies the 'thin triangle' condition. That is there exists $\delta > 0$ such that each side of any geodesic triangle is contained in the union of the $\delta$–neighbourhoods of the other two sides as shown in figure 2.6 below.

In the diagram the segment $[x, y]$ lies in the $\delta$-neighbourhood of $[y, z] \cup [x, z]$.

**Example 2.3.26.**

1. Metric trees (the prototypical example);
2. $\mathbb{R}$ - trees;
2.3 Geodesic Metric Spaces

Figure 2.6: The 'thin triangle' condition

(3) Any bounded space $X$ is $\delta$-hyperbolic with $\delta = \text{diam}(X)$;

(4) Classical hyperbolic spaces of constant curvature;

Classical hyperbolic spaces of constant curvature:

$$\mathbb{H}^n_\kappa := \{ x = (u, x_n) \in \mathbb{R}^{n+1} : u \in \mathbb{R}^n, x_n > 0 \}$$

of curvature $\kappa$ ($\kappa < 0$) with the metric derived from

$$ds^2 := \frac{1}{|\kappa| x_{n+1}^2} ||dx||^2.$$

(5) Euclidean space $\mathbb{R}^n$ with $n \geq 2$ is not hyperbolic;

2.3.5 Extending Geodesics

For some results it is necessary to know that we can extend a geodesic segment in $X$ to a geodesic line. We begin by recalling that a geodesic $\gamma : [a, \infty) \to X$ is a geodesic ray from $\gamma(a)$ and a geodesic $\gamma : \mathbb{R} \to X$ is a geodesic line. For a geodesic $\gamma$ we denote by $I_\gamma$ the interval comprising its domain: so $\gamma : I_\gamma \to X$ is an isometry. $\gamma'$ is an extension of $\gamma$, denoted by $\gamma \prec \gamma'$, if $I_{\gamma'} \supseteq I_\gamma + (-\epsilon, \epsilon)$ for some $\epsilon > 0$ and $\gamma'|_{I_\gamma} = \gamma$.

Definition 2.3.27. We say $X$ has the extension property if every closed geodesic segment (or closed ray) has an extension.

Proposition 2.3.28. If $X$ has the extension property and is complete then every geodesic segment has an extension to a geodesic line.
2.4 Conclusion

Proof. For $\gamma_0 : [a, b] \to X$, a geodesic segment, let

$$\Gamma = \{ \gamma' : \gamma' \succ \gamma_0 \}$$

$\neq \emptyset$, by the extension property.

Let $\gamma_0 \prec \gamma_1 \prec \gamma_2 \prec \ldots$ be a chain in $\Gamma$, then

$$\gamma : \bigcup_{n=1}^{\infty} I_{\gamma_n} \to X$$

with $\gamma(t)$ chosen to be $\gamma_n(t)$ where $n$ is such that $t \in I_{\gamma_n}$ is a well defined geodesic.

This means $\gamma \succeq \gamma_n$ for all $n$ and so is an upper bound. Therefore, by Zorn’s lemma, there exists a maximal $\gamma \in \Gamma$ and we claim that $I_{\gamma} = \mathbb{R}$; and so $\gamma$ is a geodesic line extending $\gamma_0$.

To see this, suppose that $I_{\gamma} = [c, d]$ with $c \neq d$. Then, using the geodesic extension property, there exists $\epsilon > 0$ and a geodesic $\bar{\gamma} : [c, d + \epsilon] \to X$ such that $\bar{\gamma}|_{[c, d]} = \gamma$.

Thus $\gamma \prec \bar{\gamma}$ which contradicts the maximality of $\gamma$. \qed

2.4 Conclusion

In this chapter we have surveyed metric and geodesic spaces. In metric spaces we covered convergence and completeness as well as the special case of normed linear spaces and in particular inner-product spaces. In geodesic spaces we paid particular attention to notions of convexity (Menger and Busemann) and gave a proof for Menger’s theorem that metric and geodesic convexity coincide.

2.5 Historical Notes

Jacques Hadamard (born in Versailles, France, 1865–1963) received the Prix Poncelet in 1898 for his research achievements over the preceding ten years. From the time he arrived in Paris in 1887 his research turned more towards mathematical
physics, but he believed he was a mathematician, not a physicist. His famous 1898 work on geodesics on surfaces of negative curvature [52] laid the foundations of symbolic dynamics.

![Image](image.jpg)

Figure 2.7: Jacques Hadamard, Maurice Fréchet and Felix Hausdorff

**Maurice Fréchet** (1878–1973), a French mathematician, was taught as a school student by Jacques Hadamard before Hadamard left Paris for Bordeaux. He developed the idea of metric spaces. In 1906 he defended what was said to be one of the most dazzling French theses in mathematics of his time. In it, Fréchet defined the concept of an abstract metric space. This marked the beginnings of analysis on appropriately structured abstract spaces, one of the principal directions analysis has taken in the 20th century.

**Felix Hausdorff** (1868–1942) was a German mathematician who is considered to be one of the founders of modern topology and who contributed significantly to set theory, descriptive set theory and functional analysis. Hausdorff developed the concepts of metric and topological spaces including their modern definitions. He sought to develop the idea of “closeness” without any dependency on the specific way the distance between these points are measured.

The concept of a normed linear space is essentially due to the Polish mathematician **Stefan Banach** (1892 –1945), who first considered the idea in his doctoral dissertation of 1920. Banach (together with his colleagues) continued to work on the theory and applications of such spaces for the rest of his life.

Inner-product spaces were implicitly studied by many mathematicians (for example
the two Germans, David Hilbert (1862–1943) and Erhard Schmidt (1876–1959) and the Hungarian Friederich Riesz (1880–1956)) during the first three decades of the twentieth century. However, the axioms were not made explicit until 1929 when they were expounded by John von Neumann as a basis for his axiomatic development of quantum mechanics.

Figure 2.8: David Hilbert and Stefan Banach

The career of Karl Menger (1902–1985) spanned sixty years, during which he published 234 papers, 65 of them before the age of thirty. Characteristic of his work in geometry and topology was the reworking of fundamental concepts from an intrinsic point of view. An early influence was his work with Hans Hahn at the University of Vienna. In the early 1930s Menger developed a general notion of curvature of an arc in a compact convex metric space. The extension from arc to higher dimensional manifolds was achieved by Menger’s student Abraham Wald who obtained a fundamentally new way of introducing Gaussian curvature. Menger’s comment was “This should make geometers realize that the fundamental notion of curvature does not depend on coordinates, equations, parametrizations or differentiability assumptions. The essence of curvature lies in the general notion of convex metric space and a quadruple of points in such a space.”

In Menger’s four Untersuchungen über allgemeine Metrik ([76] and [77], p 91), he introduced his notion of convexity and gave the following theorem concerning
geodesic arcs in complete, convex metric spaces.

**Theorem 2.5.1.** The geodesic arcs joining two points $a, b$ of a complete, convex metric space are characterised among all arcs joining $a, b$ by the following property: if $p, q$ are elements of a geodesic arc joining $a, b$ ($p, q$ both distinct from $a, b$) then either $p$ is between $a$ and $q$, or $p$ is between $q$ and $b$, or $p$ is identical with $q$.

However the first complete proof was given by Menger’s contemporary, Leonard Blumenthal (1901–1984), in a Note [17] to the Bulletin of the American Mathematical Society. In the introduction of [17] he pointed out that Menger had not proved the sufficiency of the property. That is, every arc joining $a, b$ that has the above property is a geodesic arc. Blumenthal went on to point out that a search of the literature, as well as conversation with Menger, revealed that apparently, this gap had not previously been addressed.

**Herbert Busemann** (1905–1994) specialised in convex and differential geometry. He was a winner of the Lobachevsky Medal in 1985 for his “innovative book” titled *The geometry of geodesics* [28]. He was also an accomplished painter with several public exhibitions.
3

CAT(0) spaces

3.1 Introduction

In 1957, Alexandrov introduced several equivalent definitions of what it means for a metric space to have curvature bounded above by any real number $\kappa$ [2]. Spaces satisfying these definitions were given the name CAT($\kappa$) spaces by Mikhael Gromov, [50], in honour of Cartan, Alexandrov and Toponogov. As we have already mentioned in 1.1, all these involved a comparison with a well understood model space.

These conditions have played an important role in various areas of mathematics, for instance harmonic maps and Lipschitz extensions and have played a basic role in the structure and actions of hyperbolic groups. In the context of Riemannian manifolds, the local variant of CAT($\kappa$) coincides with the assumption that the sectional curvature is at most $\kappa$. Being CAT($\kappa$) itself is a stronger global condition that additionally implies the manifold is simply-connected. However, unlike sectional curvature, the notion of a CAT($\kappa$) space makes sense in the context of arbitrary geodesic metric spaces. There are many results on metric spaces that involve such a curvature condition as an hypothesis.
3.2 CAT(κ) spaces

To discuss CAT(κ) spaces we need to be able to compare the rate at which the geodesics of the space move apart (or together) with those of model spaces of constant curvature κ. We have already discussed these model spaces in 1.1 and the following proposition expands on the earlier ideas.

**Proposition 3.2.1.** $M^n_κ$ is a geodesic metric space. If $κ ≤ 0$ then $M^n_κ$ is uniquely geodesic and all balls in $M^n_κ$ are convex. If $κ > 0$, then there is a unique geodesic segment joining $x, y ∈ M^n_κ$ if and only if $d(x, y) < \frac{π}{\sqrt{κ}}$. If $κ > 0$, closed balls in $M^n_κ$ of radius less than $\frac{π}{\sqrt{κ}}$ are convex.

The Law of Cosines in $M^n_κ$ has the following forms. Given a geodesic triangle in $M^n_κ$ with sides of positive lengths $a, b, c$ and angle $α$ at the vertex opposite the side of length $a$: for

\[
\begin{align*}
κ > 0; & \quad \cos(\sqrt{κa}) = \cos(\sqrt{κb})\cos(\sqrt{κc}) \\
& \quad + \sin(\sqrt{κb})\sin(\sqrt{κc})\cos(α). \\
κ = 0; & \quad a^2 = b^2 + c^2 - 2bc\cos(α) \\
κ < 0; & \quad \cosh(\sqrt{-κa}) = \cosh(\sqrt{-κb})\cosh(\sqrt{-κc}) \\
& \quad - \sinh(\sqrt{-κb})\sinh(\sqrt{-κc})\cos(α)
\end{align*}
\]

In particular, by fixing $b, c$ and $κ$, it can be seen that $a$ is a strictly increasing function of $α$ varying from $|b - c|$ to $b + c$ as $α$ varies from 0 to $π$.

In order to define a CAT(κ) space we need to firstly define a comparison triangle.

**Definition 3.2.2.** Let $Δ$ be a geodesic triangle in $X$, with vertices $x, y, z$ and sides the geodesic segments $γ_{xy}, γ_{yz}$ and $γ_{zx}$. A comparison triangle for $Δ$ is a geodesic triangle $\tilde{Δ}$ in $M^2_κ$ with vertices $\tilde{x}, \tilde{y}, \tilde{z}$ such that $d(\tilde{x}, \tilde{y}) = d(x, y)$, $d(\tilde{y}, \tilde{z}) = d(y, z)$ and $d(\tilde{z}, \tilde{x}) = d(z, x)$.

For $p ∈ γ_{xy}$ the comparison point in $\tilde{Δ}$ is the point $\tilde{p} ∈ [\tilde{x}, \tilde{y}]$ with $d(\tilde{x}, \tilde{p}) = d(x, p)$. Comparison points for points in $γ_{yz}$ and $γ_{zx}$ are defined in a similar way.
If $\kappa \leq 0$ then such a comparison triangle always exists and if $\kappa > 0$ then it exists provided the perimeter of $\triangle$ is less than $\frac{2\pi}{\sqrt{\kappa}}$. In both cases it is unique up to an isometry of $M^2_{\kappa}$.

**Definition 3.2.3.** A geodesic triangle satisfies the CAT($\kappa$) inequality if for every pair of points $p$ and $q$ on it we have $d(p, q) \leq d(\bar{p}, \bar{q})$ where $\bar{p}$ and $\bar{q}$ are the respective corresponding points of a comparison triangle in $M^2_{\kappa}$.

**Definition 3.2.4.** A CAT($\kappa$) space is a geodesic metric space in which every geodesic triangle satisfies the CAT($\kappa$) inequality.

For most of our work from now on we will be concerned with CAT(0) spaces (that is when $\kappa = 0$). We note that in our definition of a CAT($\kappa$) space we did not require that $X$ be complete. Complete CAT(0) spaces are often called Hadamard spaces. We also note that a CAT($\kappa$) space is also a CAT($\tau$) space for any $\tau \leq \kappa$. In particular any CAT($\kappa$) space with $\kappa \leq 0$ is a CAT(0) space. Figure 3.2 shows CAT(0) spaces in relation to other metric spaces.

**Definition 3.2.5.** A metric space $X$ is said to be of curvature less than or equal to $\kappa$ if it is locally a CAT($\kappa$) space. That is for every $x \in X$ there exists $r > 0$ such that the ball $B_r(x)$, endowed with the induced metric, is a CAT($\kappa$). If $X$ is of curvature less than or equal to 0 then we say that it is non-positively curved or of non-positive curvature.

The following proposition gives some important properties of CAT(0) spaces, most of which are direct consequences of their definition.

---

**Figure 3.1: Comparison triangles**
3.2 CAT(κ) spaces

Proposition 3.2.6. Let (X, d) be a CAT(0) space.

(1) There is a unique geodesic segment joining each pair of points x, y ∈ X and this geodesic varies continuously with its endpoints;
(2) hence, approximate mid-points are close to mid-points;
(3) Every closed ball $B_r[x] := \{y; d(y, x) \leq r\}$, $r > 0$ is metrically convex; that is, it contains the geodesic segment joining any two of its points [or equivalently if the space is complete, contains their metric midpoint] and is contractible to a point.
(4) X is Busemann convex.

Proof. For (1), (2) and (3) see the proof to Proposition 1.4 p160 in [20].

For (4) the proof follows:
We begin with the assumption that $\gamma_1(0) = \gamma_2(0)$ and consider a comparison triangle $\triangle \subset \mathbb{E}^2$ for $\triangle(\gamma_1(0), \gamma_1(1), \gamma_2(1))$. For $t \in [0, 1]$ we have $d(\gamma_1(t), \gamma_2(t)) = \ldots$
3.2 CAT(κ) spaces

\(td(\gamma_1(1), \gamma_2(1)) = td(\gamma_1(1), \gamma_2(1)).\) We have from the CAT(0) inequality that 
\(d(\gamma_1(t), \gamma_2(t)) \leq d(\gamma_1(t), \gamma_2(t))\) and so 
\(d(\gamma_1(t), \gamma_2(t)) \leq td(\gamma_1(1), \gamma_2(1)).\)

The CAT(0) inequality can be reformulated in a number of different ways.

**Proposition 3.2.7.** The following conditions are equivalent:

1. \((X, d)\) is a CAT(0) space;
2. If \(x\) is any vertex of a geodesic triangle \(\Delta(x, y, z) \subset X\) and \(p\) is any point on the opposite side then \(d(x, p) \leq d_2(\bar{x}, \bar{p})\) where \(\bar{p} \in [\bar{y}, \bar{z}]\) is the comparison point of the comparison triangle \(\bar{\Delta}(x, y, z)\) in \(E^2\); See figure 3.3
3. The Alexandrov angle at any vertex of a geodesic triangle is less than or equal to the angle at the corresponding vertex of the comparison triangle. See again figure 3.3

**Figure 3.3: The Alexandrov angle**

This is equivalent to the law of cosines. That is for a geodesic triangle in \(X\) with sides of length \(a, b\) and \(c\) we have,

\[c^2 \geq a^2 + b^2 - 2ab \cos \gamma\]

where \(\gamma\) is the Alexandrov angle between the sides of length \(a\) and \(b\).
(4) The CN (Courbure Négative or Negative Curvature)-inequality of Bruhat and Tits [25] holds. That is for any three points \( x, y_1, y_2 \in X \),

\[
d(x, \frac{1}{2}(y_1 \oplus y_2))^2 \leq \frac{1}{2}d(x, y_1)^2 + \frac{1}{2}d(x, y_2)^2 - \frac{1}{4}d(y_1, y_2)^2.
\]

Figure 3.4 illustrates this inequality.

![Figure 3.4: The CN-inequality of Bruhat and Tits](image)

Note that in \( \mathbb{E}^2 \) equality occurs courtesy of the parallelogram law. Fortunately the CN-inequality goes the right way to allow many Hilbert space arguments that depend on the parallelogram law to continue to work in CAT(0) spaces.

(5) The preceding reformulations of the CAT(0) condition concern the geometry of triangles. There is also a useful reformation concerning the geometry of quadrilaterals.

The 4-point condition holds, that is for any 4 points \( a, b, c, d \) in \( X \) let \( \bar{a}, \bar{b}, \bar{c}, \bar{d} \) be corresponding comparison points in \( \mathbb{E}^2 \); that is \( d(a, b) = d(\bar{a}, \bar{b}) \), \( d(b, c) = d(\bar{b}, \bar{c}) \), \( d(c, d) = d(\bar{c}, \bar{d}) \) and \( d(d, a) = d(\bar{d}, \bar{a}) \) then \( d(a, c) \leq d(\bar{a}, \bar{c}) \) and \( d(b, d) \leq d(\bar{b}, \bar{d}) \). See figure 3.5.

This condition first appeared in the work of Reshetnyak [84].

Note: the comparison quadrilateral is planar, but this may make no sense in \( X \), indeed \([a, c]\) and \([b, d]\) may not intersect.

**Proof.** See the proofs to Proposition 1.7, p 161 and Proposition 1.11, p 164 also in [20].
Remark 3.2.8. (Ptolemy inequality). The classical theorem of Ptolemy states that for every four points \( a, b, c, d \in \mathbb{R}^2 \) we have that
\[
\|a - b\|\|c - d\| \leq \|a - c\|\|b - d\| + \|a - d\|\|b - c\|,
\]
and the equality holds if and only if all four points lie on a common circle. Let \((X, d)\) be a geodesic space, then the inequality
\[
d(x, y)d(u, v) \leq d(x, u)d(y, v) + d(x, v)d(y, u)
\]
for all \( x, y, u, v \in X \), is called Ptolemy and metric spaces where it holds true are called Ptolemaic. Hence a geodesic metric space is CAT(0) if and only if it is Busemann and Ptolemaic.

We finish this section with the following very useful theorem. It shows in particular that the real hyperbolic space \( \mathbb{H}^n \) is a CAT(0) space.

**Theorem 3.2.9.** For a space \( X \) we have the following

1. If \( X \) is a CAT(\( \kappa \)) space, then it is a CAT(\( \kappa' \)) space for every \( \kappa' \geq \kappa \).
2. If \( X \) is a CAT(\( \kappa' \)) space for every \( \kappa' > \kappa \), then it is a CAT(\( \kappa \)) space.

**Proof.** See Theorem 1.12 page 165 in [20].  

---

Figure 3.5: The 4-point condition
We have already introduced the parallelogram law and the polarization identity for inner-product spaces (see Theorem 2.2.19). We now introduce the polarization identity for CAT(0) spaces. The CAT(0) space $X$ is characterized among geodesic metric spaces by the CN-inequality of Bruhat and Tits [25]:

$$d(p, \frac{x \oplus y}{2})^2 \leq \frac{1}{2} (d(p, x)^2 + d(p, y)^2) - \frac{1}{4}d(x, y)^2,$$

for all $p, x, y \in X$.

Here, by analogy with convex sums, for $a, b \in X$ and $t \in [0, 1]$ we denote by $ta \oplus (1-t)b$ the point in $[a, b]$, the unique geodesic segment between $a$ and $b$, whose distance from $a$ is $(1-t)d(a, b)$.

By analogy with the polarization identity for inner-product spaces, for any $p \in X$, we introduce,

$$\langle x, y \rangle_p := d(p, \frac{x \oplus y}{2})^2 - \frac{1}{4}d(x, y)^2,$$

for all $x, y \in X$.

**Properties of $\langle x, y \rangle_p$**

1. $\langle x, x \rangle_p = d(x, p)^2$.
   *Proof:* Immediate from the definition.

2. $\langle x, y \rangle_p = \langle y, x \rangle_p$, for all $x, y \in X$ [Symmetry].
   *Proof:* Immediate from symmetry of the metric and commutivity of $\oplus$.

3. $\langle x, y \rangle_p = 0$, for all $x \in X$ if and only if $y = p$.
   *Proof:* ($\Leftarrow$) Setting $y = p$ and observing that $d\left(p, \frac{x + p}{2}\right) = \frac{1}{2}d(x, p)$ (by the definition of $\oplus$) the result follows by substitution into the definition of $\langle \cdot, \cdot \rangle_p$.
   ($\Rightarrow$) In particular, taking $x$ to be $y$ yields $\langle y, y \rangle_p = 0$ and the result follows from (1).

4. ‘Cauchy-Schwarz’ inequality: $|\langle x, y \rangle_p| \leq d(x, p)d(y, p)$ for all $x, y \in X$.
   Further $\langle x, y \rangle_p = d(x, p)d(y, p)$ if and only if either $y \in [x, p]$ or $x \in [y, p]$. 
and \(-\langle x, y \rangle_p = d(x, p)d(y, p)\) if and only if \(p \in [x, y]\).

Proof:

(a) \(\langle x, y \rangle_p \leq d(x, p)d(y, p)\)

\[
\langle x, y \rangle_p = d \left( p, \frac{x \oplus y}{2} \right)^2 - \frac{1}{4}d(x, y)^2
\]

\[
\leq \frac{1}{2} (d(p, x)^2 + d(p, y)^2) - \frac{1}{2}d(x, y)^2, \quad \text{by the CN-inequality}
\]

\[
\leq \frac{1}{2} (d(x, p)^2 + d(y, p)^2 - (d(x, p) - d(y, p))^2),
\]

using \(|d(x, p) - d(p, y)| \leq d(x, y)|.

Equality in case (a)

\((\Rightarrow)\) \(|d(x, p) - d(y, p)| = d(x, y)| so either \(d(x, p) = d(x, y) + d(y, p)\) in which case [by Menger’s theorem] \(y \in [x, p]\), or \(d(y, p) = d(x, y) + d(x, p)\) and \(x \in [y, p]\).

\((\Leftarrow)\) In either case it is readily checked that we have equality in the CN-inequality and \(|d(x, p) - d(y, p)| = d(x, y)|.

(b) \(-\langle x, y \rangle_p \leq d(x, p)d(y, p)\)

\[-\langle x, y \rangle_p = \frac{1}{4}d(x, y)^2 - d \left( p, \frac{x \oplus y}{2} \right)^2
\]

\[
\leq \frac{1}{2} (d(p, x)^2 + d(p, y)^2) - 2d \left( p, \frac{x \oplus y}{2} \right)^2, \quad \text{by the CN-inequality}
\]

\[
= d(x, p)d(y, p) + \frac{1}{2} (d(x, p) - d(y, p))^2 - 2d \left( p, \frac{x \oplus y}{2} \right)^2.
\]

So the result follows provided,

\[|d(x, p) - d(y, p)| \leq 2d \left( p, \frac{x \oplus y}{2} \right),\]
but this follows since,
\[ d(x, p) - \frac{1}{2} d(x, y) = d(x, p) - d \left( x, \frac{x + y}{2} \right) \leq d \left( p, \frac{x + y}{2} \right), \]
and
\[ -d(y, p) + \frac{1}{2} d(x, y) = -d(y, p) + d \left( y, \frac{x + y}{2} \right) \leq d \left( p, \frac{x + y}{2} \right), \]
and the same with \( x \) and \( y \) interchanged.

Equality in case (b) implies either \( p \in [m, y] \subset [x, y] \) or \( p \in [x, m] \subset [x, y] \). And it is easily checked that \( p \in [x, y] \) implies equality in case (b).

### The angle defined by \( \langle x, y \rangle_p \)

For \( x \) and \( y \) both distinct from \( p \) let \( \overline{xpy} \) be the unique angle in the range \([0, \pi]\) with
\[ \cos(\overline{xpy}) := \frac{\langle x, y \rangle_p}{d(p, x)d(p, y)}; \]
that is,
\[ \cos(\overline{xpy}) = \frac{d \left( p, \frac{x + y}{2} \right)^2 - \frac{1}{4} d(x, y)^2}{d(p, x)d(p, y)}. \]

Let \( \overline{p\bar{p}\bar{y}} \) be a comparison triangle in \( \mathbb{E}^2 \); that is \( d_2(\overline{p\bar{p}}) = d(p, x) \), \( d_2(\overline{p\bar{y}}) = d(p, y) \) and \( d_2(\overline{\bar{p}\bar{y}}) = d(x, y) \), then by the CAT(0) condition
\[ d \left( p, \frac{x + y}{2} \right) \leq d_2 \left( \frac{\overline{p\bar{p}x + \bar{y}}}{} \right), \]
so
\[ \cos(\overline{\overline{p\bar{p}y}}) = \frac{d \left( p, \frac{x + y}{2} \right)^2 - \frac{1}{4} d(x, y)^2}{d(p, x)d(p, y)} \leq \frac{d_2 \left( \overline{p\bar{p}2} \right)^2 - \frac{1}{4} d_2(\overline{\bar{p}\bar{y}})^2}{d_2(\overline{p\bar{p}})d_2(\overline{p\bar{y}})} = \cos \overline{Z_p(xy)}, \]
by the polarization identity in \( \mathbb{E}^2 \),

where \( \overline{Z_p(xy)} \) is the angle at \( \overline{p} \) in the comparison triangle.
Thus, the Alexandrov angle $\angle_p(xy) := \lim_{t \to 0} \angle_p(tp \oplus (1-t)xtp \oplus (1-t)y) \leq \angle_p(xy) \leq \hat{xy}$.

However, if $X$ is locally Euclidean then, $\angle_p(xy) = \lim_{t \to 0} \hat{tp} \oplus (1-t)\hat{xtp} \oplus (1-t)y$.

3.4 Building new CAT(0) spaces from old ones

We can find new examples of CAT(0) spaces by building new spaces from old CAT(0) spaces.

1. Any nonempty convex subset of a $CAT(0)$ space is also a $CAT(0)$ space;
2. Any starlike subset of a 2 dimensional Hilbert space is a $CAT(0)$ space;
3. If $X_1, X_2$ are CAT(0) spaces then $X_1 \oplus_2 X_2$ is a CAT(0) space; that is, $X \times Y$ with the metric $d((x_1, y_1), (x_2, y_2)) = \sqrt{(d_x(x_1, x_2))^2 + (d_y(y_1, y_2))^2}$.
4. Spaces built by Reshetnyak’s theorem are CAT(0) spaces, (see next subsection).

3.4.1 Reshetnyak’s Theorem

A natural way to construct interesting new metric spaces is to take a disjoint collection of known metric spaces and glue them together. As we will see, if we glue complete CAT($\kappa$) spaces along complete, convex, isometric subspaces, the result is a CAT($\kappa$) space. We use the notation $X_1 \sqcup_f X_2$ for the quotient of the disjoint union of $X_1$ and $X_2$ by the equivalence relation $f$.

**Theorem 3.4.1.** (Reshetnyak’s Gluing Theorem.) Let $\{(X_i,d_i)\}, i = 1, 2$ be two complete locally compact spaces of curvature less than or equal to $\kappa$. Suppose that there are convex sets $C_i \in X_i, i = 1, 2$ and an isometry $f : C_1 \to C_2$. Attach these spaces together along the isometry $f$. Then $X_1 \sqcup_f X_2$ is a space of curvature less then or equal to $\kappa$.

**Proof.** In order to simplify notations we will not distinguish spaces $X_i$ and their images in $X_1 \sqcup_f X_2$ under natural projections. Let $C \in X_1 \sqcup_f X_2$ be the common
3.4 Building new CAT(0) spaces from old ones

projection of $C_i, i = 1, 2$ to $X_1 \bigcup_f X_2$. Convexity of $C$ implies that the restrictions of $d$ to $X_i$ coincide with $d_i, i = 1, 2$. In particular, every point $p \in X_1 \bigcup_f X_2$ has a neighbourhood $U \subset X_1 \bigcup_f X_2$ such that $U_i = U \cap X_i$ is a normal ball in $X_i$.

Consider a triangle $\triangle abc$ in $U$ with which we may assume that $a, b \in X_2$ and $c \in X_1$. Then there are points $c_1, c_2 \in C$ on the geodesic segments $[a, c], [b, c]$ respectively. Decompose $\triangle abc$ into three triangles $\triangle ac_1c_2$, $\triangle bac_2$ and $\triangle cc_1c_2$. Now place their comparison triangles in the plane as in Figure 3.8 below.

Figure 3.6: Pre-gluing diagram

Figure 3.7: Example of Reshetnyak’s Theorem
We can see that usual angle comparison arguments show that the angles at $c_1$ and $c_2$ in the polyhedron $\bar{a}c\bar{c}_1\bar{c}\bar{c}_2\bar{b}$ are greater than or equal to $\pi$. Loosely speaking, one can consider this polyhedron like a triangle $\bar{a}bc$ having two “concave sides”, $\bar{a}c\bar{c}_1$ and $\bar{b}_2\bar{c}$. By “straightening this triangle” (compare with Alexandrov’s Lemma 4.3.3 in [26]) we see that the angles of $\triangle abc$ are not greater than the angles of its comparison triangle.

Remark 3.4.2. The figure illustrated in Figure 3.7 is the case when $X_1$ corresponds to the triangle in $\mathbb{E}^2$ with vertices $O(0,0)$, $B(0,2)$ and $D(1,1)$ while $X_2$ corresponds to the triangle in $\mathbb{E}^2$ with vertices $A(2,0)$, $B(0,2)$ and $C(-2,0)$. This example will be significant in later work.

### 3.5 Nearest Point Projections in CAT(0) spaces

In later work we will make use of orthogonal projections onto complete convex subsets of $CAT(0)$ spaces. Nearest point projection is the name given to the map $P_C : X \to C$ constructed in the following Proposition. However before we consider this Proposition we need to define nonexpansive mapping and nonexpansive retraction.
3.5 Nearest Point Projections in CAT(0) spaces

Definition 3.5.1. For a metric space \((X,d)\), a mapping \(T : X \to X\) is nonexpansive if
\[
d(T(x), T(y)) \leq d(x, y), \quad x, y \in X.
\]
A nonexpansive retraction of a metric space \((X,d)\) onto one of its subspaces \(C\) is a nonexpansive mapping \(P\) of \(X\) onto \(C\), (that is \(d(P(x), P(y)) \leq d(x, y)\) for each \(x, y \in X\)) which leaves each point of \(C\) fixed. In this case \(C\) is said to be a nonexpansive retract of \(X\).

Proposition 3.5.2. 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 \(x \in X \setminus C\) and \(y \in C \setminus P_C(x)\) we have \(\angle_{P_C(x)}(x, y) \geq \frac{\pi}{2}\);
4. The mapping \(P_C : X \to C : x \mapsto P_C(x)\) is a nonexpansive retraction onto \(C\); the map \(H : X \times [0, 1] \to X\) sending \((x, t)\) to the point a distance \(td(x, P_C(x))\) from \(x\) on the geodesic segment \([x, P_C(x)]\) is a continuous homotopy from the identity map of \(X\) to \(P_C\).

Proof. Note - The proof for parts (1) and (3) follow that given in [20] Proposition 2.4, while the proof given here to part (4) corrects a defect in that found in the same reference.

(1) To show the existence of \(P_C(x)\) we choose a sequence of points \(y_n \in C\) such that \(d(y_n, x)\) tends to \(d(x, C)\). That this is Cauchy follows from a consideration of a comparison triangle in \(\mathbb{E}^2\) for the points \(x, y_n\) and \(y_m\). Since \(C\) is complete a limit point exists and we can take \(P_C(x)\) to be that limit point. As well the fact that every such sequence \((y_n)\) is Cauchy establishes the uniqueness of \(P_C(x)\).

(2) This follows directly from the triangle inequality.

(3) If \(\angle_{P_C(x)}(x, y)\) were to be less than \(\frac{\pi}{2}\), then we could find points \(x_1 \in\)
3.5 Nearest Point Projections in CAT(0) spaces

\[ P_C(x), x \] and \( y_1 \in [P_C(x), y] \) distinct from \( P_C(x) \) such that in the comparison triangle \( \Delta(x_1, P_C(x), y_1) \) the angle at \( P_C(x) \) would be less than \( \frac{\pi}{2} \). Thus by using the above and the CAT(0) inequality we could have for some point \( y_2 \in [P_C(x), y_1] \subset C \) the inequality \( d(x_1, y_2) < d(x_1, P_C(x)) \). However this contradicts (2) which says that \( d(x_1, P_C(x)) = d(x_1, C) \). Thus the angle cannot be less than \( \frac{\pi}{2} \).

(4) To prove closest point projections are nonexpansive (and similarly the homotopy result) consider the four points \( x, y \in X \) and \( P_C(x), P_C(y) \), their closest point projections onto \( C \) and let \( \bar{x}, \bar{y}, P_C(x) \) and \( P_C(y) \) be corresponding comparison points. From the 4-point property we have \( d(\bar{x}, \bar{y}) = d(x, y) \), \( d(\bar{y}, P_C(y)) = d(y, P_C(y)) \) etc. for the other 2 pairs of sides while for the “diagonals” we have \( d(\bar{x}, P_C(y)) \geq d(x, P_C(y)) \) and \( d(\bar{y}, P_C(x)) \geq d(y, P_C(x)) \).

Thus

\[
\angle_{P_C(x)}(\bar{x}, P_C(y)) \text{\{from the “comparison” quadrilateral\}} \geq \angle_{P_C(x)}(x, P_C(y)) \text{\{from the comparison triangle\}} \geq \frac{\pi}{2}.
\]

A similar result occurs at \( P_C(y) \) and hence

\[
d(x, y) = d(\bar{x}, \bar{y}) \geq d(P_C(x), P_C(y)) \geq d(P_C(x), P_C(y)).
\]

This is an appropriate point to define a property known as “nice projection” onto geodesics (or property (N) for short).

**Definition 3.5.3.** A complete CAT(0) space \((X, d)\) has the property of “nice projection” onto geodesics (property (N)) if, given any geodesic \( \gamma : [0, l] \to X \) and \( P_\gamma \), the metric projection onto \( \gamma \), then \( P_\gamma(m) \in [P_\gamma(x), P_\gamma(y)] \) for any \( m \in [x, y] \) and \( x, y \in X \).

We now recall Remark 3.4.2 to illustrate a case in which Property N does not hold. We note that the triangle \( BOD \) has been constructed so that \( \angle BDO = \frac{\pi}{2} \) and triangle \( ABC \) constructed so that \( \angle ABC = \frac{\pi}{2} \). We also note that \( O \) is the
midpoint of $AC$. From this it is easy to see that

$$P_{DB}(C) = P_{DB}(A) = B \quad \text{and} \quad P_{DB}(O) = D \neq \left[ \frac{1}{2} P_{DB}(C) \oplus \frac{1}{2} P_{DB}(B) \right].$$

So in this case Property N does not hold.

### 3.6 Analogues of Weak Convergence in CAT(0) spaces

One of the difficulties in extending results from Hilbert spaces into CAT(0) spaces is the seeming lack of a dual space and hence of a weak topology. However, many Hilbert space arguments involving weak-compactness can be replaced by asymptotic centre arguments or the notion of $\Delta$-convergence or more recently $\phi$-convergence.

In 1976, T. C. Lim [73] introduced a concept of convergence in a general metric space setting which he called $\Delta$-convergence. Kuczumow [71] introduced an identical notion of convergence in Hilbert spaces which he called ‘almost convergence’. W.A. Kirk and B. Panyanak in [69] adapted Lim’s concept to CAT(0) spaces. They also showed that many Hilbert space results involving weak convergence have precise analogues in CAT(0) settings, for example Opial’s property [81], the Kadec-Klee property and the demiclosedness principle for nonexpansive mappings. More recently $\Delta$-convergence has been shown to be equivalent to a variant of Sosovs $\phi$-convergence.

#### 3.6.1 Asymptotic centres

Michael Edelstein [33] appears to have been the first to introduce the notion of the asymptotic centre of a sequence in 1972.
Definition 3.6.1. Let $X$ be a metric space. For any bounded sequence $(x_n)$;

1. its asymptotic radius about $x$ is $r(x, (x_n)) := \limsup_{n \to \infty} d(x, x_n)$
2. its asymptotic radius is $r((x_n)) := \inf \{ r(x, (x_n)) : x \in X \}$ and
3. its asymptotic centre is $A((x_n)) := \{ x : r(x, (x_n)) = r((x_n)) \}$.

Proposition 3.6.2. In any complete CAT(0) space, $X$, the asymptotic centre of any bounded sequence consists of precisely one point.

Proof. Let $r$ be the asymptotic radius and let $\epsilon > 0$. Then by assumption there exists $x \in X$ such that $\limsup_{n \to \infty} d(x, x_n) < r + \epsilon$; thus for $n$ sufficiently large $d(x, x_n) < r + \epsilon$, that is, for $n$ sufficiently large $x \in B_{r+\epsilon}(x_n)$. Thus

$$C_\epsilon := \bigcup_{k=1}^{\infty} \left( \bigcap_{i=k}^{\infty} B_{r+\epsilon}(x_i) \right) \neq \emptyset.$$ 

As the ascending union of convex sets, $C_\epsilon$ is convex. Also the closure $\overline{C}_\epsilon$ of $C_\epsilon$ is convex. Therefore

$$C := \bigcap_{\epsilon > 0} \overline{C}_\epsilon \neq \emptyset.$$ 

Clearly for $u \in C$, $\limsup_{n \to \infty} d(u, x_n) \leq r$. Uniqueness of such a $u$ follows from the CN inequality, 3.2.7. [Indeed, suppose $u, v \in C$ with $u \neq v$. Then if $m$ is the midpoint of the geodesic joining $u$ and $v$,

$$d(m, x_n)^2 \leq \frac{d(u, x_n)^2 + d(v, x_n)^2}{2} - \frac{1}{4} d(u, v)^2.$$ 

This implies $\limsup_{n \to \infty} d(m, x_n)^2 \leq r^2 - \frac{1}{4} d(u, v)^2$, which is a contradiction.] \qed

3.6.2 $\triangle$-convergence

Definition 3.6.3. A sequence $(x_n)$ in $X$ is said to $\triangle$-converge to $x$, $(x_n \rightharpoonup x)$, if $x$ is the unique asymptotic centre of every subsequence $(x_{n_k})$ of $(x_n)$.

We will now make use of the asymptotic centre technique and the following notion of regularity.
Definition 3.6.4. A bounded sequence \((x_n)\) in \(X\) is said to be regular if \(r((x_n)) = r((x_{n_k}))\) for every subsequence \((x_{n_k})\) of \((x_n)\).

Remark 3.6.5. If \((x_n)\) is a regular bounded sequence with asymptotic centre \(x\) then \(x\) is the asymptotic centre for every subsequence of \((x_n)\).

It is known that every bounded sequence in a Banach space has a regular subsequence (see for example page 166 of [46]). The proof is metric in nature and carries over to this setting without change as we do in the following lemma.

Lemma 3.6.6. Let \(X\) be a complete CAT(0) space with \((x_n)\) a bounded sequence in \(X\). Then \((x_n)\) has a regular subsequence.

Proof. We will use the notation that \((x_{n_k}) \prec (x_n)\) to indicate that \((x_{n_k})\) is a subsequence of \((x_n)\) and we begin by setting

\[
r_0 = \inf \{r(x, (x_{n_k})): (x_{n_k}) \prec (x_n)\}.
\]

We now select \((x_{n_k}^1) \prec (x_n)\) such that

\[
r(x, (x_{n_k}^1)) < r_0 + 1,
\]

and let

\[
r_1 = \inf \{r(x, (x_{n_k})): (x_{n_k}) \prec (x_{n_k}^1)\}.
\]

Having defined \((x_{n_k}^i) \prec (x_{n_k}^{i-1})\), set

\[
r_i = \inf \{r(x, (x_{n_k})): (x_{n_k}) \prec (x_{n_k}^i)\}
\]

and select \((x_{n_k}^{i+1}) \prec (x_{n_k}^i)\) so that

\[
r(x, (x_{n_k}^{i+1})) \leq r_i + \frac{1}{i+1}.
\]

We note that \(r_1 \leq r_2 \leq r_3 \leq \cdots\); thus if \(r = \lim_{i \to \infty} r_i\) then \(\lim_{i \to \infty} r(x, (x_{n_k})) = r\).

Now, consider the diagonal sequence \((x_{n_k}^i)\) and let \(\varphi = r(x, (x_{n_k}^i))\). Since \((x_{n_k}^i) \prec (x_{n_k}^i)\), clearly \(\varphi \geq r_i\). On the other hand, since \((x_{n_k}^i) \prec (x_{n_k}^{i+1})\), 3.6.2.1 implies \(\varphi \leq r_i + \frac{1}{i+1}\). Thus \(\varphi = r\) and, since any subsequence \((u_n)\) of \((x_{n_k}^i)\) also satisfies
3.6 Analogues of Weak Convergence in CAT(0) spaces

\((u_n) \prec (x^i_{nk})\) and \((u_n) \prec (x^{i+1}_{nk})\) from which \(r(x, (u_n)) = r\), we conclude \((x^i_k)\) is regular. \(\square\)

From Lemma 3.6.6 and Remark 3.6.5 Theorem 3.6.7 follows immediately.

**Theorem 3.6.7.** Every bounded sequence of a complete CAT(0) space has a \(\triangle\)-convergent subsequence, so closed bounded sets are \(\triangle\)-sequentially compact.

We note at this stage that the very definition of asymptotic centre shows that CAT(0) spaces enjoy a property that in the context of Banach spaces is known as Opial’s property after the Polish mathematician Zdzislaw Opial [81]: that is for \((x_n) \subset X\) such that \((x_n)\) is \(\triangle\)-convergent to \(x\) and for \(y \in X\) with \(y \neq x\) we have

\[
\liminf_{n \to \infty} d(x_n, x) < \liminf_{n \to \infty} d(x_n, y).
\]

We also note that in a Hilbert space, \(\triangle\)-convergence and weak convergence coincide. This is a simple consequence of closed balls being weak compact and Opial’s property, that for any weakly null sequence \((x_n)\) and \(x \neq 0\) we have

\[
\liminf_{n \to \infty} ||x_n|| < \liminf_{n \to \infty} ||x + x_n||.
\]

Accordingly, in the context of CAT(0) spaces, we will henceforth refer to \(\triangle\)-convergence simply as weak convergence and write \(x_n \rightharpoonup x\) or \(x_n \triangleleft x\).

### 3.6.3 \(\phi\) - convergence

In 2004, E N Sosov [88] introduced two different notions of convergence in geodesic metric spaces, which coincide with \(\triangle\)-convergence and hence weak convergence in Hilbert spaces. In 2009, it was shown in [35] that \(\triangle\)-convergence is also equivalent to a specialization of Sosov’s notion of \(\phi\)-convergence in CAT(0) spaces.

Again, let \(X\) be a CAT(0) space, with \(x\) a given point in \(X\). For \(\Gamma\), any geodesic through \(x\), we define the function \(\phi_{\Gamma} : X \to R\) as

\[
\phi_{\Gamma}(x_n) := d(x, P_{\Gamma}(x_n))
\]
where \( P_\Gamma \) is the closest point projection onto \( \Gamma \) (a convex set). This is illustrated in Figure 3.9.

![Figure 3.9: A specialization of Sosov’s notion of \( \phi \)-convergence](image)

**Definition 3.6.8.** A bounded sequence \((x_n) \subseteq X\) is said to \( \phi \)-converge to a point \( x \in X \) if

\[
\lim_{n \to \infty} \phi_\Gamma(x_n) = 0 \quad \text{for all geodesics } \Gamma \ni x.
\]

The following proposition gives the connection between \( \Delta \)-convergence and \( \phi \)-convergence in CAT(0) spaces mentioned above.

**Proposition 3.6.9.** A sequence \((x_n) \subseteq X\) is \( \Delta \)-convergent to \( x \in X \) if and only if it is \( \phi_\Gamma \)-convergent to \( x \in X \).

**Proof.** \( \Rightarrow \): Let \( \Gamma \) be a geodesic through \( x \), and \( P_\Gamma(x_n) \) the projection of \( x_n \) onto \( \Gamma \). Since \( x \in \Gamma \), \( (x_n) \phi_\Gamma \)-converges to \( x \) if, and only if, \( P_\Gamma(x_n) \to x \) as \( n \to \infty \). So if \((x_n)\) does not \( \phi_\Gamma \)-converge to \( x \) there exists a \( \Gamma_1 \) such that \( P_{\Gamma_1}(x_n) \) does not converge to \( x \) in a strong sense. Indeed there exists a subsequence of \( P_{\Gamma_1}(x_n) \) which we denote the same and \( p \in \Gamma_1 \) with \( p \neq x \) such that \( P_{\Gamma_1}(x_n) \to p \). Now, since \( P_{\Gamma_1}(x_n) \) is the projection of \((x_n)\) onto \( \Gamma_1 \), taking subsequences if necessary, we have that

\[
\lim d(x_n, p) \leq \lim d(x_n, x)
\]

which is a contradiction of the uniqueness of the \( \Delta \)-limit.

\( \Leftarrow \): If \((x_n)\) does not \( \Delta \)-converge to \( x \) then there exists a subsequence of \((x_n)\) which
we denote the same and a point $p \neq x$ such that

$$\lim d(x_n, p) < \lim d(x_n, x).$$

Now it is enough to consider the segment determined by $x$ and $p$ to get a contradiction to the fact that $(x_n) \phi_r$-converges to $x$. \qed

In a Hilbert space this form of $\phi-$convergence is readily seen to be equivalent to weak convergence. This can be seen by considering a geodesic, $\Gamma$, through $x$ parallel to a unit vector $a$; viz $\{x + \lambda a : \lambda \in \mathbb{R}\}$. Then the projection of $x_n$ onto $\Gamma$ is $P_{\Gamma}(x_n) = x + \langle x_n - x, a \rangle a$ and after noting that $||a|| = 1$ it follows that $\phi_r(x_n) = |\langle x_n - x, a \rangle|$. Hence $\phi_r(x_n) \to 0$, for all $a$, if and only if $x_n$ converges weakly to $x$.

**Remark 3.6.10.** Clearly $d(x_n, x) \to 0 \Rightarrow x_n \rightharpoonup x$.

### 3.7 Conclusion

In this chapter we have dealt with the properties of CAT(0) spaces. We have a section on polarization in CAT(0) spaces which includes new work on angles. From this the following questions arise:

1. For $a, b, y$ and $p \in X$ is $\langle ta \oplus (1 - t)b, y \rangle_p \leq (\text{ or } \geq) t \langle a, y \rangle_p + (1 - t) \langle b, y \rangle_p$ for $t \in [0, 1]$; that is, is $\langle ta \oplus (1 - t)b, y \rangle_p$ a convex (or possibly concave) function of $t$ on $[0, 1]$? In Hilbert spaces it is of course an affine function for $t \in \mathbb{R}$.

2. How, if at all, is $\langle x_n, y \rangle_x \longrightarrow 0$ related to the sequence $(x_n)$ converging weakly to $x$ (here weak convergence means $\Delta$ or equivalently $\phi$ convergence in $X$)? Again, in Hilbert spaces they are equivalent.

We considered ways of building new CAT(0) spaces from old ones including Reshetnyak’s gluing theorem which we illustrated by constructing an example illustrating the failure of Property N. Then we concluded the chapter with an examination of analogues of weak convergence in CAT(0) spaces.
3.8 Historical Notes

Michael Edelstein (1917 –2003) was born in Poland, near the border with what was then East Prussia. He moved to Israel as a young man and served in the British army during the Second World War. For this reason, he came rather late to mathematics and academic life. His mathematical work was mainly of a geometrical nature relating particularly to the geometry of metric spaces and Banach spaces. He is regarded as a founder of metric fixed point theory.

Edelstein came to Dalhousie University Canada in 1964 and remained there until his retirement in 1982. During this time he was an outstanding presence in his department and it was mainly under his influence that it grew as a research department and became prominent on the Canadian mathematical scene. For many years he acted as editor of the Canadian Mathematical Bulletin.

![Figure 3.10: Michael Edelstein and Zdzislaw Opial](image)

Zdzislaw Opial (1930 –1974) was born in Poland and was a mathematics professor at the Jagiellonian University, Kraków, where he served for a period as vice-rector. His greatest successes were achieved in the theory of differential equations and inequalities as well as analysis and functional analysis. He also dealt with the history of mathematics, the didactics and methodology of mathematics.
education as well as the popularization of mathematics.

The Opial property (see 3.6.2) is named after him. This is an abstract property of Banach spaces that plays an important role in the study of weak convergence for iterates on mappings of Banach spaces and the asymptotic behaviour of nonlinear semigroups.

François Bruhat (1929–2007) was born in Paris. Bruhat was among the first to understand the importance of distributions, as introduced by Laurent Schwartz, in the theory of Lie groups. In 1957 he published two papers dealing with real analytic manifolds with singularities following a collaboration with Henri Cartan.

Together with Jacques Tits, Bruhat developed the theory of simple algebraic groups over local fields. The Bruhat order of a Weyl group, the Bruhat decomposition, and the Schwartz-Bruhat functions are named after him. He is also known for the CN (Courbure Négative or Negative Curvature)-inequality in association with Jacques Tits (see (4) in proposition 3.2.7). Finally we should mention that Bruhat was a third generation member of the Nicolas Bourbak group.

Figure 3.11: François Bruhat, Jacques Tits and Yurii Reshetnyak

Jacques Tits (b 1930) is a Belgian / French mathematician who works on group theory and geometry. In 1974 he changed his citizenship from Belgian to French in order to teach at the Collège de France, which at that point required French
citizenship. He introduced the theory of buildings (sometimes known as Tits buildings), which are combinatorial structures, some of which provide further instances of CAT(0) spaces. The action of groups on these structures is of importance in the algebraic theory of groups. In collaboration with François Bruhat he developed the theory of affine buildings, and later classified all irreducible buildings of affine type and rank at least four.

Tits was also an “honourary” member of the Nicolas Bourbaki group and as such he helped popularize Harold Scott MacDonald Coxeter’s work, introducing terms such as Coxeter number, Coxeter group and Coxeter graph. Tits received the Wolf Prize in 1993 and in 2008 he was awarded the Abel Prize, along with John Thompson, for “their profound achievements in algebra and in particular for shaping modern group theory”.

**Yurii Reshetnyak** (b 1929) is a Russian mathematician from Saint Petersburg. He works in geometry and the theory of functions of a real variable. He is mostly known for the Reshetnyak gluing theorem (see 3.4.1 and [84]) and the Reshetnyak majorization theorem. In 2000 he received the Lobachevsky Prize from the Russian Academy of Sciences.
4

Convex Analysis in CAT(0) spaces

4.1 Introduction

Victor Klee (1925–2007), in his journal article entitled “What is a convex set” [70] describes the study of convex sets as follows “The study of convex sets is a branch of geometry, analysis and linear algebra that has numerous connections with other areas of mathematics and serves to unify many apparently diverse mathematical phenomena.” The early developments of convexity theory were finite-dimensional and directed mainly toward the solution of quantitative problems.

In this chapter we will begin with a section on convex sets in CAT(0) spaces followed by sections on convex functions and then the differentiability of convex functions in CAT(0) spaces.
4.2 Convex sets in CAT(0) spaces

We begin by recalling our definition for convex sets.

**Definition 4.2.1.** In a geodesic space \((X, d)\), \(C \subseteq X\) is said to be a convex subset if \((C, d|_C)\) is a convex metric space; that is any geodesic joining any two of its points lies entirely in \(C\).

The following lemma is an analogue to one in Banach spaces.

**Lemma 4.2.2.** Let \(X\) be a CAT(0) space. Then a convex set \(C \subseteq X\) is closed if and only if it is weakly sequentially closed.

**Proof.** \(\Leftarrow\) Follows immediately. \(\Rightarrow\) For \((x_n) \subset C\) and \(x_n \rightharpoonup x \in X\) we assume \(x \not\in C\). Let \(\gamma\) be the geodesic from \(x\) to \(P_C(x)\) and denote the geodesic segment \([x, P_C(x)]\) by \(\Gamma\). We claim that \(P_\Gamma(x_n) = P_C(x)\) for all \(n \in \mathbb{N}\). Indeed, if for some \(m \in \mathbb{N}\) we had \(P_\Gamma(x_m) \neq P_C(x)\) then by Proposition 3.5.2 we would have

\[
\angle_{P_C(x)}(x_m, P_\Gamma(x_m)) \geq \frac{\pi}{2} \quad \text{and} \quad \angle_{P_\Gamma(x_m)}(x_m, P_C(x)) \geq \frac{\pi}{2}
\]

which is impossible.

Finally \(d(P_\Gamma(x_n), x) = d(P_C(x), x) \not\to 0\) as \(n \to \infty\) which by Proposition 3.6.9 contradicts \(x_n \rightharpoonup x\).

\[\square\]

### 4.2.1 Hyperplanes and half-spaces in CAT(0) spaces

The notions of hyperplanes and half-spaces is well known in Hilbert spaces. In this subsection we will define analogous notions in CAT(0) spaces.

**Definition 4.2.3.** In a CAT(0) spaces \((X, d)\) for a geodesic \(\gamma : [0, l] \to X\) and \(x_0 = \gamma(t_0)\) where \(0 < t_0 < l\), the associated **hyperplane** is defined as

\[
H_\gamma(x_0) = \{x \in X : P_\gamma(x) = x_0\}.
\]
4.2 Convex sets in CAT(0) spaces

In Hilbert space, $H_{\gamma}(x_0)$ is the hyperplane perpendicular to $\gamma$ and containing $x_0$. Having defined hyperplanes we can now define half-spaces.

**Definition 4.2.4.** In a CAT(0) space $(X,d)$ for a geodesic $\gamma : [0,l] \rightarrow X$ with $t_0 \in (0,l)$ and $x_0 = \gamma(t_0)$, we define open half-spaces as

$$
\overset{\circ}{H}_{\gamma}^+(x_0) = \{ x \in X : P_{\gamma}(x) = \gamma(t) \text{ for } t > t_0 \} \\
\overset{\circ}{H}_{\gamma}^-(x_0) = \{ x \in X : P_{\gamma}(x) = \gamma(t) \text{ for } t < t_0 \},
$$

while closed half-spaces are defined as

$$
H_{\gamma}^+(x_0) = \{ x \in X : P_{\gamma}(x) = \gamma(t) \text{ for } t \geq t_0 \} \\
H_{\gamma}^-(x_0) = \{ x \in X : P_{\gamma}(x) = \gamma(t) \text{ for } t \leq t_0 \}.
$$

**Remark 4.2.5.**

1. Here “open” and “closed” refer to these being open or closed in the metric topology;

2. From the above definition we can see that $H_{\gamma}(x_0) = H_{\gamma}^+(x_0) \cap H_{\gamma}^-(x_0)$;

3. $H_{\gamma}^+(x_0) = H_{\gamma}^-(x_0)$ where $\gamma^-(t) = \gamma(l-t)$.

It is clear that the half-spaces defined above are convex if and only if $X$ has property $N$ (see section 3.5.3). Figure 4.1 illustrates how convexity can fail in the absence of property $N$.

4.2.2 A Separation Theorem in CAT(0) spaces

In normed linear spaces the Hahn-Banach theorem, in its geometrical form, states that a closed and convex set can be separated from any external point by means of a hyperplane. This principle has numerous uses in convexity, geometry, optimization theory and economics to mention only a few.

We now look at the theorem in a CAT(0) setting.
4.2 Convex sets in CAT(0) spaces

Figure 4.1: Potential lack of convexity of $H^+_\gamma(x_0)$ in the absence of property N

Theorem 4.2.6. Let $(X, d)$ be a CAT(0) space and $C$ a closed convex subset of $X$. For $x \not\in C$ there exists a hyperplane $H$ with $x \in H^-$ and $C \subseteq H^+$.

Figure 4.2: A separation theorem

Proof. Now for $x \not\in C$ we write $m = \frac{1}{2}x + \frac{1}{2}P_C(x)$ where $P_C(x)$ is the nearest point projection of $x$ on $C$. Then from the previous subsection we have $x \in H^-_{[x, P_C(x)]}(m)$.

Now suppose there exists $c \in C$ with $y = P_{[x, P_C(x)]}(c) \in [x, m]$. From 3.5.2 part(2) we know $P_C(y) = P_C(x)$. Now we also know that $d(y, c) < d(c, P_C(x))$ otherwise
4.2 Convex sets in $\text{CAT}(0)$ spaces

$d(c, P_C(x))$ would be the shortest distance from $c$ to $[x, P_C(x)]$. So

$$d(y, P_C(x)) < d(y, c) < d(c, P_C(x)).$$

From Proposition 3.5.2 part (3) and the comparison triangle for $\Delta(c, y, P_C(x))$ we have

$$\angle_{P_C(x)}(\bar{c}, \bar{y}) \geq \angle_{P_C(x)}(c, y) \geq \frac{\pi}{2},$$

which implies that

$$d(y, c) \geq d(c, P_C(x)), d(y, P_C(x))$$

and so we have a contradiction. So $C \subseteq H^+_{[x, P_C(x)]}(m)$. \qed

From this theorem we have the following Corollary

**Corollary 4.2.7.** For $C$ a closed convex subset of a $\text{CAT}(0)$ space $X$ we have $C = \bigcap H^+_\gamma(z)$ where the intersection is taken over all $\gamma$ and $z$ such that $C \subseteq H^+_{\gamma}(z)$.

4.2.3 Hyperplane characterization of weak convergence

![Hyperplane characterization of weak convergence](image)

**Figure 4.3:** Hyperplane characterization of weak convergence

For each $y \in X$, $y \neq x$ and writing $m = \left(\frac{1}{2}x \oplus \frac{1}{2}y\right)$, let $\gamma_y : [0, d(x, y)] \to X$ be the geodesic with $\gamma(0) = x$ and $\gamma(d(x, y)) = y$ then

$$x \in H^+_\gamma(m) \text{ and } x_n \to x \quad (\Rightarrow P_\gamma(x_n) \to x)$$
4.3 Convex functions in CAT(0) spaces

implies eventually

\[ x_n \in H^+_{\gamma_y}(m) \]

and conversely (see Proposition 3.6.9).

4.3 Convex functions in CAT(0) spaces

4.3.1 Background

Convex functions play an important role in many areas of mathematics. They are especially important in the study of optimization problems where they present a number of convenient properties. For example, a (strictly) convex function on an open set has no more than one minimum and enjoy a number of generic continuity and differentiability conditions. Even in infinite-dimensional spaces, under suitable additional hypotheses, convex functions continue to satisfy such properties and, as a result, they are the most understood functionals in the calculus of variations.

**Definition 4.3.1.** Let \( X \) be a convex set in a real vector space then a function \( f : X \to (-\infty, \infty] \) is called convex if:

\[
f(tx_1 + (1 - t)x_2) \leq tf(x_1) + (1 - t)f(x_2) \quad \text{for all} \quad x_1, x_2 \in X \quad \text{and} \quad t \in [0, 1].
\]

4.3.2 Convex functions in CAT(0) spaces

For \( X \) a geodesic metric space we say \( f : X \to (-\infty, \infty] \) is convex if for \( x, y \in X \) and \( \gamma_{xy} \) a geodesic joining them, we have

\[
f(\gamma_{xy}(1-t)d(x,y)) \leq tf(x) + (1-t)f(y),
\]

so in the case \( X \) is a CAT(0) space

\[
f(tx \oplus (1-t)y) \leq tf(x) + (1-t)f(y) \quad \text{for all} \quad x, y \in X.
\]
4.3 Convex functions in CAT(0) spaces

The effective domain of $f$ is defined as

$$\text{dom}(f) = \{x : f(x) < \infty\}$$

and the epigraph of $f$ is defined as

$$\text{epi}(f) = \{(x, \xi) : x \in X, \xi \geq f(x)\}.$$ 

Sublevel sets of $f$ are defined as $\text{sub}_\beta(f) = \{x \in X : f(x) \leq \beta\}$.

The following are readily seen

1. A function $f : X \to (-\infty, \infty]$ is convex if and only if $\text{epi}(f)$ is a convex subset of $X \times (-\infty, \infty)$,
2. If $f$ is convex then $\text{sub}_\beta(f)$ is convex for all $\beta$,
3. If $f$ is convex and continuous then $\text{epi}(f)$ is closed,
4. $\text{epi}(f)$ is closed if and only if $\text{sub}_\beta(f)$ is closed for all $\beta$ (as projections in CAT(0) spaces are nonexpansive).

**Definition 4.3.2.** A function $f : X \to (-\infty, \infty]$ is lower semicontinuous (lsc) if for $x_n \to x$ then

$$\liminf_{n \to \infty} f(x_n) \geq f(x).$$

**Lemma 4.3.3.** $\text{sub}_\beta(f)$ is closed for all $\beta$ if and only if $f$ is lower semicontinuous.

**Proof.** ($\Leftarrow$) Let $f$ be lower semicontinuous and $x_n \in \text{sub}_\beta(f)$ and $x_n \to x$. Then

$$\beta \geq \liminf_{n \to \infty} f(x_n) \geq f(x)$$

then $x \in \text{sub}_\beta(f)$, so $\text{sub}_\beta(f)$ is closed.

($\Rightarrow$) We will show that $f$ not lower semicontinuous implies there exists $\beta$ such that $\text{sub}_\beta(f)$ is not closed.

If $f$ is not lower semicontinuous then there exists $x_n \to x$ with

$$l = \liminf_{n \to \infty} f(x_n) < f(x).$$

We choose $\beta$ such that $l < \beta_1 < \beta_2 < f(x)$. 

65
4.3 Convex functions in CAT(0) spaces

So there exists a subsequence $(x_{n_k})$ with

$$\lim_{k \to \infty} f(x_{n_k}) = \liminf_{n \to \infty} f(x_n) = l < \beta_1.$$ 

So there also exists $K$ such that $k \geq K$ and $f(x_{n_k}) < \beta_2$ so $x_{n_k} \in \text{sub}_{\beta_2}(f)$, but $(x_{n_k})$ for $k > K$ is a subset of $(x_n)$, so $x_{n_k} \to x$, but $x \not\in \text{sub}_{\beta_2}(f)$.

So $\text{sub}_{\beta_2}(f)$ is not closed.

Example 4.3.4.

In a CAT(0) space:

(i) $f(x) = d(x, x_0)$ where $x_0$ is a point in $X$ is a convex function. More generally,

(ii) the distance function $d_C : X \to \mathbb{R}$ to a closed convex nonempty set $C$ is a convex function. This is a corollary of Proposition 3.5.2, see Proposition 4.3.5 below,

(iii) for geodesics $\gamma_1 : [0, l_1] \to X$, $\gamma_2 : [0, l_2] \to X$ let $f(t) = d(\gamma_1(tl_1), \gamma_2(tl_2))$ then $f$ is a convex function on $[0, 1]$.

(iv) The indicator function $(\text{ind}_C(x) = 0$ if $x \in C$ and $\infty$ otherwise) is convex precisely when $C$ is a closed convex set.

(v) Displacement functions for isometries [[20], Definition II.6.1]. Let $T : X \to X$ be an isometry. The displacement function of $T$ is the function $d_T : X \to [0, \infty)$ defined by $d_T(x) = d(x, Tx)$. It is convex and Lipschitz.

(vi) Busemann functions [[20], Definition II 8.7]. Let $\gamma : [0, \infty) \to X$ be a geodesic ray. The function $b_\gamma : X \to \mathbb{R}$ defined by

$$b_\gamma(x) = \lim_{t \to \infty} [d(x, \gamma(t)) - t], \quad x \in X$$

is called the Busemann function associated to the ray $\gamma$. Busemann functions are convex and 1-Lipschitz. Concrete examples of Busemann functions are given in [[20], p. 273].
4.3 Convex functions in CAT(0) spaces

We now present a corollary to Proposition 3.5.2 in which Example 1 is proved to be convex.

**Proposition 4.3.5.** For $X$ a CAT(0) space, let $C$ be a complete convex subset in $X$ and $d_C$ the distance function to $C$, that is $d_C(x) = d(x, C)$. Then:

1. $d_C$ is a convex function;
2. $|d_C(x) - d_C(y)| \leq d(x, y)$;
3. the restriction of $d_C$ to a sphere with centre $x$ and radius $r \leq d_C(x)$ attains its infimum at a unique point $y$ and $d_C(x) = d_C(y) + r$.

**Proof.** For each part let $P_C$ be the projection of $X$ onto $C$.

1. Let $\gamma_1 : [0, l_1] \to X$ be a geodesic segment and $\gamma_2 : [0, l_2] \to C$ be the geodesic segment between $P_C(\gamma_1(0))$ and $P_C(\gamma_1(l_1))$. Since the distance function is convex, (Proposition 2.3.20), we have

   \[
   d_C(\gamma_1(tl_1)) \leq d(\gamma_1(tl_1), \gamma_2(tl_2)) \\
   \leq (1 - t)d(\gamma_1(0), \gamma_2(0)) + td(\gamma_1(l_1), \gamma_2(l_2)) \\
   = (1 - t)d_C(\gamma_1(0)) + td_C(\gamma_1(l_1)).
   \]

2. For $x, y \in X$ we have

   \[
   d_C(x) \leq d(x, P_C(y)) \\
   \leq d(x, y) + d(y, P_C(y)) \\
   = d(x, y) + d_C(y)
   \]

   and part (2) follows.

3. Let $y$ be the point on the geodesic segment $[x, P_C(x)]$ such that $d(x, y) = r$ (see Figure 4.4). Then by Proposition 3.5.2 part (2) we have $d_C(x) = d_C(y) + r$. If we assume there exists $y'$ such that $d(x, y') = r$ and
4.3 Convex functions in CAT(0) spaces

Figure 4.4: Infimum at a unique point

\[ d_C(y') \leq d_C(y) \text{ then} \]
\[ d(x, P_C(y')) \leq d(x, y') + d(y', P_C(y')) \]
\[ = r + d_C(y') \]
\[ = d(x, P_C(x)) \]

so \( P_C(y') = P_C(x) \) and \( y' = y \).

\[ \square \]

Lemma 4.3.6. The distance function \( d_C \) (4.3.5) is weakly lower semicontinuous.

Proof. Let \( (x_n) \subset X, x \in X \) and \( x_n \to x \). Suppose that
\[ \liminf_{n \to \infty} d_C(x_n) < d_C(x). \]
That is, there exists a subsequence \( (x_{n_k}) \), \( k_0 \in \mathbb{N} \) and \( \delta > 0 \) such that \( d_C(x_{n_k}) < d_C(x) - \delta \) for all \( k > k_0 \). By continuity and convexity of the distance function, we have
\[ d_C(y) \leq d_C(x) - \delta \]
for all \( y \in \overline{\{x_{n_k} : k > k_0\}} \). But this, via Lemma 4.2.2, gives a contradiction to \( x_n \to x \). \( \square \)
4.3.3 Tangent cones to points of a geodesic metric space

Our definition of the tangent cone comes from [[20], Def. I 5.6 and Def. II 3.18]. The Euclidean cone $C_0 M$ over a metric space $(M, \rho)$ is by definition the quotient of $[0, \infty) \times M$ by the equivalence

$$(t_1, y_1) \equiv (t_2, y_2) \quad \text{if} \quad [t_1 = t_2 = 0] \text{ or } [t_1 = t_2, y_1 = y_2].$$

We denote the equivalence class $(t, y)$ by $ty$ and the class $(0, x)$ is denoted $0$ and is called the vertex of the cone. Let $\rho_\pi(y_1, y_2) = \min \{\pi, \rho(y_1, y_2)\}$ for $y_1, y_2 \in M$. The distance $D$ between two points $x_1 = t_1 y_1$ and $x_2 = t_2 y_2$ in $C_0 M$ is defined so that $D(x_1, x_2) = t_1$ if $x_2 = 0$ and so that the comparison angle $\bar{\angle}(x_1, x_2) = \rho_\pi(y_1, y_2)$. This is achieved by putting

$$D(x_1, x_2)^2 = t_1^2 + t_2^2 - 2t_1 t_2 \cos(\rho_\pi(y_1, y_2)).$$

Think of $(M, \rho)$ as isometrically embedded in a Euclidean sphere (with “great circle” metric) and extended by dilation to $(X, D)$. Then $(C_0 M, D)$ is a metric space which is complete if and only if the original metric space $(M, d)$ is complete. (see [[20], I 5.6])

We are now able to define the tangent cone.

Let $(X, d)$ be a geodesic metric space and define two non-trivial geodesics $\gamma_1$ and $\gamma_2$ issuing from the same point $p \in X$ to have the same direction at $p$ if the Alexandrov angle between them is zero, in which case we write $\gamma_1 \sim \gamma_2$, then $\sim$ is an equivalence relation on the set of all geodesics emanating from $p$ and the Alexandrov angle between any two of them is a metric $\rho$ on the equivalence classes of $\sim$. We denote this metric space by $S_p(X)$ and call it the space of directions at $p$. We note that the space of directions need not be complete even if $X$ is. The Euclidean cone $C_0 S_p(X)$ over $S_p(X)$ is called the tangent cone at $p$. We will use the notation $T_p(X) = C_0 S_p(X)$. If $x \in X$ then $\tau(x) = d(x, p)[\gamma_{p_2}]_\sim$ will denote the corresponding element in the tangent cone.
4.3 Convex functions in CAT(0) spaces

Theorem 4.3.7 (I.G. Nikolaev, [20], Theorem II 3.19). Let $X$ be a CAT(0) space, then the completion of the space of directions at each point $p \in X$ is a CAT(1) space, and the completion of the tangent cone at $p$ is a CAT(0) space.

Given $x_1 = t_1 u_1$ and $x_2 = t_2 u_2 \in T_p(X)$, we define their scalar product as

$$\langle x_1, x_2 \rangle_p = t_1 t_2 \cos(\angle_p(x_1, x_2)).$$

This natural definition first appeared in [[72], Definition 2.2]. It is nice to have $d(x, 0) = \sqrt{\langle x_1, x_2 \rangle_p}$ for $x \in T_p(X)$ and

$$\langle \lambda x_1, x_2 \rangle_p = \langle x_1, \lambda x_2 \rangle_p = \lambda \langle x_1, x_2 \rangle_p \quad \text{for} \quad \lambda \geq 0.$$

The sum $x_1 + x_2$ is defined as the midpoint of $[x_1, x_2]$.

4.3.4 Differentiability of Convex Functions

In Subsection 2.3.5 we introduced the geodesic extension property (GEP) and we assume here that our CAT(0) spaces have the GEP. Later, we will weaken this assumption. Further, we will see in the next example that convex continuous functions on CAT(0) spaces need not be locally Lipschitz, so we will often have to tacitly assume this property.

Example 4.3.8. In a CAT(0) space $X$, let $\gamma_n$, with $n \in \mathbb{N}$, be a set of geodesic rays issuing from a common point $x_0$. This forms a complete CAT(0) space $X_0$ whose tangent cone at $x_0$ can be identified with $X$ itself. Define $f : X \to \mathbb{R}$ so that $f(x_0) = 0$ and $f$ is linear (in an obvious sense) on each $\gamma_n$, for $n \in \mathbb{N}$, having the slope $n$ on $\gamma_n$ for $n = 1, 2, \cdots$. Then $f$ is continuous and convex, but not locally Lipschitz at $x_0$. 

4.3 Convex functions in CAT(0) spaces

**Lemma 4.3.9.** Suppose a CAT(0) space $X$ has the geodesic extension property. Let $f : X \to \mathbb{R}$ be convex continuous, $x_0 \in X$ and $\gamma : [0, l] \to X$ a geodesic issuing from $x_0$. Then the one-sided derivative of $f$ at $x_0$ along $\gamma$ defined as

$$d^+ f(x_0, \gamma) = \inf_{t > 0} \frac{f(\gamma(t)) - f(\gamma(0))}{t}$$

exists.

**Lemma 4.3.10.** Suppose a CAT(0) space $X$ has the geodesic extension property. Let $f : X \to \mathbb{R}$ be convex continuous and locally Lipschitz at $x_0 \in X$. Let $\gamma_1 : [0, l_1] \to X$ and $\gamma_2 : [0, l_2] \to X$ be two geodesic rays, both issuing from $x_0$ with the Alexandrov angle $\angle_{x_0}(\gamma_1, \gamma_2) = 0$. Then

$$d^+ f(x_0, \gamma_1) = d^+ f(x_0, \gamma_2).$$

**Proof.** For the proof we refer to [[20], Proposition II 3.1].

We note here that we are able to use Lemma 4.3.10 to define directional derivatives.

**Definition 4.3.11.** We say a convex continuous function $f : X \to \mathbb{R}$ is differentiable at a point $x_0 \in X$ in direction $v \in S_{x_0}(X)$ if $f$ has a one-sided derivative at $x_0$ along some $\gamma \in v$. We note that by Lemma 4.3.10 this does not depend on the choice of $\gamma$.

**Definition 4.3.12.** A convex continuous function $f : X \to R$ is Gâteaux differentiable at a point $x_0$ if

$$\lim_{t \to 0^+} \frac{f(\gamma(t)) + f(\gamma(-t)) - 2f(x_0)}{t} = 0$$

for any geodesic $\gamma : [-l, l] \to X$ with $\gamma(0) = x_0$. If the limit is uniform with respect to $\gamma$, then $f$ is Fréchet differentiable at $x_0$. 

71
Proposition 4.3.13. Suppose a convex continuous function \( f : X \to \mathbb{R} \) is Gâteaux differentiable at a point \( x_0 \). Then the (two-sided) limit

\[
\lim_{t \to 0} \frac{f(\gamma(t)) - f(\gamma(0))}{t}
\]

exists for any geodesic \( \gamma : [-l, l] \to X \) with \( \gamma(0) = x_0 \).

Proposition 4.3.14. Let \( f : X \to \mathbb{R} \) be a convex continuous function and \( x_0 \in X \). If \( f \) is Gâteaux differentiable at \( x_0 \), then it is differentiable at \( x_0 \) in each direction \( v \in S_{x_0}(X) \). If, moreover, \( f \) is Fréchet differentiable, then the directional derivatives are uniform with respect to \( v \in S_0(X) \).

4.3.5 Subdifferentials

Definition 4.3.15. Let \( X \) be a CAT(0) space and let \( f : X \to \mathbb{R} \) be a convex function and \( x_0 \in X \). The subdifferential of \( f \) at \( x_0 \) is the set

\[
\partial f(x_0) = \{ v \in T_{x_0}(X) : \langle v, \tau(x) \rangle_{x_0} \leq f(x) - f(x_0) \text{ for all } x \in X \}.
\]

Theorem 4.3.16. Let \( X \) be a complete CAT(0) space, \( x_0 \in X \) and let \( f : X \to (-\infty, \infty] \) be a locally Lipschitz convex function. Then \( \partial f(x_0) \neq \emptyset \)

Proof. We first prove that, for any \( \epsilon > 0 \), the \( \epsilon \)-subdifferential, of \( f \) at \( x_0 \) defined by

\[
\partial_\epsilon f(x_0) = \{ v \in T_{x_0}(X) : \langle v, \tau(x) \rangle_{x_0} \leq f(x) - f(x_0) + \epsilon \text{ for all } x \in X \}
\]

is nonempty. To do so, we let \( p = (x_0, f(x_0) - \epsilon) \in X \times \mathbb{R} \) and \( q = P_{\text{epi}(f)}(p) \in X \times \mathbb{R} \), then

\[
\tau(x_\epsilon) \in \partial_\epsilon f(x_0) \text{ where } q = (x_\epsilon, f(x_\epsilon)),
\]

in particular, \( \partial_\epsilon f(x_0) \neq \emptyset \).

Although \( \partial_\epsilon f(x_0) \) needn’t be convex in general (see Example 4.3.18 below) each
4.3 Convex functions in CAT(0) spaces

connected component of $\partial f(x_0)$ is closed and convex and so also weakly closed. Furthermore, the convexity of each component of $\partial f(x_0)$ also implies boundedness and hence each component of $\partial f(x_0)$ is weakly compact. We can thus conclude that

$$\partial f(x_0) = \bigcap_{\epsilon > 0} \partial_\epsilon f(x_0)$$

is nonempty. \hfill \qed

We note that a convex function $f$ attains its minimum at $x_0$ if and only if $0 \in \partial f(x_0)$.

**Example 4.3.17.** Let $\gamma_n$ with $n \in \mathbb{N}$ be geodesic rays issuing from a common point $x_0 = \gamma_n(0)$. This forms a complete $CAT(0)$ space $X$ whose tangent cone at $x_0$ can be identified with $X$ itself. Define

$$f : X \to \mathbb{R} \text{ so } f(\gamma_n(t)) = \begin{cases} -t, & n = 1, \\ nt, & n = 2, 3, \ldots. \end{cases}$$

Thus $f$ is continuous and convex on $X$. Clearly,

$$\partial f(x_0) = \bigcup_{n=2}^{\infty} I_n,$$

where $I_n$ is the interval $[1, n]$ on $\gamma_n$, for each $n = 2, \ldots$. Then we can see that $\partial f(x_0)$ is not convex, not bounded and not weakly closed.

**Example 4.3.18.** Let $\gamma_n$, with $n \in \mathbb{N}$, be geodesic rays issuing from a common point $x_0$. This forms a complete $CAT(0)$ space $X$ whose tangent cone at $x_0$ can be identified with $X$ itself. Define $f : X \to \mathbb{R}$ so that $f(x_0) = 0$ and $f$ is linear (in an obvious sense) on each $\gamma_n$, for $n \in \mathbb{N}$, having the slope $-1$ on $\gamma_1$ and the slope $2$ on $\gamma_n$ for $n = 2, \ldots$. Then $f$ is continuous, convex and Lipschitz. Clearly,

$$\partial f(x_0) = \bigcup_{n=2}^{\infty} I_n,$$

where $I_n$ is the interval $[1, n]$ on $\gamma_n$, for each $n = 2, \ldots$. In summary $\partial f(x_0)$ is bounded, but not convex nor weakly closed. Similarly for an $\epsilon$-subdifferential.
4.4 Conclusion

In this chapter we have studied convex sets and convex functions in CAT(0) spaces and presented a number of new concepts. We began by introducing notions of hyperplanes and half spaces in CAT(0) spaces and then introduced a separation theorem (theorem 4.2.6). This was followed by a hyperplane characterization of weak convergence in CAT(0) spaces. We finished with a consideration of the differentiability of convex functions and the concept of subdifferentials in CAT(0) spaces.

4.5 Historical Notes

Hermann Minkowski (1864–1909) made the first systematic study of convexity. His work in [79] contains, at least in germinal form, most of the important ideas of the topic. He was born in what is now Lithuania, but his parents were German. They settled in Königsberg where while still at the Gymnasium he was reading the work of Dedekind, Dirichlet and Gauss. He studied at the University of Königsberg where he became close friends with Hilbert who was an undergraduate there at the same time. Minkowski settled in Göttingen in 1902 where he accepted a chair that he held for the rest of his life (cut short when he died suddenly from a ruptured appendix).

It was at Göttingen that he became interested in mathematical physics, gaining enthusiasm from Hilbert and his associates. He developed a new view of space and time and laid the mathematical foundation of the theory of relativity. By 1907 Minkowski realised that the work of Lorentz and Einstein could be best understood in a non-euclidean space. He considered space and time, which were formerly thought to be independent, to be coupled together in a four-dimensional ‘space-time continuum’ of negative curvature.

The research of Victor Klee (1926–2007) covered a wide range of interests that included convexity, functional analysis, optimisation, the theory of algorithms and various aspects of combinatorics and geometry. Klee was born in San Francisco, USA, and spent almost his entire career at the University of Washington, Seattle.
His bibliography makes it clear that by the late 1960’s he had already a career’s worth of papers in continuous and infinite convexity. He received many honours, including a Guggenheim Fellowship, the Ford Award (1972) and the Allendoerfer Award (1980 and 1999). He was a Fellow of the American Academy of Arts and Sciences and Fellow of the American Association for the Advancement of Science.

In bibliographic notes from [85], Wayne Roberts and Dale Varberg say that the recognition of convex functions as a class of functions to be studied is generally traced to J. L. Jensen (1859–1925), [57] and [58], but as is sometimes the case, earlier work can be cited that anticipated what was to come. In this case O. Hölder (1859–1937), in [54], proved that if $f'' \geq 0$, then $f$ satisfied what later came to be known as Jensen’s inequality.

Johan Ludwig Jensen was born in Denmark, but undertook his early schooling in Sweden while his father worked there. After the position in Sweden finished the family returned to Denmark and Jensen completed his schooling in Copenhagen. He attended the College of Technology where mathematics began to dominate his life. Jensen was essentially self taught in research level mathematics and never held an academic post. He accepted a position working with a telephone company so he could support himself while he continued his research. He continued working there until the year before he died. He contributed to the Riemann Hypothesis as well as studying infinite series, the gamma function and inequalities for convex functions.
In 1906, Jensen published a paper, [58], in which he proved an inequality for convex functions which had a whole host of classical inequalities as special cases.

Figure 4.6: J. L. Jensen and O. Hölder

Otto Hölder was a German mathematician born in Stuttgart. Hölder first studied at the University of Stuttgart and then in 1877 went to Berlin where he was a student of Leopold Kronecker, Karl Weierstraße and Ernst Kummer. He is famous for many things including Hölder’s inequality and the Jordan-Hölder theorem.
5

Fixed Point Theory in CAT(0) spaces

5.1 Introduction

CAT(0) spaces have many properties in common with uniformly convex Banach spaces. For example, closed convex sets are uniquely proximinal, descending sequences of nonempty bounded closed convex sets have nonempty intersection and "asymptotic centre" techniques apply. They also share many properties with Hilbert spaces, for example nearest point projections onto closed convex sets are nonexpansive and a notion of angle exists for which there is a law of cosines. As well the family of all bounded closed convex subsets of a given CAT(0) space is normal in the sense of [67].

5.2 Fixed point theory in CAT(0) spaces

Although we are lacking a “weak topology”, we do have the notion of weak sequential convergence (see subsections 3.6.2 and 3.6.3) which as we shall see throughout this chapter proves sufficient for us to develop a fixed point theory for nonexpansive type mapping which parallels that in Hilbert space. We begin by recalling our
definition of nonexpansive mapping:

**Definition 5.2.1.** For a metric space \((X,d)\), a mapping \(T : X \to X\) is nonexpansive if

\[ d(T(x), T(y)) \leq d(x, y), \quad x, y \in X. \]

It is a contraction if the inequality is strict for \(x \neq y\) and a uniform contraction if

\[ d(Tx, Ty) \leq kd(x, y) \]

for all \(x, y \in X\) and some \(k \in [0, 1)\).

The first metric fixed point theorem, generally known as the Banach Contraction Principle, appeared in explicit form in Banach’s 1922 thesis where it was used to establish the existence of a solution for an integral equation.

**Theorem 5.2.2.** *(Banach’s Contraction Principle)* Let \((X,d)\) be a complete metric space and let \(T : X \to X\) be a uniform contraction. Then \(T\) has a unique fixed point in \(X\) and for each \(x_0 \in X\) the sequence of iterates \(\{T^n(x_0)\}\) converges to this fixed point.

Felix Browder’s 1968 notion of demiclosedness at \(w\) for a Banach space, that is \(x_n \rightharpoonup x\) and \((I-T)x_n \to w\) in norm implies \(Tx = w\), is basic to much of metric fixed point theory (especially the case when \(w = 0\)). While “\((I-T)x_n \to w\) in norm” does not extend to a geodesic metric space, when \(w = 0\) it can be rephrased as \(d(x_n, Tx_n) \to 0\). From which follows the analogue of demiclosed with \(w = 0\) for nonexpansive maps.

**Proposition 5.2.3.** If \(X\) is a CAT(0) space and \(T : C \to C\) is a nonexpansive mapping, then \(x_n \rightharpoonup x\) and \(d(x_n, T(x_n)) \to 0\) implies \(Tx = x\).

**Proof.**

\[
\liminf_{n \to \infty} d(x_n, T(x)) \leq \liminf_{n \to \infty} d(x_n, T(x_n)) + \liminf_{n \to \infty} d(T(x_n), T(x)) \\
\leq \liminf_{n \to \infty} d(x_n, x).
\]

So by Opial’s property (see subsection 3.6.2) we must have \(T(x) = x\). 

Lemma 5.2.4. Let \( C \) be a bounded closed convex subset of a complete \( \text{CAT}(0) \) space \( X \). Suppose \( T \) is a nonexpansive mapping so that

\[
\inf \{ d(x, T(x)) : x \in C \} = 0
\]

then \( T \) has a fixed point in \( C \), that is there exists \( x \in C \) such that \( T(x) = x \).

Proof. From 5.2.0.1 there exists a sequence \( \{ x_n \} \subseteq C \) such that \( d(x_n, T(x_n)) \to 0 \). Since \( C \) is closed convex bounded there exists \( (x_{n_k}) \) such that \( x_{n_k} \to x \in C \) and still \( d(x_{n_k}, T(x_{n_k})) \to 0 \). The demiclosedness principle implies then that \( Tx = x \). \( \square \)

5.2.1 Approximate fixed point property

Definition 5.2.5. A subset \( C \) of a metric space is said to have the Approximate Fixed Point Property (AFPP for short) for nonexpansive mappings if for any nonexpansive \( T : C \to C \) we have \( \inf \{ d(x, T(x)) : x \in C \} = 0 \).

Bounded Sets. We first show closed bounded convex subsets of a \( \text{CAT}(0) \) space have the approximate fixed point property thereby providing a method for approximating fixed points of nonexpansive mappings in \( \text{CAT}(0) \) spaces. Let \( C \) be a bounded closed convex subset of a complete \( \text{CAT}(0) \) space and let \( T : C \to C \) be nonexpansive. Let us choose \( u \in C \) and for \( t \in (0, 1) \) define the mapping \( T_t : C \to C \) by taking \( T_t(x) = (1-t)u \oplus tT(x) \). The convexity of the metric leads to \( d(T_t(x), T_t(y)) \leq td(x, y) \), for \( x, y \in C \), so \( T_t \) is a contraction mapping of \( C \) into \( C \). Since \( C \) is complete, Banach’s contraction mapping theorem assures the existence of a unique point \( u_t \) such that;

\[
u_t \in [u, T(u_t)] \text{ and } d(u, u_t) = td(u, T(u_t)). \tag{5.2.1.1}\]

However, more can be said. Since \( T_t \to T \) uniformly on \( C \) as \( t \to 1 \) it follows that \( \lim_{t \to 1^-} d(u_t, T(u_t)) = 0 \). In particular \( u_n = u_{1-\frac{1}{n}} \) is an approximate fixed point sequence so we have the following proposition;
Proposition 5.2.6. Closed bounded convex subsets of a CAT(0) space have the approximate fixed point property.

Theorem 5.2.7 (Theorem 26, [64]). Let $C$ be a bounded closed convex subset of a complete CAT(0) space $X$, let $T : C \to C$ be nonexpansive, fix $u \in C$ and for each $t \in [0,1)$ let $u_t$ be the point of $[u, T(u_t)]$ satisfying 5.2.1.1. Then $\lim_{t \to 1^-} u_t$ converges to the unique fixed point of $T$ which is nearest $u$.

Proof. Fix $0 < k \leq l \leq 1$ and consider $\triangle(\bar{u}_0, \bar{T}(u_k), \bar{T}(u_l))$, the comparison triangle of $\triangle(u_0, T(u_k), T(u_l))$ in $\mathbb{H}^2$. By the CAT(0) inequality we have
\[
\|\bar{T}(u_k) - \bar{T}(u_l)\| = d(T(u_k), T(u_l)) \leq d(u_k, u_l)
\]
\[
\leq \|\bar{u}_k - \bar{u}_l\| = \|k^{-1}\bar{T}(u_k) - l^{-1}\bar{T}(u_l)\|.
\]
We now follow the argument of Halpern, [53], to conclude that
\[
\|\bar{u}_0 - \bar{u}_l\|^2 \geq \|\bar{u}_0 - \bar{u}_k\|^2 + \|\bar{u}_k - \bar{u}_l\|^2.
\]
Since the sequence $(\|\bar{u}_0 - \bar{u}_l\|^2)$ is monotone increasing we have
\[
d(u_k, u_l)^2 \leq \|\bar{u}_k - \bar{u}_l\|^2 \leq \|\bar{u}_0 - \bar{u}_l\|^2 - \|\bar{u}_0 - \bar{u}_k\|^2 \to 0 \text{ as } k, l \to \infty.
\]
Thus $(u_k)$ converges to some point $z \in C$. By continuity
\[
d(z, T(z)) = \lim_{k \to 1^-} d(u_k, T(u_k)) = (1 - k)d(u_0, T(u_k)) = 0,
\]
so $z$ is a fixed point of $T$.

Now let $p$ be any other fixed point of $T$. By repeating the preceding argument taking $u_l = u_1 = p$, we conclude
\[
d(u_0, p)^2 = \|\bar{u}_0 - \bar{p}\|^2 \geq \|\bar{u}_0 - \bar{u}_k\|^2 + \|\bar{u}_k - \bar{p}\|^2
\]
\[
\geq \|\bar{u}_0 - \bar{u}_k\|^2 + d(u_k, p)^2
\]
\[
= kd(u_0, u_k)^2 + d(u_k, p)^2.
\]
5.2 Fixed point theory in CAT(0) spaces

from which (letting $k \to 1^-$)

$$d(u_0, p)^2 \geq d(u_0, z)^2 + d(z, p)^2.$$ 

This proves that $z$ is the unique fixed point of $T$ which is nearest $u_0$. □

As a consequence of the previous theorem, if $\text{fix}(T)$ denotes the fixed point set of a nonexpansive map $T$ then $\text{fix}(T)$ is closed convex, and given any $x \in C$,

$$\lim_{t \to 1^-} x_t = P(x) \in \text{fix}(T),$$

where the mapping $P : x \mapsto P(x)$ is the nearest point projection of $C$ onto $\text{fix}(T)$ and so nonexpansive by Proposition 3.5.2. Therefore $\text{fix}(T)$ is a nonexpansive retract of $C$.

Now let $C$ be a closed convex subset of a CAT(0) space $X$ and let $T : C \to C$ be nonexpansive. For each $t \in (0, 1]$ we define the mapping $F_t : C \to C$ by taking $F_t(u) = u_t$ which leads to the next theorem.

**Theorem 5.2.8.** Suppose $X$ is a complete CAT(0) space. Then :

(i) Each of the mappings $F_t$ is nonexpansive.

(ii) $F_t$ and $T$ have the same fixed points.

(iii) If the fixed point set $\text{fix}(T)$ of $T$ is nonempty, then $\lim_{t \to 1^-} F_t(x)$ exists for each $x \in C$ and this limit is the point of $\text{fix}(T)$ which is nearest $x$.

**Proof.** Parts (i) and (ii) are direct computations and (iii) follows from Theorem 5.2.7. □

**Unbounded Sets.** Next we turn to considerations of the approximate fixed point property when the closed convex set is unbounded.
5.3 Averaged maps

**Definition 5.2.9.** For a metric space $X$ we say a curve $\gamma : X \to [0, \infty)$ is directional (with constant $b$) if there is $b \geq 0$ such that

$$t - s - b \leq d(\gamma(s), \gamma(t)) \leq t - s$$

for all $t \geq s \geq 0$ as $b \to 0$ or for $|s - t| >> 0$, that is $\gamma$ approximates a geodesic. A subset of $X$ is said to be (geodesically) directionally bounded if it does not contain a directional curve.

Shafrir, in [86], proved that a closed convex subset of a complete hyperbolic metric space has the approximate fixed point property for nonexpansive mappings if and only if it is directionally bounded. It follows as an immediate corollary of Shafrir’s result that a closed convex subset of a complete CAT(0) space with the geodesic extension property (subsection 2.3.5) has the approximate fixed point property if and only if it is directionally bounded and so we have the following theorem.

**Theorem 5.2.10** (Theorem 25 in [64]). A closed convex subset of a complete CAT(0) space with the geodesic extension property has the approximate fixed point property for nonexpansive mappings if and only if it does not contain a geodesic ray.

While in Hilbert space a closed convex set has the fixed point property if and only if it is bounded (see Ray [83] and Theorem 32.2 in [49]), we note that this is no longer true in a complete CAT(0) space; for example it is shown in [37] that a closed convex subset of an $R$-tree has the fixed point property for nonexpansive mappings if (and only if) it is geodesically bounded.

### 5.3 Averaged maps

Given a convex subset $C$ of a unique geodesic metric space and a mapping $T : C \to C$, the **averaged mapping** is

$$V(x) := \frac{1}{2} x \oplus \frac{1}{2} T x,$$
5.3 Averaged maps

which we sometimes suggestively write as $V = \frac{1}{2}I \oplus \frac{1}{2}T$. More generally one may consider

$$V_\alpha = (1 - \alpha)I \oplus \alpha T \quad \text{for} \quad \alpha \in (0, 1).$$

The exploitation of gains in regularity resulting from averaging can easily be traced back to Krasnosel’skii (1920–1997). Averaged maps play a significant role in the algorithms we will shortly introduce in connection with the feasibility problem: Find $x \in A \cap B$ where $A$ and $B$ are constraint sets.

Clearly the averaged map $V = \frac{1}{2}I \oplus \frac{1}{2}T$ and $T$ share the same fixed point set $[Tx = x \Leftrightarrow Vx = x]$, the same approximate fixed point sequences and if $T$ is nonexpansive so too is $V$.

5.3.1 Asymptotic regularity

**Definition 5.3.1.** A mapping $T : C \to C$ is said to be asymptotically regular if $d(x_n, Tx_n) \to 0$ as $n \to \infty$, where for a given $x_0$, $x_{n+1} = Tx_n$ ($n = 0, 1, 2, \cdots$); that is, the orbit $x_n = T^nx_0$ is an approximate fixed point sequence for $T$.

That the averaged map is asymptotically regular whenever $T$ is nonexpansive is a result due to Ishikawa in 1976, [56], and was independently discovered by Goebel and Kirk, [48], and Edelstein and O’Brien, [34]. The proof we give is an adaption of that found in [46] or [61].

**Theorem 5.3.2.** Let $C$ be a closed bounded convex subset of a CAT(0) space and suppose $T : C \to C$ is nonexpansive. Then the mapping $V(x) := \frac{1}{2}x \oplus \frac{1}{2}Tx$ is asymptotically regular.

**Proof.** We begin by selecting $x_0 \in C$, then having defined $x_n$ we take $y_n = V(x_n)$ and set $x_{n+1} = \frac{1}{2}(x_n \oplus y_n)$ for $n = 1, 2, \cdots$. The sequences $\{x_n\}$, $\{y_n\}$ satisfy

$$x_{n+1} = \frac{1}{2}(x_n + y_n) \quad (5.3.1.1)$$

$$d(x, y) \leq d(y_{n+1}, y_n). \quad (5.3.1.2)$$

83
and we \textbf{claim} for each \(i, n \in \mathbb{N}\)

\[(1 + \frac{n}{2})d(y_i, x_i) \leq 2^n \left[ d(y_i, x_i) - d(y_{i+n}, x_{i+n}) \right] + d(y_{i+n}, x_i), \quad (5.3.1.3)\]

from which the result is obtained as follows.

Since \(\{d(x_n, y_n)\} = \{d(x_n, V(x_n))\}\) is monotone nonincreasing, there exists a number \(r \geq 0\) such that \(\lim_{n \to \infty} d(x_{n+1}, V(x_{n+1})) = r\). Now letting \(i \to \infty\) in the above inequality, we have

\[(1 + \frac{n}{2})r \leq \text{diam}(C).\]

This clearly implies \(r = 0\), so \((x_n)\) is an approximate fixed point sequence for \(V\), hence the theorem holds.

\textbf{Proof of claim.} We use induction on \(n\). Note the claim is trivially true for all \(i\) if \(n = 0\), and we make the inductive assumption that the theorem holds for a given \(n \in \mathbb{N}\) and all \(i\). After replacing \(i\) with \(i + 1\) in the inequality we have

\[
\left(1 + \frac{n}{2}\right)d(y_{i+1}, x_{i+1}) \leq 2^n \left[ d(y_{i+1}, x_{i+1}) - d(y_{i+n+1}, x_{i+n+1}) \right]
+ d(y_{i+n+1}, x_{i+1}).
\]

We can now observe that by 5.3.1.1, the convexity of the metric and 5.3.1.2 respectively we have:

\[
d(y_{i+n+1}, x_{i+1}) \leq \frac{1}{2} \left[ d(y_{i+n+1}, x_i) + d(y_{i+n+1}, y_i) \right]
\leq \frac{1}{2} \left[ d(y_{i+n+1}, x_i) + \sum_{k=0}^{n} d(y_{i+k+1}, y_{i+k}) \right]
\leq \frac{1}{2} \left[ d(y_{i+n+1}, x_i) + \sum_{k=0}^{n} d(x_{i+k+1}, x_{i+k}) \right].
\]

Combined with the previous inequality, this gives

\[
\left(1 + \frac{n}{2}\right)d(y_{i+1}, x_{i+1}) \leq 2^n \left[ d(y_{i+1}, x_{i+1}) - d(y_{i+n+1}, x_{i+n+1}) \right]
+ \frac{1}{2} \left[ d(y_{i+n+1}, x_i) + \sum_{k=0}^{n} d(x_{i+k+1}, x_{i+k}) \right].
\]
5.3 Averaged maps

Thus

\[ 2 \left( 1 + \frac{n}{2} \right) d(y_{i+1}, x_{i+1}) \leq 2^{n+1} \left[ d(y_{i+1}, x_{i+1}) - d(y_{i+n+1}, x_{i+n+1}) \right] \]

\[ + \left[ d(y_{i+n+1}, x_i) + \sum_{k=0}^{n} d(x_{i+k+1}, x_{i+k}) \right]. \]

Next we observe that 5.3.1.1 and 5.3.1.2 imply

\[ d(y_{n+1}, x_{n+1}) \leq d(y_{n+1}, y_n) + d(y_n, x_{n+1}) \]

\[ \leq d(x_{n+1}, x_n) + d(y_n, x_{n+1}) \]

\[ = d(y_n, x_n). \]

This means that the sequence \( \{d(y_n, x_n)\} \) is monotone nonincreasing. Thus

\[ \sum_{k=0}^{n} d(x_{i+k+1}, x_{i+k}) = \frac{1}{2} \sum_{k=0}^{n} d(y_{i+k}, x_{i+k}) \leq \frac{n+1}{2} d(y_i, x_i), \]

and we have

\[ 2 \left( 1 + \frac{n}{2} \right) d(y_{i+1}, x_{i+1}) \leq 2^{n+1} \left[ d(y_{i+1}, x_{i+1}) - d(y_{i+n+1}, x_{i+n+1}) \right] \]

\[ + \left[ d(y_{i+n+1}, x_i) + \frac{(n+1)}{2} d(y_i, x_i) \right] \]

\[ = 2^{n+1} \left[ d(y_i, x_i) - d(y_{i+n+1}, x_{i+n+1}) \right] \]

\[ + 2^{n+1} d(y_{i+1}, x_{i+1}) \]

\[ + \left[ \frac{(n+1)}{2} - 2^{n+1} \right] d(y_i, x_i) + d(y_{i+n+1}, x_i). \]

Since \( 2 \left( 1 + \frac{n}{2} \right) - 2^{n+1} \leq 0, \)

\[ \left[ 2 \left( 1 + \frac{n}{2} \right) - 2^{n+1} \right] d(y_{i+1}, x_{i+1}) + \left[ 2^{n+1} - \frac{(n+1)}{2} \right] d(y_i, x_i) \]

\[ \geq \left[ 2 \left( 1 + \frac{n}{2} \right) - 2^{n+1} \right] d(y_i, x_i) + \left[ 2^{n+1} - \frac{(n+1)}{2} \right] d(y_i, x_i) \]

\[ = \left[ 1 + \frac{(n+1)}{2} \right] d(y_i, x_i), \]
and we have
\[
\left[ 1 + \left( \frac{n+1}{2} \right) \right] d(y_i, x_i) \leq 2^{n+1} \left[ d(y_{i+1}, x_{i+1}) - d(y_{i+n+1}, x_{i+n+1}) \right] + d(y_{i+n+1}, x_i),
\]
which completes the induction.

**Remark 5.3.3.**

1. To simplify the numerics we have presented the result for \( \alpha = \frac{1}{2} \), but it remains true for \( V_\alpha = (1 - \alpha)I \oplus \alpha T, (\alpha \in (0, 1)) \), with the proof following the same steps.

2. By paying greater attention to detail it can be seen that the rate of convergence of \( d(x_n, V(x_n)) \), or \( d(x_n, T(x_n)) \), to 0 is uniform with respect to both the choice of \( x_0 \in C \), [34], and of the particular nonexpansive self map \( T \) of \( C \) which is chosen. [46].

In the case \( X \) is an R-tree, the orbit \( (x_n) \) actually converges to a fixed point of \( T \).

**Proposition 5.3.4** (Theorem 4.7, [37]). Let \((X, d)\) be a complete R-tree. Suppose \( C \) is a closed convex geodesically bounded subset of \( X \) and suppose \( T : C \to C \) is nonexpansive. Fix \( t \in (0, 1) \) and define \( V_t : C \to C \) by
\[
V_t(x) = (1 - t)x \oplus tT(x),
\]
then \( \{V_t^n(x)\} \) converges to a fixed point of \( T \) for each \( x \in C \).

### 5.3.2 Firmly nonexpansive maps

For \( C \) a closed convex subset of a Hilbert space \( H \), a mapping \( T : C \to H \) is said to be **firmly nonexpansive** if any one (and hence all of) the following conditions hold;

\begin{itemize}
  \item[(a)] \( \|T(x) - T(y)\|^2 + \|(I - T)x - (1 - T)y\|^2 \leq \|x - y\|^2 \) for all \( x, y \in C \);
  \item[(b)] \( \langle T(x) - T(y), x - y \rangle \geq \|T(x) - T(y)\|^2 \) for all \( x, y \in C \);
\end{itemize}
5.3 Averaged maps

(c) \[ \|T(x) - T(y)\| \leq \|(1 - t)x + tTx) - ((1 - t)y + tTy)\| \] for all \( t \in [0, 1] \) and all \( x, y \in C \);

(d) \[ T = \frac{1}{2}(I + V) \], where \( V : C \to H \) is nonexpansive;

(e) \[ 2T - I \] is nonexpansive.

Clearly such maps are nonexpansive and as we show below closest point projections onto closed convex sets are an example. From (e) it follows that reflections in closed convex sets are nonexpansive (a result key to the underlying theory for Douglas-Rachford methods). Firmly nonexpansive maps also play a key role in the theory of monotone operators on Hilbert spaces. See for example Bauschke and Combettes [12].

In the setting of general CAT(0) space only (c) and (d) make sense and we choose to use (c) (first introduced in the Banach space setting by R. Bruck, [24]) as the definition.

**Definition 5.3.5.** For \( X \) a CAT(0) space, \( T : X \to X \) is firmly nonexpansive if

\[ d(T(x), T(y)) \leq d(((1 - t)x \oplus tT(x)), ((1 - t)y \oplus tT(y))) \]

for all \( t \in [0, 1] \) and all \( x, y \in X \).

And, we have

**Theorem 5.3.6.** The metric projection onto a closed convex subset \( C \) of a CAT(0) space \( X \) is firmly nonexpansive.

**Proof.** Let \( x, y \) and \( t (0, 1) \). From Proposition 3.5.2 we have \( P_C(x) = P_C((1 - t)x \oplus tP_C(x)) \), then

\[ d(P_C(x), P_C(y)) = d(P_C((1 - t)x \oplus tP_C(x)), P_C((1 - t)y \oplus tP_C(y))) \]

\[ \leq d((1 - t)x \oplus tP_C(x), (1 - t)y \oplus tP_C(y)). \]

\[ \square \]
5.4 Fixed point Property in product spaces

We also have the following proposition, when \( F_t \) is the map introduced in subsection 5.2.1, defined by \( F_t : x \mapsto u_t \).

**Proposition 5.3.7.** \( F_t \) is a firmly nonexpansive mapping having the same set of fixed points as \( T \).

For the proof see Proposition 5.5 in [3].

We also have that (c) implies (d). Unfortunately this version of firmly nonexpansiveness does not allow us to deduce that reflections are nonexpansive in CAT(0) spaces. Could it, however, provide the beginnings of a theory of monotone operators on CAT(0) spaces?

5.4 Fixed point Property in product spaces

We now consider the product \( X \times Y \) of two CAT(0) spaces \( (X, d_X) \) and \( (Y, d_Y) \). If the metric, \( d \), which we assign to \( X \times Y \) is

\[
d((x,y),(u,v))^2 = d_X(x,u)^2 + d_Y(y,v)^2,
\]

then \( (X \times Y, d) \) is a CAT(0) space and so if \( X \) and \( Y \) are bounded, \( X \times Y \) has the fixed point property with respect to mappings that are nonexpansive mappings relative to \( d \).

We now consider the case \( (X \times Y)_\infty \) where the metric \( d_\infty \) is defined as

\[
d_\infty((x,y),(u,v)) = \max\{d_X(x,u),d_Y(y,v)\}.
\]

Here the product space \( (X \times Y) \) is no longer a CAT(0) space.

**Theorem 5.4.1** (Theorem 20, [65]). Let \( (X, d_X) \) be a metric space with the fixed point property for nonexpansive mappings and let \( Y \) be a complete and bounded CAT(0) space. Then every nonexpansive mapping \( T : (X \times Y)_\infty \to (X \times Y)_\infty \) has a fixed point.

If \( Y \) is an \( R \)-tree in the previous theorem, then we can relax the boundedness
assumption on $Y$, by requiring that $Y$ is geodesically bounded, that is, does not contain a geodesic ray.

We can now proceed to the next theorem which uses the result in [37].

**Theorem 5.4.2.** [Theorem 21 in [65]] Let $(X,d_X)$ be a metric space with the fixed point property for nonexpansive mappings and suppose $(Y,d_Y)$ is a complete $R$-tree which is geodesically bounded. Then every nonexpansive mapping $T : (X \times Y)_{\infty} \to (X \times Y)_{\infty}$ has a fixed point.

We also have a corollary from this theorem.

**Corollary 5.4.3.** Suppose $(X_i,d_i), i = 1, \cdots, N$, are complete geodesically bounded $R$-trees. Then every nonexpansive mapping

$$T : (X_1 \times \cdots \times X_N)_{\infty} \to (X_1 \times \cdots \times X_N)_{\infty}$$

has a fixed point.

**Proof.** The proof comes by induction using theorem 5.4.2. \qed

In this next result the CAT(0) space is not assumed to be complete.

**Remark 5.4.4.** Suppose $X$ is a metric space which has the approximate fixed point property for nonexpansive mappings and suppose $Y$ is a bounded convex subset of a CAT(0) space. Then $(X \times Y)_{\infty}$ has the approximate fixed point property. See Theorem 25 in [65] which mimics the one given in [38].

### 5.5 Locally nonexpansive mapping

A mapping $T$ defined on a metric $X$ is said to be locally nonexpansive if each point $x \in X$ has a neighbourhood $N_x$ such that the restriction of $T$ to $N_x$ is nonexpansive. The asymptotic regularity of $T$ at $x$ means that $\lim_{n \to \infty} d(T^n(x), T^{n+1}(x)) = 0$. Theorem 17 in [65] shows that the assumption that $T$ be nonexpansive (Theorem 16 in [65]) can be relaxed to $T$ locally nonexpansive; that is $T$ asymptotically regular is equivalent to its orbits being an approximate fixed point sequence. Given a
mapping $T : C \to C$, where $C$ is a subset of a metric space $X$, and a number $\epsilon > 0$, the $\epsilon$-fixed point set of $T$ is the set

$$F_\epsilon(T) = \{ x \in C : d(x, T(x)) \leq \epsilon \}.$$

**Theorem 5.5.1.** (Theorem 17, [65].) Let $D$ be a connected bounded open set in a complete CAT(0) space $X$, and let $T : \overline{D} \to X$ be a mapping that is continuous on $\overline{D}$ and locally nonexpansive on $D$. Then the following alternative holds:

1. $T$ has a fixed point in $D$, or;
2. $\inf \{ d(x, T(x)) : x \in \partial D \} \leq \inf \{ d(x, T(x)) : x \in D \}$.

**Proof.** For convenience we assume the $\text{diam}(D) \leq 1$, and we show that the assumption $\inf \{ d(x, T(x)) : x \in D \} < \inf \{ d(x, T(x)) : x \in \partial D \}$ implies that $T$ has a fixed point. Let $\zeta = \inf \{ d(x, T(x)) : x \in \partial D \}$. Thus by assumption $F_\epsilon(T) \neq \emptyset$ for some $\epsilon < \zeta$. The following is easy to verify.

1. If $z \in F_\epsilon(T)$ for $\epsilon < \zeta$ then the segment $[z, T(z)]$ lies entirely in $D \cap F_\epsilon(T)$ and is bounded away from the boundary of $D$.

Fix $\epsilon > 0, \epsilon < \zeta$, such that $F_\epsilon(T) \neq \emptyset$, for each $z \in F_\epsilon(T)$ let $g(z) = \frac{1}{2} z \oplus \frac{1}{2} T(z)$. Clearly $g : F_\epsilon(T) \to F_\epsilon(T)$. Now fix $z \in F_\epsilon(T)$, and observe that the sequence $\{z_n\} := \{ g^n(z) \}$ is well defined. By applying Theorem 15, [65] to $g$ we can conclude that $\lim_{n \to \infty} d(z_{n+1}, z_n) = 0$. It follows that $\lim_{n \to \infty} d(z_{n+1}, T(z_n)) = 0$. Now let $\epsilon_0 < \epsilon$ be chosen small enough that if $u$ and $v$ lie in a connected component $F_0$ of $F_{\epsilon_0}(T)$ then $\text{seg}[u, v] \subset D$, and define the sequence $\{ \epsilon_j \}$ by the relation $\epsilon_{j+1} = \sqrt{\epsilon_j^2 + 2 \epsilon_j}, j = 1, 2, \ldots$. (Thus $\epsilon_j = \sqrt{\epsilon_{j-1}^2 + 1 - 1}.$) For the remainder of the argument $F_0$ is a fixed component of $F_{\epsilon_0}(T)$. Thus if $z \in F_0$, then by connectedness $\{ g^n(z) \} \subset F_0$. Therefore by what we have just seen

$$F_0 \cap F_{\epsilon_n}(T) \neq \emptyset$$

for each $n = 0, 1, 2, \ldots$.

Next observe that if $u, v \in F_0 \cap F_{\epsilon_n}(T)$ for $n$ sufficiently large, then $\text{seg}[u, v] \subset F_0 \cap F_{\epsilon_n}(T)$. This proof can now be completed as in the proof of Theorem 13 in [65], upon replacing $KC$ with $F_0$. \qed
5.6 Asymptotically nonexpansive mapping

We look now at a class of mappings known as asymptotically nonexpansive maps. A mapping $T : M \to M$ of a metric space $X$ with $M \subseteq X$ is said to be asymptotically nonexpansive if there exists a sequence of real numbers $\{k_n\}$ with $k_n \to 1$ for which

$$d(T^n(x), T^n(y)) \leq k_n d(x, y) \quad \text{for each} \quad x, y \in M.$$  

Whereas nonexpansive mappings might be seen to arise naturally in the class of CAT(0) spaces (in particular as isometries or local isometries), the significance of the class of asymptotically nonexpansive mappings is not entirely clear. Theorem 6.5 in [36] shows that if $X$ is a bounded hyperconvex metric space and if $T : X \to X$ is asymptotically nonexpansive, then $F_\epsilon(T) \neq \emptyset$ for any $\epsilon > 0$, where $F_\epsilon(T) = \{x \in X : d(x, T(x)) \leq \epsilon\}$. In the special case below when $X$ is a complete R-tree, it is possible to conclude that $T$ actually has a fixed point. The facts needed are

(i) $X$ is an R-tree if and only if $X$ is CAT($\kappa$) for all $\kappa$ ([20], p. 167) and

(ii) the metric ultrapower ([61], p. 249 and [62], pp. 177–199) of a CAT($\kappa$) space is again a CAT($\kappa$) space ([20], p. 187).

Therefore the metric ultrapower of a complete R-tree is a complete R-tree. Our final theorem gives a much more general result.

**Theorem 5.6.1.** (Theorem 28 in [65].) Let $C$ be a bounded closed and convex subset of a complete CAT(0) space $X$. Then every asymptotically nonexpansive mapping $T : C \to C$ has a fixed point.

**Proof.** We begin with the following lemma

**Lemma 5.6.2** (Lemma 29 in [65]). Under the assumptions of Theorem 5.6.1, $F_\epsilon(T) \neq \emptyset$ for every $\epsilon > 0$.

The remainder of the proof follows theorem 28 in [65].
Corollary 5.6.3. Let $C$ be a nonempty bounded closed convex subset of an $R$-tree. Then every asymptotically nonexpansive mapping $T : C \to C$ has a fixed point.

5.7 Conclusion

In this chapter we vindicated the notion of weak sequential convergence introduced in Chapter 3 by demonstrating that it allows the development in CAT(0) spaces of a fixed point theory for nonexpansive type mappings analogous to that available for Hilbert spaces. We have also shown that closed bounded convex sets of a CAT(0) space have the approximate fixed point property which provided a method for approximating fixed points of nonexpansive mappings in such spaces. We studied averaged mappings and established that the averaged map is asymptotically regular and firmly nonexpansive whenever $T$ is nonexpansive. Whilst we established that metric projection onto a closed convex subset of a CAT(0) space is firmly nonexpansive in an appropriate sense, we were unable to deduce from this that reflections in CAT(0) spaces are nonexpansive, something which as we shall see fails in general.

5.8 Historical Notes

The origins of metric fixed point theory, which date to the later part of the nineteenth century, rest in the use of successive approximations to establish the existence and uniqueness of solutions to differential and integral equations. However, it is the Polish mathematician Stefan Banach [see section (2.5)] who is credited with placing the underlying ideas into an abstract framework suitable for broad applications well beyond the scope of elementary analysis. Around 1922, Banach recognised the fundamental role of ‘metric completeness’; a property shared by all of the spaces commonly exploited in analysis and for many years activity was limited to minor extensions of Banach’s contraction mapping principle [7] and its manifold applications.
Mark Krasnosel’skii was an Ukrainian mathematician renowned for his work on nonlinear functional analysis and its applications. He presented many new general principles on solvability of a large variety of nonlinear equations, including one-sided estimates, cone stretching and contractions, fixed-point theorems for monotone operators and a combination of the Schauder fixed point and contraction mapping theorems that was the genesis of condensing operators.

Metric fixed point theory gained new impetus largely as a result of the pioneering work of Felix Browder in the mid-nineteen sixties and the development of nonlinear functional analysis as an active and vital branch of mathematics. Pivotal in this development were the 1965 existence theorems of Browder, ([21] and [22]), and Kirk, [63], and the early metric results of Edelstein [see section (3.8)].

The original theorem of Browder exploited special convexity properties of the norm in certain Banach spaces, while Kirk identified the underlying property of ‘normal structure’ and the role played by weak compactness. The early phases of the development centred around the identification of spaces whose bounded convex sets possessed normal structure and it was soon discovered that certain weakenings and variants of normal structure also sufficed. By the mid nineteen seventies it
was apparent that normal structure was a substantially stronger condition than
needed. Felix Browder (b 1927) was born in Moscow, capital of the then Soviet
Union. While his father (Earl Browder (1891–1973)) was born in Wichita, USA,
his mother, Raissa Berkmann, was born in St Petersburg. Earl joined the Socialist
party at 16 and opposed World War 1. In the 1920s he often visited Russia as a
representative of the Communist Trade Unions in the United States where he met
Raissa at a Soviet social function in Moscow on one of his visits. Earl was married
previously, but following his divorce he married Raissa and their first two children,
Felix and Andrew, were born in Moscow. Earl returned to the USA in 1929, but
Raissa remained in Moscow until 1933.
Felix was a child prodigy who entered MIT in 1944 and graduated in 1946 with his
first degree and in 1948 obtained his doctorate from Princeton University when
he was only 20. His doctoral adviser was Solomon Lefschetz and his thesis was
entitled *The Topological Fixed Point Theory and its Applications in Functional
Analysis*.
Since 1948 Browder had successive appointments at MIT, then Boston University.
In 1953 he was awarded a Guggenheim Fellowship, but in the same year he was
drafted into the army as the Korean war was being fought, the Vietnamese were
fighting the French with US involvement looming and there was increasing Cold
War tension. However in the army he was classified as a security risk (because of
his name and his father’s activities) and was eventually put on trial, but was fi-
nally acquitted. However universities were frightened of his name and thought they
might lose sponsorships if they employed him despite his brilliant mathematical
skills. After leaving the army in 1955 he did obtain a position at Brandeis Univer-
sity followed by Yale and finally Professor of Mathematics at Chicago University
where he remained until he retired in 1985. After retiring he became a professor
at Rutgers University where he was also Vice-President for Research from 1986 to
1991. He is currently a Professor of Mathematics at Rutgers.
Nonlinear functional analysis had its origins in the study of nonlinear ordinary and
partial differential equations, but it came to encompass a wider range of questions
in all branches of analysis and in differential geometry, in theoretical physics, and
in economics. Felix Browder has been the dominant figure in this field since the
early 1950s. The introduction of monotone and, later, accretive operator theory
led to the solution of problems that had previously been out of reach.
**William Arthur ("Art") Kirk** [b 1936] is an American mathematician. He attended Indiana State University (1954–1956) followed by DePauw University (1956–1958) where he obtained his A. B. (Mathematics). He completed his PhD, entitled "Metrization of Surface Curvature", at the University of Missouri in August 1962 under the supervision of Leonard Blumenthal [see section (2.5)]. From 1962 to 1967 he was an Assistant Professor at the University of California (Riverside). He then moved to the University of Iowa where he became a full Professor in 1971 and the Department Chair from 1985 to 1991. He also holds an honourary doctorate from Maria Curie-Sklodowska University, Lublin, Poland, an institution which was an early centre of study for fixed point theory of metric spaces. In his speech presenting Art Kirk to his University, Kazimierz Goebel (one of Kirk’s long time collaborators) used terms such as “metric fixed point theory, nonexpansive, Lipshitzian and uniformly Lipshitzian transformations, normal stucture of convex sets, weak and weak star topologies, hyperconvex spaces and ultra-filter methods” to describe the breadth of Kirk’s contributions to fixed point theory. In 2003 and 2004 he published two ground breaking papers on geodesic geometry and fixed point theory ([64] and [65]), returning his interest close to those of his 1962 PhD thesis.
PAGE LEFT BLANK INTENTIONALLY
6

Projection and Reflection
Algorithms in CAT(0) spaces

6.1 Introduction

The convex feasibility problem associated with the nonempty closed convex sets $A, B$ in a Hilbert space $H$ is to

“find some $x \in A \cap B$”.

Projection algorithms in general aim to compute a point in this intersection of convex sets and in this chapter and the following chapter we will consider three such algorithms and their extension into CAT(0) spaces. This allows us to treat feasibility problems where the sets are metrically, but not necessarily algebraically convex, for example star shaped sets in $E^2$.

The figure 6.1 shows a comparison between between the three algorithms that we will study. This example is set in $E^2$ with the closed convex sets represented by the affine manifolds $A = \{(x, y) : 5x + 8y - 20 = 0\}$ and $B = \{(x, y) : 3x - 8y + 12 = 0\}$. The green graph (dotted) represents the Project Project algorithm, blue (dashed) represents the Project Project Average algorithm and red (dot and dash) represents the Reflect Reflect Average algorithm (Douglas-Rachford).
6.1 Introduction

The method of alternating projection into convex sets (sometimes known as the "project, project" algorithm) has emerged following initial work by John von Neumann who, in the 1930s, proved that when $A$ and $B$ were closed subspaces of a Hilbert space the iterative scheme

$$x_{n+1} = P_B P_A(x_n), \quad n \in \mathbb{N},$$

converged in norm for any initial starting point $x_0$ to $P_{A \cap B} x_0$ [80]. This method has flourished and has given rise to both a beautiful theory and a number of useful algorithms, see for example [8], [9], [10], [11], [13] and [14].

**Theorem 6.1.1.** (von Neumann). Let $H$ be a Hilbert space and $A, B \subset H$ be closed affine sets (translates of subspaces). For any starting point $x_0 \in H$, the sequence defined by 6.1.0.1 converges in norm to the closest point of $A \cap B$ to $x_0$.

**Proof.** See Theorem 3.1 in [14].

In 1965, weak convergence was established by L. M. Bregman (b 1941) [19] for the
case when $A, B \in H$ are closed convex sets in a Hilbert space with $A \cap B \neq \emptyset$.

**Theorem 6.1.2.** (Bregman). Let $H$ be a Hilbert space and $A, B \subset H$ be closed convex sets with $A \cap B \neq \emptyset$. Assume $x_0 \in H$ is a starting point and $(x_n) \subset H$ the sequence generated by 6.1.0.1. Then $(x_n)$ weakly converges to a point from $A \cap B$.

**Proof.** See [19] or Theorem 3.13 in [14]. \hfill \Box

An old problem as to whether or not the convergence of 6.1.0.1 has to be in norm was answered quite recently in the negative by H. S. Hundal in [55].

**Example 6.1.3.** [Hundal] There exists a hyperplane $A \subset l_2$, a convex cone $B \subset l_2$ and a point $x_0 \in l_2$ such that the sequence generated by 6.1.0.1 from the starting point $x_0$ converges weakly to a point in $A \cap B$, but not in norm.

However in many situations commonly found in applications it can be shown that norm convergence pertains [16].

The alternating projection algorithm plays a key role in optimization work and has many applications outside mathematics, for example in medical imaging [11]. An extension into CAT(0) spaces [6] allows the use of the alternating projection method in a much more general setting where there may be no natural linear structure. There are a great number of situations such as tree spaces in phylogenomics, some models of cognition, configuration spaces in robotics to name a few. Further information on these can be found in [5], [15], [41], [42], [43], [44].

A variant of the alternating projection algorithm is **project-project-average**; Here the iteration scheme is

$$x_{n+1} = \frac{1}{2} (I + P_B P_A)(x_n).$$

Similar convergence results to those obtained for the above follow from slightly modified arguments.

The **Douglas-Rachford algorithm** when specialised to the above convex feasibility problem is another method for finding a point in the intersection of two
6.1 Introduction

closed convex sets $A, B \in H$. The result first appeared in a 1956 paper by Jim Douglas Jr and Henry H. Rachford Jr [32] and is often referred to as the "reflect, reflect, average" algorithm as it uses reflections for its mappings.

In a Hilbert space $H$ we define the reflection map in a closed convex set $A \in H$ of a point $x$ as

$$R_A x = P_A x + (P_A x - x) = (2P_A - I)x$$

where $P_A$ is the closest point projection onto $A$ as shown in Figure 6.2.

![Figure 6.2: Reflection of $x$ in a convex set in Hilbert space](image)

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

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

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 [74] in 1979 using the fact that closest point projections are firmly nonexpansive and hence $T$ is also nonexpansive and asymptotically regular. That its iterates weakly converge to a fixed point of $T$ then followed by the well known result initially due to Opial and mentioned already in 3.6.2.

We now give some of the details of the Lions-Mercier proof.
Proof. For $C$ closed, convex and nonempty we see that the metric projection onto $C$, $P_C$, is uniquely determined by the variational inequality,

$$\langle x - P_C(x), c - P_C(x) \rangle \leq 0, \quad \text{for all } c \in C.$$

Using this, a simple calculation shows that $P_C$ is firmly nonexpansive, that is, for all $x, y$ in $X$,

$$\langle P_C(x) - P_C(y), x - y \rangle \geq \|P_C(x) - P_C(y)\|^2$$

and it follows that $P_C$ is also both nonexpansive and monotone. Further

$$\|R_Cx - R_Cy\|^2 = \|2P_C(x) - x - (2P_Cy - y)\|^2$$
$$= \|2(P_Cx - P_Cy) - (x - y)\|^2$$
$$= 4\|P_C(x) - P_C(y)\|^2 - 4\langle P_C(x) - P_C(y), x - y \rangle + \|x - y\|^2$$
$$\leq \|x - y\|^2.$$

Thus, $R_C$ is nonexpansive, and so by Ishikawa, $P_C = \frac{1}{2}(R_C + I)$ is asymptotically regular, that is, orbits under $P_C, x_n = T^n x_0$ are approximate fixed point sequences, $\|Tx_n - x_n\| \to 0$.

More importantly, if both $A$ and $B$ are convex then $R_A$ and $R_B$ and hence $R_AR_B$
are nonexpansive. So again by Ishikawa, in the Douglas-Rachford Algorithm \( T := \frac{1}{2}(R_A R_B + I) \) is nonexpansive and asymptotically regular. Thus, provided \( A \cap B \) is nonempty, by the well known result of Opial, for any initial point \( x_0 \) the orbit

\[ x_{n+1} := T(x_n), \]

converges weakly to a fixed point of \( T \).

Finally, noting that,

\[
T := \frac{1}{2}(R_A R_B + I) = \frac{1}{2} \left( (2P_A - I)(2P_B - I) + I \right),
\]

we have for any \( x \) in \( \text{Fix}(T) \),

\[ P_B x + (I - P_B)x = x = Tx = P_A(2P_B - I)x + (I - P_B)x, \]

So,

\[ P_B x = P_A((2P_B - I)x), \]

and conclude that \( P_B(x) \) is in both \( A \) and \( B \) as claimed.

We note that it can also be shown that \( P_B(x_n) \) converges weakly to \( P_Bx \). However an example where the Douglas-Rachford does not converge in norm is not known. The Douglas-Rachford algorithm may be seen as a refinement of von Neumann’s alternating projection algorithm and it proves to be more stable and to have better convergence properties.

### 6.2 Alternating Projections in CAT(0) spaces

The work in this section formed part of my research and the results presented were obtained jointly with Miroslav Bacak, Ian Searston and Brailey Sims. They were published in J. Math. Anal. Appl. ([6]). We begin by defining our problem in the following way:
Let $X$ be a complete CAT(0) space with $A, B \subset X$ closed convex sets. Then the alternating projection method produces the sequence

$$x_{2n-1} = P_A(x_{2n-2}), \quad x_{2n} = P_B(x_{2n-1}) \quad n \in \mathbb{N} \quad (6.2.0.2)$$

where $x_0 \in X$ is a given starting point.

### 6.2.1 Preliminaries

For the remainder of this section we let $X$ be a CAT(0) space and $C \subset X$ a closed convex set.

We now need a series of definitions before moving to our main result.

**Definition 6.2.1.** Regularity of sets in a CAT(0) space.

(i) $A, B \subset X$ are **boundedly regular** if for any bounded set $S \subset X$ and any $\varepsilon > 0$ there exists $\delta > 0$ such that if $x \in S$ and $\max\{d(x, A), d(x, B)\} \leq \delta$ then $d(x, A \cap B) < \varepsilon$.

(ii) $A, B \subset X$ are **boundedly linearly regular** if for any bounded set $S \subset X$ there exists $\kappa > 0$ such that for any $x \in S$ we have

$$d(x, A \cap B) \leq \kappa \max\{d(x, A), d(x, B)\}.$$  

(iii) $A, B \subset X$ are **linearly regular** if there exists $\kappa > 0$ such that for any $x \in X$ we have

$$d(x, A \cap B) \leq \kappa \max\{d(x, A), d(x, B)\}.$$  

**Definition 6.2.2.** A sequence $(x_n)$ is **linearly convergent** to a point $x \in X$ if there exists $K \geq 0$ and $\beta \in [0, 1)$ such that

$$d(x, x_n) \leq K \beta^n, \quad n \in \mathbb{N}.$$  

The parameter $\beta$ is called a **rate of linear convergence**.
6.2 Alternating Projections in CAT(0) spaces

**Definition 6.2.3.** A sequence \((x_n) \subset X\) is **Fejér monotone** with respect to \(C\) if, for any \(c \in C\),
\[
d(x_{n+1}, c) \leq d(x_n, c), \quad n \in \mathbb{N}.
\]

We now have the following proposition

**Proposition 6.2.4.** Let \((X, d)\) be a complete CAT(0) space and let \((x_n) \subset X\) be a Fejér monotone sequence with respect to \(C\). Then

(i) \((x_n)\) is bounded,
(ii) \(d_C(x_{n+1}) \leq d_C(x_n)\) for each \(n \in \mathbb{N}\),
(iii) \((x_n)\) weakly converges to some \(x \in C\) if and only if all weak limit points of \((x_n)\) belong to \(C\),
(iv) \((x_n)\) converges to some \(x \in C\) if and only if \(d(x_n, C) \to 0\),
(v) \((x_n)\) converges linearly to some \(x \in C\), provided there exists \(\theta \in [0, 1)\) such that \(d(x_{n+1}, C) \leq \theta d(x_n, C)\) for each \(n \in \mathbb{N}\).

**Proof.**

(i) and (ii) follow easily.

(iii) We look at the non-trivial part of (iii). If we assume that all weak limit points of \((x_n)\) lie in \(C\) then it will suffice to show that \((x_n)\) has a unique limit point.

We assume \(c_1, c_2 \in C\) with \(c_1 \neq c_2\) are weak limit points of \((x_n)\) and we will show this leads to a contradiction. That is there are subsequences \((x_{n_k})\) and \((x_{m_k})\) such that \(x_{n_k} \rightharpoonup c_1\) and \(x_{m_k} \rightharpoonup c_2\).

Without loss of generality, assume \(r(x_{n_k}) \leq r(x_{m_k})\). For any \(\varepsilon > 0\) there exists \(k_0 \in \mathbb{N}\) such that \(d(x_{n_k}, c_1) < r(x_{n_k}) + \varepsilon\) for all \(k \geq k_0\).

By Fejér monotonicity we also have \(d(x_{m_k}, c_1) < r(x_{m_k}) + \varepsilon\) for all \(m_k \geq n_{k_0}\). Hence, there exists \(k_1 \in \mathbb{N}\) such that \(d(x_{m_k}, c_1) < r(x_{m_k}) + \varepsilon\) for all \(k \geq k_1\).

However this contradicts the original assumption that \(c_2\) is the unique asymptotic centre of \((x_{m_k})\).
Suppose \( d(x_n, C) \rightarrow 0 \). Since for all \( k \in \mathbb{N} \) we have

\[
d(x_{n+k}, x_n) \leq d(x_{n+k}, P_C(x_n)) + d(x_n, P_C(x_n)) \quad (6.2.1.1)
\]

and hence, by Fejér monotonicity,

\[
d(x_{n+k}, x_n) \leq d(x_n, P_C(x_n)) + d(x_n, P_C(x_n)) \leq 2d(x_n, C), \quad (6.2.1.2)
\]

hence \( (x_n) \) is Cauchy and so converges to a point from \( C \). The proof of the converse follows easily.

(v) From 6.2.1.1 and 6.2.1.2 we obtain

\[
d(x_{n+k}, x_n) \leq 2d(x_n, C) \leq 2\theta^n d(x_0, C)
\]

for all \( n, k \in \mathbb{N} \). The sequence \( (x_n) \) converges to some \( x \in C \) and so letting \( k \rightarrow \infty \) we have

\[
d(x, x_n) \leq 2d(x_n, C) \leq 2\theta^n d(x_0, C) \quad n \in \mathbb{N}.
\]

That is \( x_n \rightarrow x \) linearly with rate \( \theta \).

The following lemma completes this subsection.

**Lemma 6.2.5.** Let \( A, B \subset X \) be closed convex sets with \( A \cap B \neq \emptyset \). Then the sequence generated by Algorithm 6.2.0.2 is Fejér monotone with respect to \( A \cap B \).

**Proof.** Pick \( c \in A \cap B \). Fix \( n \in \mathbb{N} \) and without loss of generality assume that \( x_n \in A \). Recalling that \( x_{n+1} = P_B(x_n) \) and if \( x_{n+1} = c \) we have finished. Otherwise by Proposition 3.5.2 (iii) we have \( \angle_{x_{n+1}}(c, x_n) \geq \frac{\pi}{2} \). It follows that we have in the comparison triangle \( \angle_{x_{n+1}}(c, x_n) \geq \frac{\pi}{2} \) hence \( d(x_{n+1}, c) \leq d(x_n, c) \). 

\( \square \)
6.2 Alternating Projections in CAT(0) spaces

6.2.2 Convergence Results

We now introduce the main theorem for this section.

Theorem 6.2.6. (Convergence Result) 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 \((x_n) \subset X\) be the sequence generated by Algorithm 6.2.0.2. Then

(i) \((x_n)\) weakly converges to a point \( x \in A \cap B \);

(ii) If \( A \) and \( B \) are boundedly regular, then \( x_n \to x \);

(iii) If \( A \) and \( B \) are boundedly linearly regular, then \( x_n \to x \) linearly;

(iv) If \( A \) and \( B \) are linearly regular, then \( x_n \to x \) linearly with a rate independent of the starting point.

Proof. We begin by establishing two results which we will use in the proofs.

Let us fix \( n \in \mathbb{N} \) and without loss of generality assume \( x_n \in A \) and \( x_{n+1} \notin A \cap B \). By Proposition 3.5.2 we have

\[
\angle_{x_{n+1}}(x_n, PA \cap B(x_n)) \geq \angle_{x_{n+1}}(x_n, PA \cap B(x_n)) \geq \frac{\pi}{2}.
\]

From the comparison triangle we have

\[
d^2(x_n, PA \cap B(x_n)) \geq d^2(x_n, x_{n+1}) + d^2(PA \cap B(x_n), x_{n+1})
\]

which we can rewrite as

\[
d^2(x_n, A \cap B) \geq d^2(x_n, B) + d^2(x_{n+1}, A \cap B)
\]

and with some rearrangement we have

\[
\max\{d^2(x_n, A), d^2(x_n, B)\} \leq d^2(x_n, A \cap B) - d^2(x_{n+1}, A \cap B). \tag{6.2.2.1}
\]

Further, by Fejér monotonicity (due to Lemma 6.2.5), Proposition 6.2.4 (ii) and
6.2 Alternating Projections in CAT(0) spaces

6.2.2.1 we have

\[ \max \{d(x_n, A), d(x_n, B)\} \to 0 \quad \text{as} \quad n \to \infty. \quad (6.2.2.2) \]

We are now ready to complete the proofs.

(i) By using Fejér monotonicity (Lemma 6.2.5) we obtain that \((x_n)\) is bounded (Proposition 6.2.4 (i)) and hence has a weak limit point \(x \in X\). Taking a subsequence \((x_{n_k})\) which weakly converges to \(x\) and using Lemma 4.3.6 and 6.2.2.2 we have \(d(x, A) = d(x, B) = 0\). Hence \(x \in A \cap B\) and we can conclude that \(x_n \rightharpoonup x\), by Proposition 6.2.4 (iii).

(ii) We note that the bounded regularity of \(A\) and \(B\) with 6.2.2.2 gives \(d(x_n, A \cap B) \to 0\) as \(n \to \infty\). Then Lemma 6.2.5 and Proposition 6.2.4 (iv) give the required result.

(iii) We recall that \((x_n)\) is bounded. Hence by bounded linear regularity, there exists \(\kappa > 0\) such that, for every \(n \in \mathbb{N}\),

\[ d(x_n, A \cap B) \leq \kappa \max \{d(x_n, A), d(x_n, B)\}. \]

Using 6.2.2.1 we arrive at

\[ d^2(x_n, A \cap B) \leq \kappa^2 \left( d^2(x_n, A \cap B) - d^2(x_{n+1}, A \cap B) \right) \]

\[ d(x_{n+1}, A \cap B) \leq \sqrt{1 - \frac{1}{\kappa^2}} d(x_n, A \cap B) \]

Applying Proposition 6.2.4(v) completes the proof of (iii)

(iv) The proof of part (iv) is similar to that of part (iii).

\[ \square \]

6.2.3 Applications of results

In this subsection we present an application of the alternating projection method to convex optimization in CAT(0) spaces. We begin by giving a very different example of a convex function from those mentioned in 4.3. It is provided by the
energy functional on a special CAT(0) space that is very important in many areas of analysis and geometry ([59] and [60]). We will give an outline, but a detailed version is provided in [6].

Let $M$ and $N$ be compact Riemannian manifolds with $N$ having nonpositive sectional curvature. For $f \in L^2(M, N)$ and $h > 0$ the energy of $f$ is defined by functionals $E_h$ which are convex and continuous, whereas the energy function $E(f) = \lim_{h \to 0} E_h(f)$ is convex and lower semicontinuous. In many situations (like Theorem 7.5.2 in [59]), instead of considering the energy functionals on $L^2(M, N)$ it is more convenient to extend it to the CAT(0) space of equivalent maps between the universal covers $\tilde{M}$ and $\tilde{N}$ of $M$ and $N$ respectively. With $\rho$ as defined in example 5.1 of [6] the space $L^2_{\rho}(M, N)$ is a complete CAT(0) space.

We can now consider the energy functionals $E$ and $E_h$ as functionals on the space $L^2_{\rho}(M, N)$. Then $E_h$ is convex and continuous on $L^2_{\rho}(M, N)$, and the energy functional $E$ is convex and lower semicontinuous. Moreover this space is very different from other examples already mentioned. It is different from Hilbert spaces since it is not flat and it is different from Riemannian manifolds since it is not locally compact.

We now consider the following optimization problem. Let $X$ be a complete CAT(0) space. We are given a function $F : X \to (-\infty, \infty]$ of the form $F = \max(f, g)$, where $f, g : X \to (-\infty, \infty]$ are lsc and convex, and we wish to find a minimizer of $F$; that is some $x \in \text{sub}_\alpha(F) = \{x \in X : F(x) \leq \alpha\}$, where $\alpha = \inf_X F$. Of course we are assuming that $\inf_X F$ is finite and the set $\text{sub}_\alpha(F)$ is nonempty. Or, alternatively we may seek an approximative minimizer for $F$; that is, given some $\alpha > \inf_X F$ we want to find some $x \in \text{sub}_\alpha(F)$. Then, in case the projections onto $\text{sub}_\alpha(f)$ and $\text{sub}_\alpha(g)$ are easy to compute, we can find the desired $x \in \text{sub}_\alpha(F)$ as the limit of the alternating sequence since

$$\text{sub}_\alpha(F) = \text{sub}_\alpha(f) \cap \text{sub}_\alpha(g).$$

In general when the functions $f$ and $g$ are only lsc and convex, we have weak convergence of the alternating sequence by Theorem 6.2.6 (i), and this is the best
we can hope for. If, however, we impose additional assumptions on the functions \( f \) and \( g \), we get strong convergence, as we shall see in Proposition 6.2.7 below.

We will first recall that a function \( h : X \to (-\infty, \infty] \) is uniformly convex if there exists \( \lambda > 0 \) such that for any \( x, y \in X \) and \( u \in [x, y] \) we have

\[
h(u) \leq (1 - t)h(x) + th(y) - \lambda t(1 - t)d^2(x, y),
\]

where \( t = \frac{d(x, u)}{d(x, y)} \).

The following proposition provides the promised sufficient conditions on the functions \( f \) and \( g \) to ensure the sets \( \text{sub}_\alpha(f) \) and \( \text{sub}_\alpha(g) \) are ‘more regular’, and hence allows us to obtain, via Theorem 6.2.6, stronger types of convergence for the alternating sequence to an (approximative) minimizer of the functional \( F = \max(f, g) \).

**Proposition 6.2.7.** Let \( X \) be a complete CAT(0) space, and \( F : X \to (-\infty, \infty] \) be a functional of the form \( F = \max(f, g) \), where \( f, g : X \to (-\infty, \infty] \) are lsc convex. Let \( \alpha \geq \inf_X F > -\infty \), and \( \text{sub}_\alpha(F) \) be nonempty. If the function \( f \) is both uniformly convex and uniformly continuous on bounded sets of \( X \), then the sets \( \text{sub}_\alpha(f) \) and \( \text{sub}_\alpha(g) \) are boundedly regular.

**Proof.** Assume \( S \subset X \) is a given bounded set and \( \epsilon > 0 \). We will look for \( \delta > 0 \) such that if one picks \( x \in S \) with

\[
\max [d(x, \text{sub}_\alpha(f)), d(x, \text{sub}_\alpha(g))] < \delta, \hspace{1cm} (6.2.3.1)
\]

then

\[
d(x, \text{sub}_\alpha(F)) < \epsilon. \hspace{1cm} (6.2.3.2)
\]

Let \( b = P_{\text{sub}_\alpha(g)}(x) \), the projection of \( x \) onto the set \( \text{sub}_\alpha(g) \). If \( b \in \text{sub}_\alpha(f) \), we can take \( \delta = \epsilon \) in 6.2.3.1 to fulfill 6.2.3.2. If \( b \notin \text{sub}_\alpha(f) \), then denoting the projections \( P_{\text{sub}_\alpha(F)}(b) \) and \( P_{\text{sub}_\alpha(f)}(b) \) by \( c \) and \( a \) respectively and taking \( m \) to be the midpoint of the geodesic \([b, c]\) we have, by the uniform convexity of \( f \), that there exists \( \lambda > 0 \) such that

\[
f(m) \leq \frac{1}{2} [f(b) + f(c)] - \lambda d^2(b, c),
\]

109
and hence,

\[ d^2(b, c) \leq \frac{1}{\lambda} \left[ \frac{f(b) + f(c)}{2} - f(m) \right]. \]  

(6.2.3.3)

By uniform continuity of \( f \), there exists \( \delta' > 0 \) such that

\[ |f(b) - f(x')| < \frac{\epsilon^2 \lambda}{2}, \]

whenever \( d(b, x') < \delta' \). Therefore, if \( d(b, a) < \delta' \), from 6.2.3.3 we further have

\[ d^2(b, c) < \frac{1}{\lambda} \left[ \frac{f(a) + f(c)}{2} + \frac{\epsilon^2 \lambda}{4} - f(m) \right] \leq \frac{\epsilon^2}{4}, \]

which yields

\[ d(b, c) = d(b, \text{sub}_\alpha(F)) < \frac{\epsilon}{2}, \]

and hence, if we choose \( \delta < \frac{1}{2} \max(\epsilon, \delta') \) in 6.2.3.1 we obtain

\[ d(x, \text{sub}_\alpha(f) \cap \text{sub}_\alpha(g)) \leq d(x, b) + d(b, c) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \]

That is, the sets \( \text{sub}_\alpha(f) \) and \( \text{sub}_\alpha(g) \) are boundedly regular. \( \square \)

We note that in the above Proposition 6.2.7 we only make additional assumptions on the function \( f \), whereas the function \( g \) is arbitrary lsc convex.

**Remark 6.2.8.** CAT(0) spaces and the iterative methods considered in this chapter have been found to have important applications to minimizing the energy functional as a means of obtaining smooth (harmonic) solutions to various flows. See for example [59] and [60], and more recently [75] and [89]. Some further details may be found in [6].

### 6.3 Averaged projection method in CAT(0) spaces

Another popular method for solving convex feasibility problems is the averaged projection method. Let \( A \) and \( B \) be closed convex subsets of a complete CAT(0)
space \((X, d)\) with \(A \cap B \neq \emptyset\). We choose a starting point \(x_0 \in X\) and put

\[
x_n := \frac{1}{2} P_A(x_{n-1}) \oplus \frac{1}{2} P_B(x_{n-1}), \quad n \in \mathbb{N}.
\]  

(6.3.0.4)

By this algorithm the averaged projection method updates \(x_{n-1}\) to \(x_n\) at each step by projecting \(x_{n-1}\) onto \(A\) and \(B\) respectively, and then taking the “average” of the two projections \(P_A(x_{n-1})\) and \(P_B(x_{n-1})\).

**Theorem 6.3.1.** (Averaged projections). Let \((x_n)\) be a sequence generated by (6.3.0.4). Then the following hold:

(i) The sequence \((x_n)\) weakly converges to a point \(x \in A \cap B\).

(ii) If \(A\) and \(B\) are boundedly regular, then \(x_n \to x\).

(iii) If \(A\) and \(B\) are boundedly linearly regular, then \(x_n \to x\) linearly.

(iv) If \(A\) and \(B\) are linearly regular, then \(x_n \to x\) linearly with a rate independent of the starting point.

**Proof.** The proof is similar to the proof of Theorem 6.2.6 so we only give a brief proof here. The sequence \((x_n)\) is Fejér with respect to \(A \cap B\). Let \(c \in X\) be a weak limit point of \((x_n)\). Since

\[
d(x_n, A)^2 \leq d(x_n, A \cap B)^2 - d(P_A(x_n), A \cap B)^2
\]

and after interchanging \(A\) and \(B\) we also obtain

\[
d(x_n, B)^2 \leq d(x_n, A \cap B)^2 - d(x_{n+1}, A \cap B)^2,
\]

which together with the weak lower semicontinuity of the metric yields \(c \in A\). By Proposition 6.2.4, we finally obtain \(x_n \to x\). This proves (i).

The remaining parts follow in a very similar way to the proof of Theorem 6.2.6. \(\Box\)

**Remark 6.3.2.** Convex feasibility problems can be formulated in the language of fixed point theory. For \(P_B \circ P_A\) we note that \(x \in A \cap B\) if and only if \((P_B \circ P_A)x = x\).
We also note that $x \in A \cap B$ if and only if \( \left( \frac{1}{2} P_A \oplus \frac{1}{2} P_B \right) x = x \).

### 6.4 Reflections in CAT(0) spaces

In Chapter 3, proposition 3.5.2, we saw that projections are nonexpansive in CAT(0) spaces and in subsection 6.2.6 and section 6.3 we proved the convergence of alternating projections and averaged projections in CAT(0) spaces. So it is appropriate to investigate reflections and hence the possible convergence of the Douglas-Rachford algorithm in CAT(0) spaces.

#### 6.4.1 Geodesics and reflections.

To discuss reflections in CAT(0) spaces we require geodesics to be extendable as in Figure 6.4. That is, for $x, y \in X$ and some $\varepsilon > 0$, there exists a segment $[x, z]$ where $d(y, z) = \varepsilon$ and $[x, y] \subseteq [x, z]$.

![Figure 6.4: Extendable geodesics](image)

We recall that all geodesic segments extendable implies they are all infinitely extendable in both directions (see Subsection 2.3.5). That is, for all $x, y \in X$ there exists an isometry $\gamma : R \to X$ with $\gamma(0) = x$ and $\gamma(d(x, y)) = y$.

We also require that the extension is unique. That is the geodesic does not bifurcate. Figure 6.5 illustrates a geodesic which is bifurcating. In this figure $\Gamma_1 \cup \Gamma_2$ and $\Gamma_1 \cup \Gamma_3$ are geodesics.

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 therefore important to us.

Under 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 the 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 \( P_Cx \) is the projection of \( x \) onto the set \( C \) (see Figure 6.6).

\[\text{Figure 6.6: Reflection in a closed convex set } C \text{ of a CAT(0) space}\]

**Note**: If extensions are not required to be unique (Eg in an R-tree) then \( R_C(x) \) may be multivalued.

### 6.4.2 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. While projections are
6.4 Reflections in CAT(0) spaces

also firmly nonexpansive in an appropriate sense in CAT(0) spaces, [4], it no longer follows that reflections are nonexpansive.

R-trees in which all geodesics are unbounded are also an example of a CAT(0) space in which reflections while potentially multivalued are nonexpansive and a proof follows.

\[ d(R_c(x), R_c(y)) = d(x, P_c(x)) + d(P_c(x), P_c(y)) + d(P_c(y), R_c(y)) = d(x, y). \]

**Figure 6.7: Reflection in R-trees**

Proof. We note that for points \( x, y \) in the R-tree and any convex set \( C \) we have \( d(R_c(x), P_c(x)) = d(x, P_c(x)) \) and \( d(R_c(y), P_c(y)) = d(y, P_c(y)) \).

The possible cases are illustrated schematically in figure 6.7

In the first case we have
6.4 Reflections in CAT(0) spaces

For the more general second case we initially note that

\[ d(R_c y, P_c y) = d(y, P_c y) \]
\[ = d(R_c y, a) + \delta \quad (\delta > 0) \]
\[ d(R_c y, a) = d(y, P_c y) - \delta \]

so

\[ d(R_c x, R_c y) = d(x, P_c x) + d(P_c y, P_c x) - \delta + d(y, P_c y) - \delta \]
\[ = d(x, y) - 2\delta \]
\[ < d(x, y). \]

\[ \square \]

Reflections in the space \( M_{\kappa}^n \) of constant curvature have been proved to be nonexpansive. The proof in the case of constant negative curvature follows.

**Proposition 6.4.1** (Proposition 3.1 in [39]). Let \( \kappa \in R \) and \( n \in \mathbb{N} \). Suppose \( C \) is a nonempty closed and convex subset of \( M_{\kappa}^n \) and \( x, y \in M_{\kappa}^n \). Then,

\[ d(R_c x, R_c y) \leq d(x, y). \]

**Proof.** For simplicity we consider the case when \( \kappa = -1 \). The proof in other model spaces follows a similar pattern.

We start with the geodesic triangles \( \triangle(x, P_c x, P_c y) \) and \( \triangle(R_c x, P_c x, P_c y) \) and for convenience we write \( \alpha = \angle_{P_c x}(x, P_c y) \) and \( \alpha' = \angle_{P_c y}(y, P_c x) \) noting that from Proposition 3.5.2 part(3) we have \( \alpha, \alpha' \geq \frac{\pi}{2} \).

Using the hyperbolic cosine law we have

\[ \cosh d(x, P_c y) = \cosh d(x, P_c x) \cosh d(P_c x, P_c y) \]
\[ - \sinh d(x, P_c x) \sinh d(P_c x, P_c y) \cos \alpha \]

115
6.4 Reflections in CAT(0) spaces

\[ \cosh d(R_c x, P_c y) = \cosh d(R_c x, P_c x) \cosh d(P_c x, P_c y) - \sinh d(R_c x, P_c x) \sinh d(P_c x, P_c y) \cos(\pi - \alpha). \]

From the facts that \( d(x, P_c x) = d(R_c x, P_c x) \) and \( \cos(\pi - \alpha) \geq 0 \) it follows that \( \cosh d(x, P_c y) \geq \cosh d(R_c x, P_c y) \) and thus \( \cosh d(x, P_c y) \geq d(R_c x, P_c y) \). In a similar fashion we obtain \( \cosh d(y, P_c x) \geq d(R_c y, P_c x) \).

We now consider the geodesic triangles \( \triangle(x, P_c x, y) \) and \( \triangle(R_c x, P_c x, y) \) and write \( \beta = \angle P_c x(x, y) \). Using the hyperbolic laws as above we obtain

\[ \cosh d(x, y) = \cosh d(x, P_c x) \cosh d(P_c x, y) - \sinh d(x, P_c x) \sinh d(P_c x, y) \cos \beta \]

and

\[ \cosh d(R_c x, y) = \cosh d(R_c x, P_c x) \cosh d(P_c x, y) - \sinh d(R_c x, P_c x) \sinh d(P_c x, y) \cos(\pi - \beta). \]

Adding the two equalities and noting that \( \cos(\pi - \beta) = - \cos \beta \) we get

\[ \cosh d(x, y) + \cosh d(R_c x, y) = 2 \cosh d(x, P_c x) \cosh d(P_c x, y). \quad (6.4.2.1) \]

Similarly

\[ \cosh d(R_c x, R_c y) + \cosh d(x, R_c y) = 2 \cosh d(P_c x, x) \cosh d(P_c x, R_c y), \quad (6.4.2.2) \]

\[ \cosh d(x, y) + \cosh d(x, R_c y) = 2 \cosh d(P_c y, x) \cosh d(P_c y, y), \quad (6.4.2.3) \]

\[ \cosh d(R_c x, R_c y) + \cosh d(R_c x, y) = 2 \cosh d(P_c y, y) \cosh d(P_c y, R_c x). \quad (6.4.2.4) \]
Suppose \( d(R_c x, R_c y) > d(x, y) \). From 6.4.2.1 and 6.4.2.2

\[
\cosh d(R_c x, R_c y) + \cosh d(x, R_c y) = 2 \cosh d(P_c x, x) \cosh d(P_c x, R_c y) \\
\leq 2 \cosh d(P_c x, x) \cosh d(P_c x, y) \\
< \cosh d(R_c x, R_c y) + \cosh d(R_c x, y),
\]

from which it follows that \( d(x, R_c y) < d(R_c x, y) \). Applying 6.4.2.3 and 6.4.2.4 in a similar way we get that \( d(x, R_c y) > d(R_c x, y) \) which is a contradiction to the previous result and our assumption is false, meaning that \( d(R_c x, R_c y) \leq d(x, y) \) as required.

This result supersedes an earlier observation by David Ariza-Ruiz, Ian Searston and Brailey Sims that the result is true in the special case when \( C \) is a geodesic segment or a half-space.

Using this proposition, Fernández-Leon and Nicolae, authors of [39], establish convergence of the Douglas-Rachford algorithm in \( M^\kappa_n \). Strong convergence is obtained since the model spaces are proper metric spaces. A space is proper if every closed ball is compact.

**Theorem 6.4.2** (Theorem 3.4 in [39]). Let \( \kappa \leq 0 \) and \( n \in \mathbb{N} \). Suppose \( A \) and \( B \) are two nonempty closed and convex subsets of \( M^\kappa_n \) with \( A \cap B \neq \emptyset \) and \( T : X \to X \) defined by \( T = \frac{I + R_A R_B}{2} \). Let \( x_0 \in M^\kappa_n \) and \( (x_n) \) be the sequence starting at \( x_0 \) generated by the Douglas-Rachford algorithm. Then

1. \( (x_n) \) converges to some fixed point \( x \) of the mapping \( T \) and \( P_B(x) \in A \cap B \).
2. The "shadow" sequence \( (P_B(x_n)) \) is convergent and its limit belongs to \( A \cap B \).

**Proof.** (1) By proposition 6.4.1, \( R_A R_B \) is nonexpansive which yields that \( T \) is also nonexpansive because \( M^\kappa_n \) with \( \kappa \leq 0 \) is Busemann convex. In addition, \( \text{Fix}(T) \neq \emptyset \) since \( A \cap B \neq \emptyset \). Thus, all orbits of \( T \) are bounded. Using Proposition 2, [47] it follows that \( T \) is asymptotically regular. Applying [Proposition 6.3, [4]], we get that \( (x_n) \) \( \triangle \)-converges to a fixed point of \( T \), which implies that \( (x_n) \) converges to a fixed point of \( T \). Let \( x \) be the limit.
of \((x_n)\). Since \(x = T(x)\) it follows that \(x = R_AR_B(x)\). Moreover, because 
\[ P_B(x) = \frac{x + R_B(x)}{2}, \]
\[ P_B(x) = \frac{R_AR_B(x) + R_B(x)}{2} = P_AR_B(x) \in A. \]
Hence \(P_B(x) \in A \cap B\).

(2) This is immediate since the metric projection \(P_B\) is continuous in CAT(0) spaces.

\[ \square \]

However, as we will see in the next chapter, reflections in CAT(0) spaces need not in general be nonexpansive.

6.5 Conclusions

In this chapter we have presented some of our major work. We have proved that the projection algorithms known as “Project Project” and “Project Project Average” in Hilbert space can be extended to CAT(0) spaces. Our work on the project-project algorithm in CAT(0) spaces has been published in [6]. We also investigated the nonexpansiveness of reflections in CAT(0) spaces. A substantial example of an application of the alternating projection method to convex optimization in CAT(0) spaces is included. We developed a proof that reflections are nonexpansive in \(R\)–trees and have also provided a proof to show reflections are nonexpansive in \(M^\kappa\), the classical model spaces of constant curvature.

6.6 Historical Notes

John von Neumann (1903–1957) was a Hungarian-American pure and applied mathematician, physicist and polymath. He was born János Neumann in Budapest, Hungary. In 1913 his father, Max Neumann, purchased a title, but did not change his name. His son, however, used the German form von Neumann.
where the "von" indicated the title. Von Neumann won a place at the University of Budapest in 1921 to study mathematics, but did not attend lectures. Instead he also entered the University of Berlin in the same year to study Chemistry. He achieved outstanding results at the University of Budapest despite not attending classes.

Von Neumann was invited to Princeton, USA, in 1929 to lecture on quantum theory and was appointed professor there in 1931. In 1933 he became one of the original six mathematics professors at the newly founded Institute for Advanced Study in Princeton, a position he kept for the remainder of his life.

![Figure 6.8: John von Neumann (1903-1957)](image)

He made major contributions to many fields, including mathematics (functional analysis, ergodic theory, geometry, topology, numerical analysis and linear programming), physics (quantum mechanics, hydrodynamics and fluid dynamics), economics (game theory), computing (von Neumann architecture, self-replicating machines and stochastic computing) and statistics. He was a pioneer of the application of operator theory to quantum mechanics. In 1956 he received the Albert Einstein Commemorative Award and the Enrico Fermi Award.

**Jim Douglas, Jr.**, received his degrees in civil engineering from the University of Texas in 1947 and his PhD from Rice University in mathematics in 1952. For the next five years he was employed by the Humble (now Exxon) Production Research Laboratory. It was during this time that his work on the numerical solution of
6.6 Historical Notes

Partial differential equations and the simulation of fluid flow in porous media began. After that he worked at Rice University and the University of Chicago.

Figure 6.9: Jim Douglas Jr. and Henry H. Rachford Jr.

**Henry H. Rachford, Jr.** received his degrees in chemical engineering and chemistry from Rice University. He completed his graduate work at the Massachusetts Institute of Technology where he received a ScD degree in Chemical engineering in 1950. He also joined the Production Research Division of Humble Oil (now Exxon) company in 1949 where he worked in the development and application of computer oriented methods for the simulation of petroleum reservoir processes. In 1964 he joined the faculty of Rice University where his research and teaching was directed towards numerical analysis, especially numerical methods for partial differential equations.

**Pierre-Louis Lions** (b 1956) is a very well known mathematician. He received his doctorate from the University of Pierre and Marie Curie (Paris) in 1979. His main work is on the theory of nonlinear partial differential equations. Lions was the first to give a complete solution to the Boltzmann equation with proof. He received the Fields medal for his work in 1994 as well as many other awards. In the paper ”Viscosity solutions of Hamilton-Jacobi equations” (1983), written with Michael Crandall, he introduced the notion of viscosity solutions. This has had a great effect on the theory of partial differential equations. More recent work includes the theory of mean-field games which promises to be highly influential in economics.
Bertrand Mercier (b 1949) is an applied mathematician, born in Orleans, France. He was a student of Jacques-Louis Lions (father of Pierre-Louis) and received his doctorate in 1977. Since 1971 he has worked in many research positions in places such as the National Institute for Research in Computer Science and Control, the Centre for Applied Mathematics at the École Polytechnique, the Military Applications Division of the CEA (Commission for Atomic Energy and Alternative Energies), the Department of Studies and Techniques Automobiles PSA Peugeot-Citroën and the Technologies Directorate for Nuclear Reactors.
6.6 Historical Notes

PAGE LEFT BLANK INTENTIONALLY
7

A Prototype CAT(0) space of non-constant curvature

7.1 Introduction

In this chapter we construct and investigate a prototypical instance of a CAT(0) space of non-constant curvature by using Reshetnyak’s theorem 3.4.1 to glue two CAT(κ) spaces together. After analysing this prototype we develop a counterexample which shows that reflections in spaces of non-constant curvature are generally not nonexpansive and we use the prototype to explore the behaviour of the Douglas Rachford algorithm in such a space of non-constant curvature.

7.2 The Poincaré upper-half plane

We now move to a particular example of a CAT(0) space of constant negative curvature which will be basic to the considerations of this chapter, viz, the Poincaré upper half plane, which is a model of \( \mathbb{H}^2 \).

The Poincaré upper half-plane is \( \mathcal{H} := \{ z \in \mathbb{C} : \Im z > 0 \} \) equipped with the metric
7.2 The Poincaré upper-half plane

given by
\[ d_p(z_1, z_2) = \int_{z_1}^{z_2} \frac{|dz|}{\Im z} = \cosh^{-1} \left( 1 + \frac{|z_1 - z_2|^2}{2\Im z_1 \Im z_2} \right). \]

The geodesics of \( \mathcal{H} \) are vertical lines and semicircles with centres on the real axis.

In particular, given any \( z_1 \) and \( z_2 \in \mathcal{H} \) the unique geodesic through them is
\[ \{ z \in \mathbb{C} : \Re z = \Re z_1, \Im z > 0 \} \]
when \( \Re z_1 = \Re z_2 \), and
\[ \{ z \in \mathbb{C} : |z - c| = r, \Im z > 0 \}, \]
otherwise, where \( c = \frac{|z_1|^2 - |z_2|^2}{2\Re (z_1 - z_2)} \) and \( r = |z_1 - c| \). Figure 7.1 shows a semicircle as a geodesic.

Later we will need to find the midpoint of a geodesic segment in \( \mathcal{H} \). We recall that the midpoint of the geodesic segment \([x, y]\) would be a point \( z \) such that \([x, z] + [z, y] = [x, y]\). In our Poincaré upper halfplane let \( x = x_1 + ix_2 \) and \( y = y_1 + iy_2 \) then the midpoint of the segment \([x, y]\) is
\[
\frac{1}{2} x \oplus \frac{1}{2} y = \frac{x_1y_2 + x_2y_1}{x_2 + y_2} + i \sqrt{x_2y_2 \left[ 1 + \left( \frac{x_1 - y_1}{x_2 + y_2} \right)^2 \right]}.
\]

Proposition 6.4.1 shows that reflections in closed convex sets in the model space \( M^n_\kappa \) of constant curvature are nonexpansive. Later we will be particularly interested in the case when the convex set is a geodesic and we now consider that case.

Metric projection onto a complete convex subset of a CAT(0) space was dealt with in section 3.5 and we now consider the special case when the convex set is a geodesic (or equivalently a half-space) in the Poincaré upper-half plane.

When \( U \) is the upper half unit circle \( |z| = 1 \) and \( \Im z > 0 \), the metric projection onto \( U \) is \( P_U(z) = x_1 + ix_2 \) where \( x_1 = \frac{2\Re(z)}{1 + |z|^2} \) and \( x_2 = \sqrt{1 - x_1^2} \). So more generally, for \( C \) the geodesic \( |z - a| = r \) we have \( P_C(z) = rP_C\left( \frac{z - a}{r} \right) + a \).
7.2 The Poincaré upper-half plane

That is, for \( z = x_1 + i x_2 \)

\[
P_C(x_1 + i x_2) = \frac{2x_1 r^2 - \tau a}{2r^2 - \tau} + i \sqrt{r^2 - \left( \frac{2r^2(x_1 - a)}{2r^2 - \tau} \right)^2},
\]

where \( \tau := 2ax_1 + r^2 - (x_1^2 + x_2^2 + a^2). \)

If \( C \) is a vertical line with equation \( x_1 = a \), then the metric projection on \( C \) is

\[
P_C(x_1 + i x_2) = a + i \sqrt{(x_1 - a)^2 + x_2^2}.
\]

The reflection map in the Poincaré upper half-plane is also important to us. For \( z \in \mathcal{H} \) there exists a unique point \( R_C(z) \) such that

\[
d_{\mathcal{H}}(z, P_C(z)) = d_{\mathcal{H}}(R_C(z), P_C(z))
\]

and

\[
d_{\mathcal{H}}(z, R_C(z)) = 2d_{\mathcal{H}}(z, P_C(z)).
\]

In the case that \( C \) is the geodesic \( |z - a| = r \), \( \Im z > 0 \), the reflection on \( C \) is given
by
\[
R_C(x_1 + ix_2) = \frac{r^2}{(x_1 - a)^2 + x_2^2}(x_1 - a + ix_2) + (a + i0).
\]
If \(C\) is a vertical line through \((a, 0)\) then the reflection on \(C\) is
\[
R_C(x_1 + ix_2) = (2a - x_1 + ix_2).
\]
In both cases the reflection is an isometry.

### 7.3 The spaces \(\Phi\) and \(X\)

#### 7.3.1 Definitions

We begin with the CAT(0) space \(\Phi\) consisting of the geodesic \(|z| = 1\) in the Poincaré upper half-plane (noting that it is a convex subset) 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 \(H(\phi)\); that is,
\[
H(\phi) := \ln(\sec(\phi) + \tan(\phi)), \quad \text{for} \quad -\pi/2 < \phi < \pi/2.
\]
The \(\ell^2\)-direct product of \(\Phi\) with \(\mathbb{E}^1\), the 1-dimensional Euclidean space, produces the complete CAT(0) space
\[
X := \Phi \otimes \mathbb{E}^1
\]
and similarly the \(\ell^2\)-direct product of \(\Phi\) with \(|\mathbb{E}^1|\), the positive cone in 1-
dimensional Euclidean space, produces the complete CAT(0) space

\[ X_+ := \Phi \otimes_2 |E|^1, \]

where the metric is given by

\[
d_X((\phi_1, h_1), (\phi_2, h_2)) = \sqrt{(d_P(\phi_1, \phi_2))^2 + (h_1 - h_2)^2} \quad \text{(7.3.1.1)}
\]

\[
= \sqrt{(H(\phi_2) - H(\phi_1))^2 + (h_1 - h_2)^2} \quad \text{(7.3.1.2)}
\]

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

\[ \includegraphics[width=\textwidth]{figure7.2.png} \]

**Figure 7.2: Space** \( X := \Phi \otimes_2 \mathbb{R}^1 \)

**NOTE:** Both \( X \) and \( X_+ \) are flat (constant curvature of 0) CAT(0) spaces.

### 7.3.2 Geodesics in \( X \)

Since \( X \) and \( X_+ \) are CAT(0) spaces their geodesics are unique and their path has a minimum length with respect to the metric given by 7.3.1.2. We will examine the properties of these geodesics.
The geodesic $\Gamma_{P_1P_2}$ in $X (X_\perp)$ passing through the two distinct points $P_1 : (\phi_1, h_1)$ and $P_2 : (\phi_2, h_2)$ is the ‘vertical’ line $\{ (\phi_1, h) : h \in \mathbb{R} \}$ (‘vertical’ half line $\{ (\phi_1, h) : h > 0 \}$) when $\phi_1 = \phi_2$. Otherwise, when $\phi_1 \neq \phi_2$, $\Gamma_{P_1P_2}$ must have the functional form $h = h(\phi)$, for $-\frac{\pi}{2} < \phi < \frac{\pi}{2}$, where $h(\phi_i) = h_i$ for $i = 1, 2$.

Now, for any curve $\Gamma : h = h(\phi)$ passing through $P_1$ and $P_2$ the length of the segment between $P_1$ and $P_2$ is given by

$$l(\Gamma) = \int_{\phi_1}^{\phi_2} f(\phi, h'(\phi)) \, d\phi,$$

where $f(\phi, h'(\phi)) = \sqrt{\frac{1}{\cos^2(\phi)} + h'(\phi)^2}.$

We wish to determine $\Gamma$ such that $l(\Gamma)$ is a minimum. We use the Euler-Lagrange equation to minimize the length of the above curve which in this case reduces to the requirement, $\frac{d}{d\theta} \left( \frac{\partial f}{\partial h'} \right) = 0$, and so we obtain,

$$\frac{h''(\phi)}{h'(\phi)} = \tan(\phi),$$

$$\ln h'(\phi) = -\ln(\cos(\phi)) + C,$$

$$h'(\phi) = \frac{A}{\cos(\phi)}, \quad \text{so}$$

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

Note: Since $H(\phi)$ is one-to-one, the constants $A$ and $B$ are uniquely determined from the condition that $P_1, P_2 \in \Gamma$. In particular $A = \frac{h_1 - h_2}{H(\phi_1) - H(\phi_2)}$.

Figure 7.3: A geodesic in $X := \Phi \otimes E_1$
7.3.3 The length of geodesic segments in \( X (X_+) \)

The length of the segment of the above geodesic between \( P_1 : (\phi_1, h_1) \) and \( P_2 : (\phi_2, h_2) \) can be determined directly as \( d_P(P_1, P_2) \), or by calculating the arc-length of the geodesic segment, either way,

\[
l(\Gamma_{P_1P_2}) = \sqrt{1 + A^2 |H(\phi_1) - H(\phi_2)|} \\
= \frac{\sqrt{1 + A^2}}{|A|} |h_1 - h_2|,
\]

when \( \phi_1 \neq \phi_2 \), and

\[
l(\Gamma_{P_1P_2}) = |h_1 - h_2|,
\]

when the geodesic is vertical.

7.3.4 Mid-points of geodesic segments in \( X (X_+) \)

Again, we consider the geodesic segment between \( P_1 : (\phi_1, h_1) \) and \( P_2 : (\phi_2, h_2) \). The midpoint of this segment is \( P_m(\phi_m, h_m) \in \Gamma_{P_1P_2} \) such that \( l(\Gamma_{P_1P_m}) = l(\Gamma_{P_mP_2}) \). That is,

\[
\frac{\sqrt{1 + A^2}}{A} (h(\phi_1) - h(\phi_m)) = \frac{\sqrt{1 + A^2}}{A} (h(\phi_m) - h(\phi_2)) \quad \text{so,} \quad h(\phi_m) = \frac{h(\phi_1) + h(\phi_2)}{2}.
\]

To determine \( \phi_m \) we note that \( h(\phi_m) = AH(\phi_m) + B \), so setting \( H(\phi_m) = (h(\phi_m) - B)/A \) we have \( \ln(\sec(\phi_m) + \tan(\phi_m)) = H(\phi_m) \). Using this and the substitutions \( \sec(\phi) = \frac{1 + t^2}{1 - t^2} \) and \( \tan(\phi) = \frac{2t}{1 - t^2} \) where \( t = \tan \frac{\phi}{2} \), we reach the following quadratic equation in \( t \);

\[
(1 + \exp(H(\phi_m))) t^2 + 2t + (1 - \exp(H(\phi_m))) = 0,
\]

from which we find the solution for \( \phi_m \) within our range to be,

\[
\phi_m = 2 \tan^{-1} \left( \frac{\exp(H(\phi_m)) - 1}{\exp(H(\phi_m)) + 1} \right).
\]
### 7.3.5 Projections onto Geodesics in X

We begin by considering when two geodesics in $X$ are orthogonal to each other. Consider the geodesics $h_1(\phi) = A_1 H(\phi)$ (taking $B = 0$ without loss of generality) and $h_2(\phi) = A_2(\phi) = A_2 H(\phi) + B$.

In $X$, “locally” at $(\phi, h)$ the metric is given by

$$d^2 = \left( \frac{\delta \phi}{\cos(\phi)} \right)^2 + (\delta h)^2.$$  

So the inner product

$$k^2 \delta \phi_1 \delta \phi_2 + \delta h_1 \delta h_2 = 0 \text{ (where } k = \sec(\phi))$$  

if and only if

$$\frac{\delta h_2}{\delta \phi_2} = -k^2 \frac{\delta h_1}{\delta \phi_1}.$$  

Hence $h_2$ is orthogonal to $h_1$ at $\phi_0$ if and only if

$$h'_2(\phi_0) = A_2 \sec(\phi_0)$$

$$= -\sec^2(\phi_0)/h'_1(\phi_0)$$

$$= -\sec^2(\phi_0)/(A_1 \sec(\phi_0)).$$

That is these geodesics are orthogonal if and only if $A_1 A_2 = -1$.

Figure 7.4 shows the geodesics $h = 2H(\phi) + 4$ and $h = -0.5H(\phi) + 2$ intersecting orthogonally. We note that the representation of the angle here is skewed because the horizontal distance is calculated by the Poincaré metric while vertically it is the Euclidean metric. The Alexandrov angle is $\frac{\pi}{2}$.

We recall that the geodesic $h = H(\phi) + B$ is itself a convex set, so there is a unique nearest point projection (see section 3.5.2) of $P_1(\phi_1, h_1)$ onto $h = AH(\phi) + B$. This may be found by minimising the distance expression

$$\sqrt{(H(\phi) - H(\phi_1))^2 + (h - h_1)^2}.$$
7.3 The spaces $\Phi$ and $X$

Figure 7.4: Two geodesics intersecting orthogonally in $X$

and seen to be $P_2(\phi_2, h_2)$, where

$$H(\phi_2) = \frac{H(\phi_1) + A(h_1 - B)}{1 + A^2}$$

and

$$h_2 = AH(\phi_2) + B.$$

Since $X$ is a CAT(0) space and from Proposition 4.1 we also know projections in $X$ are nonexpansive.

7.3.6 Reflections in Geodesics in $X$

Since $X$ is a CAT(0) space with curvature bounded below the geodesic from $P_1$ to its projection $P_2$ can be uniquely extended to $P_3(\phi_3, h_3)$ such that $l(\Gamma_{P_1P_2}) = l(\Gamma_{P_2P_3})$, so $P_3$ is the reflection of $P_1$ in the geodesic $\Gamma : h = AH(\phi) + B$.

Noting that this requires $h_3 - h_2 = h_2 - h_1$ we have

$$h_3 = 2h_2 - h_1 = \frac{2AH(\phi_1) + h_1(A^2 - 1) + 2B}{1 + A^2}.$$

In the same way as above we can also write $H(\phi_2) = \frac{H(\phi_3) + A(h_3 - B)}{1 + A^2}$. By
7.4 The spaces $Y$ and $Y - X_+$

rearrangement and substitution for $H(\phi_2)$ we obtain

$$H(\phi_3) = \frac{(1 - A^2)H(\phi_1) + 2Ah_1 - 2AB}{1 + A^2}.$$ 

We now show that although $X$ is not one of the classical spaces $M^2_\kappa$ covered by theorem 6.4.1 none-the-less, reflections in $\Gamma$ are nonexpansive.

Let $P_1(\phi_{P1}, h_{P1})$ and $Q_1(\phi_{Q1}, h_{Q1})$ be two points in $X$ with their respective reflections $P_3(\phi_{P3}, h_{P3})$ and $Q_3(\phi_{Q3}, h_{Q3})$ in $h = AH(\phi) + B$ as above. Then

$$d(P_3, Q_3) = \sqrt{(H(\phi_{P3}) - H(\phi_{Q3}))^2 + (h_{P3} - h_{Q3})^2}$$

$$= \sqrt{((1 - A^2)^2 + 4A^2)(H(\phi_{P1}) - H(\phi_{Q1}))^2 + (4A^2 + (A^2 - 1)^2)(h_{P1} - h_{Q1})^2}$$

$$= \sqrt{(H(\phi_{P1}) - H(\phi_{Q1}))^2 + (h_{P1} - h_{Q1})^2}$$

$$= d(P_1, Q_1).$$

Therefore in the space $X$ reflections in geodesics are not only nonexpansive, but also isometries (as expected).

It then follows from Proposition 6.4.2 that the Douglas-Rachford algorithm converges in space $X$.

7.4 The spaces $Y$ and $Y - X_+$

In the upper half plane model of the hyperbolic space $\mathbb{H}^2_1$ let $Y = \{ z : \Im z > 0, |z| \leq 1 \}$ equipped with the metric $d_P$ inherited from $\mathbb{H}^2_{-1}$ then $Y$ is a closed, convex subset of $\mathbb{H}^2_{-1}$ and hence a $CAT(0)$ space of constant curvature $-1$.

Let $C$ be the geodesic in $Y$ given by

$$C = \{ e^{i\theta} : 0 < \theta < \pi \}.$$

132
Then $C$ is also a closed, convex subset of $Y$ and under the mapping

$$\phi \mapsto e^{i\left(\frac{\pi}{2} - \phi\right)},$$

$\Phi$ is isometric to $C$ where $\Phi$ has been defined in subsection 7.3.1.

$Y - X_+$ is obtained by gluing $X_+$ to $Y$ under the above identification of $\Phi$ with $C$ (see Figure 7.5) which by Reshetnyak’s gluing theorem (subsection 3.4.1) is a CAT(0) space of non-constant curvature, bounded below by -1 and above by 0. Hence, geodesics in $Y - X_+$ are uniquely extendable and so reflection in closed convex sets of $Y - X_+$ is well defined.

From its construction it is clear that the geodesic segment between two points of $Y - X_+$ both of which lie in $Y$ ($X_+$) is wholly contained in $Y$ ($X_+$) and the functional forms of such geodesics have already been discussed (section 7.2 and subsection 7.3.2). In determining geodesics in $Y - X_+$ two situations are of interest and are separately dealt with in following subsections.
7.4 The spaces $Y$ and $Y - X_+$

Figure 7.6: A model for $Y - X_+$ as a submanifold of $\mathbb{E}^3$.

(1) The case when we seek the geodesic from a point in $X_+$ to a point in $Y$. Of course once the functional forms of it in $X_+$ and in $Y$ have been found then its extension to half-rays in each space is also known. Here the problem is to determine the point $c \in \Phi$ (or $C$) at which the geodesic passes from $X_+$ into $Y$.

(2) The case where we seek to extend a given geodesic from a point in $Y$ ($X_+$) to a point in $\Phi$ into the other space $X_+$ ($Y$). This would be the situation for example where determining reflections in $\Phi$. Here, some sort of “transition condition” from $Y$ to $X_+$ (or $X_+$ to $Y$) is necessary.

7.4.1 Geodesic extensions between $Y$ and $X_+$

The extension of geodesics from $X_+$ to $Y$ or vice versa may be determined from the requirement that its gradient be continuous at the point where it meets $C$. We also know that geodesics follow the shortest path between two points. Since the metrics differ in $X_+$ and $Y$ because $X_+$ uses the Poincaré metric horizontally only and the Euclidean metric vertically we need to make an adjustment as the geodesic $\Gamma$ moves between these spaces to ensure that this is the case.
Figure 7.7: From $X_+$ to $Y$
Let us suppose $\Gamma_{P_1P_2}$ meets $C$ at $\phi = \phi_0$. Figure 7.7 shows the geodesic and its extension “locally” at the point where it crosses from $X_+$ into $Y$. For $0 \leq a \leq 1$ and using the respective metrics for $X_+$ and $Y$ (and writing $k = \frac{1}{\cos(\phi_0)}$ for convenience) we can write

$$\Delta l(a) = k\left(a^2\Delta \phi^2 + \Delta y^2\right)^{\frac{1}{2}} + \left(k^2(1-a)^2\Delta \phi^2 + \Delta y^2\right)^{\frac{1}{2}}.$$ 

By solving $\Delta l'(a) = 0$ we obtain the value of $a$ for which this distance is a minimum, that is $a = \frac{k}{1 + k}$. It then follows (see again Figure 7.7) that

$$\frac{\tan(\psi_1)}{\tan(\psi_2)} = \frac{1}{k} = \frac{1}{\cos(\phi_0)}.$$ 

We note that for $\phi_0 = 0$ we have $\psi_1 = \psi_2$ and for $\phi_0 \to \pi/2$ (or $-\pi/2$) we have $\psi_2 \to \pi/2$.

![Figure 7.8: Calculation to allow for change in metric](image)

Further we recall that geodesics in $X_+$ have equations of the form $h = AH(\phi) + B$ and the gradient of these geodesics is $A\sec(\phi)$ so the gradient of a geodesic at $\phi = \phi_0$ is given by $A\sec(\phi_0)$. We can also see that the gradient of the geodesic
in $X_+$ in Figure 7.8 can also be given by $\tan \alpha_1$ where $\alpha_1 = \frac{\pi}{2} - \psi_1$, that is $A \sec(\phi_0) = \tan(\alpha_1)$.

We also note from Figure 7.8 that $\alpha_2 = \frac{\pi}{2} - (\psi_2 + \phi_0)$ where $\tan(\alpha_2)$ is the gradient of the geodesic extension in $Y$. With this information we can show that

$$\tan(\alpha_2) = \cot(\psi_2 + \phi_0) = \frac{A - \tan(\phi_0)}{1 + A \tan(\phi_0)}$$

(7.4.1.1)

and by rearranging,

$$A = \frac{\tan(\alpha_2) + \tan(\phi_0)}{1 - \tan(\alpha_2) \tan(\phi_0)}$$

$$A = \tan(\alpha_2 + \phi_0).$$

(7.4.1.2)

If we have a geodesic in either $X_+$ or $Y$ then we can determine its geodesic extension into the adjacent space by using either 7.4.1.1 or 7.4.1.2. If the geodesic is extending from $X_+$ into $Y$ then we know $A$ and $\phi_0$ and the formula 7.4.1.1 gives $\alpha_2$ which enables us to calculate the equation of the geodesic extension into $Y$, we already know this has the form of the segment of a circle. If the extension is from $Y$ into $X_+$ then we know $\alpha_2$ and $\phi_0$ and formula 7.4.1.2 gives $A$ from which we can calculate the geodesic extension into $X_+$.

We use as an example the geodesic shown in Figure 7.10. In this case the geodesic $h = H(\phi) + 0.5$ in $X_+$ meets $C$ when $h = 0$, that is $\phi_0 = 0.480381$. So using 7.4.1.1 we have

$$\tan(\alpha_2) = \frac{A - \tan(\phi_0)}{1 + A \tan(\phi_0)}$$

$$= 0.314842,$$

from which we obtain the geodesic extension in $Y$ as

$$\{ z \in \mathbb{C} : |z - 0.741325| = 0.929734, \Im z > 0, |z| \leq 1 \}.$$

7.4.2 An upper-half plane model for $Y - X_+$.

In order to present a 2-dimensional representation of $Y - X_+$ we introduce an alternative representation of $X_+$ which we will call $W$. 

137
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

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

This naturally identifies $\Phi$ 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 $R(\theta) := H(\pi/2 - \theta) = \ln(\csc(\theta) + \cot(\theta)).$ (see Figure 7.9)

![Figure 7.9: An Upper-half plane model for $Y - X_+$](image)

### 7.4.3 Geodesics in the upper half-plane $Y - W$ model for $Y - X_+$

Geodesics in $W$ are radial rays: $z = \rho e^{i\theta}$, where $\rho \geq 1$ and $\theta$ is constant; or curves of the form,

$$z = (A R(\theta) + B)e^{i\theta}.$$ 

The geodesic segment joining $P_1 \in W$ to $P_2 \in Y$ is $[P_1, Q] \cup [Q, P_2]$ where $Q = e^{i\theta_0} \in C$ is chosen so that $d_W(P_1, Q) + d_\rho(Q, P_2)$ is a minimum.

The extension of a geodesic $\Gamma$ in $W$ across the “boundary” $C$ into $Y$ (or vice versa) may be determined from the requirement that $\frac{d\rho}{d\theta}$ be continuous at the point where $\Gamma$ meets $C$. However, in calculating the derivative on either side of the boundary we must also make an allowance in our calculations for the fact that

138
radially the metric changes between the spaces $Y$ and $W$ by a factor of $\frac{1}{\sin(\theta)}$. All these calculations have been implemented in Maple (see Appendix, Program 1). See Figures 7.10, 7.11 and 7.12 for some examples.

7.4.4 Using the $Y - W$ model

We use the $Y - W$ model instead of $Y - X_+$ when we want to present results pictorially because of its 2-dimensionality as figures 7.10, 7.11 and 7.12 illustrate. The results in the previous part can be converted for use in $W$ by using the conversion $\theta = \phi - \pi/2$ or equivalently $\phi = \theta - \pi/2$.

We now look at the problem of finding the geodesic between two points; one in $W$ and one in $Y$. Here the equation of neither the segment of the geodesic in $W$
or in $Y$ is known. The problem is to find a point on $C$ such that the distance between these points is a minimum from which we will obtain the geodesic and its extension in both $Y$ and $W$.

Let us consider the points $P_1(\rho_1, \theta_1)$ in $Y$ and $P_2(\rho_2, \theta_2)$ in $W$ where $0 < \rho_1 < 1$, $\rho_2 > 1$ and $0 < \theta_1, \theta_2 < \pi$. We wish to find the point $P(1, \theta)$ on $C$ such that $\Gamma_{P_1P} + \Gamma_{PP_2}$ is a minimum. This is a calculus minimization problem which we write as follows where $P$ is $(1, \theta)$:

$$
\text{Minimize } l(\Gamma_{P_1P_2})(\theta) = \sqrt{\left(\cosh^{-1}\left(1 + \frac{1 - \cos(\theta_1 - \theta_2)}{\sin(\theta) \sin(\theta_2)}\right)\right)^2 + (\rho_2 - 1)^2} \\
+ \cosh^{-1}\left(1 + \frac{\rho_1^2 + 1 - 2\rho_1 \cos(\theta_1 - \theta)}{2\rho_1 \sin(\theta_1) \sin(\theta)}\right).
$$

This is an expression in one variable ($\theta$) only and so a value of $\theta$ which minimizes it can be found by standard methods.

As an example let us consider the geodesic and its extension from $P_1(0.5, \pi/4)$ in $Y$ to $P_2(2, \pi/6)$ in $W$. Using Maple to minimise the above expression gives $\theta = 0.755987$ (see Figure 7.13 for this geodesic). We can now calculate the equation of the geodesic in $W$ to be $\rho = 2.542194 R(\theta) - 1.347962$ and from there we are able to calculate the equation of the extension into $Y$ to be $\{z \in \mathbb{C} : |z - 1.002563| = 0.739063, |z| \leq 1, \Im z > 0\}$. For the Maple calculation see Appendix, Program 2.
Figure 7.13: Graph of geodesic and its extension
7.4 The spaces $Y$ and $Y - X_+$

7.4.5 Reflections in $C$ need not be nonexpansive

To exploit the structure of $Y - X_+$ we consider reflections in $C$ (identified with $\Phi$). The projection of a point $z$ onto $C$ is determined by the condition that the geodesic segment $[z, P_C(z)]$ meets $C$ orthogonally (in terms of Alexandrov angles).

The reflection of $z$ is the point $R_C(z)$ on the extension of $[z, P_C(z)]$ such that $P_C(z)$ is the midpoint of $[z, R_C(z)]$.

From the uniqueness of geodesics it follows that,

$$R_C|_Y : Y \to W \quad \text{and} \quad R_C|_W : W \to Y,$$

are inverses.

![Figure 7.14: Reflections in C.](image)

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

$$d_{Y - X_+}(R_C(Q_1), R_C(Q_2)) = d_Y(P_1, P_2) > d_{X_+}(Q_1, Q_2).$$

Figure 7.14 illustrates these reflections.
A similar example was independently given in [39], but we believe our prototypical space \( Y - X_+ \) (or its \( Y - W \) model) is more amenable to calculations of the type explored here.

### 7.5 The Douglas-Rachford algorithm.

We have seen that reflections are not, in general, nonexpansive in spaces of non-constant curvature, for example in \( Y - X_+ \). We now will examine how the Douglas-Rachford algorithm behaves in a space such as \( Y - X_+ \) by applying it to a range of examples in this space. Again we will use \( W \) instead of \( X_+ \), so point specifications will be of the form \((\rho, \theta)\) where \(0 < \theta < \pi\) and \(0 < \rho \leq 1\) in \(Y\) and \(\rho \geq 1\) in \(W\).

So that the feasibility problem is not localised in either \(Y\) or \(X_+(W)\), we take as our two sets the closed half-rays \(C_1 = \{e^{i\theta} : \theta_0 \leq \theta \leq \pi\}\) and \(C_2 = \{e^{i\theta} : 0 \leq \theta \leq \theta_0\}\) of \(C\), with \(C_1 \cap C_2 = \{e^{i\theta_0}\} = (1, \theta_0)\). For the purpose of illustration we will use \(\theta_0 = \frac{3\pi}{4}\) (see Figure 7.15).

Reflections in \(C\) have already been discussed in 7.4.5 and so we have the necessary techniques to consider examples of the Douglas Rachford-Algorithm in this case.

![Figure 7.15: The intersecting convex sets \(C_1\) and \(C_2\)](image-url)
First Case: $Y$ to $W$

We begin with the point $x_1 = 0.733799e^{0.735867i}$, see Table 1 (Figure 7.16) which is represented by A in Figure 7.17 and project to $C_2$ to obtain $P_{C_2}x_1$. The reflection of $x_1$ in $C_2$ gives $R_{C_2}x_1$ (represented by B in the diagram). The projection of $R_{C_2}x_1$ to $C_1$ is $P_{C_1}R_{C_2}x_1$ which is the point $(1, 3\pi/4)$ (the nearest point of $C_1$) and the reflection of $R_{C_2}x_1$ is $R_{C_1}R_{C_2}x_1$ which is indicated by C in the diagram.

The average of $x_1$ and $R_{C_1}R_{C_2}x_1$ lies on the geodesic joining these points and from the midpoint formula in section 7.2 we obtain $x_2$ (indicated by D in the diagram). Table 1 shows this cycle and two more iterations (E shows the first reflection of the second cycle). The iterates appear to be rapidly stabilizing with $P_{C_2}x_n$ converging to the feasible point, $(1, 2.35619449)$.

<table>
<thead>
<tr>
<th>$n$</th>
<th>$x_n$</th>
<th>$P_{C_2}x_n$</th>
<th>$R_{C_2}x_n$</th>
<th>$P_{C_1}R_{C_2}x_n$</th>
<th>$R_{C_1}R_{C_2}x_n$</th>
<th>$x_{n+1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$(0.733799, 0.735867)$</td>
<td>$(1.0, 0.785398)$</td>
<td>$(1.452881, 0.785598)$</td>
<td>$(1.452881, 0.785598)$</td>
<td>$(1.452881, 0.785598)$</td>
<td>$(0.872208, 2.341632)$</td>
</tr>
<tr>
<td>2</td>
<td>$(0.872208, 2.341632)$</td>
<td>$(1.232662)$</td>
<td>$(1.190054, 2.332662)$</td>
<td>$(1.190054, 2.332662)$</td>
<td>$(1.190054, 2.332662)$</td>
<td>$(0.874992, 2.365024)$</td>
</tr>
<tr>
<td>3</td>
<td>$(0.874992, 2.365024)$</td>
<td>$(1.356656)$</td>
<td>$(1.189967, 2.356056)$</td>
<td>$(1.189967, 2.356056)$</td>
<td>$(1.189967, 2.356056)$</td>
<td>$(0.875008, 2.365161)$</td>
</tr>
</tbody>
</table>

Figure 7.16: Table 1

Table 2 (Figure 7.18) shows iterates from an alternative starting point $y_1 = (0.5, \pi)$ in $Y$ and we can see that it behaves similarly to the previous case.
7.5 The Douglas-Rachford algorithm.

<table>
<thead>
<tr>
<th>$y_n$</th>
<th>$n = 1$</th>
<th>$n = 2$</th>
<th>$n = 3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$y_{C_1}$</td>
<td>(0.5, $\frac{\pi}{4}$)</td>
<td>(0.703465, 2.380345)</td>
<td>(0.712493, 2.414531)</td>
</tr>
<tr>
<td>$y_{C_2}$</td>
<td>(1.232045)</td>
<td>(1.345678)</td>
<td>(1.456789)</td>
</tr>
<tr>
<td>$y_{C_1}$</td>
<td>(1.232045)</td>
<td>(1.345678)</td>
<td>(1.456789)</td>
</tr>
<tr>
<td>$y_{C_2}$</td>
<td>(1.232045)</td>
<td>(1.345678)</td>
<td>(1.456789)</td>
</tr>
<tr>
<td>$y_{n+1}$</td>
<td>(0.703465, 2.380345)</td>
<td>(0.712493, 2.414531)</td>
<td>(0.712825, 2.415799)</td>
</tr>
</tbody>
</table>

Table 3 (7.19) and Table 4 (7.21) show iterations starting in space $W$ and moving clockwise. Both these stabilise after one round to a fixed point which projects to the intersection $C_1 \cap C_2$. The diagram 7.20 traces the movement in Table 3. A is reflected in $C_2$ to B which in turn is reflected in $C_1$ to C. The average of $x_1$ and $C_1 R_C x_1$ is represented by D. The reflection of $x_2$ in $C_2$ is represented by E and the reflection of this point in $C_1$ is represented by D. So we have a fixed point which projects to the intersection $C_1 \cap C_2$.

Finally we note that if any of the initial points $x_1$ were to be projected initially to the closed convex set $C_1$ then that initial point would become immediately a fixed point of the mapping $T := \frac{1}{2}(I + R_C R_{C_1})$ and the projection of this fixed point to $C_1$ obviously belongs to $C_1 \cap C_2$. 

These two tables also show that in these instances the iterated map $T := \frac{1}{2}(I + R_C R_{C_1})$ is nonexpansive, even though some intermediary steps are not.

Specifically:

$$d_Y(Tx_1, Ty_1) = 0.3098 \leq d_Y(x_1, y_1) = 1.0476$$

$$d_Y(Tx_2, Ty_2) = 0.3056 \leq d_Y(x_2, y_2) = 0.3098$$

$$d_Y(Tx_3, Ty_3) = 0.3056 \leq d_Y(x_3, y_3) = 0.3056.$$ 

While $d(R_C R_{C_1} x_1, R_{C_1} R_{C_1} y_1) = 1.0500 > d(x_1, y_1) = 1.0476$

Thus, while reflections in CAT(0) spaces of non-constant curvature need not be nonexpansive, it appears that the averaging process in the Douglas-Rachford scheme may compensate for this.
7.5 The Douglas-Rachford algorithm.

<table>
<thead>
<tr>
<th></th>
<th>( n = 1 )</th>
<th>( n = 2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( x_n )</td>
<td>((1.75, \frac{4\pi}{16}))</td>
<td>((1.464634, 2.442068))</td>
</tr>
<tr>
<td>( PC_2 x_n )</td>
<td>((1, \frac{\pi}{16}))</td>
<td>((1, \frac{3\pi}{4}))</td>
</tr>
<tr>
<td>( RC_2 x_n )</td>
<td>((0.882893, 0.152446))</td>
<td>((0.707278, 2.320714))</td>
</tr>
<tr>
<td>( PC_1 RC_2 x_n )</td>
<td>((1, \frac{3\pi}{4}))</td>
<td>((1, \frac{3\pi}{4}))</td>
</tr>
<tr>
<td>( RC_1 RC_2 x_n )</td>
<td>((1.179269, 3.115386))</td>
<td>((1.464634, 2.442068))</td>
</tr>
<tr>
<td>( x_{n+1} )</td>
<td>((1.464634, 2.442068))</td>
<td>((1.464634, 2.442068))</td>
</tr>
</tbody>
</table>

**Figure 7.19: Table 3**

<table>
<thead>
<tr>
<th></th>
<th>( n = 1 )</th>
<th>( n = 2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( y_n )</td>
<td>((2, \frac{11\pi}{10}))</td>
<td>((1.993389, 2.383728))</td>
</tr>
<tr>
<td>( PC_2 y_n )</td>
<td>((1, \frac{11\pi}{10}))</td>
<td>((1, \frac{3\pi}{4}))</td>
</tr>
<tr>
<td>( RC_2 y_n )</td>
<td>((0.473875, 2.371483))</td>
<td>((0.540320, 2.544466))</td>
</tr>
<tr>
<td>( PC_1 RC_2 y_n )</td>
<td>((1, \frac{3\pi}{4}))</td>
<td>((1, \frac{3\pi}{4}))</td>
</tr>
<tr>
<td>( RC_1 RC_2 y_n )</td>
<td>((1.986778, 2.564898))</td>
<td>((1.993389, 2.383728))</td>
</tr>
<tr>
<td>( y_{n+1} )</td>
<td>((1.993389, 2.383728))</td>
<td>((1.993389, 2.383728))</td>
</tr>
</tbody>
</table>

**Figure 7.20: Diagram for Table 3**

**Figure 7.21: Table 4**

146
7.6 Conclusions

In this chapter we developed a prototype of a CAT(0) space of non-constant curvature. This space has curvature bounded below by -1 and above by 0. We developed a comprehensive collection of formulae for geodesics extensions, midpoints, metric projection onto geodesics (or half spaces) and reflections in our prototype space which enabled us to show that reflections are not nonexpansive in general. We investigated the behaviour of the Douglas-Rachford algorithm in this space and found that despite the lack of nonexpansiveness for some of the reflections involved the method appeared to work, converging to the point of intersection of the feasible sets or to a limit which then projected onto the point of intersection of the feasible sets. Effects of averaging seemed to make a significant contribution to convergence of the algorithm as possibly did the alternation between expansiveness and contractions of the iterated maps.
Conclusions

Following is a summary of the material in this thesis.

Chapter 1 gave an introduction and overview of our work.

Chapter 2 surveyed metric and geodesic spaces. We paid particular attention to convexity (Menger and Busemann) in geodesic spaces and gave a proof for Menger’s theorem that metric and geodesic convexity coincide.

Chapter 3 began with CAT(\(\kappa\)) spaces and then moved to CAT(0) spaces. We covered polarization in CAT(0) spaces (section 3.3) especially the new work on angles. We used Reshetnyak’s theorem to build new spaces from old and gave an example which illustrated the failure of property N. We finished with an examination of analogues of weak convergence in CAT(0) spaces.

Chapter 4 contains several items of new work;

(1) Hyperplanes and half-spaces in CAT(0) spaces (subsection 4.2.1);

(2) A separation theorem in CAT(0) spaces (subsection 4.2.2);

(3) Hyperplane characterization of weak convergence (subsection 4.2.3).

We finished with a consideration of differentiability of convex functions and the concept of subdifferentials in CAT(0) spaces.
Chapter 5 dealt with fixed point theory in CAT(0) spaces. We saw that the notion of weak sequential convergence allows the development in CAT(0) spaces of a fixed point theory for nonexpansive type mappings analogous to that available for Hilbert spaces. We have shown that closed bounded convex sets of a CAT(0) space have the approximate fixed point property. We also studied averaged mappings and established that the averaged map is asymptotically regular and firmly nonexpansive whenever $T$ is nonexpansive.

Chapter 6 contained several items of new work;

(1) Convergence results for Alternating Projections in CAT(0) spaces (subsection 6.2.2);

(2) Non-expansivity of reflections in $\mathbb{R}$–trees (subsection 6.4.2).

We investigated the nonexpansiveness of reflection in CAT(0) spaces.

Chapter 7

Section 7.3 and onwards presents our investigation of reflections and the Douglas-Rachford algorithm in a prototypical CAT(0) space of non-constant curvature.

Areas for further investigation;

(1) While reflections in CAT(0) spaces of non-constant curvature need not be nonexpansive, it appears that the averaging process in the Douglas-Rachford scheme may compensate for this. Is this the case and does it always happen?

(2) In Table 1 from section 7.5 the iterates appear to be rapidly stabilizing with $P_{C_2}x_n$ converging to the feasible point. Again is this the case and does it always happen?

(3) Extension outside of CAT(0) spaces. Investigation of positive curvature spaces and other classes of geodesic spaces for example.

Other areas could be the following;

(1) For $a, b, y$ and $p \in X$ is $\langle ta \oplus (1 - t)b, y \rangle_p \leq \langle a, y \rangle_p + (1 - t)\langle b, y \rangle_p$ for $t \in [0, 1]$; that is, $\langle ta \oplus (1 - t)b, y \rangle_p$ a convex (or a concave) function of $t$ on $[0, 1]$? In Hilbert spaces it is of course an affine function for $t \in \mathbb{R}$.
(2) How, if at all, is \( \langle x_n, y \rangle_x \rightarrow 0 \) related to the sequence \( (x_n) \) converging weakly to \( x \) (here weak convergence means \( \Delta \) or equivalently \( \phi \) convergence in \( X \))? Again, in Hilbert spaces they are equivalent.

**Remark:** Let \( X \) be the weighted metric tree with vertices \( x, y, m \) and \( p \) and edges \([x, m], [y, m] \) and \([p, m]\) of respective lengths \( \alpha, \alpha \) and \( \beta \). Then \( \langle x, y \rangle_p = \beta^2 - \alpha^2 \), while \( d(x, p)d(y, p) = (\alpha + \beta)^2 \), so here we have,

\[
|\langle x, y \rangle_p| \leq d(x, p)d(y, p).
\]

We ask: Is this true in general?
PAGE LEFT BLANK INTENTIONALLY
Appendix

The appendix contains Maple programs for some of the examples contained in Chapter 7.

(1) Program 1 is an example of a Maple program used to draw geodesics as in 7.10, 7.11 and 7.12,

(2) Program 2 is the program used to solve the example at the end of subsection 7.4.4 and then draw the geodesic,

(3) Program 3 is the program used to produce the the data for Table 3 and draw the diagram to go with it in section 7.5.

These programs are on the following pages.
\[ f := (a, b) \rightarrow a \cdot \ln(\csc(\theta) + \cot(\theta)) + b \]
\[ f := (a, b) \rightarrow a \ln(\csc(\theta) + \cot(\theta)) + b \quad (1) \]

\[ f(a, b) = 1; \]
\[ u := \text{unapply(solve(%)}, (a, b)); \]
\[ simplify \left( a \cdot \cos(u(a, b)) + \sin(u(a, b)) \right); \]
\[ c := \text{unapply(%, (a, b)}; \]
\[ simplify \left( \sqrt{1 - 2 \cdot c(a, b) \cdot \cos(u(a, b)) + c(a, b)^2} \right); \]
\[ r := \text{unapply(%, (a, b)}; \]
\[ simplify \left( \pi \cdot \sin^{-1} \left( \frac{\sin(u(a, b))}{r(a, b)} \right) \right); \]
\[ cb := \text{unapply(%, (a, b))} \]

\[ a \ln(\csc(\theta) + \cot(\theta)) + b = 1 \]
\[ u := (a, b) \rightarrow \arctan \left( \frac{2 e - \frac{b - 1}{a}}{1 + \left( \frac{b - 1}{a} \right)^2} \right) - 1 \]
\[ 1 + \left( \frac{b - 1}{a} \right)^2 \]
\[ c := (a, b) \rightarrow \frac{a \left( e - \frac{a}{a} + 1 \right)}{a e - a + 2 e} \]
\[ r := (a, b) \rightarrow 2 \sqrt{\frac{2 \left( \frac{b - 1}{a} \right)}{e - \frac{a}{\left( \frac{b - 1}{a} \right)^2}}} \]
\[ \left( \frac{2 \left( \frac{b - 1}{a} \right)}{e - \frac{a}{\left( \frac{b - 1}{a} \right)^2}} \right)^2 \]
\[ cb := (a, b) \rightarrow \pi \cdot \arcsin \left( \frac{c \left( \frac{b - 1}{a} \right)}{\left( a e - a + 2 e \right)} \right) + 1 \]

\[ A, B, l := 3.0111, -0.6540, \frac{\pi}{4}; \]
\[ 'u'(A, B) = u(A, B); \]
\[ 'c'(A, B) = c(A, B); \]

Figure 9.1: Program 1 Page 1
\[ r(A, B) = r(A, B); \]
\[ cb(A, B) = cb(A, B); \]
\[ \#r := (A, B) \rightarrow 1.3896; \]

\[ \text{refPl} := \text{plots}\{\text{polarplot}\}(1, \theta = 0..\pi, \text{ colour } = \text{ red}); \]
\[ \text{Pl} := \text{plots}\{\text{polarplot}\}(f(A, B), \theta = 1..u(A, B), \text{ colour } = \text{ black}); \]

\[ \text{if } |A + \sin(u(A, B)) \cdot \tan(u(A, B))| < .0001 \text{ then} \]
\[ \text{Pl2} := \text{plots}\{\text{polarplot}\}\left(\frac{1}{2 \cdot \cos(\theta)}, \theta = 0..u(A, B), \text{ colour } = \text{ blue}\right); \]
\[ \text{else if } c(A, B) < 0 \text{ then} \]
\[ \text{Pl2} := \text{plot}\{[c(A, B) - r(A, B) \cdot \cos(\theta), r(A, B) \cdot \sin(\theta), \theta = \text{Pi} - cb(A, B) ..\text{Pi}], \text{ colour } = \text{ blue}\}; \]
\[ \text{else Pl2} := \text{plot}\{[c(A, B) - r(A, B) \cdot \cos(\theta), r(A, B) \cdot \sin(\theta), \theta = cb(A, B) ..\text{Pi}], \text{ colour } = \text{ blue}\}; \]
\[ \text{fi; \text{else if } c(A, B) < \cos(u(A, B)) \text{ then} \text{Pl2} := \text{plot}\{[c(A, B) + r(A, B) \cdot \cos(\theta), r(A, B) \cdot \sin(\theta), \theta = \text{Pi} - cb(A, B) ..\text{Pi}], \text{ colour } = \text{ blue}\}; \]
\[ \text{else Pl2} := \text{plot}\{[c(A, B) + r(A, B) \cdot \cos(\theta), r(A, B) \cdot \sin(\theta), \theta = cb(A, B) ..\text{Pi}], \text{ colour } = \text{ blue}\}; \]
\[ \text{fi fi}; \]

\[ \text{plots}\{\text{display}\}\{[\text{refPl}, \text{Pl}, \text{Pl2}], \text{thickness } = 2\} \]

\[ A, B, l := 3.0111, -0.6540, \frac{1}{4} \pi \]
\[ u(3.0111, -0.6540) = 1.047202076 \]
\[ c(3.0111, -0.6540) = 1.269667455 \]
\[ r(3.0111, -0.6540) = 1.158618980 \]
\[ cb(3.0111, -0.6540) = \pi - 0.8442384902 \]

Figure 9.2: Program 1 Page 2
Figure 9.3: Program 1 Page 3
> restart

> rho1, rho2, theta1, theta2 := 0.5, 2, \frac{\pi}{4}, \frac{\pi}{6}:

> rho1, rho2, theta1, theta2 := 0.5, 2, \frac{\pi}{4}, \frac{\pi}{6} \\

(13)

> f := (theta) -> sqrt\left(\left(\arccosh\left(\frac{\sin(\theta) - \sin(\theta_2)}{\sin(\theta_2) - \sin(\theta)}\right)\right)^2 + (\rho_2 - 1)^2\right) + \arccosh\left(\frac{\rho_1 + 1 - 2 \rho_1 \cos(\theta_1 - \theta)}{2 \rho_1 \sin(\theta_1) \sin(\theta)}\right)

f := \theta \rightarrow \sqrt{\arccosh\left(\frac{1 - \cos(\theta - \theta_2)}{\sin(\theta) \sin(\theta_2)}\right)^2 + (\rho_2 - 1)^2} + \arccosh\left(\frac{\rho_1 + 1 - 2 \rho_1 \cos(\theta_1 - \theta)}{\rho_1 \sin(\theta_1) \sin(\theta)}\right) \\

(14)

> g := (theta) -> diff(f(theta), theta):

> g := (\theta) \rightarrow \frac{1}{2} \left(\arccosh\left(1 + \frac{2 \left(1 - \sin\left(\frac{\theta + \frac{\pi}{3}}{\sin(\theta)}\right)\right)}{\cos(\theta) \sin(\theta)}\right)\right) - \frac{2 \left(1 - \sin\left(\frac{\theta + \frac{\pi}{3}}{\sin(\theta)}\right)\right) \cos(\theta)}{\sin(\theta)^2} \sqrt{2} \right) \bigg/ \left(\sqrt{\arccosh\left(1 + \frac{2 \left(1 - \sin\left(\frac{\theta + \frac{\pi}{3}}{\sin(\theta)}\right)\right)}{\sin(\theta)}\right)^2} + 1 \sqrt{1 - \sin\left(\frac{\theta + \frac{\pi}{3}}{\sin(\theta)}\right)}\right) + \left(1.000000000 \cos\left(\frac{\frac{\pi}{4} + \theta}{\sin(\theta)}\right)\right) \sqrt{2} \\

\left(1.250000000 - 1.000000000 \cos\left(\frac{\frac{\pi}{4} + \theta}{\sin(\theta)}\right)\right) \sqrt{2} \bigg/ \left(1.000000000 \left(1.25 - 1.0 \cos\left(\frac{\frac{\pi}{4} + \theta}{\sin(\theta)}\right)\right) \sqrt{2} \sin(\theta)\right) \\

(15)

Figure 9.4: Program 2 Page 1
\[ \sqrt{2 + \frac{1.000000000 \left(1.25 - 1.0 \cos \left(-\frac{1}{4} \pi + \theta \right) \right)}{\sin (\theta)}} \]

\[ \theta_s := \text{fsolve}(g \text{theta} = 0, \text{theta}, 0..\pi) \];
\[ \theta_s := 0.7559873561 \]  \hspace{1cm} (16)

\[ c := \text{evalf} \left( \frac{1-ho_1^2}{2 \cdot (\cos(\theta) - \rho_1 \cos(\theta_1))} \right) ; \]
\[ c := 1.002563424 \]  \hspace{1cm} (17)

\[ r := \text{evalf} \left( \sqrt{(\cos(\theta_s) - c)^2 + \sin^2(\theta_s)} \right) ; \]
\[ r := 0.7390629363 \]  \hspace{1cm} (18)

\[ A := \text{evalf} \left( \frac{\rho_2 - 1}{\ln(csc(\theta_2) + cot(\theta_2)) - \ln(csc(\theta_1) + cot(\theta_1))} \right) \]
\[ A := 2.542193977 \]  \hspace{1cm} (19)

\[ B := \text{evalf} \left( 1 - A \cdot \ln(csc(\theta_1) + cot(\theta_1)) \right) \]
\[ B := -1.347962434 \]  \hspace{1cm} (20)

\[ \text{restart} \]
\[ f := (a, b) \rightarrow a \cdot \ln(csc(\theta) + cot(\theta)) + b \]
\[ f := (a, b) \rightarrow a \ln(csc(\theta) + cot(\theta)) + b \]  \hspace{1cm} (21)

\[ f(a, b) = 1; \]
\[ a := \text{unapply}(\text{solve}(%, \theta), (a, b)); \]
\[ simplify \left( \frac{a}{\cos(u(a,b)) + \sin(u(a,b))} \right) ; \]
\[ c := \text{unapply}(\%, (a, b)); \]
\[ simplify \left( \sqrt{1 - 2c(a,b) \cdot \cos(u(a,b)) + c(a,b)^2} \right) ; \]
\[ r := \text{unapply}(\%, (a, b)); \]
\[ simplify \left( \pi - \sin^{-1} \left( \frac{\sin(u(a,b))}{r(a,b)} \right) \right) ; \]
\[ cb := \text{unapply}(\%, (a, b)) \]

\[ a \ln(csc(\theta) + cot(\theta)) + b = 1 \]

Figure 9.5: Program 2 Page 2
\[ u := (a, b) \rightarrow \arctan \left( \frac{2 e^{-\frac{b-1}{a}}}{1 + \left( e^{-\frac{b-1}{a}} \right)^2} \right), \]
\[ c := (a, b) \rightarrow \frac{a \left( \frac{2 (b-1)}{a} + 1 \right)}{a e^{-\frac{2 (b-1)}{a}} - a + 2 e^{-\frac{b-1}{a}}} \]
\[ r := (a, b) \rightarrow 2 \sqrt{\frac{a e^{-\frac{2 (b-1)}{a}} - a + 2 e^{-\frac{b-1}{a}}}{e^{-\frac{2 (b-1)}{a}} - a + 2 e^{-\frac{b-1}{a}}}} \]
\[ cb := (a, b) \rightarrow \pi - \arcsin \left( \sqrt{\frac{a e^{-\frac{2 (b-1)}{a}} - a + 2 e^{-\frac{b-1}{a}}}{e^{-\frac{2 (b-1)}{a}} - a + 2 e^{-\frac{b-1}{a}}}} \right) \]

> A, B, l := 2.5422, -1.3480, \frac{\pi}{6};

'\( u'(A, B) = u(A, B) \);
'\( c'(A, B) = c(A, B) \);
'\( r'(A, B) = r(A, B) \);
'\( cb'(A, B) = cb(A, B) \);

refPl := plots[polarplot](\[1, \theta = 0 \ldots \pi, \text{color} = \text{red}\]) :
Pl := plots[polarplot](\[f(A, B), \theta = 1..u(A, B), \text{color} = \text{black}\]) :
if |A + \sin(u(A, B)) \cdot \tan(u(A, B))| < .0001 then
Pl2 := plots[polarplot]\(\left[\begin{array}{c}
\frac{1}{2 \cos(\theta)}, 
\theta = 0..u(A, B), 
\text{color} = \text{blue}
\end{array}\right]\):
else if c(A, B) < 0 then
if \( \cos(u(A, B)) < c(A, B) \) then
Pl2 := plot([c(A, B) - r(A, B) \cdot \cos(\theta), r(A, B) \cdot \sin(\theta), \theta = Pi - cb(A, B) \cdot Pi], \text{color} = \text{blue}) :
else Pl2 := plot([c(A, B) - r(A, B) \cdot \cos(\theta), r(A, B) \cdot \sin(\theta), \theta = cb(A, B) \cdot Pi], \text{color} = \text{blue}) fi:
else if c(A, B) < \cos(u(A, B)) then
Pl2 := plot([c(A, B) + r(A, B) \cdot \cos(\theta), r(A, B) \cdot \sin(\theta), \theta = Pi - cb(A, B) \cdot Pi], \text{color} = \text{blue}) fi:
else Pl2 := plot([c(A, B) + r(A, B) \cdot \cos(\theta), r(A, B) \cdot \sin(\theta), \theta = cb(A, B) \cdot Pi], \text{color} = \text{blue}) fi:
fi:
plots[display](\[\text{refPl, Pl, Pl2}, \text{thickness} = 2\])

Figure 9.6: Program 2 Page 3
\[ A, B, l := 2.5422, -1.3480, \frac{1}{6} \pi \]
\[ u(2.5422, -1.3480) = 0.7559787199 \]
\[ c(2.5422, -1.3480) = 1.002560596 \]
\[ r(2.5422, -1.3480) = 0.7390538476 \]
\[ cb(2.5422, -1.3480) = \pi - 1.18956844 \]

**Figure 9.7: Program 2 Page 4**
with plots:

\[ \rho_1, \rho_2, \theta_1, \theta_2, \theta_3 := 1.75, 1, \frac{\pi}{16}, \frac{\pi}{16}, \frac{3\pi}{4}; \]

\[ Q := \text{plots}[\text{polarplot}](1, \theta = 0..\theta_3, \text{colour} = \text{red}); \]

\[ P := \text{plots}[\text{polarplot}](1, \theta = \theta_3..\pi, \text{colour} = \text{green}); \]

\[ \text{proj1} := \text{plots}[\text{polarplot}](\rho, \theta = 1..\rho_1, \text{colour} = \text{black}); \]

\[ c_1 := \sec(\theta_2); \]

\[ r_1 := \sqrt{\left(\cos(\theta_2) - c_1\right)^2 + \sin^2(\theta_2)}; \]

\[ f(x) := \text{arccosh}\left(1 + \frac{\left(x - \cos(\theta_2)\right)^2 + \left(\sqrt{r_1^2 - (x - c_1)^2} - \sin(\theta_2)\right)^2}{2\sin(\theta_2)\cdot\sqrt{r_1^2 - (x - c_1)^2}}\right) - (\rho_1 - \rho_2); \]

\[ x_3 := \text{fsolve}(f(x) = 0, x = -1..1); \]

\[ y_3 := \sqrt{r_1^2 - (x_3 - c_1)^2}; \]

\[ \rho_r := \text{evalf}\left(\sqrt{\frac{y_3^2}{x_3^2 + y_3^2}}\right); \]

\[ \text{if } x_3 > 0 \text{ then } \theta_r := \text{evalf}\left(\tan^{-1}\left(\frac{y_3}{x_3}\right)\right) \text{ else } \theta_r := \text{evalf}\left(\pi - \tan^{-1}\left(\frac{y_3}{x_3}\right)\right) \text{ fi.} \]

\[ \text{refl1} := \text{plot}([x, \sqrt{r_1^2 - (x - c_1)^2}, x = x_3..\cos(\theta_3), \text{colour} = \text{blue}); \]

\[ c_2 := \frac{1 - (x_3^2 + y_3^2)}{2\cos(\theta_3) - x_3}; \]

\[ r_2 := \sqrt{\left(\cos(\theta_3) - c_2\right)^2 + \sin^2(\theta_3)}; \]

\[ \text{proj2} := \text{plot}([x, \sqrt{r_2^2 - (x - c_2)^2}, x = \cos(\theta_3)..x_3, \text{colour} = \text{blue}); \]

Figure 9.8: Program 3 Page 1
\[ A_1 := \text{evalf} \left( \frac{\left( c_2 - \cos (\theta_3) \right)}{\sin (\theta_3)} + \tan (\theta_3) \right) \];

\[ B_1 := \text{evalf} \left( \rho_2 - A_1 \ln \left( \csc (\theta_3) + \cot (\theta_3) \right) \right) ; \]

\[ d\left( (x_3, y_3), \left( x_3, \cos (\theta_3), \sin (\theta_3) \right) \right) := \text{evalf} \left( \cosh^{-1} \left( 1 + \frac{\left( (x_3 - \cos (\theta_3))^2 + (y_3 - \sin (\theta_3))^2 \right)}{2 \cdot y_3 \cdot \sin (\theta_3)} \right) \right); \]

\[ \rho_3 := \text{evalf} \left( \frac{d\left( (x_3, y_3), \left( x_3, \cos (\theta_3), \sin (\theta_3) \right) \right) \cdot A_1}{\sqrt{1 + A_1^2}} \right) + \rho_2; \]

\[ R := \text{evalf} \left( \frac{\left( \rho_3 - B_1 \right)}{A_1} \right); \]

\[ \theta_4 := \text{evalf} \left( 2 \cdot \arctan \left( \exp (-R) \right) \right); \]

\[ \text{refl2} := \text{plots}[\text{polarplot}](A_1 \ln \left( \csc (\theta) + \cot (\theta) \right) + B_1, \theta = \theta_3 .. \theta_4, \text{color} = \text{black}); \]

\[ \rho_m := \frac{\left( \text{rho}_4 + \rho_3 \right)}{2}; \]

\[ R(\theta) := \ln \left( \csc (\theta) + \cot (\theta) \right); \]

\[ A_2 := \text{evalf} \left( \frac{\left( \rho_3 - \rho_1 \right)}{R(\theta_4) - R(\theta_1)} \right); \]

\[ B_2 := \text{evalf} \left( \rho_3 - A_2 R(\theta_1) \right); \]

\[ \theta_m := \text{evalf} \left( 2 \cdot \arctan \left( \exp \left( \frac{B_2}{A_2} - \frac{\rho_m}{A_2} \right) \right) \right); \]

\[ \text{av1} := \text{plots}[\text{polarplot}](A_2 \ln \left( \csc (\theta) + \cot (\theta) \right) + B_2, \theta = \theta_m .. \theta_4, \text{color} = \text{black}); \]

\[ \text{av2} := \text{plots}[\text{polarplot}](A_2 \ln \left( \csc (\theta) + \cot (\theta) \right) + B_2, \theta = \theta_1 .. \theta_m, \text{color} = \text{black}, \text{linestyle} = \text{dash}); \]

\[ A_3 := \text{evalf} \left( \frac{\left( \rho_m - \rho_2 \right)}{R(\theta_m) - R(\theta_1)} \right); \]

Figure 9.9: Program 3 Page 2
$$B_3 := \rho_2 - A_3 R \left( \theta_3 \right);$$

$$proj^3 := plots\[ polarplot\left( A_3 \ln(\csc(\theta) + \cot(\theta)) + B_3, \theta = \theta_3 .. \theta_m, colour = black \right);$$

$$c_3 := \cos \left( \theta_3 \right) + \sin \left( \theta_3 \right) \cdot \frac{\left( A_3 - \tan(\theta_3) \right)}{\left( 1 + A_3 \tan(\theta_3) \right)};$$

$$r_3 := \sqrt{\left( \cos(\theta_3) - c_3 \right)^2 + \sin^2(\theta_3)}; f(x) := \text{arccosh}\left( \frac{1 + \sqrt{\left[ 1 + A_3^2 \right]} \left( \frac{1}{\sqrt{r_3^2 - (x - c_3)^2}} \right)}{2 \sin(\theta_3) \cdot \sqrt{r_3^2 - (x - c_3)^2}} \right) - \frac{\sqrt{\left[ 1 + A_3^2 \right]} \left( \cos(\theta_3) - c_3 \right)}{A_3};$$

$$x_4 := \text{fsolve}(f(x) = 0, x = 1 .. 1.1);$$

$$y_4 := \sqrt{r_3^2 - (x_4 - c_3)^2}; refl^3 := plots\left( \left[ x, \sqrt{r_3^2 - (x - c_3)^2} \right], x = x_4 \cdot \cos(\theta_3) \right);$$

$$colour = \text{blue};$$

$$\rho_3 := \text{evalf}\left( \sqrt{x_4^2 + y_4^2} \right);$$

$$\text{if } x_4 > 0 \text{ then } \theta_3 := \text{evalf}\left( \tan^{-1}\left( \frac{y_4}{x_4} \right) \right) \text{ else } \theta_3 := \text{evalf}\left( \text{Pi} - \tan^{-1}\left( \frac{y_4}{x_4} \right) \right) \text{ fi;}$$

---

Figure 9.10: Program 3 Page 3
\begin{align*}
z_1 & := \text{textplot}
\left( \begin{array}{c}
1.716, 0.341,
typeset \ "A" \\
\end{array} \right),
align = \text{right}; \\
z_2 & := \text{textplot}
\left( \begin{array}{c}
x_3, y_3,
typeset \ "B" \\
\end{array} \right),
align = \text{below}; \\
z_3 & := \text{textplot}
\left( \begin{array}{c}
-1.179, 0.0309,
typeset \ "C" \\
\end{array} \right),
align = \text{above}; \\
z_4 & := \text{textplot}
\left( \begin{array}{c}
-1.120, 0.943,
typeset \ "D" \\
\end{array} \right),
align = \text{left}; \\
z_5 & := \text{textplot}
\left( \begin{array}{c}
-0.482, 0.518,
typeset \ "E" \\
\end{array} \right),
align = \text{right};
\end{align*}

\text{plots(display)}([P, Q, proj1, refl1, proj2, refl2, av1, av2, proj3, refl3, z1, z2, z3, z4, z5], \text{thickness} = 2)

\rho_1, \rho_2, \theta_1, \theta_2, \theta_3 := 1.75, 1, \frac{1}{16} \pi, \frac{1}{16} \pi, \frac{3}{4} \pi

c_1 := \sec \left( \frac{1}{16} \pi \right)

r_1 := \sqrt{\left( \cos \left( \frac{1}{16} \pi \right) - \sec \left( \frac{1}{16} \pi \right) \right)^2 + \sin \left( \frac{1}{16} \pi \right)^2}

Figure 9.11: Program 3 Page 4
\[ f := x \rightarrow \text{arccosh} \left( 1 + \frac{1}{2} \left( \frac{x - \cos(\theta_2)}{\sqrt{r_1^2 - (x - c_1)^2}} \right)^2 + \frac{1}{2} \left( \frac{\sin(\theta_2)}{\sqrt{r_1^2 - (x - c_1)^2}} \right)^2 \right) - \rho_1 + \rho_2 \]

\[ x_3 := 0.8726535459 \]

\[ y_3 := \sqrt{\left( \cos\left( \frac{1}{16} \pi \right) - \sec\left( \frac{1}{16} \pi \right) \right)^2 + \sin\left( \frac{1}{16} \pi \right)^2 - \left( 0.8726535459 - \sec\left( \frac{1}{16} \pi \right) \right)^2} - \rho_r := 0.8828927903 \]

\[ \theta_r := 0.1524458430 \]

\[ \text{refl}_1 := \text{PLOT}(...) \]

\[ c_2 := \frac{1}{\sqrt{2} - 1.745307092} \left( 0.2384757888 - \cos\left( \frac{1}{16} \pi \right) - \sec\left( \frac{1}{16} \pi \right) \right)^2 
- \sin\left( \frac{1}{16} \pi \right)^2 + \left( 0.8726535459 - \sec\left( \frac{1}{16} \pi \right) \right)^2 \]

\[ r_2 := \frac{1}{2} \left( 4 \left( -\frac{1}{2} \sqrt{2} - \frac{1}{2} \sqrt{2} - 1.745307092 \right) \left( 0.2384757888 \right) 
- \left( \cos\left( \frac{1}{16} \pi \right) - \sec\left( \frac{1}{16} \pi \right) \right)^2 - \left( \frac{1}{16} \pi \right)^2 \right)^{1/2} + 2 \]

\[ A_1 := -0.05191007372 \]

\[ B_1 := 0.9542478321 \]

\[ d \left( 0.8726535459, \right) \left( \cos\left( \frac{1}{16} \pi \right) - \sec\left( \frac{1}{16} \pi \right) \right)^2 + \sin\left( \frac{1}{16} \pi \right)^2 - \left( 0.8726535459 \right) 
- \sec\left( \frac{1}{16} \pi \right)^2 \right)^{1/2} \cdot \frac{1}{2} \sqrt{2} \cdot \frac{1}{2} \sqrt{2} := 3.458097556 \]

\[ \rho_3 := 1.179268728 \]

**Figure 9.12: Program 3 Page 5**
\[ R := -4.334821351 \]
\[ \theta_3 := 3.115385724 \]
\[ \rho_m := 1.464634364 \]
\[ R := \theta \ln (\csc (\theta) + \cot (\theta)) \]
\[ A_2 := 0.08579060947 \]
\[ B_2 := 1.551155726 \]
\[ \theta_m := 2.442067952 \]
\[ A_3 := -3.654386884 \]
\[ B_3 := 1 + 3.654386884 \ln (\sqrt{2} - 1) \]
\[ c_3 := -0.7851489305 \sqrt{2} \]
\[ r_3 := 0.8140146345 \]
\[ f := x \rightarrow \text{arccosh} \left( 1 + \frac{1}{2} \left( x - \cos (\theta_3) \right)^2 + \left( \frac{\sqrt{r_3^2 - (x - c_3)^2} - \sin (\theta_3)}{\sin (\theta_3)} \frac{\sqrt{r_3^2 - (x - c_3)^2}}{r_3} \right) \right) \]
\[ x_4 := -0.4820652626 \]
\[ y_4 := \sqrt{0.6626198252 - \left( -0.4820652626 + 0.7851489305 \sqrt{2} \right)^2} \]
\[ \text{refl3} := \text{PLOT(...)} \]
\[ \rho_3 := 0.7072779358 \]
\[ \theta_3 := 2.320713641 \]
\[ z1 := \text{PLOT(...)} \]
\[ z2 := \text{PLOT(...)} \]
\[ z3 := \text{PLOT(...)} \]
\[ z4 := \text{PLOT(...)} \]
\[ z5 := \text{PLOT(...)} \]

Figure 9.13: Program 3 Page 6
Figure 9.14: Program 3 Page 7
Bibliography


170


[64] W.A. Kirk, (Geodesic geometry and fixed point theory in: Seminar of Mathematical Analysis, Malaga/Seville, *Colecc. Abierta*, **64** (2003), 195–225. 80, 82, 95


174


