Contents

Preface 1

Brief history 1

Satellites and successors 4

Acknowledgments 5

Part 1. Preliminaries

Chapter 1. Model plane 9

§A. Trigonometry 9

§B. Hemisphere lemma 12

Chapter 2. Metric spaces 13

§A. Metrics and their relatives 13

§B. Notations 15

§C. Length spaces 17

§D. Convex sets 20

§E. Quotient spaces 21

§F. Gluing and doubling 23

§G. Kuratowsky embedding 24

Chapter 3. Maps and functions 25

§A. Submaps 25

§B. Lipschitz conditions 25

§C. Isometries and submetries 27

§D. Speed of curves 28

§E. Convex real-to-real functions 29
§F. Convex functions on a metric space 32

Chapter 4. Ultralimits 35
§A. Ultrafilters 35
§B. Ultralimits of points 37
§C. Ultralimits of spaces 38
§D. Ultralimits of sets 41
§E. Ultralimits of functions 41

Chapter 5. Space of spaces 43
§A. Convergence of subsets 44
§B. Convergence of spaces 45
§C. Gromov's selection theorem 47
§D. Convergence of compact spaces. 47
§E. Ultralimits revisited 51

Chapter 6. The ghost of Euclid 53
§A. Geodesics, triangles and hinges 53
§B. Model angles and triangles 54
§C. Angles and the first variation 56
§D. Space of directions 58
§E. Tangent space 59
§F. Velocity of curves 60
§G. Differential 61
§H. Ultratangent space 62
§I. Ultradifferential 64
§J. Remarks 64

Chapter 7. Dimension theory 65
§A. Definitions 65
§B. Linear dimension 67

Part 2. Fundamentals

Chapter 8. Fundamentals of curvature bounded below 71
§A. Four-point comparison 71
§B. Geodesics 73
§C. More comparisons 75

§D. Function comparison 78
§E. Development 79
§F. Local definitions and globalization 82
§G. Properties of geodesics and angles 89
§H. On positive lower bound 93
§I. Remarks 96

Chapter 9. Fundamentals of curvature bounded above 99
§A. Four-point comparison. 99
§B. Geodesics 101
§C. More comparisons 102
§D. Thin triangles 105
§E. Function comparison 107
§F. Development 109
§G. Patchwork globalization 110
§H. Angles 112
§I. Reshetnyak gluing theorem 114
§J. Space of geodesics 116
§K. Lifting globalization 120
§L. Reshetnyak majorization 124
§M. Hadamard–Cartan theorem 131
§N. Convex sets 135
§O. Remarks 137

Chapter 10. Kirszbraun revisited 141
§A. Short map extension definitions 141
§B. $(1+n)$-point comparison 145
§C. Helly's theorem 147
§D. Kirszbraun's theorem 148
§E. Remarks 153

Chapter 11. Warped products 157
§A. Definition 157
§B. Extended definitions 159
§C. Examples 160
§D. 1-dimensional base 161
§E. Remarks 163

Chapter 12. Polyhedral spaces 165
§A. Definitions 165
§B. Curvature bounds 166
§C. Flag complexes 168
§D. Remarks 171

Part 3. Structure and tools

Chapter 13. First order differentiation 175
§A. Ultratangent space 175
§B. Length property of tangent space 176
§C. Rademacher theorem 179
§D. Differential 182
§E. Definition of gradient 185
§F. Calculus of gradient 186
§G. Semicontinuity of |gradient| 189
§H. Polar vectors 191
§I. Linear subspace of tangent space 192
§J. Comments 194

Chapter 14. Dimension of CAT spaces 195
§A. The case of complete geodesic spaces 195
§B. The case of CAT spaces 200
§C. Dimension 203
§D. Finite-dimensional spaces 207
§E. Remarks 208

Chapter 15. Dimension of CBB spaces 211
§A. Struts and rank 211
§B. Right-inverse theorem 212
§C. Dimension theorem 215
§D. Inverse function theorem 219
§E. Finite-dimensional spaces 221
§F. One-dimensional spaces 224

Chapter 16. Gradient flow 225
§A. Gradient-like curves 226
§B. Gradient curves 228

§C. Distance estimates 230
§D. Existence and uniqueness 231
§E. Convergence 233
§F. Gradient flow 236
§G. Line splitting theorem 236
§H. Radial curves 239
§I. Radial comparisons 241
§J. Gradient exponent 246
§K. Remarks 247

Appendix A. Semisolutions 255

Bibliography 277

Index 285

Preface

Alexandrov spaces are defined via axioms similar to those given by Euclid. The Alexandrov axioms replace certain equalities with inequalities. Depending on the signs of the inequalities, we obtain Alexandrov spaces with *curvature bounded above* (CBA) and *curvature bounded below* (CBB). The definitions of the two classes of spaces are similar, but their properties and known applications are quite different.

The goal of this book is to give a comprehensive exposition of the structure theory of Alexandrov spaces with curvature bounded above and below. It includes all the basic material as well as selected topics inspired by considering the two contexts simultaneously. We only consider the intrinsic theory, leaving applications aside. Our presentation is linear, with a few exceptions where topics are deferred to later chapters to streamline the exposition. This book includes material *up to the definition of dimension*. Another volume still in preparation will cover further topics.

Brief history

The first synthetic description of curvature is due to Abraham Wald [**157**]; it was given in a lone publication on a "coordinateless description of Gauss surfaces" published in 1936. In 1941, similar definitions were rediscovered by Alexandr Alexandrov [**16**].

In Alexandrov's work the first fruitful applications of this approach were given. Mainly: *Alexandrov's embedding theorem* [**11**,**12**], which describes closed convex surfaces in Euclidean 3-space, and the *gluing theorem* [**13**], which gave a flexible tool to modify non-negatively curved metrics on a sphere. These two results together gave an intuitive geometric tool to study embeddings and bending of surfaces in Euclidean space and changed the subject dramatically. They formed the foundation of the branch of geometry now called *Alexandrov geometry*.

Curvature bounded below. The theory grew out of studying intrinsic and extrinsic geometry of convex surfaces without the smoothness condition. It was developed by Alexandr Alexandrov and his school. Here is a very incomplete list of contributors to the subject: Yuriy Borisov, Yuriy Burago, Boris Dekster, Iosif Liberman, Sergey Olovyanishnikov, Aleksey Pogorelov, Yuriy Reshetnyak, Yuriy Volkov, Viktor Zalgaller.

The first result in higher dimensional Alexandrov spaces was the splitting theorem. It was proved by Anatoliy Milka [**117**] and appeared in 1967. Milka used a global definition similar to the one used in this book.

In the 80's the interest in convergence of Riemannian manifolds spurred by *Gromov's compactness theorem* [**71**] turned attention toward the singular spaces that can occur as limits of Riemannian manifolds. Immediately it was recognized that if the manifolds have a uniform lower sectional curvature bound, then the limit spaces have a lower curvature bound in the sense of Alexandrov. There followed during the 90's an explosion of work on intrinsic theory of Alexandrov spaces starting with papers of Yuriy Burago, Grigory Perelman, and Michael Gromov [**44**, **125**]. Similar ideas were developed independently by Karsten Grove and Peter Petersen, whose work was not converted into a publication, and also by Conrad Plaut [**135**].

Around the same time an implicit application of higher-dimensional Alexandrov geometry was given by Michael Gromov in his bound on Betti numbers [**75**]. Another implicit application, which essentially used Alexandrov geometry before it was was actually introduced, given later by Wu-Yi Hsiang and Bruce Kleiner in their paper on non-negatively curved 4-manifolds with infinite symmetry groups [**83**]. The work of Hsiang and Kleiner and its extension by Karsten Grove and Burkhard Wilking [**78**] are some of the most beautiful applications of this branch of Alexandrov geometry.

The above activity was very much related to so-called *comparison geometry*, a branch of differential geometry that compares Riemannian manifolds to spaces of constant curvature. In addition to the already-mentioned *Gromov's compactness theorem*, the following results had a big influence on the development of Alexandrov geometry: *Toponogov comparison theorem* [**154**], which is a generalization of the theorem of Alexandrov [**14**]; *Toponogov splitting theorem* [**154**], which is a generalization of Cohn-Vossen's theorem [**55**]; *Finiteness* *theorems* of Cheeger and Grove–Petersen [**52**, **77**]; Gromov's bound on the number of generators of the fundamental group [**73**, 1.5]; and *Yamaguchi fibration theorem* [**160**].

Let us give a list of available introductory texts on Alexandrov spaces with curvature bounded below:

- The first introduction to Alexandrov geometry is given in the original paper of Yuriy Burago, Michael Gromov, and Grigory Perelman [**44**] and its extension [**125**] written by Perelman.
- A brief and reader-friendly introduction was written by Katsuhiro Shiohama [**148**, Sections 1–8].
- [**37**, Chapter 10] ~~gives~~ another reader-friendly introduction, written by Dmiti Burago, Yuriy Burago, and Sergei Ivanov.

In addition, let us mention two surveys, one by Conrad Plaut [**137**] and the other by the third author [**130**].

Curvature bounded above. The study of spaces with curvature bounded above started later, inspired by analogy with the theory of curvature bounded below. The first paper on the subject was written by Alexandrov [**18**], appearing in 1951. An analogous weaker definition was considered earlier by Herbert Busemann [**45**].

Contributions to the subject were made by Valerii Berestovskii, Arne Beurling, Igor Nikolaev, Dmitry Sokolov, Yuriy Reshetnyak, Samuel Shefel; this list is not complete as well. The most fundamental results were obtained by Yuriy Reshetnyak. They include his *majorization theorem* and *gluing theorem*. The gluing theorem states that if two non-positively curved spaces have isometric convex sets, then the space obtained by gluing these sets along an isometry is also non-positively curved.

The development of Alexandrov geometry was greatly influenced by the *Hadamard–Cartan theorem*. Its original formulation states that the exponential map at any point of a complete Riemannian manifold with nonpositive sectional curvature is a covering. In particular, it implies that the universal cover is diffeomorphic to Euclidean space of the same dimension. See further discussion below (9.65).

An influential implicit application of Alexandrov spaces with curvature bounded above can be seen in *Euclidean buildings*, introduced by Jacques Tits as a means to study algebraic groups.

Here is a list of available texts covering the basics of Alexandrov spaces with curvature bounded above:

- The book of Martin Bridson and André Haefliger [**34**] gives the most comprehensive introduction available today.

- The lecture notes of Werner Ballmann [**21**,**22**] include a brief and clear introduction.
- [**37**, Chapter 9] gives another reader-friendly introduction, by Yuriy Burago, Dmitry Burago, and Sergei Ivanov.
- A book by the three authors of the present volume [**10**] gives an introduction aiming at reaching interesting applications and theorems with a minimum of preparation.
- The book of Jürgen Jost [**89**] gives a more analytic viewpoint to the subject.

One of the most striking applications of CAT(0) spaces was given by Dmitry Burago, Sergei Ferleger, and Alexey Kononenko [**38**], who used them to study *billiards*; this idea was developed further in [**39–43**]. Another beautiful application is the construction of *exotic aspherical manifolds* by Michael Davis [**59**]; related results are surveyed in [**50**, **60**]. Both of these topics are discussed in [**10**]. The study of group actions on CAT(0) spaces and CAT(0) cube complexes played a key role in the proof of the *virtually fibered conjecture* that a finite cover of every closed hyperbolic 3-manifold fibers over the circle.

Satellites and successors

Surfaces with *bounded integral curvature* were studied by Alexandrov's school. An excellent book on the subject was written by Alexandr Alexandrov and Viktor Zalgaller [**15**]; see also a more up-to-date survey by Yuriy Reshetnyak [**139**].

Spaces with *two-sided bounded curvature* is another subject already studied by Alexandrov's school; a good survey is written by Valerij Berestovskij and Igor Nikolaev [**24**].

A spin-off of the idea of synthetically defining upper curvature bounds was given by Michael Gromov [**76**]. He defined so-called δ*-hyperbolic spaces*, which satisfy a coarse version of the negative curvature condition, applying in particular to discrete metric spaces. This notion and its various generalizations such as semi-hyperbolicity (a coarse version of non-positive curvature) and relative hyperbolicity have led to the emergence of the subject of *geometric group theory*, which relates geometric properties of groups to their algebraic ones. This is a well-developed subject with a large number of subfields and applications, such as the theory of small cancellation groups, automatic groups, mapping class groups, automorphisms of free groups, isoperimetric inequalities on groups, actions on $\mathbb{R}$-trees, Gromov's boundaries of groups.

The so-called *curvature dimension condition* introduced by John Lott, Cédric Villani, and Karl-Theodor Sturm gives a synthetic description of Ricci curvature bounded below; see the book of Villani [**156**] and references therein. A striking application of this theory to geodesic flow in CBB spaces was found recently by Elia Bruè, Andrea Mondino, and Daniele Semola [**35**].

Alexandrov geometry influenced the development of *analysis on metric spaces*. An excellent book on the subject was written by Juha Heinonen, Pekka Koskela, Nageswari Shanmugalingam, and Jeremy Tyson [**82**].

Acknowledgments

We thank Semyon Alesker, Valerii Berestovskii, I. David Berg, Richard Bishop, Yuriy Burago, Alexander Christie, Nicola Gigli, Sergei Ivanov, Bernd Kirchheim, Bruce Kleiner, Rostislav Matveyev, John Lott, Greg Kuperberg, Nikolai Kosovsky, Nina Lebedeva, Wilderich Tuschmann, and Sergio Zamora Barrera. A special thanks to Alexander Lytchak, who contributed deeply to this book during the long process of writing, but refused to be one of us.

We thank the mathematical institutions where we worked on this book: BIRS, MFO, Henri Poincaré Institute, IHES, University of Cologne, Max Planck Institute for Mathematics, EIMI.

During the long writing of this book, we were partially supported by the following grants: Stephanie Alexander—Simons Foundation grant 209053; Vitali Kapovitch—NSF grant DMS-0204187, NSERC Discovery grants, and Simons Foundation grant 390117; Anton Petrunin—NSF grants DMS-0406482, DMS-0905138, DMS-1309340, DMS-2005279, Simons Foundation grants 245094 and 584781, and Minobrnauki, grant 075-15-2022-289.

Part 1

Preliminaries

Chapter 1

Model plane

A. Trigonometry

Given a real number κ, the *model κ-plane* will be a complete simply connected 2-dimensional Riemannian manifold of constant curvature κ.

The model κ-plane will be denoted by $\mathbb{M}^2(\kappa)$.

- If $\kappa > 0$, $\mathbb{M}^2(\kappa)$ is isometric to a sphere of radius $\frac{1}{\sqrt{\kappa}}$; the unit sphere $\mathbb{M}^2(1)$ will be also denoted by $\mathbb{S}^2$.
- If $\kappa = 0$, $\mathbb{M}^2(\kappa)$ is the Euclidean plane, which is also denoted by $\mathbb{E}^2$.
- If $\kappa < 0$, $\mathbb{M}^2(\kappa)$ is the Lobachevsky plane with curvature κ.

Let $\varpi\kappa = \operatorname{diam}\mathbb{M}^2(\kappa)$, so $\varpi\kappa = \infty$ if $\kappa \leqslant 0$ and $\varpi\kappa = \pi/\sqrt{\kappa}$ if $\kappa > 0$; ϖ is just a cursive form of π.

The distance between points $x, y \in \mathbb{M}^2(\kappa)$ will be denoted by $|x - y|$, and $[xy]$ will denote the geodesic segment connecting x and y. The segment $[xy]$ is uniquely defined for $\kappa \leqslant 0$ and for $\kappa > 0$ it is defined uniquely if $|x - y| < \varpi\kappa = \pi/\sqrt{\kappa}$.

A triangle in $\mathbb{M}^2(\kappa)$ with vertices x, y, z will be denoted by $[xyz]$. Formally, a triangle is an ordered set of its sides, so $[xyz]$ is just a short notation for the triple $([yz], [zx], [xy])$.

The angle of $[xyz]$ at x will be denoted by $\measuredangle\left[x\,{}^{y}_{z}\right]$.

By $\tilde{\triangle}^{\kappa}\{a, b, c\}$ we denote a triangle in $\mathbb{M}^2(\kappa)$ with side lengths a, b, c, so $[xyz] = \tilde{\triangle}^{\kappa}\{a, b, c\}$ means that $x, y, z \in \mathbb{M}^2(\kappa)$ are such that

$$|x - y| = c, \quad |y - z| = a, \quad |z - x| = b.$$

For $\tilde{\triangle}^{\kappa}\{a,b,c\}$ to be defined, the sides a, b, c must satisfy the triangle inequality. If $\kappa > 0$, we require in addition that $a+b+c < 2\cdot\varpi\kappa$; otherwise $\tilde{\triangle}^{\kappa}\{a,b,c\}$ is considered to be undefined.

Trigonometric functions. We will need three *trigonometric functions* in $\mathbb{M}^2(\kappa)$: *cosine*, *sine*, and *modified distance*, denoted by cs^κ, sn^κ, and md^κ respectively.

They are defined as the solutions of the following initial value problems respectively:

$$\begin{cases} x'' + \kappa\cdot x = 0, \\ x(0) = 1, \\ x'(0) = 0. \end{cases} \qquad \begin{cases} y'' + \kappa\cdot y = 0, \\ y(0) = 0, \\ y'(0) = 1. \end{cases} \qquad \begin{cases} z'' + \kappa\cdot z = 1, \\ z(0) = 0, \\ z'(0) = 0. \end{cases}$$

Namely, we set $\mathrm{cs}^\kappa(t) = x(t)$, $\mathrm{sn}^\kappa(t) = y(t)$, and

$$\mathrm{md}^\kappa(t) = \begin{cases} z(t) & \text{if } 0 \leqslant t \leqslant \varpi\kappa, \\ \frac{2}{\kappa} & \text{if } t > \varpi\kappa. \end{cases}$$

Here are the tables which relate our trigonometric functions to the standard ones, where we take $\kappa > 0$:

$$\mathrm{sn}^{\pm\kappa} = \tfrac{1}{\sqrt{\kappa}}\cdot \mathrm{sn}^{\pm 1}(x\cdot\sqrt{\kappa}); \qquad \mathrm{cs}^{\pm\kappa} = \mathrm{cs}^{\pm 1}(x\cdot\sqrt{\kappa}); \qquad \mathrm{md}^{\pm\kappa} = \tfrac{1}{\kappa}\cdot \mathrm{md}^{\pm 1}(x\cdot\sqrt{\kappa});$$

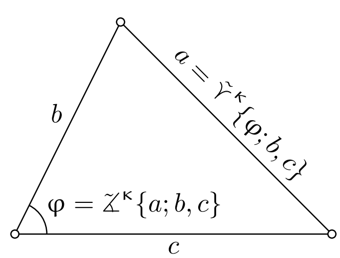
$$\mathrm{sn}^{-1} x = \sinh x; \qquad \mathrm{cs}^{-1} x = \cosh x; \qquad \mathrm{md}^{-1} x = \cosh x - 1;$$

$$\mathrm{sn}^{0} x = x; \qquad \mathrm{cs}^{0} x = 1; \qquad \mathrm{md}^{0} x = \tfrac{1}{2}\cdot x^2;$$

$$\mathrm{sn}^{1} x = \sin x; \qquad \mathrm{cs}^{1} x = \cos x; \qquad \mathrm{md}^{1} x = \begin{cases} 1 - \cos x & \text{for } x \leqslant \pi, \\ 2 & \text{for } x > \pi. \end{cases}$$

Note that

$$\mathrm{md}^\kappa(x) = \int_0^x \mathrm{sn}^\kappa(t)\cdot \mathbf{d}t \quad \text{for} \quad x \leqslant \varpi\kappa.$$



Let ϕ be the angle of $\tilde{\triangle}^{\kappa}\{a,b,c\}$ opposite to a. In this case, we will write

$$a = \tilde\gamma^\kappa\{\phi; b, c\} \quad \text{or} \quad \phi = \tilde\measuredangle^\kappa\{a; b, c\}.$$

The functions $\tilde\gamma^\kappa$ and $\tilde\measuredangle^\kappa$ will be called respectively the *model side* and the *model angle*. Let

$$\tilde\gamma^\kappa\{\phi; b, -c\} = \tilde\gamma^\kappa\{\phi; -b, c\} := \tilde\gamma^\kappa\{\pi - \phi; b, c\};$$

in this way we define $\tilde\gamma^\kappa\{\phi; b, c\}$ when one of the numbers b and c is negative.

1.1. Properties of standard functions.

(a) *For fixed a and ϕ, the function $y(t) = \mathrm{md}^\kappa\left(\tilde{\gamma}^\kappa\{\phi; a, t\}\right)$ satisfies the ~~following~~ differential equation~~:~~*

$$y'' + \kappa \cdot y = 1.$$

(b) *Let $\alpha\colon [a, b] \to \mathbb{M}^2(\kappa)$ be a unit-speed geodesic, and A be the image of a complete geodesic. If $f(t)$ is the distance from $\alpha(t)$ to A, the function $y(t) = \mathrm{sn}^\kappa(f(t))$ satisfies the ~~following~~ differential equation~~:~~*

$$y'' + \kappa \cdot y = 0$$

for $y \neq 0$.

(c) *For fixed κ, b, and c, the function*

$$a \mapsto \tilde\measuredangle^\kappa\{a; b, c\}$$

is increasing and defined on a real interval. Equivalently, the function

$$\phi \mapsto \tilde{\gamma}^\kappa\{\phi; b, c\}$$

is increasing and defined if $b, c < \varpi^\kappa$, and $\phi \in [0, \pi]$. (Formally speaking, if $\kappa > 0$ and $b + c \geqslant \varpi^\kappa$, it is defined only for $\phi \in [0, \pi)$, but $\tilde{\gamma}^\kappa\{\phi; b, c\}$ can be extended to $[0, \pi]$ as a continuous function.)

(d) *For fixed ϕ, a, b, c, the function*

$$\kappa \mapsto \tilde\measuredangle^\kappa\{a; b, c\} \quad \textit{and} \quad \kappa \mapsto \tilde{\gamma}^\kappa\{\phi; b, c\}$$

are respectively nondereasing (in fact, increasing, if $|b - c| < a < b + c$) and nonincreasing (in fact, increasing, if $0 < \phi < \pi$).

(e) *(Alexandrov's lemma) Assume that for real numbers a, b, a', b', x, and κ, the following two expressions are defined:*

$$\tilde\measuredangle^\kappa\{a; b, x\} + \tilde\measuredangle^\kappa\{a'; b', x\} - \pi, \qquad \tilde\measuredangle^\kappa\{a'; b + b', a\} - \tilde\measuredangle^\kappa\{x; a, b\},$$

Then they have the same sign.

All the properties except Alexandrov's lemma (*e*) can be shown by direct calculation. Alexandrov's lemma is reformulated in 6.3 and is proved there.

Cosine law. The formulas $a = \tilde{\gamma}^\kappa\{\phi; b, c\}$ and $\phi = \tilde\measuredangle^\kappa\{a; b, c\}$ can be rewritten using the cosine law in $\mathbb{M}^2(\kappa)$:

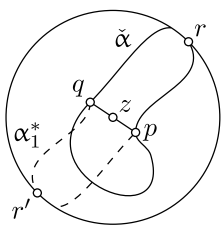
$$\cos\phi = \begin{cases} \dfrac{b^2 + c^2 - a^2}{2 \cdot b \cdot c} & \text{if } \kappa = 0, \\ \dfrac{\mathrm{cs}^\kappa a - \mathrm{cs}^\kappa b \cdot \mathrm{cs}^\kappa c}{\kappa \cdot \mathrm{sn}^\kappa b \cdot \mathrm{sn}^\kappa c} & \text{if } \kappa \neq 0. \end{cases}$$

However, rather than using these explicit formulas, we mainly will use the properties of $\tilde\measuredangle^\kappa$ and $\tilde{\gamma}^\kappa$ listed in 1.1.

B. Hemisphere lemma

1.2. Hemisphere lemma. *For $\varkappa > 0$, any closed path of length less than $2 \cdot \varpi\varkappa$ (respectively, at most $2 \cdot \varpi\varkappa$) in $\mathbb{M}^2(\varkappa)$ lies in an open (respectively, closed) hemisphere.*

Proof. Applying rescaling, we may assume that $\varkappa = 1$, and thus $\varpi\varkappa = \pi$ and $\mathbb{M}^2(\varkappa) = \mathbb{S}^2$. Let α be a closed curve in $\mathbb{S}^2$ of length $2 \cdot \ell$.



Assume $\ell < \pi$. Let $\check{\alpha}$ be a subarc of α of length ℓ, with endpoints p and q. Since $|p - q| \leqslant \ell < \pi$, there is a unique geodesic $[pq]$ in $\mathbb{S}^2$. Let z be the midpoint of $[pq]$. We claim that α lies in the open hemisphere centered at z. If not, α intersects the boundary great circle of this hemisphere; let r be a point in the intersection. Without loss of generality, we may assume that $r \in \check{\alpha}$. The arc $\check{\alpha}$ together with its reflection in z form a closed curve of length $2 \cdot \ell$ that contains r and its antipodal point r'. Thus

$$\ell = \operatorname{length} \check{\alpha} \geqslant |r - r'| = \pi,$$

a contradiction.

If $\ell = \pi$, then either α is a local geodesic, and hence a great circle, or α may be strictly shortened by substituting a geodesic arc for a subarc of α whose endpoints p^1, p^2 are arbitrarily close to a point p on α. In both cases α lies in a closed hemisphere; the former case is trivial, and in the latter case, α lies in a closed hemisphere obtained as a limit of closures of open hemispheres containing the shortened curves as p^1, p^2 approach p. $\square$

1.3. Exercise. *Give a proof of ~~the~~ hemisphere lemma* (1.2) *based on Crofton's formula.*

Chapter 2

Metric spaces

In this chapter we fix conventions and notations. We are assuming that the reader is familiar with basic notions in metric geometry.

A. Metrics and their relatives

Definitions. Let $\mathbb{I}$ be a subinterval of $[0,\infty]$. A function ρ defined on $\mathcal{X}\times\mathcal{X}$ is called an $\mathbb{I}$-valued metric if the following conditions hold:

- $\rho(x,x)=0$ for any x;
- $\rho(x,y)\in\mathbb{I}$ for any pair $x\ne y$;
- $\rho(x,y)+\rho(x,z)\geqslant\rho(y,z)$ for any triple of points x,y,z.

The value $\rho(x,y)$ is also called the *distance* between x and y.

The above definition will be used for four choices of interval $\mathbb{I}$: $(0,\infty)$, $(0,\infty]$, $[0,\infty)$, and $[0,\infty]$. Any $\mathbb{I}$-valued metric can be referred to briefly as a metric; the interval should be apparent from context but by default, a metric is $(0,\infty)$-valued. If we need to be more specific we may also use the following names:

- a $(0,\infty)$-valued metric may be called a *genuine metric*.
- a $(0,\infty]$-valued metric may be called an ∞-*metric*.
- a $[0,\infty)$-valued metric may be called a *genuine pseudometric*.
- A $[0,\infty]$-valued metric may be called a *pseudometric* or ∞-*pseudometric*.

A metric space is a set equipped with a metric. The distance between points x and y in a metric space $\mathcal{X}$ will usually be denoted by

$$|x-y| \qquad \text{or} \qquad |x-y|_{\mathcal{X}};$$

the latter will be used if we need to emphasize that we are working in the space $\mathcal{X}$.

The function $\operatorname{dist}_x : \mathcal{X} \to \mathbb{R}$ defined as
$$\operatorname{dist}_x : y \mapsto |x - y|$$
will be called the *distance function* from x.

Any subset A in a metric space $\mathcal{X}$ will be also considered as a *subspace*; that is, a metric space with the metric defined by restricting the metric of $\mathcal{X}$ to $A \times A \subset \mathcal{X} \times \mathcal{X}$.

The *direct product* $\mathcal{X} \times \mathcal{Y}$ of two metric spaces $\mathcal{X}$ and $\mathcal{Y}$ is defined as the metric space carrying the metric
$$|(p, \phi) - (q, \psi)| = \sqrt{|p - q|^2 + |\phi - \psi|^2}$$
for $p, q \in \mathcal{X}$ and $\phi, \psi \in \mathcal{Y}$.

A map between two metric spaces is called an *isometry* if it is a bijection and preserves distances between points.

Zero and infinity. Genuine metric spaces are the main objects of study in this book. However, the generalizations above are useful in various definitions and constructions. For example, the construction of length metric (see Section 2C) uses infinite distances. The following definition gives another example.

2.1. Definition. *Assume $\{\mathcal{X}_\alpha\}_{\alpha \in \mathcal{A}}$ is a collection of ∞-metric spaces. The disjoint union*
$$\mathbf{X} = \bigsqcup_{\alpha \in \mathcal{A}} \mathcal{X}_\alpha$$
has a natural ∞-metric on it defined as follows: given two points $x \in \mathcal{X}_\alpha$ and $y \in \mathcal{X}_\beta$, let
$$\begin{aligned} |x - y|_{\mathbf{X}} &= \infty && \text{if } \alpha \neq \beta, \\ |x - y|_{\mathbf{X}} &= |x - y|_{\mathcal{X}_\alpha} && \text{if } \alpha = \beta. \end{aligned}$$
The resulting ∞-metric space $\mathbf{X}$ will be called the disjoint union *of $\{\mathcal{X}_\alpha\}_{\alpha \in \mathcal{A}}$, denoted by*
$$\bigsqcup_{\alpha \in \mathcal{A}} \mathcal{X}_\alpha.$$

Now let us give examples showing that vanishing and infinite distance between distinct points can appear naturally and useful in constructions.

Suppose a set $\mathcal{X}$ comes with a set of metrics $|* - *|_\alpha$ for $\alpha \in \mathcal{A}$. Then
$$|x - y| = \sup\{\, |x - y|_\alpha : \alpha \in \mathcal{A} \,\}$$
is in general only an ∞-metric; that is, even if the metrics $|* - *|_\alpha$ are genuine, then $|* - *|$ might be $(0, \infty]$-valued.

Let $\mathcal{X}$ be a set, $\mathcal{Y}$ be a metric space, and $\Phi\colon\ \mathcal{X} \to \mathcal{Y}$ be a map. If Φ is not injective, then the *pullback*

$$|x - y|_{\mathcal{X}} = |\Phi(x) - \Phi(y)|_{\mathcal{Y}}$$

defines only a pseudometric on $\mathcal{X}$.

Corresponding metric space and metric component. The following two observations show that nearly any question about metric spaces can be reduced to a question about genuine metric spaces.

Assume $\mathcal{X}$ is a pseudometric space. Set $x \sim y$ if $|x - y| = 0$. Note that if $x \sim x'$, then $|y - x| = |y - x'|$ for any $y \in \mathcal{X}$. Thus, $|{*} - {*}|$ defines a metric on the quotient set $\mathcal{X}/{\sim}$. This way we obtain a metric space $\mathcal{X}'$. The space $\mathcal{X}'$ is called the *corresponding metric space* for the pseudometric space $\mathcal{X}$. Often we do not distinguish between $\mathcal{X}'$ and $\mathcal{X}$.

Set $x \approx y$ if and only if $|x - y| < \infty$; this is another equivalence relation on $\mathcal{X}$. The equivalence class of a point $x \in \mathcal{X}$ will be called the *metric component* of x; it will be denoted by $\mathcal{X}_x$. One could think of $\mathcal{X}_x$ as $\mathrm{B}(x,\infty)_{\mathcal{X}}$, the open ball centered at x and radius ∞ in $\mathcal{X}$; see definition below.

It follows that any ∞-metric space is a *disjoint union* of genuine metric spaces, the metric components of the original ∞-metric space; see Definition 2.1

To summarize this discussion: Given a $[0,\infty]$-valued metric space $\mathcal{X}$, we may pass to the corresponding $(0,\infty]$-valued metric space $\mathcal{X}'$ and break the latter into a disjoint union of metric components, each of which is a genuine metric space.

B. Notations

Balls. Given $R \in [0,\infty]$ and a point x in a metric space $\mathcal{X}$, the sets

$$\mathrm{B}(x,R) = \{\, y \in \mathcal{X} : |x - y| < R \,\},$$

$$\overline{\mathrm{B}}[x,R] = \{\, y \in \mathcal{X} : |x - y| \leqslant R \,\}$$

are called respectively the *open* and the *closed balls* of radius R with center x.

If we need to emphasize that these balls are taken in the space $\mathcal{X}$, we write

$$\mathrm{B}(x,R)_{\mathcal{X}} \quad \text{and} \quad \overline{\mathrm{B}}[x,R]_{\mathcal{X}}$$

respectively.

Since in the model space $\mathbb{M}^m(\kappa)$ all balls of the same radius are isometric, often we will not need to specify the center of the ball, and may write

$$\mathrm{B}(R)_{\mathbb{M}^m(\kappa)} \quad \text{and} \quad \overline{\mathrm{B}}[R]_{\mathbb{M}^m(\kappa)}$$

respectively.

A set $A \subset \mathcal{X}$ is called *bounded* if $A \subset \mathrm{B}(x,R)$ for some $x \in \mathcal{X}$ and $R < \infty$.

Distances to sets. For subset $A \subset \mathcal{X}$, let us denote the distance from A to a point x in $\mathcal{X}$ by $\operatorname{dist}_A x$; that is,

$$\operatorname{dist}_A x \coloneqq \inf\{\, |a - x| \,:\, a \in A \,\}.$$

For any subset $A \subset \mathcal{X}$, the sets

$$\mathrm{B}(A, R) = \{\, y \in \mathcal{X} \,:\, \operatorname{dist}_A y < R \,\},$$
$$\overline{\mathrm{B}}[A, R] = \{\, y \in \mathcal{X} \,:\, \operatorname{dist}_A y \leqslant R \,\}$$

are called respectively the *open* and *closed R-neighborhoods* of A.

Diameter, radius, and packing. Let $\mathcal{X}$ be a metric space. Then the *diameter* of $\mathcal{X}$ is defined as

$$\operatorname{diam} \mathcal{X} = \sup\{\, |x - y| \,:\, x, y \in \mathcal{X} \,\}.$$

The *radius* of $\mathcal{X}$ is defined as

$$\operatorname{rad} \mathcal{X} = \inf\{\, R > 0 \,:\, \mathrm{B}(x, R) = \mathcal{X} \text{ for some } x \in \mathcal{X} \,\}.$$

The *ε-pack* of $\mathcal{X}$ (or *packing number*) is the maximal number (possibly infinite) of points in $\mathcal{X}$ at distance $> \varepsilon$ from each other; it is denoted by $\operatorname{pack}_\varepsilon \mathcal{X}$. If $m = \operatorname{pack}_\varepsilon \mathcal{X} < \infty$, then a set $\{x^1, x^2, \dots, x^m\}$ in $\mathcal{X}$ such that $|x^i - x^j| > \varepsilon$ is called a *maximal ε-packing* in $\mathcal{X}$.

G-delta sets. Recall that an arbitrary union of open balls in a metric space is called an *open set*. A subset of a metric space is called a *G-delta set* if it can be presented as an intersection of a countable number of open subsets.

2.2. Baire's theorem. *Let $\mathcal{X}$ be a complete metric space and $\{\Omega_n\}$, $n \in \mathbb{N}$, be a collection of open dense subsets of $\mathcal{X}$. Then $\bigcap_{n\in\mathbb{N}} \Omega_n$ is dense in $\mathcal{X}$.*

Proper spaces. A metric space $\mathcal{X}$ is called *proper* if all closed bounded sets in $\mathcal{X}$ are compact. This condition is equivalent to each of the following statements:

(1) For some (and therefore any) point $p \in \mathcal{X}$ and any $R < \infty$, the closed ball $\overline{\mathrm{B}}[p, R] \subset \mathcal{X}$ is compact.

(2) The function $\operatorname{dist}_p : \mathcal{X} \to \mathbb{R}$ is proper for some (and therefore any) point $p \in \mathcal{X}$.

We will also often use the following two classical statements:

2.3. Proposition. *Proper metric spaces are separable and second countable.*

2.4. Proposition. *Let $\mathcal{X}$ be a metric space. Then the following are equivalent*

(a) $\mathcal{X}$ is compact;

(*b*) *$\mathcal{X}$ is* sequentially compact*; that is, any sequence of points in $\mathcal{X}$ contains a convergent subsequence;*

(*c*) *$\mathcal{X}$ is complete and for any $\varepsilon > 0$ there is a* finite ε-net *in $\mathcal{X}$; that is, there is a finite collection of points $p_1, \dots, p_N$ such that $\bigcup_i \mathrm{B}(p_i, \varepsilon) = \mathcal{X}$.*

(*d*) *$\mathcal{X}$ is complete and for any $\varepsilon > 0$ there is a* compact ε-net *in $\mathcal{X}$; that is, $\mathrm{B}(K, \varepsilon) = \mathcal{X}$ for a compact set $K \subset \mathcal{X}$.*

C. Length spaces

A *curve* in a metric space $\mathcal{X}$ is a continuous map $\alpha\colon \mathbb{I} \to \mathcal{X}$, where $\mathbb{I}$ is a *real interval* (that is, an arbitrary convex subset of $\mathbb{R}$).

2.5. Definition. *Let $\mathcal{X}$ be a metric space. Given a curve $\alpha\colon \mathbb{I} \to \mathcal{X}$, we define its* length *as*

$$\operatorname{length}\alpha \coloneqq \sup\left\{\sum_{i\geqslant 1} |\alpha(t_i) - \alpha(t_{i-1})| \;:\; t_0, \dots, t_n \in \mathbb{I}, t_0 \leqslant \dots \leqslant t_n\right\}.$$

The following lemma is an easy exercise.

2.6. Lower semicontinuity of length. *Assume $\alpha_n\colon \mathbb{I} \to \mathcal{X}$ is a sequence of curves that converges pointwise to a curve $\alpha_\infty\colon \mathbb{I} \to \mathcal{X}$. Then*

$$\operatorname{length}\alpha_\infty \leqslant \varliminf_{n\to\infty} \operatorname{length}\alpha_n.$$

Given two points x and y in a metric space $\mathcal{X}$, consider the value

$$\|x - y\| = \inf_\alpha\{\operatorname{length}\alpha\},$$

where infimum is taken for all paths α from x to y.

It is easy to see that $\|{*} - {*}\|$ defines a $(0, \infty]$-valued metric on $\mathcal{X}$; it will be called the *induced length metric* on $\mathcal{X}$. Clearly

$$\|x - y\| \geqslant |x - y|$$

for any $x, y \in \mathcal{X}$.

It easily follows from the definition that the length of a curve α with respect to $\|{*}-{*}\|$ is equal to the length of α with respect to $|{*}-{*}|$. In particular, iterating the construction produces the same metric $\|{*} - {*}\|$.

2.7. Definition. *If $\|x - y\| = |x - y|$ for any pair of points x, y in a metric space $\mathcal{X}$, then $|{*} - {*}|$ is called* length metric*, and $\mathcal{X}$ is called a* length space*.*

In other words, a metric space $\mathcal{X}$ is a *length space* if for any $\varepsilon > 0$ and any two points $x, y \in \mathcal{X}$ with $|x - y| < \infty$ there is a path $\alpha\colon [0, 1] \to \mathcal{X}$ connecting x to y such that

$$\operatorname{length}\alpha < |x - y| + \varepsilon.$$

In this book, most of the time we consider length spaces. If $\mathcal{X}$ is a length space~~,~~ and $A \subset \mathcal{X}$, then the set A comes with the inherited metric from $\mathcal{X}$, which might be not a length metric. The corresponding length metric on A will be denoted by $|{*}-{*}|_A$.

Variations of the definition. We will need the following variations of Definition 2.7:

- Assume $R > 0$. If $\|x-y\| = |x-y|$ for any pair $|x-y| < R$, then $\mathcal{X}$ is called an *R-length space*.
- If any point in $\mathcal{X}$ admits a neighborhood Ω such that $\|x-y\| = |x-y|$ for any pair of points $x, y \in \Omega$ then $\mathcal{X}$ is called a *locally length space*.
- A metric space is called *geodesic* if for any two points x, y with $|x-y| < \infty$ there is a geodesic $[xy]$ in $\mathcal{X}$
- Assume $R > 0$. A metric space is called *R-geodesic* if for any two points x, y such that $|x-y| < R$ there is a geodesic $[xy]$ in $\mathcal{X}$.

Note that the notions of ∞-length spaces and length spaces are the same. Clearly, any geodesic space is a length space and any R-geodesic space is R-length.

2.8. Example. *Consider a metric space $\mathcal{X}$ obtained by gluing a countable collection of disjoint intervals $\mathbb{I}_n$ of length $1 + \frac{1}{n}$ where for each $\mathbb{I}_n$ one end is glued to p and the other to q. Then $\mathcal{X}$ carries a natural complete length metric such that $|p-q| = 1$, but there is no geodesic connecting p to q.*

2.9. Exercise. *Let $\mathcal{X}$ be a metric space and $\|{*}-{*}\|$ be the length metric on it. Show the following:*

(*a*) *If $\mathcal{X}$ is complete, then $(\mathcal{X}, \|{*}-{*}\|)$ is complete.*

(*b*) *If $\mathcal{X}$ is compact, then $(\mathcal{X}, \|{*}-{*}\|)$ is geodesic.*

2.10. Exercise. *Give an example of a complete length space such that no pair of distinct points can be joined by a geodesic.*

2.11. Exercise. *Let $\mathcal{X}$ be a complete length space. Show that for any compact subset K in $\mathcal{X}$ there is a compact path-connected subset K' that contains K.*

2.12. Definition. *Consider two points x and y in a metric space $\mathcal{X}$.*

(i) *A point $z \in \mathcal{X}$ is called a* midpoint *between x and y if*

$$|x-z| = |y-z| = \tfrac{1}{2}\cdot|x-y|.$$

(ii) *Assume $\varepsilon \geqslant 0$. A point $z \in \mathcal{X}$ is called an ε-*midpoint *between x and y if*

$$|x-z| \leqslant \tfrac{1}{2}\cdot|x-y| + \cancel{\varepsilon,} \quad \textit{and} \quad |y-z| \leqslant \tfrac{1}{2}\cdot|x-y| + \varepsilon.$$

Note that a 0-midpoint is the same as a midpoint.

The following lemma was essentially proved by Karl Menger [**115**, Section 6].

2.13. Lemma. *Let $\mathcal{X}$ be a complete metric space.*

(*a*) *Assume that for any pair of points $x, y \in \mathcal{X}$, and any $\varepsilon > 0$ there is an ε-midpoint z. Then $\mathcal{X}$ is a length space.*

(*b*) *Assume that for any pair of points $x, y \in \mathcal{X}$ there is a midpoint z. Then $\mathcal{X}$ is a geodesic space.*

(*c*) *If for some $R > 0$, the assumptions (a) or (b) hold only for pairs of points $x, y \in \mathcal{X}$ such that $|x - y| < R$, then $\mathcal{X}$ is an R-length or an R-geodesic space respectively.*

Proof. Fix a pair of points $x, y \in \mathcal{X}$. Let $\varepsilon_n = \frac{\varepsilon}{2^{2 \cdot n}}$.

Set $\alpha(0) = x$, and $\alpha(1) = y$. Let $\alpha(\frac{1}{2})$ be an ε_1-midpoint between $\alpha(0)$ and $\alpha(1)$. Further, let $\alpha(\frac{1}{4})$ and $\alpha(\frac{3}{4})$ be ε_2-midpoints between the pairs $(\alpha(0), \alpha(\frac{1}{2})$ and $(\alpha(\frac{1}{2}), \alpha(1)$ respectively. Applying the above procedure recursively, on the n-th step we define $\alpha(\frac{k}{2^n})$, for every odd integer k such that $0 < \frac{k}{2^n} < 1$, to be an ε_n-midpoint between the already defined $\alpha(\frac{k-1}{2^n})$ and $\alpha(\frac{k+1}{2^n})$.

This way we define $\alpha(t)$ for all dyadic rationals t in $[0, 1]$. If $t \in [0, 1]$ is not a dyadic rational, consider a sequence of dyadic rationals $t_n \to t$ as $n \to \infty$. By completeness of $\mathcal{X}$, the sequence $\alpha(t_n)$ converges; let $\alpha(t)$ be its limit. It is easy to see that $\alpha(t)$ does not depend on the choice of the sequence t_n, and $\alpha\colon [0, 1] \to \mathcal{X}$ is a path from x to y. Moreover,

$$\begin{aligned}\text{length}\,\alpha &\leqslant |x - y| + \sum_{n=1}^{\infty} 2^{n-1} \cdot \varepsilon_n \\ &\leqslant |x - y| + \tfrac{\varepsilon}{2}.\end{aligned} \tag{1}$$

Since $\varepsilon > 0$ is arbitrary, we have (*a*).

To prove (*b*), one should repeat the same argument taking midpoints instead of ε_n-midpoints. In this case, 1 holds for $\varepsilon_n = \varepsilon = 0$.

The proof of (*c*) is obtained by a straightforward modification of the proofs above. ☐

Since in a compact set a sequence of $\frac{1}{n}$-midpoints (z_n) contains a convergent subsequence, Lemma 2.13 implies the following.

2.14. Corollary. *A proper length space is geodesic.*

2.15. Hopf–Rinow theorem. *Any complete, locally compact length space is proper.*

Proof. Let $\mathcal{X}$ be a locally compact length space. Given $x \in \mathcal{X}$, denote by $\rho(x)$ the supremum of all $R > 0$ such that the closed ball $\overline{\mathrm{B}}[x, R]$ is compact. Since $\mathcal{X}$ is locally compact,

(2) $$\rho(x) > 0 \quad \text{for any} \quad x \in \mathcal{X}.$$

It is sufficient to show that $\rho(x) = \infty$ for some (and therefore any) point $x \in \mathcal{X}$.

Assume the contrary; that is, $\rho(x) < \infty$.

(3) *$B = \overline{\mathrm{B}}[x, \rho(x)]$ is compact for any x.*

Indeed, $\mathcal{X}$ is a length space; therefore for any $\varepsilon > 0$, the set $\overline{\mathrm{B}}[x, \rho(x) - \varepsilon]$ is a compact ε-net in B. Since B is closed and hence complete, it is compact by Proposition 2.4. △

(4) *$|\rho(x) - \rho(y)| \leqslant |x - y|_{\mathcal{X}}$ for any $x, y \in \mathcal{X}$; in particular $\rho : \mathcal{X} \to \mathbb{R}$ is a continuous function.*

Indeed, assume the contrary; that is, $\rho(x) + |x - y| < \rho(y)$ for $x, y \in \mathcal{X}$. Then $\overline{\mathrm{B}}[x, \rho(x) + \varepsilon]$ is a closed subset of $\overline{\mathrm{B}}[y, \rho(y)]$ for $\varepsilon > 0$. Then compactness of $\overline{\mathrm{B}}[y, \rho(y)]$ implies compactness of $\overline{\mathrm{B}}[x, \rho(x) + \varepsilon]$, a contradiction. △

Set $\varepsilon = \min_{y \in B}\{\rho(y)\}$; the minimum is defined since B is compact. From 2, we have $\varepsilon > 0$.

Choose a finite $\frac{\varepsilon}{10}$-net $\{a_1, a_2, \dots, a_n\}$ in B. The union W of the closed balls $\overline{\mathrm{B}}[a_i, \varepsilon]$ is compact. Clearly $\overline{\mathrm{B}}[x, \rho(x) + \frac{\varepsilon}{10}] \subset W$. Therefore $\overline{\mathrm{B}}[x, \rho(x) + \frac{\varepsilon}{10}]$ is compact, a contradiction. □

2.16. Exercise. *Construct a geodesic space that is locally compact, but whose completion is neither geodesic nor locally compact.*

D. Convex sets

2.17. Definition. *Let $\mathcal{X}$ be a geodesic space and $A \subset \mathcal{X}$.*

A is convex *if for every two points $p, q \in A$ any geodesic $[pq]$ lies in A.*

A is weakly convex *if for every two points $p, q \in A$ there is a geodesic $[pq]$ that lies in A.*

We say that A is totally convex *if for every two points $p, q \in A$, every local geodesic from p to q lies in A.*

If for some $R \in (0, \infty]$ these definitions are applied only for pairs of points such that $|p - q| < R$ and only for the geodesics of length $< R$, then A is called respectively R-convex, weakly R-convex, *or* totally R-convex.

A set $A \subset \mathcal{X}$ is called locally convex *if every point $a \in A$ admits an open neighborhood $\Omega \ni a$ such that for every two points $p, q \in A \cap \Omega$ every geodesic $[pq] \subset \Omega$ lies in A. Similarly one defines locally weakly convex and locally totally convex sets.*

Remarks. Let us state a few observations that easily follow from the definition.

- The notion of (*weakly*) *convex set* is the same as (*weakly*) ∞*-convex set*.
- The inherited metric on a weakly convex set coincides with its length metric.
- Any open set is locally convex by definition.

The following proposition states that weak convexity survives under ultralimits. An analogous statement about convexity does not hold; for example, there is a sequence of convex discs in $\mathbb{S}^2$ that converges to a hemisphere, which is not convex.

2.18. Proposition. *Let $\mathcal{X}_n$, be a sequence of geodesic spaces. Let ω be an ultrafilter on $\mathbb{N}$ (see Definition 4.1). Assume that $A_n \subset \mathcal{X}_n$ is a sequence of weakly convex sets, $\mathcal{X}_n \to \mathcal{X}_\omega$, and $A_n \to A_\omega \subset \mathcal{X}_\omega$ as $n \to \omega$. Then A_ω is a weakly convex set of $\mathcal{X}_\omega$.*

Proof. Fix $x_\omega, y_\omega \in A_\omega$. Consider sequences $x_n, y_n \in A_n$ such that $x_n \to x_\omega$ and $y_n \to y_\omega$ as $n \to \omega$.

Denote by α_n a geodesic path from x_n to y_n that lies in A_n. Let

$$\alpha_\omega(t) = \lim_{n\to\omega} \alpha_n(t).$$

It remains to observe that α_ω is a geodesic path that lies in A_ω. $\square$

E. Quotient spaces

Quotient spaces. Assume $\mathcal{X}$ is a metric space with an equivalence relation $\sim$. Note that given a family of pseudometrics ρ_α on $\mathcal{X}/\sim$, their least upper bound

$$\rho(x, y) = \sup_\alpha \{\, \rho_\alpha(x, y)\, \}$$

is also a pseudometric. If ~~for~~ the projections $\mathcal{X} \to (\mathcal{X}/\sim, \rho_\alpha)$ are *short* (that is, *distance non-increasing*), then so is $\mathcal{X} \to (\mathcal{X}/\sim, \rho)$.

It follows that the quotient space $\mathcal{X}/\sim$ admits a natural quotient pseudometric; this is the maximal pseudometric on $\mathcal{X}/\sim$ that makes the quotient map $\mathcal{X} \to \mathcal{X}/\sim$ short. The corresponding metric space will be also denoted as $\mathcal{X}/\sim$ and will be called the *quotient space* of $\mathcal{X}$ by the equivalence relation $\sim$.

In general, the points of the metric space $\mathcal{X}/\sim$ are formed by equivalence classes in $\mathcal{X}$ for a wider equivalence relation. However, in most of the cases we will consider, the set of equivalence classes will coincide with the set of points in the metric space $\mathcal{X}/\sim$.

2.19. Proposition. *Let $\mathcal{X}$ be a length space and $\sim$ be an equivalence relation on $\mathcal{X}$. Then $\mathcal{X}/\sim$ is a length space.*

Proof. Let $\mathcal{Y}$ be an arbitrary metric space. Since $\mathcal{X}$ is a length space, a map $f\colon \mathcal{X} \to \mathcal{Y}$ is short if and only if

$$\operatorname{length}(f \circ \alpha) \leqslant \operatorname{length} \alpha$$

for any curve $\alpha\colon \mathbb{I} \to \mathcal{X}$. Denote by $\|{*} - {*}\|$ the length metric on $\mathcal{Y}$. It follows that if $f\colon \mathcal{X} \to \mathcal{Y}$ is short, then so is $f\colon \mathcal{X} \to (\mathcal{Y}, \|{*} - {*}\|)$.

Consider the quotient map $f\colon \mathcal{X} \to \mathcal{X}/\sim$. Recall that the space $\mathcal{X}/\sim$ is defined by the maximal pseudometric that makes f short.

Denoting by $\|{*} - {*}\|$ the length metric on $\mathcal{X}/\sim$, it follows that

$$f\colon \mathcal{X} \to (\mathcal{X}/\sim, \|{*} - {*}\|)$$

is also short.

Note that

$$\|x - y\| \geqslant |x - y|_{\mathcal{X}/\sim}$$

for any $x, y \in \mathcal{X}/\sim$. From maximality of $|{*} - {*}|_{\mathcal{X}/\sim}$, we get

$$\|x - y\| = |x - y|_{\mathcal{X}/\sim}$$

for any $x, y \in \mathcal{X}/\sim$; that is, $\mathcal{X}/\sim$ is a length space. □

Group actions. Assume a group G acts on a metric space $\mathcal{X}$. Consider a relation $\sim$ on $\mathcal{X}$ defined by $x \sim y$ if there is $g \in G$ such that $x = g \cdot y$. Note that $\sim$ is an equivalence relation.

In this case, the quotient space $\mathcal{X}/\sim$ will also be denoted by $\mathcal{X}/G$, and can be regarded as the space of G-orbits in $\mathcal{X}$.

Assume that a group G acts on $\mathcal{X}$ by isometries. Then the distance between orbits $G \cdot x$ and $G \cdot y$ in $\mathcal{X}/G$ can be defined directly:

$$|G \cdot x - G \cdot y|_{\mathcal{X}/G} = \inf\big\{\, |x - g \cdot y|_{\mathcal{X}} = |g^{-1} \cdot x - y|_{\mathcal{X}} : g \in G \,\big\}.$$

If the G-orbits are closed, then $|G \cdot x - G \cdot y|_{\mathcal{X}/G} = 0$ if and only if $G \cdot x = G \cdot y$. In this case, the quotient space $\mathcal{X}/G$ is a genuine metric space.

The following proposition follows from the definition of a quotient space:

2.20. Proposition. *Assume $\mathcal{X}$ is a metric space and a group G acts on $\mathcal{X}$ by isometries. Then the projection $\pi\colon \mathcal{X} \to \mathcal{X}/G$ is a* submetry*; that is, $\pi(\mathrm{B}(p, r)) = \mathrm{B}(\pi(p), r)$ for any $p \in \mathcal{X}, r > 0$ (see Definition* 3.7*).*

F. Gluing and doubling

Gluing. Recall that the disjoint union of metric spaces can be also considered as a metric space; see Definition 2.1. Therefore the quotient space construction works as well for an equivalence relation on the disjoint union of metric spaces.

Consider two metric spaces $\mathcal{X}_1$ and $\mathcal{X}_2$ with subsets $A_1 \subset \mathcal{X}_1$ and $A_2 \subset \mathcal{X}_2$, and a bijection $\phi\colon A_1 \to A_2$. Consider the minimal equivalence relation on $\mathcal{X}_1 \sqcup \mathcal{X}_2$ such that $a \sim \phi(a)$ for any $a \in A_1$. In this case, the corresponding quotient space $(\mathcal{X}_1 \sqcup \mathcal{X}_2)/{\sim}$ will be called the *gluing of $\mathcal{X}$ and $\mathcal{Y}$ along ϕ* and denoted by

$$\mathcal{X}_1 \sqcup_\phi \mathcal{X}_2.$$

Note that if the map $\phi\colon A_1 \to A_2$ is distance-preserving, then the inclusions $\iota_i\colon \mathcal{X}_i \to \mathcal{X}_1 \sqcup_\phi \mathcal{X}_2$ are also distance-preserving, and

$$|\iota_1(x_1) - \iota_2(x_2)|_{\mathcal{X}_1 \sqcup_\phi \mathcal{X}_2} = \inf_{a_2 = \phi(a_1)} \{\, |x_1 - a_1|_{\mathcal{X}_1} + |x_2 - a_2|_{\mathcal{X}_2} \,\}$$

for any $x_1 \in \mathcal{X}_1$ and $x_2 \in \mathcal{X}_2$.

Doubling. Let $\mathcal{V}$ be a metric space and $A \subset \mathcal{V}$ be a closed subset. A metric space $\mathcal{W}$ glued from two copies of $\mathcal{V}$ along A is called the *doubling of $\mathcal{V}$ in A*.

The space $\mathcal{W}$ is completely described by the following properties:

- The space $\mathcal{W}$ contains $\mathcal{V}$ as a subspace; in particular the set A can be treated as a subset of $\mathcal{W}$.
- There is an isometric involution of $\mathcal{W}$ which is called *reflection in A*; it will be denoted by $x \mapsto x'$.
- For any $x \in \mathcal{W}$ we have $x \in \mathcal{V}$ or $x' \in \mathcal{V}$ and

  $$|x' - y|_{\mathcal{W}} = |x - y'|_{\mathcal{W}} = \inf_{a \in A} \{|x - a|_{\mathcal{V}} + |a - y|_{\mathcal{V}}\}$$

  for any $x, y \in \mathcal{V}$.

The image of $\mathcal{V}$ under reflection in A will be denoted by $\mathcal{V}'$. The subspace $\mathcal{V}'$ is an isometric copy of $\mathcal{V}$. Clearly $\mathcal{V} \cup \mathcal{V}' = \mathcal{W}$ and $\mathcal{V} \cap \mathcal{V}' = A$. Moreover, $a = a' \iff a \in A$.

The following proposition follows directly from the definitions.

2.21. Proposition. *Assume $\mathcal{W}$ is the doubling of the metric space $\mathcal{V}$ in its closed subset A. Then:*

(*a*) *If $\mathcal{V}$ is a complete length space, then so is $\mathcal{W}$.*

(*b*) *If $\mathcal{V}$ is proper, then so is $\mathcal{W}$. In this case, for any $x, y \in \mathcal{V}$ there is $a \in A$ such that*

$$|x - a|_{\mathcal{V}} + |a - y|_{\mathcal{V}} = |x - y'|_{\mathcal{W}}.$$

(c) *Given $x \in \mathcal{W}$, let $\bar{x} = x$ if $x \in \mathcal{V}$, and $\bar{x} = x'$ otherwise. The map $\mathcal{W} \to \mathcal{V}$ defined by $x \mapsto \bar{x}$ is short and length-preserving. In particular, if γ is a geodesic in $\mathcal{W}$ with ends in $\mathcal{V}$, then $\bar{\gamma}$ is a geodesic in $\mathcal{V}$ with the same ends.*

G. Kuratowsky embedding

Given a metric space $\mathcal{X}$, let us denote by $\mathrm{Bnd}(\mathcal{X}, \mathbb{R})$ the space of all bounded functions on $\mathcal{X}$ equipped with the sup-norm

$$\|f\| = \sup_{x \in \mathcal{X}} \{|f(x)|\}.$$

Kuratowski embedding. Given a point $p \in \mathcal{X}$, consider the map $\mathrm{kur}_p : \mathcal{X} \to \mathrm{Bnd}(\mathcal{X}, \mathbb{R})$ defined by $\mathrm{kur}_p x = \mathrm{dist}_x - \mathrm{dist}_p$. The map kur_p will be called the *Kuratowski map at p.*

From the triangle inequality, we have

$$\| \mathrm{kur}_p x - \mathrm{kur}_p y\| = \sup_{z \in \mathcal{X}} \{||x - z| - |y - z||\} = |x - y|.$$

Therefore, for any $p \in \mathcal{X}$, the Kuratowski map gives a distance-preserving map $\mathrm{kur}_p : \mathcal{X} \hookrightarrow \mathrm{Bnd}(\mathcal{X}, \mathbb{R})$. Thus we can (and often will) consider the space $\mathcal{X}$ as a subset of $\mathrm{Bnd}(\mathcal{X}, \mathbb{R})$.

2.22. Exercise. *Show that any compact metric space is isometric to a subspace in a compact length space.*

Chapter 3

Maps and functions

Here we introduce several classes of maps between metric spaces and develop a language to describe various notions of convexity/concavity of real-valued functions on general metric spaces.

A. Submaps

We will often need maps and functions defined on subsets of a metric space. We call them *submaps* and *subfunctions*. Thus, given metric spaces $\mathcal{X}$ and $\mathcal{Y}$, a submap $\Phi\colon \mathcal{X} \multimapinv \mathcal{Y}$ is a map defined on a subset $\operatorname{Dom}\Phi \subset \mathcal{X}$.

A submap $\Phi\colon \mathcal{X} \multimapinv \mathcal{Y}$ is *continuous* if the inverse image of any open set is open. Since $\operatorname{Dom}\Phi = \Phi^{-1}(\mathcal{Y})$, the domain $\operatorname{Dom}\Phi$ of a continuous submap is open. The same holds for upper and lower semicontinuous functions $f\colon \mathcal{X} \multimapinv \mathbb{R}$ since they are continuous functions for a special topology on $\mathbb{R}$.

(Continuous partially defined maps could be defined via closed sets; namely, one could require that inverse images of closed sets are closed. While this condition is equivalent to continuity for functions defined on the whole space, it is different for partially defined functions. In particular, with this definition the domain of a continuous submap would have to be closed.)

B. Lipschitz conditions

3.1. Lipschitz maps. *Suppose $\mathcal{X}$ and $\mathcal{Y}$ are metric spaces, $\Phi\colon \mathcal{X} \multimapinv \mathcal{Y}$ is a continuous submap, and $\ell \in \mathbb{R}$.*

(*a*) *The submap Φ is called ℓ-*Lipschitz *if*

$$|\Phi(x) - \Phi(y)|_{\mathcal{Y}} \leqslant \ell \cdot |x - y|_{\mathcal{X}}$$

for any two points $x, y \in \operatorname{Dom}\Phi$.

- 1-*Lipschitz maps will be also called* short.

(*b*) *We say that Φ is* Lipschitz *if it is ℓ-Lipschitz for a constant ℓ. The minimal such constant is denoted by* $\operatorname{lip}\Phi$.

(*c*) *We say that Φ is* locally Lipschitz *if any point $x \in \operatorname{Dom}\Phi$ admits a neighborhood $\Omega \subset \operatorname{Dom}\Phi$ such that the restriction $\Phi|_\Omega$ is Lipschitz.*

(*d*) *Given $p \in \operatorname{Dom}\Phi$, we denote by $\operatorname{lip}_p\Phi$ the infimum of the real values ℓ such that p admits a neighborhood $\Omega \subset \operatorname{Dom}\Phi$ such that the restriction $\Phi|_\Omega$ is ℓ-Lipschitz.*

Note that $\Phi\colon \mathcal{X} \to \mathcal{Y}$ is ℓ-Lipschitz if and only if

$$\Phi(\mathrm{B}(x,R)_{\mathcal{X}}) \subset \mathrm{B}(\Phi(x),\ell\cdot R)_{\mathcal{Y}}$$

for any $R \geqslant 0$ and $x \in \mathcal{X}$. A dual version of this property is considered in the following definition.

3.2. Definitions. *Let $\Phi\colon \mathcal{X} \to \mathcal{Y}$ be a map between metric spaces, and $\ell \in \mathbb{R}$.*

(*a*) *The map Φ is called* ℓ-co-Lipshitz *if*

$$\Phi(\mathrm{B}(x,\ell\cdot R)_{\mathcal{X}}) \supset \mathrm{B}(\Phi(x),R)_{\mathcal{Y}}$$

for any $x \in \mathcal{X}$ and $R > 0$.

(*b*) *The map Φ is called* co-Lipschitz *if it is ℓ-co-Lipschitz for some constant ℓ. The minimal such constant is denoted by* $\operatorname{colip}\Phi$.

From the definition of co-Lipschitz maps we get the following:

3.3. Proposition. *Any co-Lipschitz map is open and surjective.*

In other words, ℓ-co-Lipschitz maps can be considered as a quantitative version of open maps. For that reason they are also called ℓ-open [**44**]. Also, be aware that some authors refer to our ℓ-co-Lipschitz maps as $\frac{1}{\ell}$-co-Lipschitz.

3.4. Proposition. *Let $\mathcal{X}$ and $\mathcal{Y}$ be metric spaces such that $\mathcal{X}$ is complete, and let $\Phi\colon \mathcal{X} \to \mathcal{Y}$ be a continuous co-Lipschitz map. Then $\mathcal{Y}$ is complete.*

Proof. Choose a Cauchy sequence y_n in $\mathcal{Y}$. Passing to a subsequence if necessary, we may assume that $|y_n - y_{n+1}|_{\mathcal{Y}} < \frac{1}{2^n}$ for each n. It is sufficient to show that y_n converges in $\mathcal{Y}$.

Denote by ℓ a co-Lipschitz constant for Φ. Note that there is a sequence x_n in $\mathcal{X}$ such that

$$\Phi(x_n) = y_n \quad\text{and}\quad |x_n - x_{n+1}|_{\mathcal{X}} < \tfrac{\ell}{2^n} \tag{1}$$

for each n. Indeed, such a sequence can be constructed recursively. Assuming that the points $x_1, \dots, x_{n-1}$ are already constructed, the existence of a sequence x_n satisfying 1 follows since Φ is ℓ-co-Lipschitz.

Notice that the sequence x_n is Cauchy. Since $\mathcal{X}$ is complete, x_n converges in $\mathcal{X}$; denote its limit by x_∞ and set $y_\infty = \Phi(x_\infty)$. Since Φ is continuous, $y_n \to y_\infty$ as $n \to \infty$. Hence the result. $\square$

3.5. Lemma. *Let $\mathcal{X}$ be a metric space and $f \colon \mathcal{X} \to \mathbb{R}$ be a continuous function. Then for any $\varepsilon > 0$ there is a locally Lipschitz function $f_\varepsilon \colon \mathcal{X} \to \mathbb{R}$ such that $|f(x) - f_\varepsilon(x)| < \varepsilon$ for any $x \in \mathcal{X}$.*

Proof. Assume that $f \geqslant 1$. Construct a continuous positive function $\rho \colon \mathcal{X} \to \mathbb{R}_{>0}$ such that

$$|x - y| < \rho(x) \quad \Rightarrow \quad |f(x) - f(y)| < \varepsilon.$$

Consider the function

$$f_\varepsilon(y) = \sup\left\{ f(x) \cdot \left(1 - \tfrac{|x-y|}{\rho(x)}\right) : x \in \mathcal{X} \right\}.$$

It is straightforward to check that each f_ε is locally Lipschitz and $0 \leqslant f_\varepsilon - f < \varepsilon$.

Since any continuous function can be presented as the difference of two continuous functions bounded below by 1, the result follows. $\square$

C. Isometries and submetries

3.6. Isometry. *Let $\mathcal{X}$ and $\mathcal{Y}$ be metric spaces and $\Phi \colon \mathcal{X} \to \mathcal{Y}$ be a map.*

(*a*) *The map Φ is* distance-preserving *if*

$$|\Phi(x) - \Phi(x')|_{\mathcal{Y}} = |x - x'|_{\mathcal{X}}$$

for any $x, x' \in \mathcal{X}$.

(*b*) *A distance-preserving bijection Φ is called an* isometry.

(*c*) *The spaces $\mathcal{X}$ and $\mathcal{Y}$ are called* isometric (*briefly $\mathcal{X} \overset{iso}{=\!=} \mathcal{Y}$*) *if there is an isometry $\Phi \colon \mathcal{X} \to \mathcal{Y}$.*

3.7. Submetry. *A map $\sigma \colon \mathcal{X} \to \mathcal{Y}$ between the metric spaces $\mathcal{X}$ and $\mathcal{Y}$ is called a* submetry *if*

$$\sigma(\mathrm{B}(p, r)_{\mathcal{X}}) = \mathrm{B}(\sigma(p), r)_{\mathcal{Y}}$$

for any $p \in \mathcal{X}$ and $r \geqslant 0$.

Note $\sigma \colon \mathcal{X} \to \mathcal{Y}$ is a submetry if it is 1-Lipschitz and 1-co-Lipschitz.

Note also that any submetry is an onto map.

The main source of examples of submetries comes from isometric group actions. Namely, assume $\mathcal{X}$ is a metric space and G is a subgroup of isometries of $\mathcal{X}$. Denote by $[x] = G \cdot x$ the G-orbit of $x \in \mathcal{X}$ and $\mathcal{X}/G$ the set of all G-orbits; let us equip it with the pseudometric defined by

$$|[x] - [y]|_{\mathcal{X}/G} = \inf\{\, |g \cdot x - h \cdot y|_{\mathcal{X}} : g, h \in G \,\}.$$

Note that if all the G-orbits form closed sets in $\mathcal{X}$, then $\mathcal{X}/G$ is a genuine metric space.

3.8. Proposition. *Let $\mathcal{X}$ be a metric space. Assume that a group G acts on $\mathcal{X}$ by isometries and in such a way that every G-orbit is closed. Then the projection map $\mathcal{X} \to \mathcal{X}/G$ is a submetry.*

Proof. We need to show that the map $x \mapsto [x] = G \cdot x$ is 1-Lipschitz and 1-co-Lipschitz. The co-Lipschitz part follows directly from the definitions of Hausdorff distance and co-Lipschitz maps.

Assume $|x - y|_{\mathcal{X}} < r$; equivalently $\mathrm{B}(x, r)_{\mathcal{X}} \ni y$. Since the action $G \curvearrowright \mathcal{X}$ is isometric, $\mathrm{B}(g \cdot x, r)_{\mathcal{X}} \ni g \cdot y$ for any $g \in G$.

In particular, the orbit $G \cdot y$ lies in the open r-neighborhood of the orbit $G \cdot x$. In the same way we can prove that the orbit $G \cdot x$ lies in the open r-neighborhood of the orbit $G \cdot y$. That is, the Hausdorff distance between the orbits $G \cdot x$ and $G \cdot y$ is less than r or, equivalently, $|[x] - [y]|_{\mathcal{X}/G} < r$. Since x and y are arbitrary, the map $x \mapsto [x]$ is 1-Lipschitz. $\square$

3.9. Proposition. *Let $\mathcal{X}$ be a length space and $\sigma \colon \mathcal{X} \to \mathcal{Y}$ be a submetry. Then $\mathcal{Y}$ is a length space.*

Proof. Fix $\varepsilon > 0$ and a pair of points $x, y \in \mathcal{Y}$.

Since σ is 1-co-Lipschitz, there are points $\hat{x}, \hat{y} \in \mathcal{X}$ such that $\sigma(\hat{x}) = x$, $\sigma(\hat{y}) = y$, and $|\hat{x} - \hat{y}|_{\mathcal{X}} < |x - y|_{\mathcal{Y}} + \varepsilon$.

Since $\mathcal{X}$ is a length space, there is a curve γ joining $\hat{x}$ to $\hat{y}$ in $\mathcal{X}$ such that

$$\operatorname{length} \gamma \leqslant |x - y|_{\mathcal{Y}} + \varepsilon.$$

The curve $\sigma \circ \gamma$ joins x to y. Since σ is 1-Lipschitz, and by the above,

$$\begin{aligned} \operatorname{length} \sigma \circ \gamma &\leqslant \operatorname{length} \gamma \\ &\leqslant |x - y|_{\mathcal{Y}} + \varepsilon. \end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, $\mathcal{Y}$ is a length space. $\square$

D. Speed of curves

Let $\mathcal{X}$ be a metric space. Recall that a *curve* in $\mathcal{X}$ is a continuous map $\alpha \colon \mathbb{I} \to \mathcal{X}$, where $\mathbb{I}$ is a real interval. A curve is called *Lipschitz* or *locally Lipschitz* if α is Lipschitz or locally Lipschitz respectively. Length of curves is defined in 2.5.

The following theorem follows from [**37**, 2.7].

3.10. Theorem. *Let $\mathcal{X}$ be a metric space and $\alpha\colon \mathbb{I}\to\mathcal{X}$ be a locally Lipschitz curve. Then the speed function*

$$\operatorname{speed}_{t_0}\alpha=\lim_{\substack{t\to t_0+\\ s\to t_0-}}\frac{|\alpha(t)-\alpha(s)|}{|t-s|}$$

is defined for almost all $t_0\in\mathbb{I}$, and

$$\operatorname{length}\alpha=\int\limits_{\mathbb{I}}\operatorname{speed}_t\alpha\cdot\mathbf{d}t,$$

where $\int$ denotes the Lebesgue integral.

A curve $\alpha\colon\mathbb{I}\to\mathcal{X}$ is *unit-speed* if

$$b-a=\operatorname{length}(\alpha|_{[a,b]})$$

for any subinterval $[a,b]\subset\mathbb{I}$. If α is Lipschitz, then, according to ~~the above~~ theorem, this is equivalent to

$$\operatorname{speed}\alpha\overset{a.e.}{=\!=}1.$$

The following generalization of the standard Rademacher theorem on differentiability almost everywhere of Lipschitz maps between smooth manifolds [**37**, 5.5.2] was proved by Bernd Kirchheim [**94**].

The conclusion of the standard Rademacher theorem does not make sense for maps to a metric space since the target might have no linear structure. But the theorem does not hold even if we assume that the target is a Banach space. For example, consider the map $[0,1]\to L^1[0,1]$, defined by $x\mapsto\chi_{[0,x]}$, where χ_A denotes the characteristic function of A. This map is distance-preserving and in particular Lipschitz, but its differential is undefined at any point.

3.11. Theorem. *Let $\mathcal{X}$ be a metric space and $f\colon\mathbb{R}^n\multimap\mathcal{X}$ be 1-Lipschitz. Then for almost all $x\in\operatorname{Dom}f$ there is a pseudonorm $\|{*}\|_x$ on $\mathbb{R}^n$ such that*

$$|f(y)-f(z)|_{\mathcal{X}}=\|z-y\|_x+o(|y-x|+|z-x|).$$

Given f, the (pseudo)norm $\|{*}\|_x$ in the above theorem will be called its *differential of the induced metric* at x, or *metric differential* at x.

E. Convex real-to-real functions

In this section we will discuss generalized solutions of the ~~following~~ differential inequalities

$$y''+\kappa\cdot y\geqslant\lambda\quad\text{and respectively}\quad y''+\kappa\cdot y\leqslant\lambda \tag{1}$$

for fixed $\kappa,\lambda\in\mathbb{R}$. The solutions $y\colon\mathbb{R}\multimap\mathbb{R}$ are only assumed to be upper (respectively lower) semicontinuous subfunctions.

~~The~~ inequalities 1 are understood in the sense of distributions. That is, for any smooth function ϕ with compact support $\operatorname{Supp}\phi \subset \operatorname{Dom} y$ the following inequality should be satisfied:

$$\int\limits_{\operatorname{Dom} y} [y(t) \cdot \phi''(t) + \varkappa \cdot y(t) \cdot \phi(t) - \lambda] \cdot \mathbf{d}t \geqslant 0, \tag{2}$$

$$\text{respectively} \quad \leqslant 0.$$

The integral is understood in the sense of Lebesgue; in particular ~~the~~ inequality 2 makes sense for any Borel-measurable subfunction y. The proofs of the following propositions are straightforward.

3.12. Proposition. *Let $\mathbb{I} \subset \mathbb{R}$ be an open interval and $y_n : \mathbb{I} \to \mathbb{R}$ be a sequence of solutions of one of the inequalities in 1. Assume $y_n(t) \to y_\infty(t)$ as $n \to \infty$ for any $t \in \mathbb{I}$. Then y_∞ is a solution of the same inequality in 1.*

Assume y is a solution of one of the inequalities in 1. For $t_0 \in \operatorname{Dom} y$, let us define the *right (left) derivative* $y^+(t_0)$ $(y^-(t_0))$ at t_0 by

$$y^\pm(t_0) = \lim_{t \to t_0 \pm} \frac{y(t) - y(t_0)}{|t - t_0|}.$$

Note that our sign convention for y^- is not standard—for $y(t) = t$ we have $y^+(t) = 1$ and $y^-(t) = -1$.

3.13. Proposition. *Let $\mathbb{I} \subset \mathbb{R}$ be an open interval and $y : \mathbb{I} \to \mathbb{R}$ be a solution of an inequality in 1. Then y is locally Lipschitz; its right and left derivatives $y^+(t_0)$ and $y^-(t_0)$ are defined for any $t_0 \in \mathbb{I}$. Moreover*

$$y^+(t_0) + y^-(t_0) \geqslant 0 \quad \text{or respectively} \quad y^+(t_0) + y^-(t_0) \leqslant 0.$$

The next theorem gives a number of equivalent ways to define such generalized solutions.

3.14. Theorem. *Let $\mathbb{I}$ be an open real interval and $y : \mathbb{I} \to \mathbb{R}$ be a locally Lipschitz function. Then the following conditions are equivalent:*

(a) *$y'' \geqslant \lambda - \varkappa \cdot y$ (respectively $y'' \leqslant \lambda - \varkappa \cdot y$).*

(b) *(barrier inequality) For any $t_0 \in \mathbb{I}$, there is a solution $\bar{y}$ of the ordinary differential equation $\bar{y}'' = \lambda - \varkappa \cdot \bar{y}$ with $\bar{y}(t_0) = y(t_0)$ such that $\bar{y} \geqslant y$ (respectively $\bar{y} \leqslant y$) for all $t \in [t_0 - \varpi\varkappa, t_0 + \varpi\varkappa] \cap \mathbb{I}$.*

The function $\bar{y}$ is called a lower *(respectively* upper*)* barrier *of y at t_0.*

(c) *(Jensen's inequality) For any pair of values $t_1 < t_2$ in $\mathbb{I}$ such that $|t_2 - t_1| < \varpi\varkappa$, the unique solution $z(t)$ of*

$$z'' = \lambda - \varkappa \cdot z$$

such that

$$z(t_1) = y(t_1), \quad z(t_2) = y(t_2)$$

satisfies $y(t) \leqslant z(t)$ *(respectively* $y(t) \geqslant z(t)$*) for all* $t \in [t_1, t_2]$.

Further, the following property holds:

(d) *Suppose* $y'' \leqslant \lambda - \kappa \cdot y$. *Let* $t_0 \in \mathbb{I}$, *and* $\bar{y}$ *be a solution of the ordinary differential equation* $\bar{y}'' = \lambda - \kappa \cdot \bar{y}$ *such that* $\bar{y}(t_0) = y(t_0)$ *and* $y^+(t_0) \leqslant \bar{y}^+(t_0) \leqslant -y^-(t_0)$. *(Note that such a* $\bar{y}$ *is unique if* y *is differentiable at* t_0.*)*

Then $\bar{y} \geqslant y$ *for all* $t \in [t_0 - \varpi\kappa, t_0 + \varpi\kappa] \cap \mathbb{I}$; *that is,* $\bar{y}$ *is a barrier of* y *at* t_0. *(Similarly, by reversing inequalities, for* $y'' \geqslant \lambda - \kappa \cdot y$.*)*

Note that Theorem 3.14 implies that y satisfies $y'' \geqslant \lambda$ ($y'' \leqslant \lambda$) on an interval $\mathbb{I} \subset \mathbb{R}$ if and only if $y(t) - \frac{\lambda}{2} \cdot t^2$ is convex (concave) on $\mathbb{I}$.

We will often need the following fact about convergence of derivatives of convex functions:

3.15. Two-shoulder lemma. *Let* $\mathbb{I}$ *be an open interval and* $f_n : \mathbb{I} \to \mathbb{R}$ *be a sequence of convex functions. Assume the functions* f_n *pointwise converge to a function* $f_\infty : \mathbb{I} \to \mathbb{R}$. *Then for any* $t_0 \in \mathbb{I}$,

$$f_\infty^\pm(t_0) \leqslant \varliminf_{n\to\infty} f_n^\pm(t_0).$$

Proof. Since the f_n are convex, we have $f_n^+(t_0) + f_n^-(t_0) \geqslant 0$, and for any t,

$$f_n(t) \geqslant f_n(t_0) \pm f^\pm(t_0) \cdot (t - t_0).$$

Passing to the limit, we get

$$f_\infty(t) \geqslant f_\infty(t_0) + \left[\varlimsup_{n\to\infty} f_n^+(t_0)\right] \cdot (t - t_0)$$

for $t \geqslant t_0$, and

$$f_\infty(t) \geqslant f_\infty(t_0) - \left[\varlimsup_{n\to\infty} f_n^-(t_0)\right] \cdot (t - t_0)$$

for $t \leqslant t_0$. Hence the result. □

3.16. Corollary. *Let* $\mathbb{I}$ *be an open interval and* $f_n : \mathbb{I} \to \mathbb{R}$ *be a sequence of functions such that* $f_n'' \leqslant \lambda$ *that converge pointwise to a function* $f_\infty : \mathbb{I} \to \mathbb{R}$. *Then:*

(a) *If* f_∞ *is differentiable at* $t_0 \in \mathbb{I}$, *then*

$$f_\infty'(t_0) = \pm \lim_{n\to\infty} f_n^\pm(t_0).$$

(b) *If all* f_n *and* f_∞ *are differentiable at* $t_0 \in \mathbb{I}$, *then*

$$f_\infty'(t_0) = \lim_{n\to\infty} f_n'(t_0).$$

Proof. Set $\hat{f}_n(t) = f_n(t) - \frac{\lambda}{2} \cdot t^2$ and $\hat{f}_\infty(t) = f_\infty(t) - \frac{\lambda}{2} \cdot t^2$. Note that the $\hat{f}_n$ are concave and $\hat{f}_n \to \hat{f}_\infty$ pointwise. Thus the theorem follows from ~~the~~ two-shoulder lemma (3.15). $\square$

F. Convex functions on a metric space

In this section we define different types of convexity/concavity in the context of metric spaces; it will be mostly used for geodesic spaces. The notation refers to the corresponding second-order ordinary differential inequality.

3.17. Definition. *Let $\mathcal{X}$ be a metric space. We say that an upper semicontinuous subfunction $f : \mathcal{X} \multimap (-\infty, \infty]$ satisfies the inequality*

$$f'' + \kappa \cdot f \geqslant \lambda$$

if for any *unit-speed geodesic γ in* Dom f*, the real-to-real function $y(t) = f \circ \gamma(t)$ satisfies*

$$y'' + \kappa \cdot y \geqslant \lambda$$

in the domain $\{\, t : y(t) < \infty \,\}$; see the definition in Section 3E.

We say that a lower semicontinuous subfunction $f : \mathcal{X} \multimap [-\infty, \infty)$ satisfies the inequality

$$f'' + \kappa \cdot f \leqslant \lambda$$

if the subfunction $h = -f$ satisfies

$$h'' - \kappa \cdot h \geqslant -\lambda.$$

Functions satisfying the inequalities

$$f'' \geqslant \lambda \quad \text{and} \quad f'' \leqslant \lambda$$

*are called λ-*convex *and λ-*concave *respectively.*

0-*convex and* 0-*concave subfunctions will also be called* convex *and* concave *respectively.*

If f is λ-convex for $\lambda > 0$, then f will be called strongly convex*; correspondingly, if f is λ-concave for $\lambda < 0$, then f will be called* strongly concave.

If for any point $p \in$ Dom f *there is a neighborhood $\Omega \ni p$ and a real number λ such that the restriction $f|_\Omega$ is λ-convex (or λ-concave), then f is called* semiconvex (*respectively* semiconcave).

Various authors define the class of λ-convex (λ-concave) functions differently. Their definitions may correspond to $\pm\lambda$-convex ($\pm\lambda$-concave) or $\pm\frac{\lambda}{2}$-convex ($\pm\frac{\lambda}{2}$-concave) functions in our definitions.

3.18. Proposition. *Let $\mathcal{X}$ be a metric space. Assume that $f : \mathcal{X} \multimap \mathbb{R}$ is a semiconvex subfunction and $\phi : \mathbb{R} \to \mathbb{R}$ is a nondecreasing semiconvex function. Then the composition $\phi \circ f$ is a semiconvex subfunction.*

The proof is straightforward.

Chapter 4

Ultralimits

Here we introduce ultralimits of sequences of points, metric spaces, and functions. Our presentation is based on [**96**].

Ultralimits are closely related to Gromov–Hausdorff limits. We use them only as a canonical way to pass to convergent subsequences. We could avoid using them at the cost of saying "pass to a convergent subsequence" too many times. Doing this might be cumbersome and it obscure ideas of the proof; see for example the proof of the globalization theorem for general CBB spaces. Also the use of ultralimits is convenient when dealing with CAT spaces due to the lack of compactness results.

A. Ultrafilters

We will need the existence of a selective ultrafilter ω that will be fixed once and for all. The existence follows from the axiom of choice and the continuum hypothesis.

Measure-theoretic definition. Recall that $\mathbb{N}$ denotes the set of natural numbers, $\mathbb{N} = \{1, 2, \dots\}$.

4.1. Definition. *A finitely additive measure ω on $\mathbb{N}$ is called an* ultrafilter *if it satisfies*

(*a*) *$\omega(S) = 0$ or 1 for any subset $S \subset \mathbb{N}$. An ultrafilter ω is called* nonprincipal *if in addition*

(*b*) *$\omega(F) = 0$ for any finite subset $F \subset \mathbb{N}$. A nonprincipal ultrafilter ω is called* selective *if in addition*

(c) *for any partition of $\mathbb{N}$ into sets $\{C_\alpha\}_{\alpha\in\mathcal{A}}$ such that $\omega(C_\alpha) = 0$ for each α, there is a set $S \subset \mathbb{N}$ such that $\omega(S) = 1$ and $S \cap C_\alpha$ is a one-point set for each $\alpha \in \mathcal{A}$.*

If $\omega(S) = 0$ for some subset $S \subset \mathbb{N}$, we say that S is *ω-small*. If $\omega(S) = 1$, we say that S contains *ω-almost all* elements of $\mathbb{N}$.

4.2. Advanced exercise. *Let ω be an ultrafilter and $f\colon \mathbb{N} \to \mathbb{N}$. Suppose that $\omega(S) \leqslant \omega(f^{-1}(S))$ for any set $S \subset \mathbb{N}$. Show that $f(n) = n$ for ω-almost all $n \in \mathbb{N}$.*

Classical definition. More commonly, a nonprincipal ultrafilter is defined as a collection, say $\mathfrak{F}$, of sets in $\mathbb{N}$ such that

(1) if $P \in \mathfrak{F}$ and $Q \supset P$, then $Q \in \mathfrak{F}$,
(2) if $P, Q \in \mathfrak{F}$, then $P \cap Q \in \mathfrak{F}$,
(3) for any subset $P \subset \mathbb{N}$, either P or its complement is an element of $\mathfrak{F}$,
(4) if $F \subset \mathbb{N}$ is finite, then $F \notin \mathfrak{F}$.

Setting

$$P \in \mathfrak{F} \quad \Longleftrightarrow \quad \omega(P) = 1$$

makes these two definitions equivalent.

A nonempty collection of sets $\mathfrak{F}$ that does not include the empty set and satisfies only conditions 1 and 2 is called a *filter*; if in addition $\mathfrak{F}$ satisfies Condition 3 it is called an *ultrafilter*. From Zorn's lemma, it follows that every filter is contained in an ultrafilter. Thus there is an ultrafilter $\mathfrak{F}$ contained in the filter of all complements of finite sets; clearly this $\mathfrak{F}$ is nonprincipal.

The existence of a selective ultrafilter follows from the continuum hypothesis; this was proved by Walter Rudin [**142**].

Stone–Čech compactification. Given a set $S \subset \mathbb{N}$, consider the subset Ω_S of all ultrafilters ω such that $\omega(S) = 1$. It is straightforward to check that the sets Ω_S for all $S \subset \mathbb{N}$ form a topology on the set of ultrafilters on $\mathbb{N}$. The resulting space is called the *Stone–Čech compactification* of $\mathbb{N}$; it is usually denoted by $\beta\mathbb{N}$.

There is a natural embedding $\mathbb{N} \hookrightarrow \beta\mathbb{N}$ defined by $n \mapsto \omega_n$, where ω_n is the principal ultrafilter such that $\omega_n(S) = 1$ if and only if $n \in S$. Using this embedding, we can (and will) consider $\mathbb{N}$ as a subset of $\beta\mathbb{N}$.

The space $\beta\mathbb{N}$ is the maximal compact Hausdorff space that contains $\mathbb{N}$ as an everywhere dense subset. More precisely, for any compact Hausdorff space $\mathcal{X}$ and a map $f\colon \mathbb{N} \to \mathcal{X}$, there is a unique continuous map $\bar{f}\colon \beta\mathbb{N} \to X$ such that the restriction $\bar{f}|_{\mathbb{N}}$ coincides with f.

B. Ultralimits of points

Fix an ultrafilter ω. Assume x_n is a sequence of points in a metric space $\mathcal{X}$. Define an *ω-limit* of x_n to be a point x_ω such that for any $\varepsilon > 0$, ω-almost all elements of x_n lie in $\mathrm{B}(x_\omega, \varepsilon)$; that is,

$$\omega\{\, n \in \mathbb{N} : |x_\omega - x_n| < \varepsilon \,\} = 1.$$

In this case, we write

$$x_\omega = \lim_{n\to\omega} x_n \quad \text{or} \quad x_n \to x_\omega \quad \text{as} \quad n \to \omega.$$

Also, if $\mathcal{X} = \mathbb{R}$ we write $\lim_{n\to\omega} x_n = \pm\infty$ if

$$\omega\{\, n \in \mathbb{N} : \pm x_n > L \,\} = 1$$

for any $L \in \mathbb{R}$.

It easily follows from the definition that ω-limits are unique if they exist. For example, if ω is the principal ultrafilter such that $\omega(\{n\}) = 1$ for some $n \in \mathbb{N}$, then $x_\omega = x_n$.

Note that ω-limits of a sequence and its subsequences may differ. For example, in general

$$\lim_{n\to\omega} x_n \neq \lim_{n\to\omega} x_{2\cdot n}.$$

The sequence x_n can be regarded as a map $\mathbb{N} \to \mathcal{X}$. If $\mathcal{X}$ is compact, then this map can be uniquely extended to a continuous map to the Stone–Čech compactification $\beta\mathbb{N}$ of $\mathbb{N}$. Then x_ω is the image of ω.

4.3. Proposition. *Let ω be a nonprincipal ultrafilter. Assume x_n is a sequence of points in a metric space $\mathcal{X}$ and $x_n \to x_\omega$ as $n \to \omega$. Then there is a subsequence of x_n that converges to x_ω in the usual sense.*

Moreover, if ω is selective, then the subsequence $(x_n)_{n\in S}$ can be chosen so that $\omega(S) = 1$.

Proof. Given $\varepsilon > 0$, let $S_\varepsilon = \{\, n \in \mathbb{N} : |x_n - x_\omega| < \varepsilon \,\}$.

Note that $\omega(S_\varepsilon) = 1$ for any $\varepsilon > 0$. Since ω is nonprincipal, the set S_ε is infinite. Therefore we can choose an increasing sequence n_k such that $n_k \in S_{\frac{1}{k}}$ for each $k \in \mathbb{N}$. Clearly $x_{n_k} \to x_\omega$ as $k \to \infty$.

Now assume that ω is selective. Consider the sets

$$C_k = \left\{\, n \in \mathbb{N} : \tfrac{1}{k} < |x_n - x_\omega| \leqslant \tfrac{1}{k-1} \,\right\},$$

where we assume $\frac{1}{0} = \infty$, and the set

$$C_\infty = \{\, n \in \mathbb{N} : x_n = x_\omega \,\}.$$

Note that $\omega(C_k) = 0$ for any $k \in \mathbb{N}$.

If $\omega(C_\infty) = 1$, we can take the subsequence consisting of the x_n, $n \in C_\infty$.

Otherwise, discarding all empty sets among C_k and C_∞ gives a partition of $\mathbb{N}$ into a countable collection of ω-small sets. Since ω is selective, we can choose a set $S \subset \mathbb{N}$ such that S meets each set of the partition at one point and $\omega(S) = 1$. Clearly the subsequence consisting of the x_n, $n \in S$ converges to x_ω in the usual sense. □

The following proposition is analogous to the statement that any sequence in a compact metric space has a convergent subsequence; it can be proved in the same way.

4.4. Proposition. *Let $\mathcal{X}$ be a compact metric space. Then any sequence of points x_n in $\mathcal{X}$ has a unique ω-limit x_ω.*

In particular, a bounded sequence of real numbers has a unique ω-limit.

The following lemma is an ultralimit analog of the Cauchy convergence test.

4.5. Lemma. *Let x_n be a sequence of points in a complete metric space $\mathcal{X}$. If for each subsequence y_n of x_n, the ω-limit*

$$y_\omega = \lim_{n\to\omega} y_n \in \mathcal{X}$$

is defined and does not depend on the choice of a subsequence, then the sequence x_n converges in the usual sense.

Proof. Assume the contrary. Then for some $\varepsilon > 0$, there is a subsequence y_n of x_n such that $|x_n - y_n| \geqslant \varepsilon$ for all n.

It follows that $|x_\omega - y_\omega| \geqslant \varepsilon$, a contradiction. □

4.6. Exercise. *Recall that ℓ^∞ denotes the space of bounded sequences of real numbers. Show that there is a linear functional $L\colon \ell^\infty \to \mathbb{R}$ such that for any sequence $\mathbf{s} = (s_1, s_2, \dots) \in S$ the image $L(\mathbf{s})$ is a partial limit of $s_1, s_2, \dots$*

4.7. Exercise. *Suppose that $f\colon \mathbb{N} \to \mathbb{N}$ is a map such that*

$$\lim_{n\to\omega} x_n = \lim_{n\to\omega} x_{f(n)}$$

for any bounded sequence x_n of real numbers. Show that $f(n) = n$ for ω-almost all $n \in \mathbb{N}$.

C. Ultralimits of spaces

Fix a selective ultrafilter ω on the set of natural numbers.

Let $\mathcal{X}_n$ be a sequence of metric spaces. Consider all sequences $x_n \in \mathcal{X}_n$. On the set of all such sequences, define a pseudometric by the formula

$$|(x_n) - (y_n)| = \lim_{n\to\omega} |x_n - y_n|. \tag{1}$$

Note that the ω-limit on the right-hand side is always defined and takes value in $[0, \infty]$.

Let $\mathcal{X}_\omega$ be the corresponding metric space; that is, the underlying set of $\mathcal{X}_\omega$ is formed by equivalence classes of sequences of points $x_n \in \mathcal{X}_n$ defined by the relation

$$(x_n) \sim (y_n) \quad \Longleftrightarrow \quad \lim_{n\to\omega} |x_n - y_n| = 0,$$

and the distance is defined as in 1.

The space $\mathcal{X}_\omega$ is called the *ω-limit* of $\mathcal{X}_n$. Typically $\mathcal{X}_\omega$ will denote the ω-limit of a sequence $\mathcal{X}_n$; we may also write

$$\mathcal{X}_n \to \mathcal{X}_\omega \quad \text{as} \quad n \to \omega \quad \text{or} \quad \mathcal{X}_\omega = \lim_{n\to\omega} \mathcal{X}_n.$$

Given a sequence $x_n \in \mathcal{X}_n$, we will denote by x_ω its equivalence class, which is a point in $\mathcal{X}_\omega$; in this case, we may write

$$x_n \to x_\omega \quad \text{as} \quad n \to \omega \quad \text{or} \quad x_\omega = \lim_{n\to\omega} x_n.$$

4.8. Observation. *The ω-limit of any sequence of metric spaces is complete.*

Proof. Let $\mathcal{X}_n$ be a sequence of metric spaces and $\mathcal{X}_n \to \mathcal{X}_\omega$ as $n \to \omega$. Choose a Cauchy sequence x_n in $\mathcal{X}_\omega$. Passing to a subsequence, we can assume that $|x_k - x_m|_{\mathcal{X}_\omega} < \frac{1}{k}$ for any $k < m$.

Let us choose points $x_{n,m} \in \mathcal{X}_n$ such that for any fixed m we have $x_{n,m} \to x_m$ as $n \to \omega$. Note that for any $k < m$ the inequality $|x_{n,k} - x_{n,m}| < \frac{1}{k}$ holds for ω-almost all n. It follows that we can choose a nested sequence of sets

$$\mathbb{N} = S_1 \supset S_2 \supset \dots$$

such that

- $\omega(S_m) = 1$ for each m,
- $\bigcap_m S_m = \emptyset$, and
- $|x_{n,k} - x_{n,l}| < \frac{1}{k}$ for $k < l \leqslant m$ and $n \in S_m$.

Consider the sequence $y_n = x_{n,m(n)}$, where $m(n)$ is the largest value such that $n \in S_{m(n)}$. Denote by $y_\omega \in \mathcal{X}_\omega$ the ω-limit of y_n.

Observe that $|y_m - x_{n,m}| < \frac{1}{m}$ for ω-almost all n. It follows that $|x_m - y_\omega| \leqslant \frac{1}{m}$ for any m. Therefore, $x_n \to y_\omega$ as $n \to \infty$. That is, any Cauchy sequence in $\mathcal{X}_\omega$ converges. $\square$

4.9. Observation. *The ω-limit of any sequence of length spaces is geodesic.*

Proof. If $\mathcal{X}_n$ is a sequence of length spaces, then for any sequence of pairs (x_n, y_n) in $\mathcal{X}_n$ there is a sequence of $\frac{1}{n}$-midpoints z_n.

Let $x_n \to x_\omega$, $y_n \to y_\omega$, and $z_n \to z_\omega$ as $n \to \omega$. Note that z_ω is a midpoint between x_ω and y_ω in $\mathcal{X}_\omega$.

By Observation 4.8, $\mathcal{X}_\omega$ is complete. Applying Lemma 2.13 we obtain the statement. □

A geodesic space $\mathcal{T}$ is called a *metric tree* if any pair of points in $\mathcal{T}$ are connected by a unique geodesic, and the union of any two geodesics $[xy]_{\mathcal{T}}$, and $[yz]_{\mathcal{T}}$ contain the geodesic $[xz]_{\mathcal{T}}$. The latter means that any triangle in $\mathcal{T}$ is a tripod; that is, for any three points x, y, and z there is a point p such that

$$[xy] \cup [yz] \cup [zx] = [px] \cup [py] \cup [pz].$$

4.10. Exercise. *Let $\mathcal{T}$ be a metric component of the ultralimit of $\mathbb{M}^2(n)$ as $n \to \omega$.*

(a) *Show that $\mathcal{T}$ is a complete metric tree.*

(b) *Show that $\mathcal{T}$ is homogeneous; that is, given two points $s, t \in \mathcal{T}$ there is an isometry of $\mathcal{T}$ that maps s to t.*

(c) *Show that $\mathcal{T}$ has* continuum degree *at any point; that is, for any point $t \in \mathcal{T}$ the set of connected components of the complement $\mathcal{T} \setminus \{t\}$ has cardinality continuum.*

Ultrapower. If all the metric spaces in a sequence are identical, $\mathcal{X}_n = \mathcal{X}$, the ω-limit $\lim_{n\to\omega} \mathcal{X}_n$ is denoted by $\mathcal{X}^\omega$ and called the ω-power of $\mathcal{X}$.

By Theorem 5.16, there is a distance-preserving map $\iota\colon \mathcal{X} \hookrightarrow \mathcal{X}^\omega$, where $\iota(y)$ is the equivalence class of the constant sequence $y_n = y$.

The image $\iota(\mathcal{X})$ might be a proper subset of $\mathcal{X}^\omega$. For example, $\mathbb{R}^\omega$ has pairs of points at distance ∞ from each other, while each metric component of $\mathbb{R}^\omega$ is isometric to $\mathbb{R}$.

According to Theorem 5.16, if $\mathcal{X}$ is compact then $\iota(\mathcal{X}) = \mathcal{X}^\omega$; in particular, $\mathcal{X}^\omega$ is isometric to $\mathcal{X}$. If $\mathcal{X}$ is proper, then $\iota(\mathcal{X})$ forms a metric component of $\mathcal{X}^\omega$.

The embedding ι allows us to treat $\mathcal{X}$ as a subset of its ultrapower $\mathcal{X}^\omega$.

4.11. Observation. *Let $\mathcal{X}$ be a complete metric space. Then $\mathcal{X}^\omega$ is a geodesic space if and only if $\mathcal{X}$ is a length space.*

Proof. Assume $\mathcal{X}^\omega$ is a geodesic space. Then any pair of points $x, y \in \mathcal{X}$ has a midpoint $z_\omega \in \mathcal{X}^\omega$. Fix a sequence of points $z_n \in \mathcal{X}$ such that $z_n \to z_\omega$ as $n \to \omega$.

Note that $|x - z_n|_{\mathcal{X}} \to \frac{1}{2} \cdot |x - y|_{\mathcal{X}}$ and $|y - z_n|_{\mathcal{X}} \to \frac{1}{2} \cdot |x - y|_{\mathcal{X}}$ as $n \to \omega$. In particular, for any $\varepsilon > 0$, the point z_n is an ε-midpoint between x and y for ω-almost all n. It remains to apply Lemma 2.13.

The *if* part follows from Observation 4.9. □

Note that the proof above together with Lemma 4.5 imply the following:

4.12. Corollary. *Assume $\mathcal{X}$ is a complete length space and $p, q \in \mathcal{X}$ cannot be joined by a geodesic in $\mathcal{X}$. Then there are at least continuum distinct geodesics between p and q in the ultrapower $\mathcal{X}^\omega$.*

4.13. Exercise. *Let $\mathcal{X}$ be a countable set with discrete metric; that is $|x-y|_{\mathcal{X}} = 1$ if $x \neq y$. Show that*

(*a*) *$\mathcal{X}^\omega$ is not isometric to $\mathcal{X}$.*

(*b*) *$\mathcal{X}^\omega$ is isometric to $(\mathcal{X}^\omega)^\omega$.*

4.14. Exercise. *Given a nonprincipal ultrafilter ω, construct an ultrafilter ω_1 such that*

$$\mathcal{X}^{\omega_1} \stackrel{iso}{=\!=} (\mathcal{X}^\omega)^\omega$$

for any metric space $\mathcal{X}$.

4.15. Exercise. *Construct a proper metric space $\mathcal{X}$ such that $\mathcal{X}^\omega$ is not proper; that is, there is a point $p \in \mathcal{X}^\omega$ and $R < \infty$ such that the closed ball $\overline{\mathrm{B}}[p, R]_{\mathcal{X}^\omega}$ is not compact.*

D. Ultralimits of sets

Let $\mathcal{X}_n$ be a sequence of metric spaces and $\mathcal{X}_n \to \mathcal{X}_\omega$ as $n \to \omega$.

For a sequence of sets $\Omega_n \subset \mathcal{X}_n$, consider the maximal set $\Omega_\omega \subset \mathcal{X}_\omega$ such that for any $x_\omega \in \Omega_\omega$ and any sequence $x_n \in \mathcal{X}_n$ such that $x_n \to x_\omega$ as $n \to \omega$, we have $x_n \in \Omega_n$ for ω-almost all n.

The set Ω_ω is called the *open ω-limit* of Ω_n; we could also write $\Omega_n \to \Omega_\omega$ as $n \to \omega$ or $\Omega_\omega = \lim_{n\to\omega} \Omega_n$.

Applying Observation 4.8 to the sequence of complements $\mathcal{X}_n \setminus \Omega_n$, we see that Ω_ω is open for any sequence Ω_n.

This definition can be applied to arbitrary sequences of sets, but we will apply it only for sequences of open sets.

E. Ultralimits of functions

Recall that a family of submaps (see section 3A) between metric spaces $\{f_\alpha : \mathcal{X} \multimap \mathcal{Y}\}_{\alpha\in\mathcal{A}}$ is called *equicontinuous* if for any $\varepsilon > 0$ there is $\delta > 0$ such that for any $p, q \in \mathcal{X}$ with $|p - q| < \delta$ and any $\alpha \in \mathcal{A}$ we have $|f_\alpha(p) - f_\alpha(q)| < \varepsilon$.

Let $f_n : \mathcal{X}_n \multimap \mathbb{R}$ be a sequence of subfunctions.

Set $\Omega_n = \operatorname{Dom} f_n$. Consider the open ω-limit set $\Omega_\omega \subset \mathcal{X}_\omega$ of Ω_n.

Assume there is a subfunction $f_\omega : \mathcal{X}_\omega \multimap \mathbb{R}$ that satisfies the following conditions: (1) $\operatorname{Dom} f_\omega = \Omega_\omega$, (2) if $x_n \to x_\omega \in \Omega_\omega$ for a sequence of points $x_n \in \mathcal{X}_n$, then $f_n(x_n) \to f_\omega(x_\omega)$ as $n \to \omega$. In this case, the subfunction $f_\omega : \mathcal{X}_\omega \to \mathbb{R}$ is said to be the ω-limit of $f_n : \mathcal{X}_n \to \mathbb{R}$.

The following lemma gives a mild condition on a sequence of functions f_n guaranteeing the existence of its ω-limit.

4.16. Lemma. *Let $\mathcal{X}_n$ be a sequence of metric spaces and $f_n : \mathcal{X}_n \multimap \mathbb{R}$ be a sequence of subfunctions.*

Assume that for any positive integer k, there is an open set $\Omega_n(k) \subset \operatorname{Dom} f_n$ such that the restrictions $f_n|_{\Omega_n(k)}$ are uniformly bounded and equicontinuous and the open ω-limit of $\Omega_n(n)$ coincides with the open ω-limit of $\operatorname{Dom} f_n$. Then the ω-limit f_ω of f_n is defined; moreover f_ω is a continuous subfunction.

In particular, if the functions f_n are uniformly bounded and equicontinuous, then its ω-limit f_ω is defined, bounded and uniformly continuous.

The proof is straightforward.

4.17. Exercise. *Construct a sequence of compact length spaces $\mathcal{X}_n$ with a converging sequence of ℓ-Lipschitz concave (see Definition 3.17) functions $f_n : \mathcal{X}_n \to \mathbb{R}$ such that the ω-limit $\mathcal{X}_\omega$ of $\mathcal{X}_n$ is compact and the ω-limit $f_\omega : \mathcal{X}_\omega \to \mathbb{R}$ of f_n is not concave.*

If $f : \mathcal{X} \multimap \mathbb{R}$ is a subfunction, the ω-limit of the constant sequence $f_n = f$ is called the ω-power of f and is denoted by f^ω. So

$$f^\omega : \mathcal{X} \multimap \mathbb{R}, \quad f^\omega(x_\omega) = \lim_{n\to\omega} f(x_n).$$

Evidently, if f^ω is defined, then f is continuous.

Recall that we treat $\mathcal{X}$ as a subset of its ω-power $\mathcal{X}^\omega$. Note that $\operatorname{Dom} f = \mathcal{X} \cap \operatorname{Dom} f^\omega$. Moreover, if $\mathrm{B}(x,\varepsilon)_{\mathcal{X}} \subset \operatorname{Dom} f$ then $\mathrm{B}(x,\varepsilon)_{\mathcal{X}^\omega} \subset \operatorname{Dom} f^\omega$.

Chapter 5

Space of spaces

In this chapter we discuss the Gromov–Hausdorff convergence of metric spaces.

To the best of our knowledge, Hausdorff convergence of subsets of a fixed metric space was first introduced by Felix Hausdorff [**81**], and a couple of years later an equivalent definition was given by Wilhelm Blaschke [**29**]. A further refinement of this definition was introduced by Zdeněk Frolík [**66**] and then rediscovered by Robert Wijsman [**158**]. However this refinement was a step in the direction of the so-called *closed convergence* introduced by Hausdorff in the original book. For that reason we call it Hausdorff convergence instead of *Hausdorff–Blashcke–Frolík–Wijsman convergence*.

Gromov–Hausdorff convergence was first introduced by David Edwards [**62**] and rediscovered later by Michael Gromov [**72**]. It was an essential tool in Gromov's proof that any group of polynomial growth has a nilpotent subgroup of finite index. Other versions of convergence of metric spaces were considered earlier, but each time the definition was limited to very specific types of problems.

The definition of Gromov–Hausdorff convergence of metric spaces uses the notion of Hausdorff convergence. Gromov–Hausdorff convergence means that a sequence of metric spaces admits a sequence of distance-preserving embeddings into a common ambient metric space so that their images converge in the Hausdorff sense. Our definition of Gromov–Hausdorff convergence and Gromov–Hausdorff distance differ somewhat from the standard definition.

A. Convergence of subsets

Let $\mathcal{X}$ be a metric space and $A \subset \mathcal{X}$. Recall that the distance from A to a point x in $\mathcal{X}$ is given by

$$\operatorname{dist}_A x := \inf\{\,|a - x| \,:\, a \in A\,\}.$$

By this definition, we have $\operatorname{dist}_\emptyset x = \infty$ for any x.

5.1. Definition of Hausdorff convergence. *Given a sequence of closed sets A_n in a metric space $\mathcal{X}$, a closed set $A_\infty \subset \mathcal{X}$ is called the Hausdorff limit of A_n, briefly $A_n \underset{\mathrm{H}}{\to} A_\infty$, if*

$$\operatorname{dist}_{A_n} x \to \operatorname{dist}_{A_\infty} x \quad \textit{as} \quad n \to \infty$$

for any fixed $x \in \mathcal{X}$.

In this case the sequence of closed sets A_n is said to be converging *or* converging in the sense of Hausdorff.

5.2. Selection theorem. *Let $\mathcal{X}$ be a proper metric space ~~and~~ A_n be a sequence of closed sets in $\mathcal{X}$. Then A_n has a converging subsequence in the sense of Hausdorff.*

Proof. Since $\mathcal{X}$ is proper, we can choose a countable dense set $\{x_1, x_2, \dots\}$ in $\mathcal{X}$.

If the sequence $a_n = \operatorname{dist}_{A_n} x_k$ is unbounded for some k, then we can pass to a subsequence of A_n such that $\operatorname{dist}_{A_n} x_k \to \infty$ as $n \to \infty$ for *any* k. The obtained sequence converges to the empty set.

Now suppose that $a_n = \operatorname{dist}_{A_n} x_k$ is bounded for each k. In this case, passing to a subsequence of A_n, we can assume that $\operatorname{dist}_{A_n} x_k$ converges as $n \to \infty$ for any fixed k.

Note that for each n, the function $\operatorname{dist}_{A_n} : \mathcal{X} \to \mathbb{R}$ is 1-Lipschitz and nonnegative. Therefore the sequence $\operatorname{dist}_{A_n}$ converges pointwise to a 1-Lipschitz nonnegative function $f : \mathcal{X} \to \mathbb{R}$.

Set $A_\infty = f^{-1}(0)$. Since f is 1-Lipschitz, $\operatorname{dist}_{A_\infty} y \geqslant f(y)$ for any $y \in \mathcal{X}$. It remains to show that $\operatorname{dist}_{A_\infty} y \leqslant f(y)$ for any y.

Assume the contrary, that is, $f(z) < R < \operatorname{dist}_{A_\infty} z$ for $z \in \mathcal{X}$ and $R > 0$. Then for any sufficiently large n there is a point $z_n \in A_n$ such that $|x - z_n| \leqslant R$. Since $\mathcal{X}$ is proper, we can pass to a partial limit z_∞ of z_n as $n \to \infty$.

Clearly $f(z_\infty) = 0$, that is, $z_\infty \in A_\infty$. On the other hand,

$$\operatorname{dist}_{A_\infty} y \leqslant |z_\infty - y| \leqslant R < \operatorname{dist}_{A_\infty} y,$$

a contradiction. $\square$

B. Convergence of spaces

5.3. Definition. *Let* $\{\mathcal{X}_\alpha : \alpha \in \mathcal{A}\}$ *be a set of metric spaces. A metric space* $\mathbf{X}$ *is called a* common space *of* $\{\mathcal{X}_\alpha : \alpha \in \mathcal{A}\}$ *if its underlying set is formed by the disjoint union*

$$\bigsqcup_{\alpha\in\mathcal{A}} \mathcal{X}_\alpha$$

and each inclusion $\iota_\alpha : \mathcal{X}_\alpha \hookrightarrow \mathbf{X}$ *is distance-preserving.*

5.4. Definition. *Let* $\mathbf{X}$ *be a common space for proper metric spaces* $\mathcal{X}_1, \mathcal{X}_2, \dots,$ *and* $\mathcal{X}_\infty$. *Assume that* $\mathcal{X}_n$ *forms an open set in* $\mathbf{X}$ *for each* $n < \infty$ *and* $\mathcal{X}_n \underset{\mathrm{H}}{\to} \mathcal{X}_\infty$ *in* $\mathbf{X}$ *as* $n \to \infty$.

Then the topology τ *of* $\mathbf{X}$ *is called a* Gromov–Hausdorff convergence *and we write* $\mathcal{X}_n \underset{\mathrm{GH}}{\longrightarrow} \mathcal{X}_\infty$ *or* $\mathcal{X}_n \xrightarrow{\tau} \mathcal{X}_\infty$*; the latter notation is used if we need to consider the specific Gromov–Hausdorff convergence* τ. *The space* $\mathcal{X}_\infty$ *is called the* limit space *of the sequence* $\mathcal{X}_n$ *along* τ.

When we write $\mathcal{X}_n \underset{\mathrm{GH}}{\longrightarrow} \mathcal{X}_\infty$ we mean that we made a choice of a Gromov–Hausdorff convergence.

Note that for a fixed sequence $\mathcal{X}_n$ of metric spaces, one may construct different Gromov–Hausdorff convergences, say $\mathcal{X}_n \xrightarrow{\tau} \mathcal{X}_\infty$ and $\mathcal{X}_n \xrightarrow{\tau'} \mathcal{X}'_\infty$, and their limit spaces $\mathcal{X}_\infty$ and $\mathcal{X}'_\infty$ need not be isometric to each other. For example, for the constant sequence $\mathcal{X}_n \overset{iso}{=\!=} \mathbb{R}_{\geqslant 0}$, one may take $\mathcal{X}_\infty \overset{iso}{=\!=} \mathbb{R}_{\geqslant 0}$. In this case, a point in the disjoint space $\mathbf{X}$ can be regarded as a pair $(x, n) \in \mathbb{R}_{\geqslant 0} \times (\mathbb{Z}_{>} \cup \{\infty\})$ and the metric on $\mathbf{X}$ can be defined by

$$|(x,n) - (y,m)|_{\mathbf{X}} \coloneqq |\tfrac{1}{n} + \tfrac{1}{m}| + |x - y|,$$

where we assume that $0 = \frac{1}{\infty}$. On the other hand, one can take $\mathcal{X}'_\infty \overset{iso}{=\!=} \mathbb{R}$, and consider the metric

$$\begin{aligned} |(x,n) - (y,m)|_{\mathbf{X}'} &= |\tfrac{1}{n} - \tfrac{1}{m}| + |(x-n) - (y-m)|, \\ |(x,n) - (y,\infty)|_{\mathbf{X}'} &= \tfrac{1}{n} + |(x-n) - y|, \\ |(x,\infty) - (y,\infty)|_{\mathbf{X}'} &= |x - y|, \end{aligned}$$

where $n, m < \infty$.

5.5. Induced convergences. *Suppose* $\mathcal{X}_n \xrightarrow{\tau} \mathcal{X}_\infty$ *as in Definition* 5.4, *and* $\iota_n : \mathcal{X}_n \hookrightarrow \mathbf{X}$, $\iota_\infty : \mathcal{X}_\infty \hookrightarrow \mathbf{X}$ *are the corresponding inclusions.*

(*a*) *A sequence of points* $x_n \in \mathcal{X}_n$ *converges to* $x_\infty \in \mathcal{X}_\infty$ *(briefly,* $x_n \to x_\infty$ *or* $x_n \xrightarrow{\tau} x_\infty$*) if* $|x_n - x_\infty|_{\mathbf{X}} \to 0$.

(*b*) *A sequence of closed sets* $\mathfrak{C}_n \subset \mathcal{X}_n$ *converges to a closed set* $\mathfrak{C}_\infty \subset \mathcal{X}_\infty$ *(briefly,* $\mathfrak{C}_n \to \mathfrak{C}_\infty$ *or* $\mathfrak{C}_n \xrightarrow{\tau} \mathfrak{C}_\infty$*) if* $\mathfrak{C}_n \underset{\mathrm{H}}{\to} \mathfrak{C}_\infty$ *as subsets of* $\mathbf{X}$.

(*c*) *A sequence of open sets* $\Omega_n \subset \mathcal{X}_n$ *converges to an open set* $\Omega_\infty \subset \mathcal{X}_\infty$ *(briefly,* $\Omega_n \to \Omega_\infty$ *or* $\Omega_n \xrightarrow{\tau} \Omega_\infty$*) if the complements* $\mathcal{X}_n \setminus \Omega_n$ *converge to the complement* $\mathcal{X}_\infty \setminus \Omega_\infty$ *as closed sets.*

(*d*) *Let* $\mathcal{X}_n \xrightarrow{\tau} \mathcal{X}_\infty$ *and* $\mathcal{Y}_n \xrightarrow{\theta} \mathcal{Y}_\infty$. *A sequence of submaps (where a submap is a map defined on a subset; see Section* 3A) $\Phi_n : \mathcal{X}_n \multimap \mathcal{Y}_n$ *converges to a submap* $\Phi_\infty : \mathcal{X}_\infty \multimap \mathcal{Y}_\infty$ *if the following conditions holds*
- $\operatorname{Dom}\Phi_n \to \operatorname{Dom}\Phi_\infty$ *as a sequence of open sets.*
- *for any* $x_\infty \in \operatorname{Dom}\Phi_\infty$ *and any sequence* $x_n \in \mathcal{X}_n$ *such that* $x_n \to x_\infty$, *we have*

$$\mathcal{Y}_n \ni \Phi_n(x_n) \xrightarrow{\theta} \Phi_\infty(x_\infty) \in \mathcal{Y}_\infty$$

as $n \to \infty$.

(*e*) *Given a sequence of measures* μ_n *on* $\mathcal{X}_n$, *we say that* μ_n *weakly converges to a measure* μ_∞ *on* $\mathcal{X}_\infty$ *(briefly,* $\mu_n \to \mu_\infty$ *or* $\mu_n \xrightarrow{\tau} \mu_\infty$*) if the pushforward measures of* μ_n *weakly converge to the pushforward measure of* μ_∞.

In other words, if for any continuous function $\phi : \mathbf{X} \to \mathbb{R}$ *with a compact support, we have*

$$\int\limits_{\mathcal{X}_n} \phi \circ \iota_n \cdot \mu_n \to \int\limits_{\mathcal{X}_\infty} \phi \circ \iota_\infty \cdot \mu_\infty$$

as $n \to \infty$.

Liftings. Given a Gromov–Hausdorff convergence $\mathcal{X}_n \xrightarrow[\mathrm{GH}]{} \mathcal{X}_\infty$ and a point $p_\infty \in \mathcal{X}_\infty$, any sequence of points $p_n \in \mathcal{X}_n$ such that $p_n \xrightarrow[\mathrm{GH}]{} p_\infty$ will be called a *lifting* of p_∞. The point $p_n \in \mathcal{X}_n$ will be called a *lifting* of p_∞ to $\mathcal{X}_n$. We will also say that $\operatorname{dist}_{p_n} : \mathcal{X}_n \to \mathbb{R}$ is a *lifting* of the distance function $\operatorname{dist}_{p_\infty} : \mathcal{X}_\infty \to \mathbb{R}$. Clearly $\operatorname{dist}_{p_n} \xrightarrow[\mathrm{GH}]{} \operatorname{dist}_{p_\infty}$.

Note that liftings are not uniquely defined.

Similarly, we may refer to liftings of a point array $\mathbf{p}_\infty = (p^1_\infty, p^2_\infty, \dots, p^k_\infty)$ and of the corresponding distance map $\operatorname{dist}_{\mathbf{p}_\infty} : \mathcal{X}_\infty \to \mathbb{R}^k$,

$$\operatorname{dist}_{\mathbf{p}_\infty} : x \mapsto \big(|p^1_\infty - x|, |p^2_\infty - x|, \dots, |p^k_\infty - x|\big).$$

C. Gromov's selection theorem

5.6. Gromov's selection theorem. *Let $\mathcal{X}_n$ be a sequence of proper metric spaces with marked points $x_n \in \mathcal{X}_n$. Assume that for any $R > 0$, $\varepsilon > 0$, there is $N = N(R, \varepsilon) \in \mathbb{Z}_{>0}$ such that for each n the ball $\overline{\mathrm{B}}[x_n, R] \subset \mathcal{X}_n$ admits a finite ε-net with at most N points. Then a subsequence of $\mathcal{X}_n$ admits a Gromov–Hausdorff convergence such that the sequence of marked points $x_n \in \mathcal{X}_n$ converges.*

Proof. By the main assumption, there is a sequence of integers $M_1 < M_2 < \dots$ such that in each space $\mathcal{X}_n$ there is a sequence of points $z_{i,n} \in \mathcal{X}_n$ for which

$$|z_{i,n} - x_n|_{\mathcal{X}_n} \leqslant k + 1 \quad \text{if} \quad i \leqslant M_k$$

and $\{z_{1,n}, \dots, z_{M_k,n}\}$ is a $\frac{1}{k}$-net in $\overline{\mathrm{B}}[x_n, k]_{\mathcal{X}_n}$.

Passing to a subsequence, we may assume that the sequence

$$\ell_n = |z_{i,n} - z_{j,n}|_{\mathcal{X}_n}$$

converges for any i and j.

Let us consider a countable set of points $\mathcal{W} = \{w_1, w_2, \dots\}$ equipped with the pseudometric defined by

$$|w_i - w_j|_{\mathcal{W}} = \lim_{n \to \infty} |z_{i,n} - z_{j,n}|_{\mathcal{X}_n}.$$

Let $\hat{\mathcal{W}}$ be the metric space corresponding to $\mathcal{W}$. Denote by $\mathcal{X}_\infty$ the completion of $\hat{\mathcal{W}}$.

It remains to construct a metric on the disjoint union of

$$\mathbf{X} = \mathcal{X}_\infty \sqcup \mathcal{X}_1 \sqcup \mathcal{X}_2 \sqcup \dots$$

satisfying definitions 5.3 and 5.4.

Such a metric can be defined as follows. Fix a sequence $\varepsilon_k \to 0+$ and let N_k be the minimal integer such that

$$|w_i - w_j|_{\mathcal{W}} \lessgtr |z_{i,n} - z_{j,n}|_{\mathcal{X}_n} \pm \varepsilon_k$$

if $i, j \leqslant N_k$ and $n \geqslant N_k$. Let us equip $\mathbf{X}$ with the maximal metric such that all the inclusions $\iota_n : \mathcal{X}_n \to \mathbf{X}$ and $\iota_\infty : \mathcal{X}_\infty \to \mathbf{X}$ are isometric and $|z_{i,n} - w_i| \leqslant \varepsilon_k$ for $i \leqslant N_k$ and $n \geqslant N_k$. It is easy to verify that such a metric on $\mathbf{X}$ satisfies 5.3 and 5.4. ☐

D. Convergence of compact spaces.

5.7. Definition. *Let $\mathcal{X}$ and $\mathcal{Y}$ be metric spaces. A map $f : \mathcal{X} \to \mathcal{Y}$ is called an* ε-isometry *if the following two conditions hold:*

(*a*) *$\Im f$ is an ε-net in $\mathcal{Y}$.*

(*b*) *$||f(x) - f(x')|_{\mathcal{Y}} - |x - x'|_{\mathcal{X}}| \leqslant \varepsilon$ for any $x, x' \in \mathcal{X}$.*

5.8. Lemma. *Let $\mathcal{X}_1, \mathcal{X}_2, \dots,$ and $\mathcal{X}_\infty$ be metric spaces and $\varepsilon_n \to 0+$ as $n \to \infty$. Suppose that either*

(a) *for each n there is an ε_n-isometry $f_n \colon \mathcal{X}_n \to \mathcal{X}_\infty$, or*

(b) *for each n there is an ε_n-isometry $h_n \colon \mathcal{X}_\infty \to \mathcal{X}_n$.*

Then there is a Gromov–Hausdorff convergence $\mathcal{X}_n \xrightarrow[\mathrm{GH}]{} \mathcal{X}_\infty$.

Proof. To prove part (5.8*a*) let us construct a common space $\mathbf{X}$ for the spaces $\mathcal{X}_1, \mathcal{X}_2, \dots,$ and $\mathcal{X}_\infty$ by taking the metric ρ on the disjoint union $\mathcal{X}_\infty \sqcup \mathcal{X}_1 \sqcup \mathcal{X}_2 \sqcup \dots$ that is defined the following way:

$$\begin{aligned}
\rho(x_n, y_n) &= |x_n - y_n|_{\mathcal{X}_n}, \quad \rho(x_\infty, y_\infty) = |x_\infty - y_\infty|_{\mathcal{X}_\infty},\\
\rho(x_n, x_\infty) &= \inf\{\, |x_n - y_n|_{\mathcal{X}_n} + \varepsilon_n + |x_\infty - f(y_n)|_{\mathcal{X}_\infty} : y_n \in \mathcal{X}_n \,\},\\
\rho(x_n, x_m) &= \inf\{\, \rho(x_n, y_\infty) + \rho(x_m, y_\infty) : y_\infty \in \mathcal{X}_\infty \,\},
\end{aligned}$$

where we assume that $x_m \in \mathcal{X}_m$, $x_n \in \mathcal{X}_n$, and $x_\infty \in \mathcal{X}_\infty$.

It remains to observe that ρ is indeed a metric and $\mathcal{X}_n \xrightarrow[\mathrm{H}]{} \mathcal{X}_\infty$ in $\mathbf{X}$.

The proof of the second part is analogous; one only needs to change one line in the definition of ρ to the following:

$$\rho(x_n, x_\infty) = \inf\{\, |x_n - h(y_\infty)|_{\mathcal{X}_n} + \varepsilon_n + |x_\infty - y_\infty|_{\mathcal{X}_\infty} : y_\infty \in \mathcal{X}_\infty \,\}.$$

□

5.9. Definition. *Given two compact spaces $\mathcal{X}$ and $\mathcal{Y}$, we will write*

- *$\mathcal{X} \leqslant \mathcal{Y}$ if there is a noncontracting map $\Phi \colon \mathcal{X} \to \mathcal{Y}$.*
- *$\mathcal{X} \leqslant \mathcal{Y} + \varepsilon$ if there is a map $\Phi \colon \mathcal{X} \to \mathcal{Y}$ such that for any $x, x' \in \mathcal{X}$ we have*

$$|x - x'| \leqslant |\Phi(x) - \Phi(x')| + \varepsilon.$$

5.10. Lemma. *Let $\mathcal{X}$ and $\mathcal{Y}$ be two metric spaces and $\mathcal{X}$ be compact. Then*

$$\mathcal{X} \geqslant \mathcal{Y} \geqslant \mathcal{X} \quad \Longleftrightarrow \quad \mathcal{X} \overset{iso}{=\!=} \mathcal{Y}.$$

The following proof was suggested by Travis Morrison.

Proof. Let $f \colon \mathcal{X} \to \mathcal{Y}$ and $g \colon \mathcal{Y} \to \mathcal{X}$ be noncontracting mappings. It is sufficient to prove that $h = g \circ f \colon \mathcal{X} \to \mathcal{X}$ is an isometry.

Given any pair of points $x, y \in \mathcal{X}$, let $x_n = h^{\circ n}(x)$ and $y_n = h^{\circ n}(y)$. Since $\mathcal{X}$ is compact, one can choose an increasing sequence of integers n_k such that both sequences x_{n_i} and y_{n_i} converge. In particular, both of these sequences are Cauchy, that is,

$$|x_{n_i} - x_{n_j}|, |y_{n_i} - y_{n_j}| \to 0$$

as $\min\{i, j\} \to \infty$. Since h is noncontracting, we have

$$|x - x_{|n_i - n_j|}| \leqslant |x_{n_i} - x_{n_j}|.$$

It follows that there is a sequence $m_i \to \infty$ such that

$$x_{m_i} \to x \quad \text{and} \quad y_{m_i} \to y \quad \text{as} \quad i \to \infty. \tag{1}$$

Let $\ell_n = |x_n - y_n|$. Since h is noncontracting, the sequence ℓ_n is nondecreasing. On the other hand, from 1 it follows that $\ell_{m_i} \to |x - y| = \ell_0$ as $m_i \to \infty$; that is, ℓ_n is a constant sequence. In particular, $\ell_0 = \ell_1$ for any x and y in $\mathcal{X}$, so h is a distance-preserving map.

Thus $h(\mathcal{X})$ is isometric to $\mathcal{X}$. From 1, $h(\mathcal{X})$ is everywhere dense. Since $\mathcal{X}$ is compact, $h(\mathcal{X}) = \mathcal{X}$. $\square$

The *Gromov–Hausdorff distance* between isometry classes of compact metric spaces $\mathcal{X}$ and $\mathcal{Y}$, is defined by

$$D_{\mathrm{GH}}(\mathcal{X}, \mathcal{Y}) := \inf\{\, \varepsilon > 0 : \mathcal{X} \leqslant \mathcal{Y} + \varepsilon \text{ and } \mathcal{Y} \leqslant \mathcal{X} + \varepsilon \,\}.$$

The Gromov–Hausdorff distance turns the set of all isometry classes of compact metric spaces into a metric space. The following theorem shows that convergence in this space coincides with the Gromov–Hausdorff convergence defined above.

5.11. Theorem. *Let $\mathcal{X}_1, \mathcal{X}_2, \dots,$ and $\mathcal{X}_\infty$ be compact metric spaces. Then there is a convergence $\mathcal{X}_n \xrightarrow[\mathrm{GH}]{} \mathcal{X}_\infty$ if and only if $D_{\mathrm{GH}}(\mathcal{X}_n, \mathcal{X}_\infty) \to 0$ as $n \to \infty$.*

Proof.

If part. Suppose $a_n : \mathcal{X}_\infty \to \mathcal{X}_n$ and $b_n : \mathcal{X}_n \to \mathcal{X}_\infty$ are sequences of maps such that

$$|a_n(x) - a_n(y)|_{\mathcal{X}_\infty} \geqslant |x - y|_{\mathcal{X}_n} - \delta_n,$$
$$|b_n(v) - b_n(w)|_{\mathcal{X}_n} \geqslant |v - w|_{\mathcal{X}_\infty} - \delta_n$$

for any $x, y \in \mathcal{X}_n$, $v, w \in \mathcal{X}_\infty$, and a sequence $\delta_n \to 0+$.

Fix $\varepsilon > 0$ and choose a maximal ε-packing $\{x^1, x^2, \dots, x^k\}$ in $\mathcal{X}_\infty$ such that $\sum_{i<j} |x^i - x^j|$ is maximal. Note that

$$|a_n \circ b_n(x^i) - a_n \circ b_n(x^j)| \geqslant |x^i - x^j| - 2 \cdot \delta_n.$$

Since $\sum_{i<j} |x^i - x^j|$ is maximal,

$$|a_n \circ b_n(x^i) - a_n \circ b_n(x^j)| \to |x^i - x^j|$$

for all i and j as $n \to \infty$. For all large n, we have $2 \cdot \delta_n < |x^i - x^j| - \varepsilon$, and so

$$|b_n(x^i) - b_n(x^j)|_{\mathcal{X}_n} > \varepsilon \quad \text{and} \quad |a_n \circ b_n(x^i) - a_n \circ b_n(x^j)|_{\mathcal{X}_\infty} > \varepsilon$$

for all $i \neq j$. Therefore for each large n, the set $\{a_n \circ b_n(x^i)\}$ is a maximal ε-packing and hence an ε-net in $\mathcal{X}_\infty$.

Since $\{a_n \circ b_n(x^i)\}$ is an ε-net in $\mathcal{X}_\infty$, we have that for any $y_n \in \mathcal{X}_n$ there is x^i such that $|a_n \circ b_n(x^i) - a_n(y_n)| < \varepsilon$. Thus $|b_n(x^i) - y_n| < \varepsilon + \delta_n$, that is, $\{b_n(x^i)\}$ is a $(\varepsilon + \delta_n)$-net in $\mathcal{X}_n$.

Given $y \in \mathcal{X}_n$, choose x^i so that $|b_n(x^i) - y_n| < \varepsilon + \delta_n$ and define $h_n(y) = a_n \circ b_n(x^i)$. Observe that h_n is a $3 \cdot \varepsilon$-isometry for all large n. Since $\varepsilon > 0$ is arbitrary, there is a sequence of ε_n-isometries $\mathcal{X}_n \to \mathcal{X}_\infty$ such that $\varepsilon_n \to 0+$ as $n \to \infty$. It remains to apply 5.8.

Only-if part. Assume $\mathcal{X}_n \xrightarrow{\tau} \mathcal{X}_\infty$. Fix $\varepsilon > 0$, and choose a maximal ε-packing $\{x^1, x^2, \dots, x^k\}$ in $\mathcal{X}_\infty$. For each x^i, choose a sequence $x_n^i \in \mathcal{X}_n$ such that $x_n^i \to x^i$. Define a map $a_n : \mathcal{X}_n \to \mathcal{X}_\infty$ such that $a_n(x_n^i) = x_n$. Note that for all large n, we have $|x_n^i - x_n^j| > \varepsilon$. For each point $z \in \mathcal{X}_\infty$, choose x^i so that $|z - x^i| < \varepsilon$. Define a map $b_n : \mathcal{X}_\infty \to \mathcal{X}_n$ by setting $b_n(z) = x_n^i$. Observe that

$$|b_n(y) - b_n(z)|_{\mathcal{X}_n} + 3 \cdot \varepsilon > |y - z|_{\mathcal{X}_\infty}$$

for all large n.

In the same way we can construct a map $a_n : \mathcal{X}_n \to \mathcal{X}_\infty$ such that

$$|a_n(y) - a_n(z)|_{\mathcal{X}_\infty} + 3 \cdot \varepsilon > |y - z|_{\mathcal{X}_n}.$$

Hence $D_{GH}(\mathcal{X}_n, \mathcal{X}_\infty) \to 0$ as $n \to \infty$. ☐

The following theorem states that the isometry class of a Gromov–Hausdorff limit is uniquely defined if it is compact.

5.12. Theorem. *Let $\mathcal{X}_1$, $\mathcal{X}_2$, …, and $\mathcal{X}_\infty$ and $\bar{\mathcal{X}}_\infty$ be metric spaces such that $\mathcal{X}_n \xrightarrow{\tau} \mathcal{X}_\infty$, $\mathcal{X}_n \xrightarrow{\bar\tau} \bar{\mathcal{X}}_\infty$.*

Assume that $\bar{\mathcal{X}}_\infty$ is compact. Then $\mathcal{X}_\infty \overset{iso}{=\!=} \bar{\mathcal{X}}_\infty$.

Proof. For each point $x_\infty \in \mathcal{X}_\infty$, choose liftings $x_n \in \mathcal{X}_n$.

Choose a nonprincipal ultrafilter ω on $\mathbb{N}$. Define $\bar{x}_\infty \in \bar{\mathcal{X}}_\infty$ as the ω-limit of x_n with respect to $\bar{\tau}$. We claim that the map $x_\infty \to \bar{x}_\infty$ is an isometry.

Indeed, by the definition of Gromov–Hausdorff convergence,

$$|\bar{x}_\infty - \bar{y}_\infty|_{\bar{\mathcal{X}}_\infty} = \lim_{n\to\omega} |x_n - y_n|_{\mathcal{X}_n} = |x_\infty - y_\infty|_{\mathcal{X}_\infty}.$$

Thus the map $x_\infty \to \bar{x}_\infty$ gives a distance-preserving map $\Phi : \mathcal{X}_\infty \hookrightarrow \bar{\mathcal{X}}_\infty$. In particular, $\mathcal{X}_\infty$ is compact. Switching $\mathcal{X}_\infty$ and $\bar{\mathcal{X}}_\infty$ and applying the same argument, we get an isometric embedding $\bar{\mathcal{X}}_\infty \hookrightarrow \mathcal{X}_\infty$. Now the result follows from Lemma 5.10. ☐

5.13. Exercise. (a) *Show that a sequence of compact simply connected length spaces cannot converge to a circle.*

(b) *Construct a sequence of compact simply connected length spaces that converges to a compact non-simply connected space.*

5.14. Exercise. (*a*) *Show that a sequence of length metrics on the 2-sphere cannot converge to the unit disk.*

(*b*) *Construct a sequence of length metrics on the 3-sphere that converges to a unit 3-ball.*

5.15. Exercise. *Let $\mathcal{X}_n$ be a sequence of metric spaces that admits two Gromov–Hausdorff convergences τ and τ'. Assume $\mathcal{X}_n \xrightarrow{\tau} \mathcal{X}_\infty$ and $\mathcal{X}_n \xrightarrow{\tau'} \mathcal{X}'_\infty$. Show that if $\mathcal{X}_\infty$ is proper and there is a sequence of points $x_n \in \mathcal{X}_n$ that converges in both τ and τ', then $\mathcal{X}_\infty \overset{iso}{=\!=} \mathcal{X}'_\infty$.*

E. Ultralimits revisited

Recall that ω denotes an ultrafilter of the set of natural numbers.

5.16. Theorem. *Assume $\mathcal{X}_n$ is a sequence of complete metric spaces. Let $\mathcal{X}_n \to \mathcal{X}_\omega$ as $n \to \omega$, and let $\mathcal{Y}_n$ be a sequence of subspaces of ~~the~~ $\mathcal{X}_n$ such that $\mathcal{Y}_n \xrightarrow[\text{GH}]{} \mathcal{Y}_\infty$. Then there is a distance-preserving map $\iota : \mathcal{Y}_\infty \to \mathcal{X}_\omega$.*

Moreover:

(*a*) *If $\mathcal{X}_n \xrightarrow[\text{GH}]{} \mathcal{X}_\infty$ and $\mathcal{X}_\infty$ is compact, then $\mathcal{X}_\infty$ is isometric to $\mathcal{X}_\omega$.*

(*b*) *If $\mathcal{X}_n \xrightarrow[\text{GH}]{} \mathcal{X}_\infty$ and $\mathcal{X}_\infty$ is proper, then $\mathcal{X}_\infty$ is isometric to a metric component of $\mathcal{X}_\omega$.*

Proof. For each point $y_\infty \in \mathcal{Y}_\infty$ choose a lifting $y_n \in \mathcal{Y}_n$. Pass to the ω-limit $y_\omega \in \mathcal{X}_\omega$ of y_n. Clearly for any $y_\infty, z_\infty \in \mathcal{Y}_\infty$, we have

$$|y_\infty - z_\infty|_{\mathcal{Y}_\infty} = |y_\omega - z_\omega|_{\mathcal{X}_\omega};$$

that is, the map $y_\infty \mapsto y_\omega$ gives a distance-preserving map $\iota : \mathcal{Y}_\infty \to \mathcal{X}_\omega$.

(*a*) + (*b*). Fix $x_\omega \in \mathcal{X}_\omega$. Choose a sequence x_n of points in $\mathcal{X}_n$, such that $x_n \to x_\omega$ as $n \to \omega$.

Denote by $\mathbf{X} = \mathcal{X}_\infty \sqcup \mathcal{X}_1 \sqcup \mathcal{X}_2 \sqcup \dots$ the common space for the convergence $\mathcal{X}_n \xrightarrow[\text{GH}]{} \mathcal{X}_\infty$, as in the definition of Gromov–Hausdorff convergence. Note that x_n is a sequence of points in $\mathbf{X}$.

If the ω-limit x_∞ of x_n in $\mathbf{X}$ exists, it must lie in $\mathcal{X}_\infty$.

The point x_∞, if defined, does not depend on the choice of x_n. Indeed, if $y_n \in \mathcal{X}_n$ is another sequence such that $y_n \to x_\omega$ as $n \to \omega$, then

$$|y_\infty - x_\infty| = \lim_{n \to \omega} |y_n - x_n| = 0;$$

therefore, $x_\infty = y_\infty$.

This way we obtain a map $\nu : x_\omega \to x_\infty$, defined on $\operatorname{Dom} \nu \subset \mathcal{X}_\omega$. By construction of ι, we have $\iota \circ \nu(x_\omega) = x_\omega$ for any $x_\omega \in \operatorname{Dom} \nu$.

Finally note that if $\mathcal{X}_\infty$ is compact, then ν is defined on all of $\mathcal{X}_\omega$; this proves (a).

If $\mathcal{X}_\infty$ is proper, choose any point $z_\infty \in \mathcal{X}_\infty$ and set $z_\omega = \iota(z_\infty)$. For any point $x_\omega \in \mathcal{X}_\omega$ at finite distance from z_ω, for the sequence x_n as above we have that $|z_n - x_n|$ is bounded for ω-almost all n. Since $\mathcal{X}_\infty$ is proper, $\nu(x_\omega)$ is defined; in other words, ν is defined on the metric component of z_ω. Hence (b) follows. $\square$

Chapter 6

The ghost of Euclid

A. Geodesics, triangles and hinges

Geodesics and their relatives. Let $\mathcal{X}$ be a metric space and $\mathbb{I} \subset \mathbb{R}$ be an interval. A globally distance-preserving map $\gamma\colon \mathbb{I} \to \mathcal{X}$ is called a *unit-speed geodesic*. (Various authors call it differently: *shortest path*, *minimizing geodesic*.) In other words, $\gamma\colon \mathbb{I} \to \mathcal{X}$ is a unit-speed geodesic if the equality

$$|\gamma(s) - \gamma(t)|_{\mathcal{X}} = |s - t|$$

holds for any pair $s, t \in \mathbb{I}$.

A unit-speed geodesic between p and q in $\mathcal{X}$ will be denoted by $\mathrm{geod}_{[pq]}$. We will always assume $\mathrm{geod}_{[pq]}$ is parametrized starting at p; that is, $\mathrm{geod}_{[pq]}(0) = p$ and $\mathrm{geod}_{[pq]}(|p - q|) = q$. The image of $\mathrm{geod}_{[pq]}$ will be denoted by $[pq]$ and called a *geodesic*. The term *geodesic* will also be used for a linear reparametrization of a unit-speed geodesic. With a slight abuse of notation, we will use the notation $[pq]$ also for the class of all linear reparametrizations of $\mathrm{geod}_{[pq]}$.

A unit-speed geodesic $\gamma\colon \mathbb{R}_{\geqslant 0} \to \mathcal{X}$ is called a *half-line*.

A unit-speed geodesic $\gamma\colon \mathbb{R} \to \mathcal{X}$ is called a *line*.

A piecewise geodesic curve is called *polygonal line*; we may say *polygonal line* $p_1, \dots, p_n$ meaning polygonal line with edges $[p_1p_2], \dots, [p_{n-1}p_n]$. A closed polygonal line will be also called a polygon.

We may write $[pq]_{\mathcal{X}}$ to emphasize that the geodesic $[pq]$ is in the space $\mathcal{X}$. Also, we use the following short-cut notation:

$$]pq[= [pq] \setminus \{p, q\}, \qquad]pq] = [pq] \setminus \{p\}, \qquad [pq[= [pq] \setminus \{q\}.$$

In general, a geodesic between p and q need not exist and if it exists, it need not ~~to~~ be unique. However, once we write $\mathrm{geod}_{[pq]}$ or $[pq]$ we mean that we have fixed a choice of a geodesic.

A constant-speed geodesic $\gamma\colon [0,1] \to \mathcal{X}$ is called a *geodesic path*. Given a geodesic $[pq]$, we denote by $\mathrm{path}_{[pq]}$ the corresponding geodesic path; that is,

$$\mathrm{path}_{[pq]}(t) \equiv \mathrm{geod}_{[pq]}(t \cdot |p-q|).$$

A curve $\gamma\colon \mathbb{I} \to \mathcal{X}$ is called a *local geodesic* if for any $t \in \mathbb{I}$ there is a neighborhood $U \ni t$ in $\mathbb{I}$ such that the restriction $\gamma|_U$ is a constant-speed geodesic. If $\mathbb{I} = [0,1]$, then γ is called a *local geodesic path*.

6.1. Proposition. *Suppose $\mathcal{X}$ is a metric space and $\gamma\colon [0,\infty) \to \mathcal{X}$ is a half-line. Then the* Busemann function $\mathrm{bus}_\gamma\colon \mathcal{X} \to \mathbb{R}$

$$\mathrm{bus}_\gamma(x) = \lim_{t\to\infty} |\gamma(t) - x| - t \tag{1}$$

is defined and 1-Lipschitz.

Proof. By the triangle inequality, the function $t \mapsto |\gamma(t) - x| - t$ is nonincreasing. Clearly $|\gamma(t) - x| - t \geqslant -|\gamma(0) - x|$. Thus the limit in 1 is defined, and it is 1-Lipschitz as a limit of 1-Lipschitz functions. □

6.2. Example. *If $\mathcal{X}$ is a Euclidean space and $\gamma(t) = p + t\cdot v$ where v is a unit vector, then*

$$\mathrm{bus}_\gamma(x) = \langle x - p, v\rangle.$$

Triangles. For a triple of points $p, q, r \in \mathcal{X}$, a choice of a triple of geodesics $([qr],[rp],[pq])$ will be called a *triangle*, and we will use the short notation $[pqr] = ([qr],[rp],[pq])$. Again, given a triple $p, q, r \in \mathcal{X}$, there may be no triangle $[pqr]$, simply because one of the pairs of these points cannot be joined by a geodesic. Or there may be many different triangles, any of which can be denoted by $[pqr]$. Once we write $[pqr]$, it means we have chosen such a triangle; that is, made a choice of each $[qr]$, $[rp]$, and $[pq]$.

The value $|p-q|+|q-r|+|r-p|$ will be called the *perimeter of triangle* $[pqr]$; it obviously coincides with perimeter of the triple p, q, r as defined below.

Hinges. Let $p, x, y \in \mathcal{X}$ be a triple of points such that p is distinct from x and y. A pair of geodesics $([px],[py])$ will be called a *hinge*, and will be denoted by $\left[p\,{}^{x}_{y}\right] = ([px],[py])$.

B. Model angles and triangles

Let $\mathcal{X}$ be a metric space, $p, q, r \in \mathcal{X}$, and $\kappa \in \mathbb{R}$. Let us define the *model triangle* $[\tilde p\tilde q\tilde r]$ (briefly, $[\tilde p\tilde q\tilde r] = \tilde\triangle^{\kappa}(pqr)$) to be a triangle in the model plane $\mathbb{M}^2(\kappa)$ such

that
$$|\tilde p-\tilde q|=|p-q|,\quad |\tilde q-\tilde r|=|q-r|,\quad |\tilde r-\tilde p|=|r-p|.$$

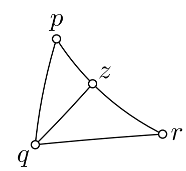
In the notation of Section 1A, $\tilde\triangle^{\kappa}(pqr)=\tilde\triangle^{\kappa}\{|q-r|,|r-p|,|p-q|\}$.

If $\kappa\leqslant 0$, the model triangle is always defined, that is, it exists and is unique up to an isometry of $\mathbb{M}^2(\kappa)$. If $\kappa>0$, the model triangle is said to be defined if in addition
$$|p-q|+|q-r|+|r-p|<2\cdot\varpi\kappa;$$
here $\varpi\kappa$ denotes the diameter of the model space $\mathbb{M}^2(\kappa)$. In this case, the model triangle also exists and is unique up to an isometry of $\mathbb{M}^2(\kappa)$. The value $|p-q|+|q-r|+|r-p|$ will be called the *perimeter of the triple* p,q,r.

If for $p,q,r\in\mathcal{X}$, $[\tilde p\tilde q\tilde r]=\tilde\triangle^{\kappa}(pqr)$ is defined and $|p-q|,|p-r|>0$, the angle measure of $[\tilde p\tilde q\tilde r]$ at $\tilde p$ will be called the *model angle* of the triple p,q,r, and will be denoted by $\tilde\measuredangle^{\kappa}\left(p\,{}^{q}_{r}\right)$.

In the notation of Section 1A,
$$\tilde\measuredangle^{\kappa}\left(p\,{}^{q}_{r}\right)=\tilde\measuredangle^{\kappa}\{|q-r|;|p-q|,|p-r|\}.$$

6.3. Alexandrov's lemma. *Let p,q,r,z be distinct points in a metric space such that $z\in\,]pr[$ and*
$$|p-q|+|q-r|+|r-p|<2\cdot\varpi\kappa.$$



Then the following expressions have the same sign:

(a) $\tilde\measuredangle^{\kappa}\left(p\,{}^{q}_{r}\right)-\tilde\measuredangle^{\kappa}\left(p\,{}^{q}_{z}\right)$,

(b) $\tilde\measuredangle^{\kappa}\left(z\,{}^{q}_{p}\right)+\tilde\measuredangle^{\kappa}\left(z\,{}^{q}_{r}\right)-\pi$.

Moreover,
$$\tilde\measuredangle^{\kappa}\left(q\,{}^{p}_{r}\right)\geqslant\tilde\measuredangle^{\kappa}\left(q\,{}^{p}_{z}\right)+\tilde\measuredangle^{\kappa}\left(q\,{}^{z}_{r}\right),$$
with equality if and only if the expressions in (a) *and* (b) *vanish.*

Proof. By the triangle inequality,
$$|p-q|+|q-z|+|z-p|\leqslant|p-q|+|q-r|+|r-p|<2\cdot\varpi\kappa.$$
Therefore the model triangle $[\tilde p\tilde q\tilde z]=\tilde\triangle^{\kappa}pqz$ is defined. Take a point $\tilde r$ on the extension of $[\tilde p\tilde z]$ beyond $\tilde z$ so that $|\tilde p-\tilde r|=|p-r|$ (and therefore $|\tilde p-\tilde z|=|p-z|$).

From monotonicity of the function $a\mapsto\tilde\measuredangle^{\kappa}\{a;b,c\}$ (1.1c), the following expressions have the same sign:

(i) $\measuredangle\left[\tilde p\,{}^{\tilde q}_{\tilde r}\right]-\tilde\measuredangle^{\kappa}\left(p\,{}^{q}_{r}\right)$;

(ii) $|\tilde p-\tilde r|-|p-r|$;

(iii) $\measuredangle\left[\tilde z\,{}^{\tilde q}_{\tilde r}\right]-\tilde\measuredangle^{\kappa}\left(z\,{}^{q}_{r}\right)$.

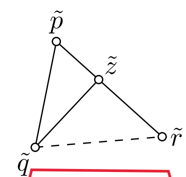
Since

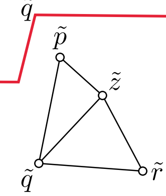
$$\measuredangle\left[\tilde p\,{}^{\tilde q}_{\tilde r}\right]=\measuredangle\left[\tilde p\,{}^{\tilde q}_{\tilde z}\right]=\tilde\measuredangle^\kappa\left(p\,{}^{q}_{z}\right)$$

and

$$\measuredangle\left[\tilde z\,{}^{\tilde q}_{\tilde r}\right]=\pi-\measuredangle\left[\tilde z\,{}^{\tilde p}_{\tilde q}\right]=\pi-\tilde\measuredangle^\kappa\left(z\,{}^{p}_{q}\right),$$

the first statement follows.





For the second statement, let us redifine $\tilde r$; construct $[\tilde q\tilde z\tilde r]=\tilde\triangle^\kappa qzr$ on the opposite side of $[\tilde q\tilde z]$ from $[\tilde p\tilde q\tilde z]$. Since

$$\begin{aligned}|\tilde p-\tilde r|&\leqslant|\tilde p-\tilde z|+|\tilde z-\tilde r|\\&=|p-z|+|z-r|\\&=|p-r|,\end{aligned}$$

we have

$$\begin{aligned}\tilde\measuredangle^\kappa\left(q\,{}^{p}_{z}\right)+\tilde\measuredangle^\kappa\left(q\,{}^{z}_{r}\right)&=\measuredangle\left[\tilde q\,{}^{\tilde p}_{\tilde z}\right]+\measuredangle\left[\tilde q\,{}^{\tilde z}_{\tilde r}\right]\\&=\measuredangle\left[\tilde q\,{}^{\tilde p}_{\tilde r}\right]\\&\leqslant\tilde\measuredangle^\kappa\left(q\,{}^{p}_{r}\right).\end{aligned}$$

Equality holds if and only if $|\tilde p-\tilde r|=|p-r|$, as required. □

C. Angles and the first variation

Given a hinge $\left[p\,{}^{x}_{y}\right]$, we define its *angle* to be

(1) $$\measuredangle\left[p\,{}^{x}_{y}\right]:=\lim_{\bar x,\bar y\to p}\tilde\measuredangle^\kappa\left(p\,{}^{\bar x}_{\bar y}\right),$$

for $\bar x\in\,]px]$ and $\bar y\in\,]py]$, if this limit exists.

Similarly to $\tilde\measuredangle^\kappa\left(p\,{}^{x}_{y}\right)$, we will use the short notation

$$\tilde\curlyvee^\kappa\left[p\,{}^{x}_{y}\right]=\tilde\curlyvee^\kappa\left\{\measuredangle\left[p\,{}^{x}_{y}\right];|p-x|,|p-y|\right\},$$

where the right-hand side is defined in Section 1A. The value $\tilde\curlyvee^\kappa\left[p\,{}^{x}_{y}\right]$ will be called the *model side* of the hinge $\left[p\,{}^{x}_{y}\right]$.

6.4. Lemma. *Let p, x, y be a triple of points in a metric space with perimeter ℓ. Then for any $\kappa,\mathrm{K}\in\mathbb{R}$,*

(2) $$\left|\tilde\measuredangle^{\mathrm{K}}\left(p\,{}^{x}_{y}\right)-\tilde\measuredangle^\kappa\left(p\,{}^{x}_{y}\right)\right|\leqslant100(|\mathrm{K}|+|\kappa|)\cdot\ell^2,$$

whenever the left-hand side is defined.

Lemma 6.4 implies that the definition of angle is independent of κ. In particular, one can take $\kappa=0$ in 1; thus the angle can be calculated from the cosine law:

$$\cos\tilde\measuredangle^0\left(p\,{}^{x}_{y}\right)=\frac{|p-x|^2+|p-y|^2-|x-y|^2}{2\cdot|p-x|\cdot|p-y|}.$$

Proof. The function $\kappa \mapsto \tilde\measuredangle^\kappa\left(p\,{}^x_y\right)$ is nondecreasing (1.1d). Thus, for $K > \kappa$, we have

$$\begin{aligned}0 \leqslant \tilde\measuredangle^K\left(p\,{}^x_y\right) - \tilde\measuredangle^\kappa\left(p\,{}^x_y\right) &\leqslant \tilde\measuredangle^K\left(p\,{}^x_y\right) + \tilde\measuredangle^K\left(x\,{}^p_y\right) + \tilde\measuredangle^K\left(y\,{}^p_x\right) \\ &\quad - \tilde\measuredangle^\kappa\left(p\,{}^x_y\right) - \tilde\measuredangle^\kappa\left(x\,{}^p_y\right) - \tilde\measuredangle^\kappa\left(y\,{}^p_x\right) \\ &= K \cdot \operatorname{area} \tilde\triangle^K(pxy) - \kappa \cdot \operatorname{area} \tilde\triangle^\kappa(pxy).\end{aligned}$$

Note that for $\kappa \geqslant 0$ a triangle of perimeter ℓ in $\mathbb{M}^2(\kappa)$ lies in a ball of radius $2 \cdot \ell$, which easily implies that $\operatorname{area} \tilde\triangle^\kappa(pxy) \leqslant 100 \cdot \ell^2$. For $\kappa < 0$ one gets the same estimate by a direct computation in the hyperbolic plane.

Therefore

$$\operatorname{area} \tilde\triangle^\kappa(pxy) \leqslant 100 \cdot \ell^2, \qquad \operatorname{area} \tilde\triangle^K(pxy) \leqslant 100 \cdot \ell^2.$$

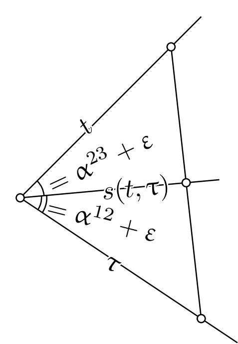
Thus 2 follows. □

6.5. Triangle inequality for angles. *Let $[px^1]$, $[px^2]$, and $[px^3]$ be three geodesics in a metric space. If all of the angles $\alpha^{ij} = \measuredangle\left[p\,{}^{x^i}_{x^j}\right]$ are defined then they satisfy the triangle inequality:*

$$\alpha^{13} \leqslant \alpha^{12} + \alpha^{23}.$$

Proof. Since $\alpha^{13} \leqslant \pi$, we can assume that $\alpha^{12} + \alpha^{23} < \pi$. Set $\gamma^i = \mathrm{geod}_{[px^i]}$. Given any $\varepsilon > 0$, for all sufficiently small $t, \tau, s \in \mathbb{R}_{\geqslant 0}$ we have

$$\begin{aligned}|\gamma^1(t) - \gamma^3(\tau)| &\leqslant |\gamma^1(t) - \gamma^2(s)| + |\gamma^2(s) - \gamma^3(\tau)| \\ &< \sqrt{t^2 + s^2 - 2 \cdot t \cdot s \cdot \cos(\alpha^{12} + \varepsilon)} \\ &\quad + \sqrt{s^2 + \tau^2 - 2 \cdot s \cdot \tau \cdot \cos(\alpha^{23} + \varepsilon)} \leqslant\end{aligned}$$



Below we define $s(t, \tau)$ so that for $s = s(t, \tau)$, this chain of inequalities can be continued as follows:

$$\leqslant \sqrt{t^2 + \tau^2 - 2 \cdot t \cdot \tau \cdot \cos(\alpha^{12} + \alpha^{23} + 2 \cdot \varepsilon)}.$$

Thus for any $\varepsilon > 0$,

$$\alpha^{13} \leqslant \alpha^{12} + \alpha^{23} + 2 \cdot \varepsilon.$$

Hence the result follows.

To define $s(t, \tau)$, consider three half-lines $\tilde\gamma^1$, $\tilde\gamma^2$, $\tilde\gamma^3$ on a Euclidean plane starting at one point, such that $\measuredangle(\tilde\gamma^1, \tilde\gamma^2) = \alpha^{12} + \varepsilon$, $\measuredangle(\tilde\gamma^2, \tilde\gamma^3) = \alpha^{23} + \varepsilon$, and $\measuredangle(\tilde\gamma^1, \tilde\gamma^3) = \alpha^{12} + \alpha^{23} + 2 \cdot \varepsilon$. We parametrize each half-line by the distance from the starting point. Given two positive numbers $t, \tau \in \mathbb{R}_{\geqslant 0}$, let $s = s(t, \tau)$ be the number such that $\tilde\gamma^2(s) \in [\tilde\gamma^1(t)\,\tilde\gamma^3(\tau)]$. Clearly $s \leqslant \max\{t, \tau\}$, so t, τ, s may be taken sufficiently small. □

6.6. Exercise. *Prove that the sum of adjacent angles is at least π.*

More precisely, suppose that the hinges $[p\,{}^x_z]$ and $[p\,{}^y_z]$ are adjacent*; that is, the side $[pz]$ is shared and the union of two sides $[px]$ and $[py]$ is a geodesic $[xy]$. Then*

$$\measuredangle[p\,{}^x_z] + \measuredangle[p\,{}^y_z] \geqslant \pi$$

whenever each angle on the left-hand side is defined.

The above inequality can be strict. For example in a metric tree angles between any two different edges coming out of the same vertex are all equal to π.

6.7. First variation inequality. *Assume that for a hinge $[q\,{}^p_x]$, the angle $\alpha = \measuredangle[q\,{}^p_x]$ is defined. Then*

$$|p - \mathrm{geod}_{[qx]}(t)| \leqslant |q - p| - t\cdot\cos\alpha + o(t).$$

Proof. Take a sufficiently small $\varepsilon > 0$. For all sufficiently small $t > 0$, we have

$$\begin{aligned}|\mathrm{geod}_{[qp]}(t/\varepsilon) - \mathrm{geod}_{[qx]}(t)| &\leqslant \tfrac{t}{\varepsilon}\cdot\sqrt{1+\varepsilon^2 - 2\cdot\varepsilon\cdot\cos\alpha} + o(t)\\ &\leqslant \tfrac{t}{\varepsilon} - t\cdot\cos\alpha + t\cdot\varepsilon.\end{aligned}$$

Applying the triangle inequality, we get

$$\begin{aligned}|p - \mathrm{geod}_{[qx]}(t)| &\leqslant |p - \mathrm{geod}_{[qp]}(t/\varepsilon)| + |\mathrm{geod}_{[qp]}(t/\varepsilon) - \mathrm{geod}_{[qx]}(t)|\\ &\leqslant |p-q| - t\cdot\cos\alpha + t\cdot\varepsilon\end{aligned}$$

for any $\varepsilon > 0$ and all sufficiently small t. Hence the result. □

D. Space of directions

Let $\mathcal{X}$ be a metric space. If the angle $\measuredangle[p\,{}^x_y]$ is defined for any hinge $[p\,{}^x_y]$ in $\mathcal{X}$, then we will say that the space $\mathcal{X}$ has *defined angles*.

Let us note that this is a strong condition. For example a Banach space which has defined angles must be Hilbert.

Let $\mathcal{X}$ be a space with defined angles. For $p \in \mathcal{X}$, consider the set $\mathfrak{S}_p$ of all nontrivial unit-speed geodesics starting at p. By 6.5, the triangle inequality holds for $\measuredangle$ on $\mathfrak{S}_p$, that is, $(\mathfrak{S}_p, \measuredangle)$ forms a pseudometric space.

The metric space corresponding to $(\mathfrak{S}_p, \measuredangle)$ is called the *space of geodesic directions* at p, denoted by Σ'_p or $\Sigma'_p\mathcal{X}$. The elements of Σ'_p are called *geodesic directions* at p. Each geodesic direction is formed by an equivalence class of geodesics starting from p for the equivalence relation

$$[px] \sim [py] \quad\Longleftrightarrow\quad \measuredangle[p\,{}^x_y] = 0;$$

the direction of $[px]$ is denoted by $\uparrow_{[px]}$.

The completion of Σ'_p is called the *space of directions* at p and is denoted by Σ_p or $\Sigma_p\mathcal{X}$. The elements of Σ_p are called *directions* at p.

E. Tangent space

The *Euclidean cone* $\mathcal{Y} = \operatorname{Cone}\mathcal{X}$ over a metric space $\mathcal{X}$ is defined as the metric space whose underlying set consists of equivalence classes in $[0,\infty)\times\mathcal{X}$ with the equivalence relation "$\sim$" given by $(0,p)\sim(0,q)$ for any points $p, q\in\mathcal{X}$, and whose metric is given by the cosine rule

$$|(s,p)-(t,q)|_{\mathcal{Y}} = \sqrt{s^2+t^2-2\cdot s\cdot t\cdot\cos\theta},$$

where $\theta = \min\{\pi, |p-q|_{\mathcal{X}}\}$. The point in $\mathcal{Y}$ that corresponds $(t,x)\in[0,\infty)\times\mathcal{X}$ will be denoted by $t\cdot x$.

The point in $\operatorname{Cone}\mathcal{X}$ formed by the equivalence class of $\{0\}\times\mathcal{X}$ is called the *tip of the cone* and is denoted by 0 or $0_{\mathcal{Y}}$. For $v\in\mathcal{Y}$ the distance $|0-v|_{\mathcal{Y}}$ is called the norm of v and is denoted by $|v|$ or $|v|_{\mathcal{Y}}$.

The *scalar product* $\langle v,w\rangle$ of two vectors $v = s\cdot p$ and $w = t\cdot q$ is defined by

$$\langle v,w\rangle \coloneqq |v|\cdot|w|\cdot\cos\theta;$$

we set $\langle v,w\rangle \coloneqq 0$ if $v=0$ or $w=0$.

6.8. Example. $\operatorname{Cone}\mathbb{S}^n$ *is isometric to* $\mathbb{R}^{n+1}$. *If* $G < \mathrm{O}(n+1)$ *is a closed subgroup, then* $\operatorname{Cone}(\mathbb{S}^n/G)$ *is isometric to* $\mathbb{R}^{n+1}/G$.

The Euclidean cone $\operatorname{Cone}\Sigma_p$ over the space of directions Σ_p is called the *tangent space* at p and denoted by T_p or $\mathrm{T}_p\mathcal{X}$. The elements of $\mathrm{T}_p\mathcal{X}$ will be called *tangent vectors* at p (despite the fact that T_p is only a cone—not a vector space).

The tangent space T_p could be also defined directly, without introducing the space of directions. To do so, consider the set $\mathfrak{T}_p$ of all geodesics starting at p, with arbitrary speed. Given $\alpha,\beta\in\mathfrak{T}_p$, set

$$(1)\qquad |\alpha-\beta|_{\mathfrak{T}_p} = \lim_{\varepsilon\to 0}\frac{|\alpha(\varepsilon)-\beta(\varepsilon)|_{\mathcal{X}}}{\varepsilon}.$$

If the angles in $\mathcal{X}$ are defined, then so is the limit in 1, and we obtain a pseudometric on $\mathfrak{T}_p$.

The corresponding metric space admits a natural isometric identification with the cone $\mathrm{T}'_p = \operatorname{Cone}\Sigma'_p$. The vectors of T'_p are the equivalence classes for the relation

$$\alpha\sim\beta \quad\iff\quad |\alpha(t)-\beta(t)|_{\mathcal{X}} = o(t).$$

The completion of T'_p is therefore naturally isometric to T_p. A vector in T'_p that corresponds to the geodesic path $\operatorname{geod}_{[pq]}$ is called *logarithm of* $[pq]$ and denoted by $\log[pq]$.

F. Velocity of curves

6.9. Definition. *Let $\mathcal{X}$ be a metric space, $a > 0$, and $\alpha\colon [0,a) \to \mathcal{X}$ be a function, not necessarily continuous, such that $\alpha(0) = p$. We say that $v \in \mathrm{T}_p$ is the* right derivative *of α at 0, briefly $\alpha^+(0) = v$, if for some (and therefore any) sequence of vectors $v_n \in \mathrm{T}'_p$ such that $v_n \to v$ as $n \to \infty$, and corresponding geodesics γ_n, we have*

$$\varlimsup_{\varepsilon\to 0+} \frac{|\alpha(\varepsilon) - \gamma_n(\varepsilon)|_{\mathcal{X}}}{\varepsilon} \to 0 \quad \text{as} \quad n \to \infty.$$

We define right and left derivatives $\alpha^+(t_0)$ and $\alpha^-(t_0)$ of α at $t_0 \in \mathbb{I}$ by

$$\alpha^{\pm}(t_0) = \check{\alpha}^+(0),$$

where $\check{\alpha}(t) = \alpha(t_0 \pm t)$.

The sign convention is not quite standard; if α is a smooth curve in a Riemannian manifold, then we have $\alpha^+(t) = -\alpha^-(t)$.

Note that if γ is a geodesic starting at p and the tangent vector $v \in \mathrm{T}'_p$ corresponds to γ, then $\gamma^+(0) = v$.

6.10. Exercise. *Assume $\mathcal{X}$ is a metric space with defined angles, and let $\alpha, \beta\colon [0,a) \to \mathcal{X}$ be two maps such that the right derivatives $\alpha^+(0)$, $\beta^+(0)$ are defined and $\alpha^+(0) = \beta^+(0)$. Show that*

$$|\alpha(t) - \beta(t)|_{\mathcal{X}} = o(t).$$

6.11. Proposition. *Let $\mathcal{X}$ be a metric space with defined angles and $p \in \mathcal{X}$. Then for any tangent vector $v \in \mathrm{T}_p\mathcal{X}$ there is a map $\alpha\colon [0,\varepsilon) \to \mathcal{X}$ such that $\alpha^+(0) = v$.*

Proof. If $v \in \mathrm{T}'_p$, then for the corresponding geodesic α we have $\alpha^+(0) = v$.

Given $v \in \mathrm{T}_p$, construct a sequence $v_n \in \mathrm{T}'_p$ such that $v_n \to v$, and let γ_n be a sequence of corresponding geodesics.

The needed map α can be found among the maps such that $\alpha(0) = p$ and

$$\alpha(t) = \gamma_n(t) \quad \text{if} \quad \varepsilon_{n+1} \leqslant t < \varepsilon_n,$$

where ε_n is a decreasing sequence converging to 0 as $n \to \infty$. In order to satisfy the conclusion of ~~the~~ proposition, one has to choose the sequence ε_n converging to 0 very fast. Note that in this construction α is not continuous. $\square$

6.12. Definition. *Let $\mathcal{X}$ be a metric space and $\alpha\colon \mathbb{I} \to \mathcal{X}$ be a curve.*

For $t_0 \in \mathbb{I}$, if $\alpha^+(t_0)$ or $\alpha^-(t_0)$ or both are defined, we say respectively that α is right *or* left *or* both-sided differentiable *at t_0. In the exceptional cases where t_0 is the left (respectively, right) end of $\mathbb{I}$, α is by definition left (respectively, right) differentiable at t_0.*

If α is both-sided differentiable at t, and

$$|\alpha^+(t)| = |\alpha^-(t)| = \tfrac{1}{2} \cdot |\alpha^+(t) - \alpha^-(t)|_{\mathrm{T}_{\alpha(t)}},$$

then we say that α is differentiable *at t.*

6.13. Exercise. *Assume $\mathcal{X}$ is a metric space with defined angles. Show that any geodesic $\gamma\colon \mathbb{I} \to \mathcal{X}$ is differentiable everywhere.*

Recall that the speed of a curve is defined in 3.10.

6.14. Exercise. *Let α be a curve in a metric space with defined angles. Suppose that $\operatorname{speed}_t \alpha$, $\alpha^+(t)$, and $\alpha^-(t)$ are defined.*

Show that α is differentiable at t.

G. Differential

6.15. Definition. *Let $\mathcal{X}$ be a metric space with defined angles, and $f\colon \mathcal{X} \multimapinv \mathbb{R}$ be a subfunction. For $p \in \operatorname{Dom} f$, a function $\phi\colon \mathrm{T}_p \to \mathbb{R}$ is called the* differential *of f at p (briefly $\phi = \mathbf{d}_p f$) if for any map $\alpha\colon \mathbb{I} \to \mathcal{X}$ such that $\mathbb{I}$ is a real interval, $\alpha(0) = p$, and $\alpha^+(0)$ is defined, we have*

$$(f \circ \alpha)^+(0) = \phi(\alpha^+(0)).$$

6.16. Proposition. *Let $f\colon \mathcal{X} \multimapinv \mathbb{R}$ be a locally Lipschitz semiconcave subfunction on a metric space $\mathcal{X}$ with defined angles. Then the differential $\mathbf{d}_p f$ is uniquely defined for any $p \in \operatorname{Dom} f$. Moreover,*

(*a*) *The differential $\mathbf{d}_p f\colon \mathrm{T}_p \to \mathbb{R}$ is Lipschitz and*

$$\operatorname{lip} \mathbf{d}_p f \leqslant \operatorname{lip}_p f;$$

that is, the Lipschitz constant of $\mathbf{d}_p f$ does not exceed the Lipschitz constant of f in any neighborhood of p.

(*b*) *$\mathbf{d}_p f\colon \mathrm{T}_p \to \mathbb{R}$ is a positive homogeneous function; that is, for any $r \geqslant 0$ and $v \in \mathrm{T}_p$ we have*

$$r \cdot \mathbf{d}_p f(v) = \mathbf{d}_p f(r \cdot v).$$

(*c*) *The differential $\mathbf{d}_p f\colon \mathrm{T}_p \to \mathbb{R}$ is the restriction of ultradifferential defined in Section 6I; that is,*

$$\mathbf{d}_p f = \mathbf{d}_p^\omega f|_{\mathrm{T}_p}.$$

Proof. Passing to a subdomain of f if necessary, we can assume that f is ℓ-Lipschitz and λ-concave for some $\ell, \lambda \in \mathbb{R}$.

Take a geodesic γ in $\operatorname{Dom} f$ starting at p. Since $f \circ \gamma$ is λ-concave, the right derivative $(f \circ \gamma)^+(0)$ is defined. Since f is ℓ-Lipschitz, we have

$$|(f \circ \gamma)^+(0) - (f \circ \gamma_1)^+(0)| \leqslant \ell \cdot |\gamma^+(0) - \gamma_1^+(0)| \tag{1}$$

for any other geodesic γ_1 starting at p.

Define $\phi\colon \mathrm{T}'_p \to \mathbb{R}\colon \gamma^+(0) \mapsto (f\circ\gamma)^+(0)$. From 1, ϕ is an ℓ-Lipschtz function defined on T'_p. Thus we can extend ϕ to all of T_p as an ℓ-Lipschitz function.

It remains to check that ϕ is the differential of f at p. Assume $\alpha\colon [0,a)\to \mathcal{X}$ is a map such that $\alpha(0)=p$ and $\alpha^+(0)=v\in\mathrm{T}_p$. Let $\gamma_n\in\Gamma_p$ be a sequence of geodesics as in the definition 6.9; that is, if

$$v_n=\gamma_n^+(0) \quad\text{and}\quad a_n=\varlimsup_{t\to 0+}|\alpha(t)-\gamma_n(t)|/t$$

then $a_n\to 0$ and $v_n\to v$ as $n\to\infty$. Then

$$\phi(v)=\lim_{n\to\infty}\phi(v_n),$$

$$f\circ\gamma_n(t)=f(p)+\phi(v_n)\cdot t+o(t),$$
$$|f\circ\alpha(t)-f\circ\gamma_n(t)|\leqslant \ell\cdot|\alpha(t)-\gamma_n(t)|.$$

Hence

$$f\circ\alpha(t)=f(p)+\phi(v)\cdot t+o(t).$$

The last part follows from the definitions of differential and ultradifferential; see Section 4E. □

H. Ultratangent space

Fix a selective ultrafilter ω on the set of natural numbers.

For a metric space $\mathcal{X}$ and $r>0$, we will denote by $r\cdot\mathcal{X}$ its *r-blowup*, which is a metric space with the same underlying set as $\mathcal{X}$ and the metric multiplied by r. The tautological bijection $\mathcal{X}\to r\cdot\mathcal{X}$ will be denoted by $x\mapsto x^r$, so

$$|x^r-y^r|=r\cdot|x-y|$$

for any $x,y\in\mathcal{X}$.

The *ω-blowup* $\omega\cdot\mathcal{X}$ of $\mathcal{X}$ is defined to be the ω-limit of the n-blowups $n\cdot\mathcal{X}$; that is,

$$\omega\cdot\mathcal{X}\coloneqq\lim_{n\to\omega}n\cdot\mathcal{X}.$$

Given a point $x\in\mathcal{X}$, we can consider the sequence x^n, where $x^n\in n\cdot\mathcal{X}$ is the image of x under n-blowup. Note that if $x\neq y$, then

$$|x^\omega-y^\omega|_{\omega\cdot\mathcal{X}}=\infty;$$

that is, x^ω and y^ω belong to different metric components of $\omega\cdot\mathcal{X}$.

The metric component of x^ω in $\omega\cdot\mathcal{X}$ is called the *ultratangent space* of $\mathcal{X}$ at x and is denoted by $\mathrm{T}^\omega_x\mathcal{X}$ or T^ω_x.

Equivalently, the ultratangent space $\mathrm{T}_x^\omega \mathcal{X}$ can be defined as follows. Consider all the sequences of points $x_n \in \mathcal{X}$ such that the sequence $(n \cdot |x - x_n|_{\mathcal{X}})$ is bounded. Define the pseudodistance between two such sequences as

$$|(x_n) - (y_n)| = \lim_{n\to\omega} n \cdot |x_n - y_n|_{\mathcal{X}}.$$

Then $\mathrm{T}_x^\omega \mathcal{X}$ is the corresponding metric space.

Tangent spaces (see section 6E) as well as ultratangent spaces generalize the notion of tangent spaces on Riemannian manifolds. In the simplest cases these two notions define the same space. However in general they are different and are both useful—often a lack of a property in one is compensated by the other. It is clear from the definition that a tangent space has a cone structure. On the other hand, in general an ultratangent space does not have a cone structure. Hilbert's cube $\prod_{n=1}^{\infty}[0, 2^{-n}]$ is an example. We remark that Hilbert's cube is a CBB(0) as well as a CAT(0) Alexandrov space.

The next theorem shows that the tangent space T_p can be (and often will be) considered as a subset of T_p^ω.

6.17. Theorem. *Let $\mathcal{X}$ be a metric space with defined angles. Then for any $p \in \mathcal{L}$, there is a distance-preserving map*

$$\iota : \mathrm{T}_p \hookrightarrow \mathrm{T}_p^\omega$$

such that for any geodesic γ starting at p we have

$$\gamma^+(0) \stackrel{\iota}{\mapsto} \lim_{n\to\omega} [\gamma(\tfrac{1}{n})]^n.$$

Proof. Given $v \in \mathrm{T}'_p$, choose a geodesic γ that starts at p and such that $\gamma^+(0) = v$. Set $v^n = [\gamma(\frac{1}{n})]^n \in n \cdot \mathcal{X}$ and

$$v^\omega = \lim_{n\to\omega} v^n.$$

Note that the value $v^\omega \in \mathrm{T}_p^\omega$ does not depend on the choice of γ; that is, if γ_1 is another geodesic starting at p such that $\gamma_1^+(0) = v$, then

$$\lim_{n\to\omega} v^n = \lim_{n\to\omega} v_1^n,$$

where $v_1^n = [\gamma_1(\frac{1}{n})]^n \in n \cdot \mathcal{X}$. The latter follows since

$$|\gamma(t) - \gamma_1(t)|_{\mathcal{X}} = o(t),$$

and therefore $|v^n - v_1^n|_{n\cdot\mathcal{X}} \to 0$ as $n \to \infty$.

Set $\iota(v) = v^\omega$. Since angles between geodesics in $\mathcal{X}$ are defined, for any v, $w \in \mathrm{T}'_p$ we have $n \cdot |v_n - w_n| \to |v - w|$. Thus $|v_\omega - w_\omega| = |v - w|$; that is, $\iota\colon\ \mathrm{T}'_p \to \mathrm{T}_p$ is a distance-preserving map.

Since T'_p is dense in T_p, we can extend ι to a distance-preserving map $\mathrm{T}_p \to \mathrm{T}^\omega_p$. □

I. Ultradifferential

Given a function $f\colon \mathcal{L} \to \mathbb{R}$, consider the sequence of functions $f_n\colon n\cdot\mathcal{L} \to \mathbb{R}$ defined by

$$f_n(x^n) = n \cdot (f(x) - f(p)),$$

where $x \mapsto x^n$ denotes the natural map $\mathcal{L} \to n \cdot \mathcal{L}$. While $n \cdot (\mathcal{L}, p) \to (\mathrm{T}^\omega, 0)$ as $n \to \omega$, the functions f_n converge to the *ω-differential* of f at p. It will be denoted by $\mathbf{d}^\omega_p f$:

$$\mathbf{d}^\omega_p f\colon\ \mathrm{T}^\omega_p \to \mathbb{R}, \quad \mathbf{d}^\omega_p f = \lim_{n\to\omega} f_n.$$

Clearly, the ω-differential $\mathbf{d}^\omega_p f$ of a locally Lipschitz subfunction f is defined and Lipschitz at each point $p \in \operatorname{Dom} f$.

J. Remarks

Spaces with defined angles include CAT and CBB spaces; see 8.14*c* and 9.15*b*.

For general metric spaces, angles may not exist, and given a hinge $[p\,{}^x_y]$ it is more natural to consider the *upper angle* defined by

$$\measuredangle^{\mathrm{up}}[p\,{}^x_y] \coloneqq \varlimsup_{\bar x,\bar y\to p} \tilde\measuredangle^\kappa(p\,{}^{\bar x}_{\bar y}),$$

where $\bar x \in\ ]px]$ and $\bar y \in\ ]py]$. The triangle inequality (6.5) holds for upper angles as well.

Chapter 7

Dimension theory

A. Definitions

In this section, we give definitions of different types of dimension-like invariants of metric spaces and state general relations between them. The proofs of most of the statements in this section can be found in the book of Witold Hurewicz and Henry Wallman [**85**]; the rest follow directly from the definitions.

7.1. Hausdorff dimension. *Let $\mathcal{X}$ be a metric space. Its Hausdorff dimension is defined as*

$$\operatorname{HausDim} \mathcal{X} = \sup\{\, \alpha \in \mathbb{R} : \operatorname{HausMes}_\alpha(\mathcal{X}) > 0 \,\},$$

where $\operatorname{HausMes}_\alpha$ *denotes the α-dimensional Hausdorff measure.*

Let $\mathcal{X}$ be a metric space and $\{V_\beta\}_{\beta\in\mathcal{B}}$ be an open cover of $\mathcal{X}$. Let us recall two notions in general topology:

- The *order* of $\{V_\beta\}$ is the supremum of all integers n such that there is a collection of $n+1$ elements of $\{V_\beta\}$ with nonempy intersection.
- An open cover $\{W_\alpha\}_{\alpha\in\mathcal{A}}$ of $\mathcal{X}$ is called a *refinement* of $\{V_\beta\}_{\beta\in\mathcal{B}}$ if for any $\alpha \in \mathcal{A}$ there is $\beta \in \mathcal{B}$ such that $W_\alpha \subset V_\beta$.

7.2. Topological dimension. *Let $\mathcal{X}$ be a metric space. The topological dimension of $\mathcal{X}$ is defined to be the minimum of nonnegative integers n such that for any open cover of $\mathcal{X}$ there is a finite open refinement with order n.*

If no such n exists, the topological dimension of $\mathcal{X}$ is infinite.

The topological dimension of $\mathcal{X}$ will be denoted by $\operatorname{TopDim} \mathcal{X}$.

The invariants satisfying the following two statements 7.3 and 7.4 are commonly called "dimension"; for that reason we call these statements axioms.

7.3. Normalization axiom. *For any $m \in \mathbb{Z}_{\geqslant 0}$,*

$$\operatorname{TopDim} \mathbb{E}^m = \operatorname{HausDim} \mathbb{E}^m = m.$$

7.4. Cover axiom. *If $\{A_n\}_{n=1}^{\infty}$ is a countable closed cover of $\mathcal{X}$, then*

$$\operatorname{TopDim} \mathcal{X} = \sup_n \{\operatorname{TopDim} A_n\},$$
$$\operatorname{HausDim} \mathcal{X} = \sup_n \{\operatorname{HausDim} A_n\}.$$

On product spaces. Recall that the direct product $\mathcal{X} \times \mathcal{Y}$ of metric spaces $\mathcal{X}$ and $\mathcal{Y}$ is defined in Section 2A. Direct product satisfies the following two inequalities:

$$\operatorname{TopDim}(\mathcal{X} \times \mathcal{Y}) \leqslant \operatorname{TopDim} \mathcal{X} + \operatorname{TopDim} \mathcal{Y}$$

and

$$\operatorname{HausDim}(\mathcal{X} \times \mathcal{Y}) \geqslant \operatorname{HausDim} \mathcal{X} + \operatorname{HausDim} \mathcal{Y}.$$

These inequalities might be strict. For topological dimension, strict inequality holds for a pair of Pontryagin surfaces [**138**]. For Hausdorff dimension, an example was constructed by Abram Besicovitch and Pat Moran [**26**].

The following theorem follows from [**85**, theorems V 8 and VII 2].

7.5. Szpilrajn's theorem. *Let $\mathcal{X}$ be a separable metric space. Assume* $\operatorname{TopDim} \mathcal{X} \geqslant m$. *Then* $\operatorname{HausMes}_m \mathcal{X} > 0$.

In particular, $\operatorname{TopDim} \mathcal{X} \leqslant \operatorname{HausDim} \mathcal{X}$.

In fact it is true that for any separable metric space $\mathcal{X}$ we have

$$\operatorname{TopDim} \mathcal{X} = \inf\{\operatorname{HausDim} \mathcal{Y}\},$$

where the infimum is taken over all metric spaces $\mathcal{Y}$ homeomorphic to $\mathcal{X}$.

7.6. Definition. *Let $\mathcal{X}$ be a metric space and $F \colon \mathcal{X} \to \mathbb{R}^m$ be a continuous map. A point $\mathbf{z} \in \Im F$ is called a* stable value *of F if there is $\varepsilon > 0$ such that $\mathbf{z} \in \Im F'$ for any ε-*close *to F continuous map $F' \colon \mathcal{X} \to \mathbb{R}^m$, that is, $|F'(x) - F(x)| < \varepsilon$ for all $x \in \mathcal{X}$.*

The next theorem follows from [**85**, theorems VI 1&2]. (This theorem also holds for non-separable metric spaces [**119**], [**63**, 3.2.10]).

7.7. Stable value theorem. *Let $\mathcal{X}$ be a separable metric space. Then* $\operatorname{TopDim} \mathcal{X} \geqslant$ *m if and only if there is a map $F \colon \mathcal{X} \to \mathbb{R}^m$ with a stable value.*

7.8. Proposition. *Suppose $\mathcal{X}$ and $\mathcal{Y}$ are metric spaces and $\Phi\colon \mathcal{X} \to \mathcal{Y}$ satisfies*

$$|\Phi(x) - \Phi(x')| \geqslant \varepsilon \cdot |x - x'|$$

for fixed $\varepsilon > 0$ and any pair $x, x' \in \mathcal{X}$. Then

$$\operatorname{HausDim} \mathcal{X} \leqslant \operatorname{HausDim} \mathcal{Y}.$$

In particular, if there is a Lipschitz onto map $\mathcal{Y} \to \mathcal{X}$, then

$$\operatorname{HausDim} \mathcal{X} \leqslant \operatorname{HausDim} \mathcal{Y}.$$

B. Linear dimension

In addition to HausDim and TopDim, we will use the so-called linear dimension. It will be applied only to Alexandrov spaces and to their open subsets (in cases both of curvature bounded below and curvature bounded above). As we shall see, in all these cases LinDim behaves nicely and is easy to work with.

Recall that a *cone map* is a map between cones respecting the cone multiplication.

7.9. Definition of linear dimension. *Let $\mathcal{X}$ be a metric space with defined angles. The* linear dimension *of $\mathcal{X}$ (denoted by* $\operatorname{LinDim} \mathcal{X}$*) is defined as the exact upper bound on $m \in \mathbb{Z}_{\geqslant 0}$ such that there is a distance-preserving cone embedding $\mathbb{E}^m \hookrightarrow \mathrm{T}_p\,\mathcal{X}$ for some $p \in \mathcal{X}$; here $\mathbb{E}^m$ denotes the m-dimensional Euclidean space and $\mathrm{T}_p\,\mathcal{X}$ denotes the tangent space of $\mathcal{X}$ at p (defined in Section* 6D*).*

Note that LinDim takes values in $\mathbb{Z}_{\geqslant 0} \cup \{\infty\}$.

The linear dimension LinDim has no immediate relations to HausDim and TopDim. Also, LinDim does not satisfy ~~the cover axiom (7.4).~~ Note that

$$\operatorname{LinDim}(\mathcal{X} \times \mathcal{Y}) = \operatorname{LinDim} \mathcal{X} + \operatorname{LinDim} \mathcal{Y} \tag{1}$$

for any two metric spaces $\mathcal{X}$ and $\mathcal{Y}$ with defined angles.

The following exercise is based on a construction of Thomas Foertsch and Viktor Schroeder [**145**]; it shows that the condition on existence of angles in 1 cannot be removed.

7.10. Exercise. *Construct metrics ρ_1 and ρ_2 on $\mathbb{R}^{10}$ defined by norms, such that $(\mathbb{R}^{10}, \rho_i)$ do* not *contain an isometric copy of $\mathbb{E}^2$ but $(\mathbb{R}^{10}, \rho_1) \times (\mathbb{R}^{10}, \rho_2)$ has an isometric copy of $\mathbb{E}^{10}$.*

Remarks. Linear dimension was first introduced by Conrad Plaut [**137**] under the name *local dimension*. *Geometric dimension*, introduced by Bruce Kleiner [**95**] is closely related; it coincides with the linear dimension for CBB and CAT spaces.

One can extend the definition to arbitrary metric spaces. To do this one should modify the definition of tangent space and take an arbitrary n-dimensional

Banach space instead of the Euclidean n-space. For Alexandrov spaces (either CBB or CAT) this modification is equivalent to our definition.