On asymmetric distances

Andrea C. G. Mennucci ⋆

May 28, 2007

Abstract

In this paper we discuss asymmetric metric spaces in an abstract setting, mimicking the usual theory of metric spaces. As a typical application, we consider asymmetric metric spaces generated by functionals in Calculus of Variation.

KEYWORDS: ASYMMETRIC METRIC, GENERAL METRIC, QUASI METRIC, OSTENSIBLE METRIC, FINSLER METRIC, PATH METRIC, LENGTH, GEODESIC CURVE, HOPF–RINOW THEOREM

1 Introduction

"Besides, one insists that the distance function be symmetric, that is, $d(x, x') = d(x', x)$*.*
(This unpleasantly limits many applications $[\cdots]$*)."* M. Gromov ([12], Intr.)

The main purpose of this paper is to study an "asymmetric metric theory"; this theory naturally generalizes the metric part of Finsler Geometry, much as symmetric metric theory generalizes the metric part of Riemannian Geometry.

In order to do this, we will define the "asymmetric metric space" as a set M equipped with a positive $b : M \times M \to \mathbb{R}^+$ which satisfy the triangle inequality, but may fail to be symmetric (see 2.1 for the formal definition). We will then state, (in sec. 2) "asymmetric definitions" in this space, such as "forward local compactness", "forward completeness", "forward boundedness", and so on.

The invention of a (possibly) asymmetric distance is quite useful in some applied fields, and mainly in Calculus of Variation, as we will exemplify in the next section §1.ii, and discuss more in depth in sec. §5.i. It is then no surprise that this theory was invented and studied many times in the past.

It was indeed shown in preceding studies that, even if a metric space does not have the "smooth" character of a Riemannian structure, still many definitions and results can be reformulated and carried on[(1)]: see for example in Busemann [5], or in Gromov's [12], or in Ambrosio *et al* [1, 2].

To exemplify this fact, in the next section we will present an example theorem that generalizes the Hopf–Rinow theorem from finite dimensional Riemannian spaces to suitable metric spaces; and a further generalization to the asymmetric case.

A different point of view is found in the theory of *quasi metric* (or *ostensible metric*) where much emphasis is given to the topological aspects of the theory; indeed, a quasi-metric will generate, in general, *three different topologies* (see §3.ii.1); this brings forth many different and non-equivalent definitions of *"completeness"* (and of the other

⋆Scuola Normale Superiore Piazza dei Cavalieri 7, 56126 Pisa, Italy

(1)and Busemann [4] attributes the first such results in this respect to Cartan, in 1928 properties commonly stated in metric theory). In our definition of the "asymmetric metric space" (M, b), we choose a different topology associated to the space than the one used in "quasi metric spaces": this makes it difficult to compare the results in the two fields. We discuss further this issue in sec. §3.ii.

Our set of hypotheses is more resembling (and slightly generalizes) Busemann theory of *general metric spaces* (as was defined in [4]); we will discuss differences and similarities in section §3.i.

Summarizing, this presentation of the theory of asymmetric metrics is more specific and more focused on "metric aspects" of what is seen in *ostensible metric theory*, and at the same time is less "geometric oriented" than what is seen in Busemann work (indeed, many of the examples presented are quite non-smooth).

§1.i The Hopf–Rinow theorem

In section I in [12], Gromov models the "metric part" of Riemannian Geometry, as follows. Consider a metric space (M, d): we can define the length $\operatorname{len}\gamma$ of a Lipschitz curve $\gamma : [\alpha, \beta] \to M$ using the total variation formula (see (2.5)); then we can define a new metric $d^g(x, y)$ as the infimum of $\operatorname{len}\gamma$ in the class of all Lipschitz curves connecting x to y. Gromov defines that a metric space is "path-metric" if $d = d^g$; he then proves in §1.11 §1.12 in [12] that [(2)]

Theorem 1.1 (symmetric Hopf-Rinow) *Suppose that (M, d) is path-metric and locally compact; then the two following facts are equivalent*

i). (M, d) is complete

ii). closed bounded sets are compact

and they imply that each pair of points can be joined by a geodesic (i.e. a minimum of $\operatorname{len}\gamma$*).*

The above is the "metric only" counterpart of the theorem of Hopf-Rinow in Riemannian Geometry: indeed, if (M, g) is a finite-dimensional Riemannian manifold, and d is the associated distance, then (M, d) is path-metric and locally compact.

Since there is a Hopf-Rinow theorem in Finsler Geometry, we would expect that there would be a corresponding theorem for "asymmetric metric spaces"; and indeed we can prove this result:

Theorem 1.2 *Suppose that the* asymmetric metric space (M, b) *is path-metric and forward-locally compact; then the following two facts are equivalent*

- *forward-bounded closed sets are compact*
- (M, b) *is forward-complete*

and they imply that any two points can be joined by a geodesic.

A first form of this theorem is due to Cohn–Vossen, according to the introduction of Busemann's [6], where though it is stated in with stronger hypothesis than the above; so a more general form is here stated and proved as Theorem 4.33 (a comparison of the two results follows that theorem).

[(2)] actually, we have added the easy implication *ii*$\Longrightarrow$*i* to the statement in [12]

§1.ii asymmetric metric theories in Calculus of Variations

For the sake of this introduction section, let M be a smooth connected differential manifold.

Consider an integrand $F : TM \to [0, \infty]$ that is Borel measurable and such that $v \mapsto F(x, v)$ is convex and positively 1-homogeneous (that is, $F(x, lv) = lF(x, v)$ for $l \geq 0$). Consider the length

$$\mathrm{len}\gamma = \int_0^1 F(\gamma(s), \dot\gamma(s))\, ds \tag{1.3}$$

where $\gamma : [0, 1] \to M$ is a locally Lipschitz curve; this generates a "distance" $b(x, y)$, as the infimum of the length of curves γ connecting x to y; b is called a *Finsler metric*, since is is easily seen to satisfy a triangular inequality; but in general b will be asymmetric.

Consider also the Hamilton-Jacobi equation

$$\Big\{ H(x, Du(x)) = 0, \quad u(a) = 0 \tag{1.4}$$

where $H : T^*M \to \mathbb{R}$ is continuous and $p \mapsto H(x, p)$ is convex, and the solution $u : M \to \mathbb{R}$ is in the viscosity sense; and $a \in M$.

In the above settings, a broad number of ideas appear naturally; we list them, in decreasing order of regularity:

- In the field of Finsler Geometry, it is assumed that F is smooth and that $v \mapsto F^2(x, v)$ is strongly convex for $v \neq 0$ (that is, the Hessian $\frac{\partial^2 F^2}{\partial v \partial v}$ is positive definite); many results of Riemannian geometry are generalized to this asymmetric settings (see [3] for a wide treaty of the subject);

- the part of Finsler Geometry that deals with existence of geodesics can be easily shown to hold when $F \in C^2$ and $v \mapsto F^2(x, v)$ is strongly convex; so this is a very common setting in books on Calculus of Variations, Control Theory and Viscosity Solutions[(3)]; it is well known that the Hamilton-Jacobi equation (1.4) is naturally associated with the distance $b(x, y)$: for example, $u(x) = b(a, x)$ is the viscosity solution of (1.4) when H is "dual" to F^2, in the sense of Legendre–Fenchel (see [16] for details).

- When we try to relax the requirements on F and H further, few results are known in the literature (see [18] for an exception).

  With appropriate hypotheses on H (and in particular on the sets $Z_x = \{p \mid H(x, p) \leq 0\}$) [16] establishes the duality between H and F and that F is continuous and locally bounded (and many other properties). The above work uses some results that are in section §5.i; for example theorem 5.9 shows that the Hopf–Rinow theorem implies a purely geometric criteria for the existence of a minimum curve for problem (1.3).

- Yet even the setting used in section §5.i is not the most general we can envision: to start this section, we just assumed that F is Borel and $F(x, \cdot)$ is convex and 1-homogeneous, and this is enough to generate the distance b, so that (M, b) is an *asymmetric metric space*. So this study of the abstract properties of *asymmetric metric spaces* may be useful in further attempts at generalizing the work that was done in [16].

(3) for example, in §2.4.2 [9]; or $L \in C^3$ and strongly convex in §1.8 in [10]

2 Asymmetric metric spaces

We review some basal concepts and results; those are similar to what is commonly found in books on metric spaces. In all of this chapter, M will be a generic set.

§2.i Asymmetric metric space

Definition 2.1 $b : M \times M \to [0,\infty)$ *is an* asymmetric distance *if* b *satisfies*

- $b \geq 0$ *and* $b(x,y) = 0$ *iff* $x = y$
- $b(x,y) \leq b(x,z) + b(z,y)\ \forall x,y,z \in M.$

We call the pair (M,b) *an* asymmetric metric space.

We agree that b defines a topology τ on M, generated by the families of *forward* and *backward* open balls

$$B^+(x,\varepsilon) \stackrel{\text{def}}{=} \{y \mid b(x,y) < \varepsilon\}, \quad B^-(x,\varepsilon) \stackrel{\text{def}}{=} \{y \mid b(y,x) < \varepsilon\}$$

that is, the topology is generated by the symmetric distance

$$d(x,y) \stackrel{\text{def}}{=} b(x,y) \vee b(y,x) \tag{2.2}$$

with balls

$$B(x,\varepsilon) \stackrel{\text{def}}{=} \{y \mid d(x,y) < \varepsilon\} = B^+(x,\varepsilon) \cap B^-(x,\varepsilon)$$

By the above definition,

Proposition 2.3 $x_n \to x$ *iff* $d(x_n,x) \to 0$*, iff both* $b(x_n,x) \to 0$ *and* $b(x,x_n) \to 0$

We will sometimes denote by τ^+ (resp. τ^-) the topology generated by forward (resp. backward) balls alone; we remark that those may differ from τ: this issue is discussed in sec. §3.ii.

Remark 2.4 *We may also consider the case of an* asymmetric distance $b : M \times M \to [0,+\infty]$ *where there may be points* x,y *such that* $b(x,y) = \infty$*; this is somewhat uncommon, so we suggest two workarounds:*

- *by using the triangular inequality, it is easily seen that the equivalence relation* $x \sim y \iff d(x,y) < \infty$ *decomposes the space* (M,b) *in equivalence classes that are all open (so two different classes are disconnected); so we may restrict the study of the space to those equivalence classes;*
- *defining* $\phi(t) = t/(1+t)$ *and* $\phi(\infty) = 1$ *and then* $\tilde{b} \stackrel{def}{=} \phi \circ b$ *and* $\tilde{d} \stackrel{def}{=} \phi \circ d$*, we may build a space* $(M,\tilde{b})$ *that is topologically equivalent, and where the distance does not assume the value* $+\infty$*.*

Unfortunately those remedies are not really useful in some cases: see 2.17.

§2.ii Lipschitz maps

Let (N, δ) be a symmetric metric space, and $f : N \to M$, $g : M \to N$. We could define that f and g are "Lipschitz w.r.t. b", or "Lipschitz w.r.t. d". That is, in the first case

$$b(f(x), f(y)) \le C\delta(x, y), \quad \delta(g(x), g(y)) \le Cb(x, y)$$

for some constant $C > 0$. In the second case

$$d(f(x), f(y)) \le C\delta(x, y), \quad \delta(g(x), g(y)) \le Cd(x, y)$$

for some constant $C > 0$.

The two concepts are equivalent for f, but not equivalent for g.

As an example, if we provide M^2 with the metric

$$d^2((x, y), (x', y')) \stackrel{\text{def}}{=} d(x, x') + d(y, y') \ ,$$

then $(x, y) \mapsto b(x, y)$ is Lipschitz w.r.t d^2, since

$$|b(x, y) - b(x', y')| \le |b(x, y) - b(x, y')| + |b(x, y') - b(x', y')| \le d(x, x') + d(y, y')$$

and this implies that b is continuous.

§2.iii Length

We induce from b the length $\text{len}^b\gamma$ of a curve $\gamma : [\alpha, \beta] \to M$, by using the total variation[(4)]

$$\text{len}^b\gamma \stackrel{\text{def}}{=} \sup_T \sum_{i=1}^n b\big(\gamma(t_{i-1}), \gamma(t_i)\big) \tag{2.5}$$

where the sup is carried out over all finite subsets $T = \{t_0, \cdots, t_n\}$ of $[\alpha, \beta]$ and $t_0 \le \cdots \le t_n$. If $\text{len}^b\gamma < \infty$, then γ is called *rectifiable*.

Remark 2.6 *The above length is* asymmetric*: if $\hat\gamma(t) \stackrel{def}{=} \gamma(-t)$, then in general $\text{len}^b\gamma$ will be different from $\text{len}^b\hat\gamma$; hence there does not exist a measure $\mathcal{H}$ on M such that the length of a curve is the measure of its image, that is*

$$\mathcal{H}(\textit{image}(\gamma)) = \text{len}^b\gamma \quad ;$$

indeed image(γ) = image$(\hat\gamma)$ which is incompatible in the general case in which $\text{len}^b\gamma \neq \text{len}^b\hat\gamma$.

We define b^g

$$b^g(x, y) = \inf \text{len}^b\gamma \tag{2.7}$$

where the inf is taken in the class of all *continuous* curves γ connecting x to y.

Note that

$$\text{len}^b\gamma \ge b(\gamma(\alpha), \gamma(\beta)) \tag{2.8}$$

so $b^g \ge b$. It is also easy to prove that b^g satisfy a triangle inequality: so b^g is an asymmetric distance (possibly taking the value ∞; see 2.4 and 2.17 on this). b^g is called *the geodesic (asymmetric) distance induced by* b. We may endow M with the distance $d^g(x, y) \stackrel{\text{def}}{=} b^g(x, y) \vee b^g(y, x)$; then $d^g \ge d$, and the topology resulting in (M, b^g) is finer than the one used in (M, b) (has more open sets and less compact sets); the two topologies may easily differ, as in this simple example:

(4) according to Busemann [4], this definition goes back to Menger (1928)

Example 2.9 *Consider*

$$M = \{x \in \mathbb{R}^2 \mid -1 \le x_1 \le 1, x_2 = 0\} \cup \bigcup_{n\ge 1} \{x \in \mathbb{R}^2 \mid -1 \le x_1 \le 1,\ \ x_2 = (x_1 - 1)/n\}$$

and b *the Euclidean distance (see fig. 1). Then* (M, b) *is compact but* (M, b^g) *is not.*

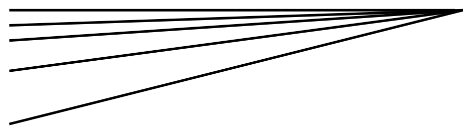


Figure 1: example 2.9

Sometimes the above infimum (2.7) is a minimum:

Definition 2.10 (Geodesics) *Let's fix* $x, y \in M$ *such that* $b^g(x, y) < \infty$*. If there is a* <u>*continuous*</u> *curve* γ *providing the minimum of the equation (2.7) that defined* $b^g(x, y)$*, then this curve is called a "(minimal) geodesic" (connecting* x *to* y*).*

Suppose that γ is rectifiable. We define the *running length* $\ell : [\alpha, \beta] \to \mathbb{R}^+$ of γ to be the length of γ restricted to $[\alpha, t]$, that is

$$\ell(t) \stackrel{\text{def}}{=} \text{len}^b \left(\gamma|_{[\alpha,t]}\right) \tag{2.11}$$

Note that:

- if γ is Lipschitz of constant L, then, by direct substitution in (2.5),

  $$\ell(t) - \ell(s) \le (t - s)L$$

  so ℓ is Lipschitz.[(5)]

  In this case the derivative $\frac{d\ell}{dt}$ exists for almost all t, and it is called the *metric derivative* of γ (as defined in [1, 2]).

- Supposing that ℓ is continuous (for simplicity), for $t \ge s$,

  $$b\big(\gamma(s), \gamma(t)\big) \le b^g\big(\gamma(s), \gamma(t)\big) \le \ell(t) - \ell(s) \tag{2.12}$$

  since the length of γ restricted to $[s, t]$ is $\ell(t) - \ell(s)$, and γ is indeed a path connecting $\gamma(s)$ to $\gamma(t)$;

- whereas the fact that ℓ is continuous does not imply that γ is continuous, as shown in this example:

Example 2.13 *Let* $M = [-1, 1]$ *and let* $f : M \times \mathbb{R} \to \mathbb{R}$ *be*

$$f(x, v) = \begin{cases} |v| & \textit{if } v > 0 \textit{ or } x \le 0 \\ |v|/x & \textit{if } v < 0 \textit{ and } x > 0 \end{cases} \tag{2.13.$\star$}$$

[(5)] when b is symmetric, if γ is rectifiable and continuous, then ℓ is continuous; this seems to be true also for asymmetric distances, but was not carefully checked

From this we induce the asymmetric distance as explained in the introduction §1.ii, to get

$$b(x,y) = \begin{cases} |x-y| & \text{if } x \le y \text{ or } y < x \le 0 \\ -\log(y) + \log(x) & \text{if } 0 < y < x \\ +\infty & \text{if } y \le 0 < x \end{cases} \tag{2.13.$\star\star$}$$

then the topology τ^+ generated by forward balls is the usual euclidean topology on $[-1,1]$; whereas the topology τ^- generated by backward balls makes $[-1,0]$ disconnected from $(0,1]$; and $\tau = \tau^-$.

$\gamma : [-1,1] \to M$ defined by $\gamma(t) = t$ has $l^\gamma(t) = t$, but is not continuous.

This is counter-intuitive since it is contrary to what is seen in symmetric metric spaces; so we will sometimes use this lemma

Lemma 2.14 *Suppose that ℓ is continuous; then γ is continuous if and only if its image is compact. (Proof follows from lemma 3.7 and (2.12)).*

§2.iv Definitions

We add some definitions

- We say that the (asymmetric) metric space (M, b) is **a path-metric space**, or that b is **intrinsic**, if $b = b^g$;
- and we say that it **admits geodesics**, or that *geodesics exist*, if, for all $x, y \in M$, there is a geodesic connecting x to y.
- A sequence $(x_n) \subset M$ is called *forward Cauchy* if

  $$\forall \varepsilon > 0,\ \ \exists N \in \mathbb{N} \text{ such that } \forall n, m,\ \ m \ge n \ge N,\ \ b(x_n, x_m) < \varepsilon \tag{2.15}$$

  We say that (M, b) is **forward complete** if any *forward Cauchy* sequence (x_n) converges to a point x (according to the topology τ: cf. 2.3 and §3.ii.1 on the notion of convergence).

  (These definitions agree with those used in Finsler Geometry (see ch. VI of [3]). See 3.4 for a different definition.)
- We say that $A \subset M$ is **forward bounded** if these two equivalent propositions hold
  - $\exists x \in M, r > 0$ s.t. $A \subset B^+(x, r)$
  - $\forall x \in M, \exists r > 0$ s.t. $A \subset B^+(x, r)$

For any *forward* definition above there is a corresponding **backward** definition, obtained by exchanging the first and the second argument of b, or by using the *conjugate distance* $\overline{b}$ defined by

$$\overline{b}(x,y) = b(y,x) \quad ; \tag{2.16}$$

and a corresponding **symmetric** definition, obtained by substituting d for b.

For the same reason, in this paper we will present mostly the *forward* versions of the theorems, since *backward* results are obtained by substituting b for $\overline{b}$.

We nonetheless emphasize these definitions.

- We say that (M,b) is **forward-locally compact** if $\forall x\ \exists\varepsilon > 0$ such that
$$D^+(x,y) \stackrel{\text{def}}{=} \{y \mid b(x,y) \le \varepsilon\}$$
is compact.
- We say that (M,b) is *backward-locally compact* if $\forall x\ \exists\varepsilon > 0$ such that
$$D^-(x,y) \stackrel{\text{def}}{=} \{y \mid b(y,x) \le \varepsilon\}$$
is compact.
- We say that (M,b) is *symmetrically-locally compact* if $\forall x\ \exists\varepsilon > 0$ such that $\{y \mid d(x,y) \le \varepsilon\}$ is compact.
- We say that (M,b) is *locally compact* if $\forall x\ \exists\varepsilon > 0$ such that both $D^-(x,y)$ and $D^+(x,y)$ are compact.

The following implications hold:

$$\begin{array}{ccc} \text{locally compact} & \longrightarrow & \text{forward locally compact} \\ \downarrow & & \downarrow \\ \text{backward locally compact} & \longrightarrow & \text{symmetrically locally compact} \end{array}$$

Remark 2.17 *In general, even if (M,b) is connected and $b < \infty$ at all points, it may be the case that $b^g(x,y) = \infty$ for some points; if we do not like this fact, we may wish to try the remedies proposed in 2.4; in particular the equivalence relation $x \sim y \iff d^g(x,y) < \infty$ decomposes the space (M,b^g) in equivalence classes that coincide with all connected components of (M,b^g) (since b^g is itself path-metric, by 4.44).*

Unfortunately the second remedy suggested in 2.17 is only a placebo when we are interested in path-metric spaces and/or in studying geodesics: suppose b is path-metric and we decide to set $\phi(t) = t/(1+t)$ and define $\tilde b \stackrel{def}{=} \phi \circ b$: then $\tilde b$ is not path-metric; so if we try to substitute $\tilde b$ by its generated path-metric distance $\tilde b^g$ we find out, by prop. 4.37, that $\tilde b^g = b^g = b$, and we are back to square one. So we may sometimes be forced to address the case when $b^g(x,y) = \infty$ for some points.

3 Comparison with related works

§3.i Comparison with Busemann's "General metric spaces"

The foundation of the theory of non-symmetric metric spaces was given by Busemann in his paper *Local Metric Geometry* [4], where it was christened as "General metric spaces".

It should be noted that, although Busemann has presented the foundation of that theory, most of his work makes the assumption that the distance be symmetric; such is the case e.g. of the renowned work on *The Geometry of Geodesics* [5]. The study of non-symmetric "General metric spaces" was carried on for example in Phadke's [17] and Zaustinsky's [19].

We report now the theory of "General metric spaces" here, in a language that is more similar to our presentation and to [19], than to [4].

- a **“General metric space”** is a space satisfying all requisites in 2.1, and moreover

$$\forall x \in M, \forall (x_n) \subset M, \quad b(x_n, x) \to 0 \text{ iff } b(x, x_n) \to 0 \tag{3.1}$$

  It is easy to prove that this is equivalent to assuming that $\tau^+ = \tau^- = \tau$, where the topology τ^+ (resp. τ^-) is generated by the families of forward (resp. backward) balls alone;

- a “General metric space” is **finitely compact** if any closed set that is contained in $B^+(x, r) \cup B^-(x, r)$ (for a choice of $x \in M$ and $r > 0$) is compact

- a “General metric space” is **weakly finitely compact** if small closed forward balls are compact. This is what was here defined as *forward local compactness*.

We add some comments that highlight the differences between the approach in the present paper and in [4]:

- the additional hypothesis (3.1) used in defining “General metric spaces” is not really restrictive; in many applications it does hold (for example, when the space is locally compact, as in [16]); still it is possible to find examples of path-metric spaces such as 4.13 and 2.13 where it does not hold (even though this latter comes from a Calculus of Variations problem of the form discussed in introduction);

- on the other hand, in the aforementioned papers it was (sometimes silently) assumed that the metric b be intrinsic; [(6)] that is, no distinction was made between b and b^g; this is restrictive, since in many simple situations we will consider the two do differ.

  A good reason to consider non-path-metric spaces is as follows: suppose (M', b') is an asymmetric metric space, and (M, b) is a subset of it, with b being the restriction of b'; then it easy to devise examples where (M', b') is path-metric but (M, b) is not [(7)] and vice versa. For example, in the forthcoming paper [8] a family of (symmetric) metric space are studied that are not (and cannot possibly) be path-metric; the reason being this theorem (adapted from 2.6 in [8]):

  Theorem 3.2 *Suppose that (M, d) is a complete path-metric space, and that $i : M \to E$ is an isometric immersion in a uniformly convex Banach space E: then $i(M)$ is convex, and any two points in M can be joined by a unique minimal geodesic (unique up to reparametrization).*

  Then (the symmetric version of) theorem 4.24 was quite useful in establishing existence of geodesics in [8].

  So in most of this paper we will carefully distinguish the two metrics b and b^g.

- In most of the cited works (with the exception of Phadke’s [17]) the space is assumed to be locally compact (or even *finitely compact*).

  According to the notes at end of the introduction of [6], some work was carried on at that time to extend known results to *forward locally compact* spaces; but the results in this paper do not appear in the announced papers.

[(6)] Note that the definition of *“intrinsic metric”* was not used at the time of [4]: it was in a sense replaced by the *“Menger convexity”* hypothesis (that we will discuss in section §4.i.1)

[(7)] many examples in this paper are built this way

§3.ii Comparison with quasi metric spaces

§3.ii.1 Topology

As already pointed out, the definition 2.1 of the *asymmetric metric* coincides with the definition of a *quasi metric* that is found in the literature, cf. Kelly [14], Reilly, Subrahmanyam and Vamanamurthy [13] (8) Fletcher and Lindgren [11, (pp 176-181)], Künzi [15]. The main difference between our theory of *asymmetric metric spaces* and *quasi metric spaces* is in the choice of the associated topology.

Indeed, we have three topologies at hand:

- the topology τ , generated by the families of forward and backward balls, or equivalently by the metric d defined in (2.2);
- the topology τ^+ generated by the families of forward balls;
- the topology τ^- generated by the families of backward balls;

it may happen that these three topologies are different.

This problem has been studied in [14]: there Kelly introduces the notion of a *bitopological space* (M, τ^+, τ^-), and extends many definition and theorems, (such as the Urysohn lemma, the Tietze's extension theorem, the Baire category theorem (9)) to these spaces. Unfortunately Kelly does not include the topology τ in his studies.

In many applications $\tau = \tau^+ = \tau^-$: see 3.8 and section §5.i. We have chosen to associate the topology τ to the "asymmetric metric space". τ is a *symmetric kind of object*: as a consequence, we have only one notion of *"open set"*, of *"compact set"*, of "*the sequence* (x_n) *converges to* x", and of "*the functions* $f : N \to M$ *and* $g : M \to N$ *are continuous*". Furthermore

Proposition 3.3 *If* $x_n \to x$ *(according to* τ*) then the sequence* (x_n) *is both a* forward Cauchy sequence *and a* backward Cauchy sequence*,*

in accordance with the symmetric case.

§3.ii.2 Cauchy sequences, and completeness

In papers on *quasi metric spaces*, the quasi-metric space (M, b) is instead usually endowed with the topology τ^+: this entails a different notion of *convergence* and *compactness*, and poses the problem to find a good (10) definition of "*Cauchy sequence*" and "*complete space*".

This problem has been studied in [13], where 7 different notions of "*Cauchy sequence*" are presented. (Indeed, the list in [13] includes the three that we defined in §2.iv: a "forward Cauchy sequence" (resp. backward) is a "left K-Cauchy sequence" (resp. right); a "symmetrical Cauchy sequence" is a "b-Cauchy sequence"). Combining these 7 definition with the τ^+ topology, [13] presents 7 different definitions of "*complete space*". Actually, by combining 7 "*Cauchy sequences*" with all the above 3 topologies, we may reach a total of 14 (!) different definitions of "*complete space*" (considering the symmetry of using the $\bar{b}$ instead of b, see eq. (2.16)). To our knowledge, no one has taken the daunting task of examining all of them.

One of the notions of "*Cauchy sequence*" and "*complete space*" from [13] has been further studied by Künzi [15]; we present it here.

(8) [13] provides also a wide discussion of the references on *quasi metrics*

(9) Another version of Baire theorem, using a better definition of completeness, is found in Thm 2 in [13].

(10) What do we mean by "good definition"? We may need that the definition would satisfy a proposition similar to 3.3; see statement iv) below

Definition 3.4 • *A sequence* $(x_n) \subset M$ *is a* "left b-Cauchy sequence" *when* $\forall \varepsilon > 0\ \exists x \in M$ *and* $\exists k \in \mathbb{N}$ *such that* $b(x, x_m) < \varepsilon$ *whenever* $m \geq k$.

- (M, b) *is a* "left b-sequentially complete space" *if any* left b-Cauchy sequence *converges to a point, according to the topology* τ^+.

It is easy to prove that

i). if $x_n \to x$ according to τ^+ then the sequence (x_n) is a "*left b-Cauchy sequence*".

ii). Any "*forward Cauchy sequence*" (as defined in (2.15)) is a "*left b-Cauchy sequence*" [11].

iii). If $\tau = \tau^+$, then any "*left b-sequentially complete space*" is a "*forward complete metric space*" as defined in this paper.

In case $\tau \neq \tau^+$, the implication may not hold.

iv). Whereas, if $x_n \to x$ according to τ^+ then the sequence (x_n) may fail to be either a "*forward Cauchy sequence*" or a "*backward Cauchy sequence*". [12] Cf. example 4.13.(iv) here.

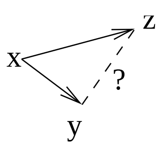
For those reasons, it is not easy to compare the results and examples in the above papers, with the result and examples here presented.

From here on, we return to our setting of "asymmetric metric spaces": we will always use the topology τ.

§3.iii Comparison with symmetric case

There are subtle but important differences between *asymmetric* and *symmetric distances*. We start with a striking remark.

Remark 3.5 *Let* $\varepsilon > 0$ *and* x, y, z *be such that* $b(x, y) < \varepsilon$ *and* $b(x, z) < \varepsilon$. *This does not imply, in general, that* $b(y, z) < 2\varepsilon$



This has consequences such as

i). In example 4.13.(iv) there exists a sequence such that $b(x_n, x) \to 0$ but $x_n \not\to x$. (Note that this example cannot be found in Busemann's "General metric space", due to property (3.1)).

ii). Fix a sequence $(x_n) \subset M$. Suppose that $\forall \varepsilon > 0$ there exists a converging sequence (y_n) such that $b(y_n, x_n) < \varepsilon$. If b is symmetric and (M, b) is complete, then (x_n) converges. If b is asymmetric and (M, b) is complete, then there is a counter-example in 4.13.(vi).

[11] just choose $n = N = k$ and $x = x_n$ in the definition of *left b-Cauchy sequence*

[12] Indeed, Kelly [14] had encountered this problem, which was a motivation of [13]

iii). In example 4.14.(iv) there is a ball $B^+(a,r)$ and points $a_i \notin B^+(a,r)$ such that $B^+(a,r) \subset \bigcup_i B^+(a_i, r/2)$: it is then difficult to find a useful definition of a "precompact set". [13]

Further examples are in §4.i.2 and §4.ii.3.

To circumvent the above problems, we will use often the following propositions.

Proposition 3.6 *If $(x_n) \subset M$ is a sequence such that $b(x, x_n) \to 0$, and M is forward-locally compact, then $x_n \to x$.*

The above may be proved by using this Lemma (that is similar to (2.3) and (2.6) in Zaustinsky's [19])

Lemma 3.7 (modulus of symmetrization) *Let $C \subset M$ be a compact set: then there exists a continuous monotonic non decreasing function $\omega : \mathbb{R}^+ \to \mathbb{R}^+$, with $\omega(0) = 0$, such that*

$$\forall x, y \in C, \quad b(x,y) \le \omega(b(y,x))$$

Proof. Define

$$f(r) = \sup_{x,y \in C,\ b(x,y) \le r} b(y,x)$$

and then f is monotone. Since C is compact, then $f < \infty$ and $\lim_{r\to 0} f(r) = 0$: otherwise we may find $\varepsilon > 0$ and x_n, y_n s.t. $b(x_n, y_n) \to 0$ while $b(y_n, x_n) > \varepsilon$: extracting converging subsequences, we obtain a contradiction.

From f we can define an ω as required, for example $\omega(r) = \frac{1}{r}\int_r^{2r} f(s)ds$ (note that $\omega \ge f$). □

This lemma may also be used as follows

Corollary 3.8 *If (M, b) is locally compact then $\forall x \in M, \varepsilon > 0\ \exists r > 0$ s.t.*

$$B^+(x,r) \subset B^-(x,\varepsilon), \quad B^-(x,r) \subset B^+(x,\varepsilon)$$

and then $\tau = \tau^+ = \tau^-$.

Therefore, in locally compact spaces, asymmetry is not important for *topological* questions (in the sense explained in §3.ii.1); but we still care for asymmetry in *metric* questions (such as completeness). Cf. example 4.14 and sec. §5.i.

4 Study of asymmetric metric spaces

§4.i Properties

Proposition 4.1 *For any $A \subset M$, the following properties are equivalent*

- $\sup_{x,y \in A} b(x,y) < \infty$
- *A is* forward bounded *and* backward bounded
- *A is* symmetrically bounded

[13] An useful definition of "precompact set" is found in [15], where, though, a different completeness hypothesis is used (see 3.4 here).

But note that in example 4.13.(v) there exists a set $A \subset M$ that is *forward bounded* but not *backward bounded* (and vice versa in 4.14.(i)).

Similarly

Proposition 4.2 *Let $(x_n) \subset M$ be a sequence; then the following are equivalent*

- (x_n) *is* forward Cauchy *and* backward Cauchy
- (x_n) *is* symmetrically Cauchy

From that we obtain: if (M, b) is either forward or backward complete, then it is symmetrically complete.

See examples 4.14, 4.15.

Proof. Suppose that (x_n) is *symmetrically Cauchy*: then $\forall \varepsilon > 0\ \exists N$ such that $\forall n, m > N$, $d(x_n, x_m) < \varepsilon$: then $b(x_n, x_m) < \varepsilon$.

Suppose that (x_n) is *forward Cauchy* and is *backward Cauchy*: $\forall \varepsilon > 0\ \exists N''$ such that $\forall n > m > N''$, $b(x_n, x_m) < \varepsilon$, and $\exists N'$ such that $\forall n > m > N'$, $b(x_m, x_n) < \varepsilon$: then we let $N = N' \vee N''$, and $\forall n, m > N$, $d(x_n, x_m) < \varepsilon$.

Suppose that (M, b) is forward complete; let (x_n) be *symmetrically Cauchy*: then it is *forward Cauchy*, and then, since (M, b) is forward complete, there is an x such that $x_n \to x$. Similarly if (M, b) is backward Cauchy. $\square$

Proposition 4.3 *Suppose that (x_n) is* forward Cauchy *and there exist subsequence n_k and a point x such that $\lim_k x_{n_k} = x$: then $\lim_n x_n = x$.*

Proof. fix $\varepsilon > 0$; since (x_n) is forward Cauchy, $\exists N$ such that $\forall m, m'$ with $m' \geq m \geq N$, $b(x_m, x_{m'}) \leq \varepsilon$; let H be such that $n_H \geq N$ and $\forall k \geq H$, $d(x, x_{n_k}) \leq \varepsilon$; for $n \geq n_H$,

$$b(x, x_n) \leq b(x, x_{n_H}) + b(x_{n_H}, x_n) \leq d(x, x_{n_H}) + b(x_{n_H}, x_n) \leq 2\varepsilon$$

at the same time, choosing a large h such that $n_h \geq n$,

$$b(x_n, x) \leq b(x_n, x_{n_h}) + b(x_{n_h}, x) \leq b(x_n, x_{n_h}) + d(x_{n_h}, x) \leq 2\varepsilon$$

so in conclusion $d(x_n, x) \leq 2\varepsilon$ $\square$

§4.i.1 Mid-point properties

In path metric spaces, for any two points x, y, there is always a curve joining them with a quasi optimal distance (that is, very near to $b(x, y)$): this is used in the following proposition, in many different but equivalent fashions, to find *approximate intermediate points* z at a prescribed distance b from x and y.

Proposition 4.4 *Suppose that the (asymmetric) metric space (M, b) is path-metric:*

i). let $\rho > 0$ and

$$S^+(a, \rho) \overset{def}{=} \{y \mid b(a, y) = \rho\}$$

$$D^+(a, \rho) \overset{def}{=} \{y \mid b(a, y) \leq \rho\}$$

(which are closed, since b is continuous) then

$$\overline{B^+(a, \rho)} = D^+(a, \rho), \quad \mathring{S}^+(a, \rho) = \emptyset \tag{4.4.$\star$}$$

ii). $\forall \theta \in (0,1)$, $\forall x, y \in M$

$$\forall \varepsilon > 0 \quad \exists z \in M \ , \quad \textit{such that}$$
$$b(x,z) < \theta b(x,y) + \varepsilon, \qquad b(z,y) < (1-\theta) b(x,y) + \varepsilon \tag{4.4.$\star\star$}$$

iii). let $x, y \in M$ *and* $\varepsilon > 0, \varepsilon \le b(x,y)$ *then*

$$\inf_{z \in S^+(x,\varepsilon)} b(z,y) = b(x,y) - \varepsilon \tag{4.4.$\diamond$}$$
$$\inf_{z \in D^+(x,\varepsilon)} b(z,y) = b(x,y) - \varepsilon \ . \tag{4.4.$\diamond\diamond$}$$

Note that, if there exists a minimum point $\overline{z} \in D^+(x,\varepsilon)$ to (4.4.$\diamond\diamond$), then $b(x,\overline{z}) = \varepsilon$, i.e. $\overline{z} \in S^+(x,\varepsilon)$. [14]

In the general case of a space (M,b), all the above applies to the metric space (M,b^g) (that is path-metric, by 4.44).

On the other hand,

Proposition 4.5 *if M is either forward or backward complete, and any one of the following properties holds:*

- $\exists \theta \in (0,1)$ *s.t.* $\forall x, y$, *property (4.4.$\star\star$) holds,*
- $\exists \theta \in (0,1)$ *s.t.* $\forall x, y$, *when* $\varepsilon = \theta b(x,y)$, *property (4.4.$\diamond$) holds,*
- $\exists \theta \in (0,1)$ *s.t.* $\forall x, y$, *when* $\varepsilon = \theta b(x,y)$, *property (4.4.$\diamond\diamond$) holds,*

then (M,b) is path-metric.

By using Zorn's Lemma, the above may be strengthened to this statement

Proposition 4.6 *if M is symmetrically complete, and any one of the following properties holds:*

- $\forall x, y$, $\exists \theta \in (0,1)$ *s.t. property (4.4.$\star\star$) holds,*
- $\forall x, y$, *there exists* $\varepsilon > 0, \varepsilon < b(x,y)$ *s.t. property (4.4.$\diamond$) holds,*
- $\forall x, y$, *there exists* $\varepsilon > 0, \varepsilon < b(x,y)$ *s.t. property (4.4.$\diamond\diamond$) holds,*

then (M,b) is path-metric.

Regarding the completeness hypothesis, see example 4.11.

All of the above results related the existence of *approximate* intermediate points to the fact that (M,b) be path-metric. If M admits geodesics, then, for any two points x, y, there is always a curve joining them with optimal distance $b^g(x,y)$ (but not necessarily equal to $b(x,y)$): this is again used to find *exact intermediate points z* at a prescribed distance b^g from x and y; this is at the base of the definition of "**Menger convexity**"; unfortunately, there are, in the literature, two different definitions:

- M is "*Menger (weakly) convex*" if given two different points $x, y \in M$, a third (different from x, y) point z exists such that $b^g(x,z) + b^g(z,y) = b^g(y,z)$;

[14] Indeed, $b(x,\overline{z}) \ge b(x,y) - b(\overline{z},y)$ by triangular inequality, and $b(x,y) - b(\overline{z},y) = \varepsilon$ by (4.4.$\diamond\diamond$); while $b(x,\overline{z}) \le \varepsilon$ since $\overline{z} \in D^+(x,\varepsilon)$

- M is "*Menger (strongly) convex*" if given any two points $x, y \in M$, and any $0 \le \lambda \le b^g(y,z)$, a third point z exists such that $b^g(x,z) = \lambda$, $b^g(z,y) = \lambda - b^g(y,z)$;

the weaker definition being used for example in Busemann [4] in proposition (1.16), the stronger in more modern treaties.

Example 4.7 *The subset $M = \mathbb{R}^3 \setminus \{(x_1,x_2,x_3), x_1 = x_2 = 0\}$ obtained by deleting a line from 3-space, endowed with the Euclidean distance $b(x,y) = |x-y|$, is a simple example of a space that is path metric, it is Menger weakly convex, but not Menger strongly convex. It is also* locally compact *but not* finitely compact *(as defined in pages 8 and 9) and not complete.*

Other similar examples are in 4.11.

It is easy to see that, if b admits geodesics then it is Menger convex (both in weaker and stronger sense). Vice versa,

Proposition 4.8 *if M is either forward or backward complete, and it is Menger strongly convex, then (M,b) admits geodesics.*

If M is symmetrically complete, and it is Menger weakly convex, then (M,b) admits geodesics.

In the case when Menger convexity is intended in the weak form, the proof needs Zorn's Lemma. The above result is similar to proposition (1.16) in [4], (where though the hypothesis was "M is Menger (weakly) convex and finitely compact").

Remark 4.9 *We may add an intermediate statement: if M is either forward or backward complete, and*

$$\begin{gathered}\exists \theta \in (0,1)\ ,\ \ \forall x,y \in M\ ,\ \ \exists z \in M\ ,\ \ \textit{such that}\\ b^g(x,z) = \theta b^g(x,y), \quad b^g(z,y) = (1-\theta) b^g(x,y)\ .\end{gathered} \tag{4.9.$\star$}$$

then (M,b) admits geodesics.

The above property (4.9.$\star$) is stronger than the weak Menger convexity, and weaker than the stronger. [15] Still, it is a condition that entails convexity, as seen in this example

Example 4.10 *Let $b(x,y) = |x-y|$, $M \subset \mathbb{R}^n$. If M is closed in $\mathbb{R}^n$ and (4.9.$\star$) holds, then M is convex (in the usual Euclidean sense).*

§4.i.2 Examples

We consider several examples, for a better understanding of the above properties and definitions.

This subset of $\mathbb{R}^2$ complements the already seen example 4.7

Example 4.11 *The set $M = (\mathbb{R} \times \mathbb{Q}) \cup (\{0\} \times \mathbb{R})$, equipped with $b(x,y) = |x-y|$, admits approximate intermediate points, in the various forms seen in proposition 4.4; and it is Menger weakly convex; and M is Lip-arc-connected; but (M,b) is not complete; and (M,b) is not path-metric, since $b^g(x,y) = |x_1| + |y_1| + |x_2 - y_2|$ when $x_2 \ne y_2$.*

[15] Indeed "*Menger weak convexity*" may be equivalently stated as: $\forall x,y \in M\ \exists \theta \in (0,1)$ s.t. $b^g(x,z) = \theta b^g(x,y), \quad b^g(z,y) = (1-\theta) b^g(x,y)\ .$

Example 4.12 *Consider $M \subset \mathbb{R}^n$ to be an open set, and b to be the Euclidean distance; then*

- *(M, b) is locally compact and locally path-metric; whereas*
- *(M, b) admits geodesics iff M is convex,*
- *and (M, b) is complete iff $M = \mathbb{R}^n$;*
- *if moreover M is equal to the interior of the closure of M in $\mathbb{R}^n$, then (M, b) is path-metric iff M is convex.*

Example 4.13 *Let M be the disjoint union of segments*

$$M \stackrel{def}{=} \bigcup_{n \in \mathbb{Z}} (\{n\} \times [0,1]) / \sim$$

with $\sim$ identifying all 0s into a class $[0]$, and 1s into $[1]$, i.e. $(n, 0) \sim (m, 0)$ and $(n, 1) \sim (m, 1)$. Let $l_n = \{n\} \times (0, 1)$. Let b be defined on segments as

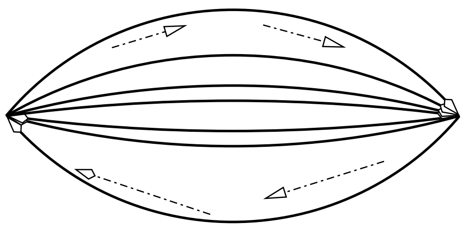
$$b(x,y) \stackrel{def}{=} \begin{cases} |x-y|(1+e^n) & \text{if } x \geq y \text{ and } x, y \in l_n \\ |x-y|(1+e^{-n}) & \text{if } x \leq y \text{ and } x, y \in l_n \end{cases}$$

and

$$b([0],[1]) = b([1],[0]) = 1$$

and extend b (geodesically) using the rule

$$b(x,y) = \inf \big\{ b(x,y),\ \ b(x,[0]) + b([0],y),\ \ b(x,[1]) + b([1],y) \big\}$$



We highlight the following properties

i). *(M, b) is a complete path-metric space.*

ii). *(M, b) is not locally compact.*

iii). *There is no geodesic connecting $[0]$ to $[1]$*

iv). *Let $x_n = (-n, 1/n)$ then*

$$b(x_n, [0]) = (1 + e^{-n})/n \to 0$$

but

$$b([0], x_n) = (1 + e^{n})/n \to \infty$$

Figure 2: **f**orward and **b**ackward **b**alls, **l**eft and **r**ight extrema (M is vertical, $r = 1/2$, a in abscissa)

v). Moreover the set $A = \{[0], x_n \mid n\}$ *is backward bounded but not forward bounded.*

vi). Moreover $\forall \varepsilon > 0$ *we can define*

$$y_n = \begin{cases} x_n & \textit{if } n\varepsilon < 2 \\ [0] & \textit{if } n\varepsilon \geq 2 \end{cases}$$

and then $b(x_n, y_n) < \varepsilon$ *and* $y_n \to [0]$*; but nonetheless* $x_n \not\to [0]$.

Example 4.14 *Let* $M = \mathbb{R}$ *and let*

$$b(x, y) = \begin{cases} e^y - e^x & \textit{if } x < y \\ e^{-y} - e^{-x} & \textit{if } x > y \end{cases} \tag{4.14.$\star$}$$

then b *generates on* $\mathbb{R}$ *the usual topology, and* (M, b) *is locally compact and path-metric (this may be proved using 5.6 since* b *comes from a Finsler structure: see in sec. 5.3.1 in [16]). The balls are the open intervals*

$$B^+(a, r) = \{y \mid b(a, y) < r\} = \big(-\log(r + e^{-a}), \log(r + e^a)\big)$$

$$B^-(a, r) = \{x \mid b(x, a) < r\} = \big(\log(e^a - r), -\log(e^{-a} - r)\big)$$

where $\log(z) = -\infty$ *if* $z \leq 0$. *(see fig. 2 and 3)*

i). We immediately note that $B^-(0, 1) = \mathbb{R}$*, that is,* M *is backward bounded (but is not forward bounded).*

Figure 3: **f**orward and **b**ackward **b**alls, **l**eft and **r**ight extrema (M is vertical, $a = 1/2$, r in abscissa)

ii). Let $x_n = n$: then for $m < n$, $b(x_n, x_m) = e^{-m} - e^{-n} < e^{-m}$ so this sequence is backward-Cauchy: then M is not backward complete.

iii). (M, b) is forward complete (and then is symmetrically complete, by 4.2). Proof: Suppose that (x_n) is an increasing forward-Cauchy sequence: then for $m > n > N$, $b(x_n, x_m) = e^{x_m} - e^{x_n} < \varepsilon$ that implies $x_m \leq \log(\varepsilon + e^{x_n})$, so x_m has a limit $x_m \to x$[16]; while if it were a decreasing forward-Cauchy sequence for $m > n > N$, $b(x_n, x_m) = e^{-x_m} - e^{-x_n} < \varepsilon$ that implies $x_m \geq -\log(\varepsilon + e^{-x_n})$, and again x_m has a limit. Suppose that x_n is a generic forward-Cauchy sequence: from any subsequence x_{nk} of x_n we may extract a monotonic sub-sub-sequence $(x_{nkh})_h$: this would be convergent to a point x; this point does not depend on the choice of subsequence: indeed, $b(x_n, x_{nkh}) \to_h b(x_n, x) < \varepsilon$ for n large. $\square$

iv). Consider the ball $B^+(0, \rho)$ of extrema $(-R, R)$ with $R(\rho) = \log(\rho + 1)$, and then the two balls

$$B^+(R,r) = \big(-\log(r+e^{-R}), \log(r+e^{R})\big), \quad B^+(-R,r) = \big(-\log(r+e^{R}), \log(r+e^{-R})\big)$$

then if $r \geq \rho/(\rho + 1)$,

$$B^+(0,\rho) \subset B^+(R,r) \cup B^+(-R,r)$$

Proof: indeed the right extrema of $B^+(-R,r)$ is positive when $\log(r + e^{-R}) \geq 0$ that is $(r + e^{-\log(\rho+1)}) \geq 1$ that is $(r + 1/(\rho + 1)) \geq 1$ $\square$

[16] $x_m \to x$ in the topological sense, i.e. $d(x_m, x) \to 0$

Example 4.15 *Consider two copies (M, b) of the above space, join them at the origin, reverse the metric on one: the resulting space $\tilde{M}, \tilde{b}$ is symmetrically complete, but is neither forward nor backward complete.*

More precisely: let $M_+ = \mathbb{R} \times \{+\}$, $M_- = \mathbb{R} \times \{-\}$,

$$\tilde{M} = M_+ \cup M_- / \sim$$

where $(0, +) \sim (0, -)$. Let

$$\tilde{b}(x, y) = \begin{cases} b(x, y) & x, y \in M_+ \\ b(y, x) & x, y \in M_- \\ \tilde{b}(x, [0]) + \tilde{b}([0], y) & \textit{otherwise} \end{cases}$$

§4.ii Geodesics

Lemma 4.16 (change of variable) *Let $\gamma : [\alpha, \beta] \to M$, $\xi : [\alpha_1, \beta_1] \to M$ be linked by*

$$\gamma = \xi \circ \varphi$$

where $\varphi : [\alpha, \beta] \to [\alpha_1, \beta_1]$ is monotone non decreasing, and continuous, and $\varphi(\alpha) = \alpha_1$, $\varphi(\beta) = \beta_1$. Then

$$\mathrm{len}^b \gamma = \mathrm{len}^b \xi$$

Proof. For any partition $T = \{t_i\}$ of $[\alpha, \beta]$ we associate the partition $S = \{s_i = \varphi(t_i)\}$ of $[\alpha_1, \beta_1]$: in this case,

$$\sum_{i=1}^{n} b\big(\gamma(t_{i-1}), \gamma(t_i)\big) = \sum_{i=1}^{n} b\big(\xi(s_{i-1}), \xi(s_i)\big) \tag{4.17}$$

Similarly, for any partition $S = \{s_i\}$ of $[\alpha_1, \beta_1]$ we choose $t_i \in \varphi^{-1}(\{s_i\})$, and associate the partition $T = \{t_i\}$ of $[\alpha, \beta]$: then again (4.17) is verified regardless of the choice of t_i: indeed if $t'_i \in \varphi^{-1}(\{s_i\})$ then $\gamma(t_i) = \gamma(t'_i)$. □

Remark 4.18 *The above lemma holds also when φ is not bijective: this is very powerful, since it simplifies the next lemma, and leads to a very short and simple proof of 4.24 and of 4.29.*

The following is a long–known but very powerful result (that may be found in section I in [4]; see also in Theorem 4.2.1 in [1], that, for the symmetric case, also discusses the *metric derivative* issue):

Lemma 4.19 (reparametrization to arc parameter) *For any rectifiable curve $\gamma : [\alpha, \beta] \to M$ of length L such that ℓ^γ is continuous, there exists an unique Lipschitz curve $\xi : [0, L] \to M$ such that*

$$\gamma(t) = \xi(\ell^\gamma(t)), \quad \ell^\xi(t) = t, \quad t \in [0, L] \tag{4.19.$\star$}$$

where ℓ^γ is the running length *of γ, and ℓ^ξ of ξ (see eq. (2.11)).*

ξ is the reparametrization to arc parameter *of γ.*

Proof. $\gamma(t) = \xi(\ell^\gamma(t))$ uniquely defines ξ: indeed, if $\ell^\gamma(t) = \ell^\gamma(t')$ then $\gamma(t) = \gamma(t')$ by (2.12). Let $\hat{l} = \ell^\gamma(\hat{t})$. If we restrict γ to $[a, \hat{t}]$ and ξ to $[0, \hat{l}]$, we can write $\gamma = \xi \circ \ell^\gamma$. By the previous lemma, $\ell^\gamma(\hat{t}) = \ell^\xi(\hat{l})$, that is $\hat{l} = \ell^\xi(\hat{l})$. □

In general we will say that a curve ξ is *parametrized by arc parameter* when $\ell^\xi(t) = t$. Combining (2.12) and (4.19.$\star$), we can state that, for any curve ξ parametrized by arc parameter, for $0 \le s < t \le L$,

$$b^g\big(\xi(s), \xi(t)\big) \le t - s \tag{4.20}$$

This leads to the following definition (due to Busemann [5], [6]):

Definition 4.21 (Segments) *A curve $\gamma : I \to M$ (where $I \subset \mathbb{R}$ is an interval) is a segment if*

$$\forall s, t \in I, s < t \ \ , \ \ t - s = b^g(\gamma(s), \gamma(t)) \tag{4.21.$\star$}$$

The *"time parameter"* is arbitrary: if $\gamma(\cdot)$ is a segment, then $\gamma(\cdot + t)$ is a segment as well.

Unfortunately, a segment may fail to be continuous: this was seen in example 2.13 where the curve γ is a segment but is not continuous. The above example cannot be found in Busemann's *general metric spaces*, since it contradicts the extra hypothesis (3.1). That example is one of the reasons why we stressed the necessity that curves be continuous in the definition 2.10 of what a geodesic is.

Segments enjoy some nice properties:

- by 4.43, a segment is always parametrized by arc parameter, and the length of a segment defined on $I = [\alpha, \beta]$ is $\beta - \alpha$;
- then a continuous segment γ is a minimal geodesic connecting $\gamma(s)$ to $\gamma(t)$, $\forall s, t \in I, s < t$.

Vice versa,

Lemma 4.22 (almost segment) *Suppose $\gamma : [0, L] \to M$ is a curve parametrized by arc parameter, connecting $\gamma(0) = x$ to $\gamma(L) = y$; and suppose that*

$$\mathrm{len}\gamma \le b^g(x, y) + \varepsilon$$

for $\varepsilon \ge 0$; then for any $0 \le s < t \le L$,

$$t - s - \varepsilon \le b^g\big(\gamma(s), \gamma(t)\big) \le t - s \tag{4.22.$\star$}$$

(that is an approximate version of (4.21.$\star$)); the rightmost inequality is (4.20); whereas

$$\begin{aligned} b^g(x, y) \quad &\le \quad b^g\big(x, \gamma(s)\big) + b^g\big(\gamma(s), \gamma(t)\big) + b^g\big(\gamma(s), y\big) \le \\ &\le \quad s + b^g\big(\gamma(s), \gamma(t)\big) + L - t \ \ . \end{aligned}$$

again by (4.20).

So (choosing $\varepsilon = 0$) we obtain that a minimal geodesic is a continuous segment.

Segments can be joined, as follows:

Lemma 4.23 (joining) *Suppose that $\gamma : [0, L] \to M$, $\gamma : [0, L'] \to M$, are geodesic segments. Suppose that $z = \gamma'(0) = \gamma(L)$: then the segments may be joined to form a continuous curve $\xi : [0, L + L'] \to M$, by defining*

$$\xi(t) = \begin{cases} \gamma(t) & \textit{for } t \in [0, L] \\ \gamma'(t - L) & \textit{for } t \in [L, L + L']. \end{cases}$$

Then ξ is a geodesic segment if and only if

$$b^g\big(\xi(0), \xi(L + L')\big) = L + L' \ \ .$$

§4.ii.1 Existence of geodesics

The following results rephrase (5),(6),(7) in section I in [6], but without assuming that (M, b) be path–metric (by properly distinguishing b by b^g):

Theorem 4.24 *Fix x. Let $\rho > 0$. Suppose*

$$D^{+g}(x,\rho) \stackrel{def}{=} \{y \mid b^g(x,y) \le \rho\}$$

is compact (in the (M, b) topology[17]). Then for any y such that $b^g(x,y) \le \rho$, there is a geodesic connecting x to y.

Note that it is not sufficient to assume that $D^+(x,\rho)$ is compact (instead of D^{+g}): see in example 4.34. It is instead enough to require that $D^+(x,\rho)$ is compact for all $\rho > 0$ (as noted also in remark 4.3.3. in [1]).

In the same spirit we state those corollaries

Corollary 4.25

- *Suppose $D^{+g}(x,\rho)$ and*

  $$D^{-g}(x,\rho) \stackrel{def}{=} \{y \mid b^g(y,x) \le \rho\}$$

  are compact, and $b^g(z,x) + b^g(x,y) \le \rho$ and $b^g(y,x) + b^g(x,z) \le \rho$, then z, y may be joined by a geodesic.
- *As a consequence, if $D^{+g}(x,2\rho)$ and $D^{-g}(x,2\rho)$ are compact, and $z, y \in D^{-g}(x,\rho) \cap D^{+g}(x,\rho)$, then z, y may be joined by a geodesic.*

We now prove the theorem; this needs some extra steps w.r.t. the proof in General Metric Spaces (where (3.1) is assumed to hold).

Proof. Fix $x, y \in M$, $\rho > 0$, as above. We will write D^{+g} instead of $D^{+g}(x,\rho)$ for brevity. By lemma 3.7, let ω be the modulus of symmetrization of $D^{+g}(x,\rho)$; let $\tilde\omega(r) \stackrel{\text{def}}{=} \sup\{r, \omega(r)\}$: then

$$d(z,y) \le \tilde\omega\big(b(z,y)\big) \tag{4.26}$$

for any $z, y \in D^{+g}$.

Let $L = b^g(x,y)$; suppose $y \ne x$ (otherwise $\gamma \equiv x$ is the geodesic). Let $\gamma_n : [0, L_n] \to M$ be a sequence of continuous paths from x to y such that $L_n \stackrel{\text{def}}{=} \mathrm{len}\gamma_n \to L$. By using the above lemma 4.19, we assume without loss of generality that γ_n are parametrized by arc parameter; by (4.26) above,

$$d\big(\gamma_n(t), \gamma_n(s)\big) \le \tilde\omega(|t-s|) \tag{4.27}$$

for all $s, t \in [0, L_n]$. To confront different paths, we extend so that $\gamma_n(t) = \gamma_n(L_n)$ for $t > L_n$.

Suppose also that $L = b^g(x,y) < \rho$ for simplicity; then definitively $L_n \le \rho$: by (2.12), we know that all of γ_n is contained in D^{+g}. Combining this argument and (4.27) we can apply the Ascoli-Arzelà theorem: we know that there is a $\gamma : [0, L] \to M$ (that again satisfies (4.27)) such that, up to a subsequence, there is uniform convergence of $\gamma_n \to \gamma$; this uniform convergence is w.r.t the distance $d(x,y) = b(x,y) \vee b(y,x)$.

[17] We recall that if $D \subset M$ is compact in (M, b^g) then it is compact in (M, b); the opposite is not true, as shown by example 2.9.

The functional $\gamma \mapsto \operatorname{len}^b \gamma$ is lower semicontinuous w.r.t. uniform convergence, since it is the supremum of the continuous functionals

$$\sum_{i=1}^{n} b\big(\gamma(t_{i-1}), \gamma(t_i)\big)$$

so we conclude that γ is a geodesic connecting x to y.

When $L = b^g(x, y) = \rho$, if γ_n is frequently wholly contained in D^{+g}, all works as above; otherwise the proof is obtained by a slight change in the above argument. Let t_n be the last of the times such that $\gamma_n([0, t]) \subset D^{+g}$, that is,

$$t_n \stackrel{\text{def}}{=} \inf\{s \mid \gamma_n(s) \notin D^{+g}(x, \rho)\} = \inf\{s \mid b^g(x, \gamma_n(s)) > \rho\}$$

then $t_n \geq \rho$, since γ_n is continuous and by arc parameter. Define

$$\tilde{\gamma}_n(t) = \begin{cases} \gamma_n(t) & t < t_n \\ \gamma_n(t_n) & t \geq t_n \end{cases} \quad ; \tag{4.28}$$

since $\tilde{\gamma}_n$ are wholly contained in D^{+g}, we can apply the above reasoning to say that there is a continuous $\tilde{\gamma}$ such that $\tilde{\gamma}_n \to \tilde{\gamma}$ uniformly. To conclude the proof we need to prove that $\tilde{\gamma}(L) = y$: since $L = \rho \leq t_n \leq L_n$, then $t_n \to L$; moreover

$$b(\gamma_n(t_n), y) \leq b^g(\gamma_n(t_n), y) = L_n - \rho$$

so by (4.26) again, $\gamma_n(t_n) \to y$: the sequence $\tilde{\gamma}_n$ is uniformly equicontinuous, this implies that $\tilde{\gamma}(L) = y$. □

Those results immediately imply what is nowadays known as *Busemann's theorem*

Theorem 4.29 *Suppose that (M, b) is compact and Lipschitz-arc-connected: then it admits geodesics.*

§4.ii.2 Hopf-Rinow Theorem

We now restate and prove theorem 1.2. To this end we prove some preliminary lemma.

Lemma 4.30 (radius of compactness) *Let $D^+(a, \rho) \stackrel{def}{=} \{y \mid b(a, y) \leq \rho\}$.*
We define the forward radius of compactness *$R : M \to \mathbb{R}^+ \cup \{\infty\}$ as*

$$R(x) \stackrel{def}{=} \sup\{\rho \geq 0 \mid D^+(x, \rho) \textit{ is compact }\} \quad .$$

Then, for any $\rho < R(x)$, $D^+(x, \rho)$ is compact.

Either $R \equiv \infty$, or $R < \infty$ and R is 1-Lipschitz w.r.t. d: indeed, for any $x, y \in M$, for any $\rho < R(x)$, we have that $D^+(y, \rho - b(x, y))$ is compact, since

$$D^+(y, \rho - b(x, y)) \subset D^+(x, \rho) \quad .$$

This implies that $R(y) \geq \rho - b(x, y)$, and then $R(y) \geq R(x) - b(x, y)$; if R is finite, the above entails $d(x, y) \geq |R(x) - R(y)|$, since x, y are arbitrary.

In general (even when $R(x) < \infty$) it is possible to find examples where $D^+(x, R(x))$ is compact, and examples where it is not. In path–metric and forward-locally compact spaces, instead, we can precisely describe the behaviour of $R(x)$ as follows

Lemma 4.31 *Suppose that (M, b) is path–metric and forward-locally compact. Choose $x \in M, \rho > 0$ such that $R(x) < \infty$ and $D^+(x, \rho)$ is compact; let $\delta \stackrel{def}{=} \min_{y \in D^+(x,\rho)} R(y)$ (note that $\delta > 0$); then $R(x) = \rho + \delta$.*

Proof. Since (M, b) is forward-locally compact, then $\delta > 0$. Choose $0 < t < s < \delta$; we want to prove that

$$V \stackrel{\text{def}}{=} \bigcup_{y \in D^+(x,\rho)} D^+(x, t)$$

is compact; indeed if $y_n \in V$, then $y_n \in D^+(x_n, t)$, for a choice of $x_n \in D^+(x, \rho)$; up to a subsequence, $x_n \to x \in D^+(x, \rho)$, so that for n large, $b(x, x_n) \le s - t$ hence $b(x, y_n) \le s$, that is, y_n is contained in the compact set $D^+(x, s)$, so we can extract a converging subsequence.

Assuming that (M, b) is path–metric, then $V \supset D^+(x, \rho + t)$ (by using (4.4.$\diamond\diamond$), with $\varepsilon = \rho$), so $V \supset D^+(x, \rho + t)$ (by (4.4.$\star$)) and then $D^+(x, \rho + t)$ is compact; and also $V = D^+(x, \rho + t)$, by existence of geodesics 4.24. This implies that $R(x) \ge \rho + t$, and we conclude by arbitrariety of t that $R(x) \ge \rho + \delta$. The opposite inequality is easily inferred from the previous lemma, where it was proved that $R(y) \ge R(x) - b(x, y)$, that implies $\delta \ge R(x) - \rho$. $\square$

A corollary of the above lemma is that (in the above hypothesis) $D^+(x, R(x))$ is not compact; so the above lemma is the quantitative version of the argument shown in (8) in section I in [6].

Lemma 4.32 *Suppose that (M, b) is path–metric, for simplicity*[18]. *Suppose that $0 < R(x) < \infty$, let $\rho = R(x)$; then for any y with $b(x, y) = \rho$ there is a continuous geodesic segment $\gamma : [0, \rho) \to M$ such that $\gamma(0) = x$ and*

$$b(\gamma(t), y) = \rho - t \ \ \forall t \in [0, \rho]. \tag{4.32.$\star$}$$

(or, equivalently, γ is segment on $[0, \rho]$ but it is continuous only on $[0, \rho)$).

Proof. Let $z_1 = x$, and $\rho_n = (1 - 2^{-n})\rho$; we would like to define z_{n+1} as a minimum point for the problem

$$\min_{z \in D^+(z_n, 2^{-n}\rho)} b(z, y) \ \ ; \tag{4.32.$\star\star$}$$

iteratively, using 4.4, we may prove those facts: $b(x, z_n) \le \rho_{n-1}$ and

$$D^+(z_n, 2^{-n}\rho) \subset D^+(x, \rho_n) \tag{4.32.$\diamond$}$$

where the r.h.s is compact: so the above problem has a minimum z_{n+1} that satisfy $b(z_n, z_{n+1}) = 2^{-n}\rho$, $b(x, z_{n+1}) \le \rho_n$ and

$$b(z_{n+1}, y) = 2^{-n}\rho \ \ . \tag{4.32.$\diamond\diamond$}$$

Using (4.32.$\diamond$) and 4.24, there exists a geodesic segment γ_n connecting z_n to z_{n+1} (of length $2^{-n}\rho$). By using (4.32.$\star\star$),(4.32.$\diamond\diamond$) and the triangle inequality for $h > n$

$$b(z_n, z_h) + 2^{-h+1}\rho = b(z_n, z_h) + b(z_h, y) \ge b(z_n, y) = 2^{-n+1}\rho$$

[18] It is possible to state a more general lemma without assuming that (M, b) is path–metric; but the added complexity is not useful in the following theorem

but on the other side

$$b(z_n, z_h) \le \sum_{j=n}^{h-1} b(z_j, z_{j+1}) = \sum_{j=n}^{h-1} \rho 2^{-j}$$

we obtain that

$$b(z_n, z_h) = (2^{-n+1} - 2^{-h+1})\rho$$

so by the joining lemma 4.23 we can join all the segments γ_n to build the required segment γ. By (4.32.$\diamond\diamond$), we obtain (4.32.$\star$). $\square$

The example 2.13 shows that we cannot expect that $\gamma(t)$ be continuous at ρ in general: indeed the "identity curve" $\gamma : [0,1] \to [0,1]$ is a segment but is not continuous at 0.

We now use the above lemmas to prove the asymmetric Hopf–Rinow theorem (in a more detailed version than what was seen in the introduction):

Theorem 4.33 (Hopf-Rinow) *Suppose that (M, b) is path-metric and forward-locally compact; then the following are equivalent*

i). forward-bounded and closed sets are compact,

ii). (M, b) is forward complete,

iii). any continuous geodesic segment $\gamma : I \to M$ defined on $I = [\alpha, \beta)$ may be completed to a continuous geodesic segment on $[\alpha, \beta]$,

and all imply that (M, b) admits geodesics.

In general, the implications i$\Longrightarrow$ii$\Longrightarrow$iii *hold for any asymmetric metric space.*

Proof.

- Suppose that forward-bounded closed sets are compact. If x_n is a forward-Cauchy sequence, then there exists N s.t. $b(x_N, x_m) \le 1$ for $m > N$, that is, $x_m \in D^+(x_N, 1)$ that is compact; then we can extract a converging subsequence, and use lemma 3.7 to obtain the result.

- Suppose that (M, b) is forward complete, let γ be the segment; then as $t_n \uparrow \beta$, the sequence $\gamma(t_n)$ is forward–Cauchy, so there is an unique limit $\lim_n \gamma(t_n)$ that we use to define $\gamma(\beta)$.

- Suppose that any segment $\gamma : I \to M$ defined on $I = [\alpha, \beta)$ may be completed: we will prove that forward–bounded closed sets are compact (that is, that the *radius of compactness* $R(x) \equiv \infty$).

  We proceed by contradiction, and suppose that there is a x s.t. $\rho = R(x) < \infty$: we will show that $D^+(x, \rho)$ is compact, contradicting 4.31. To this end let $\{y_n\} \subset D^+(x, \rho)$. Let $L_n = b(x, y_n)$.

  If $\liminf_n L_n < \rho$, we can extract a subsequence n_k s.t. $L_{n_k} \le t < \rho$, that is, $\{y_{n_k}\} \subset D^+(x, t)$ that is compact: so we can extract a converging subsequence.

  Suppose now that $\lim_n L_n = \rho$. For any y_n s.t. $L_n < \rho$, we use 4.24 to obtain the minimal geodesic segments $\gamma_n : [0, L_n] \to M$ connecting x to y_n; we extend γ_n constantly to $[L_n, \rho]$ (as was done in eqn. (4.28)). When instead $L_n = \rho$ we use 4.32 to define $\gamma_n : [0, \rho) \to M$, and define $\gamma_n(\rho) = y_n$. We use Ascoli–Arzelà theorem and a diagonal argument to find a subsequence n_k and a curve $\gamma : [0, \rho) \to M$ such that for each $0 < s < \rho$, $\lim_k \gamma_{n_k}(t) = \gamma(t)$

uniformly for $t \in [0, s]$. Since the length functional is lower-semicontinuous, we obtain that γ is a continuous geodesic segment on $[0, \rho)$; by hypothesis, it can be completed to a continuous segment on $[0, \rho]$. Let $y = \gamma(\rho)$ Fix now $\varepsilon > 0$; since γ is continuous, there is a t s.t. $b(y, \gamma(s)) < \varepsilon$ for all $s, t \leq s \leq \rho$; fix $s = t \vee \rho - \varepsilon$; we apply the triangular inequality to prove that

$$b(y, y_n) \leq b(y, \gamma(s)) + b(\gamma(s), \gamma_n(s)) + b(\gamma_n(s), y_n) \;\; ;$$

this is less than 3ε, definitively in n; so $b(y, y_n) \to 0$, and then $y_n \to y$ (by prop. 3.6).

- Existence of geodesics is guaranteed by theorem 4.24.

□

We remark that the proof of the above equivalence cannot simply follow from the proof for metric spaces §1.11 in [12], since that proof uses the property (ii) in sec. §3.iii; neither it does follow from the proof in Finsler Geometry (see section VI of [3]), since the latter uses the exponential map.

Hence we devised a different proof, that combines the idea of the diagonalization of a sequence of subsequences, as in [12], and the idea of repeated application of (4.4.$\diamond\diamond$) as in [3]; it is similar to the proof for *general metric spaces* in §I in [6], but it does not use the hypothesis (3.1), and it exploits the special ingredients of the *radius of compactness* Lemma and of 4.3.

§4.ii.3 More examples

We show some examples to highlight the impact of the hypotheses in the above theorems 4.33 (a.k.a. 1.2) and 4.24.

We tweak the example 2.9 a bit, to build this:

Example 4.34 *Consider*

$$\begin{aligned} M = \{x \in \mathbb{R}^2 \mid -1 \leq x_1 \leq 1, x_1 \neq 0, x_2 = 0\} \cup \\ \{x \in \mathbb{R}^2 \mid -1 \leq x_1 \leq 1, \;\; 1/4 \geq x_2 > 0\} \cup \\ \bigcup_n \{x \in \mathbb{R}^2 \mid -1 \leq x_1 \leq 1, \;\; x_2 = (x_1 - 1)/n\} \end{aligned}$$

and b the Euclidean distance (see fig. 4 on the next page). Let $A = (-1, 0)$, $B = (1, 0)$. (M, b) is locally compact, but is not path-metric and is not complete; (M, b) does not admit geodesics locally around A, that, the points A and $x_n = (-1, -2/n)$ do not admit a minimal connecting geodesic.

The disc $\{y \mid b(B, y) \leq 1/2\}$ is compact for b; but the discs $\{y \mid b^g(B, y) \leq \varepsilon\}$ are never compact for b^g.

Example 4.35 *Let*

$$M \stackrel{def}{=} \bigcup_n \{x \in \mathbb{R}^2 \mid x_1^2 + n^2 x_2^2 = 1, x_2 \geq 0\}$$

and b be the geodesic distance induced on M by the Euclidean distance; loosely speaking, M is the disjoint union of countable segments l_n, with length $\sim 2 + 1/n$, and with common end points.

Then (M, b) is path-metric, is complete, and is bounded, and it locally admits geodesics; but is not locally compact (and then it is not compact), and there is no geodesic curve connecting $(-1, 0)$ to $(1, 0)$.

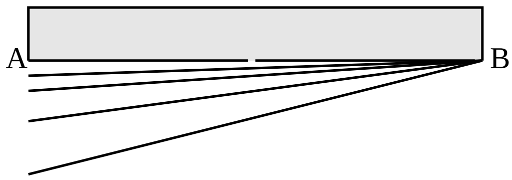


Figure 4: example 4.34

§4.iii Complements

§4.iii.1 Generating (path) metrics

For any asymmetric metric b, we may build many other asymmetric metrics $\tilde b$, as follows

Proposition 4.36 *Let $\varphi : \mathbb{R}^+ \to \mathbb{R}^+$ be continuous and concave, $\varphi(x) = 0$ only for $x = 0$; let $\tilde b \stackrel{def}{=} \varphi \circ b$: then $\tilde b$ is an asymmetric distance.*

The above family $\tilde b \stackrel{\text{def}}{=} \varphi \circ b$ is quite large; hence the following question arises: what happens if we induce $\tilde b^g$ from $\tilde b$: can we build a large family of path metrics on M, by this simple procedure? The answer is no.

Proposition 4.37 *Set everything as in the previous proposition. If moreover the derivative of φ exists and is finite at 0, then*

$$\varphi'(0)\mathrm{len}^b\gamma = \mathrm{len}^{\tilde b}\gamma$$

and

$$b^g(x,y)\varphi'(0) = \tilde b^g(x,y) \quad .$$

Proof. Let $\varepsilon > 0$. In the definition (2.5) of $\mathrm{len}^b\gamma$ it is not restrictive to use only subsets T of $[\alpha,\beta]$ such that $b(\gamma(t_i),\gamma(t_{i+i})) \le \varepsilon \ \forall i \in \{1,\dots,n-1\}$.

Let now $\delta > 0$; then there exists a $\varepsilon > 0$ such that

$$x(a-\delta) \le \phi(x) \le x(a+\delta) \quad \forall x \in [0,\varepsilon]$$

with $a = \varphi'(0)$; we obtain that

$$(a-\delta)\mathrm{len}^b\gamma \le \mathrm{len}^{\tilde b}\gamma \le (a+\delta)\mathrm{len}^b\gamma$$

hence the conclusion □

If $\varphi'(0) = \infty$, wild things may happen: see for example 1.4.b in [12].

It is also possible to prove that we cannot generate other path metrics by iterating the operation (2.7): indeed $b^g = (b^g)^g$; see 4.44.

§4.iii.2 Length structure

We hereby define a *length structure* $\mathrm{len} : \mathcal{C} \to \mathbb{R}^+$, where $\mathcal{C}$ is a family of curves $\gamma : [\alpha,\beta] \to M$ (and α,β may vary); we say that [19]

[19] the following conditions are not independent: some of them imply some others; we do not detail

- $\mathcal{C}$ *connects points*: $\forall x, y \in M$ there is at least one $\gamma \in \mathcal{C}$, $\gamma : [\alpha, \beta] \to M$ such that $\gamma(\alpha) = x, \gamma(\beta) = y$
- $\mathcal{C}$ is *reversible*: if $\gamma \in \mathcal{C}$ and $\hat{\gamma}(t) \stackrel{\text{def}}{=} \gamma(-t)$ then $\hat{\gamma} \in \mathcal{C}$
- len is *independent of reparametrization*: if $\gamma \in \mathcal{C}$ then, $\forall h > 0, k \in \mathbb{R}$, its linear reparametrization $\gamma'(s) = \gamma(hs + k)$ is in $\mathcal{C}$, and $\text{len}\gamma = \text{len}\gamma'$,
- len is *monotonic*: if $\gamma \in \mathcal{C}$ then its restriction $\gamma' = \gamma|_{[c,d]}$ is in $\mathcal{C}$, and $\text{len}\gamma' \leq \text{len}\gamma$
- len is *additive*: if $\gamma', \gamma'' \in \mathcal{C}$, $\gamma' : [\alpha', \beta'] \to M$, $\gamma : [\alpha'', \beta''] \to M$, and $\beta' = \alpha''$, then we may join the two curves and obtain γ: then γ is in $\mathcal{C}$, and

$$\text{len}\gamma = \text{len}\gamma' + \text{len}\gamma''$$

- len is *run-continuous*: that is, the running length $t \mapsto \text{len}(\gamma|_{[\alpha,t]})$ is continuous for all $\gamma \in \mathcal{C}$
- $\text{len}\gamma = 0$ when γ is constant (and constants are in $\mathcal{C}$)

We then define $b(x, y)$ on M to be the infimum of this length $\text{len}\gamma$ in the class of all $\xi \in \mathcal{C}$, $\xi : [0, 1] \to M$ with given extrema $\xi(0) = x, \xi(1) = y$.

Proposition 4.38 *b is an asymmetric semidistance: b satisfies the triangle inequality, since* len *is* independent of reparametrization *and* additive.

$b \geq 0$, and $b(x, x) = 0$ since len *is zero on constant curves.*

b may fail to be an asymmetric distance, since we cannot be sure that $b(x, y) = 0 \Longrightarrow x = y$

We also induce $\text{len}^b\gamma$ from b using the total variation, as in (2.5).

Lemma 4.39 *Suppose* len *and $\mathcal{C}$ satisfy all the hypotheses above. If* len *is lower semi continuous w.r.t. uniform convergence, then* $\text{len}\gamma = \text{len}^b\gamma$ *for any $\gamma \in \mathcal{C}$*

The proof for the asymmetric case is not different from the symmetric case (cf. 1.6 [12]). We provide here a detailed proof, for convenience of the reader.

Proof. Fix $\gamma \in \mathcal{C}$. We insert len into the definition (2.5) of $\text{len}^b\gamma$:

$$\text{len}^b\gamma = \sup_T \sum_{i=1}^n \inf_{\xi_i} \text{len}\xi_i \tag{4.40}$$

where the inf is done on $\xi_i : [0, 1] \to M, \xi_i \in \mathcal{C}$, connecting $\gamma(t_{i-1})$ to $\gamma(t_i)$; equivalently (since len is independent of reparametrization) we may choose $\xi_i : [t_{i-1}, t_i] \to M$; so we can join all the ξ_i and write

$$\text{len}^b\gamma = \sup_T \inf_{\xi \in \Xi_T} \text{len}\xi \tag{4.41}$$

where the sup is done on all choices of T finite subset of $[0, 1]$ and the inf is done on the class Ξ_T of $\xi \in \mathcal{C}$, $\xi : [0, 1] \to M$, such that $\xi(t) = \gamma(t)$ for $t \in T$.

Since we could choose $\xi = \gamma$, we immediately obtain that

$$\text{len}^b\gamma \leq \text{len}\gamma \tag{4.42}$$

Fix a T: by (4.41),

$$\inf_{\xi \in \Xi_T} \mathrm{len}\xi \le \mathrm{len}^b \gamma$$

Let $\eta > 0$. Define $\Xi_{T,\eta}$ as the class of ξ in Ξ_T such that $\mathrm{len}\xi \le \mathrm{len}^b\gamma + \eta$ and, for any interval $[t_i, t_{i+1}]$ defined by T,

$$\mathrm{len}(\xi|_{[t_i,t_{t+1}]}) \le \mathrm{len}(\gamma|_{[t_i,t_{t+1}]})$$

Then

$$\inf_{\xi \in \Xi_T} \mathrm{len}\xi = \inf_{\xi \in \Xi_{T,\eta}} \mathrm{len}\xi$$

We drop η for simplicity.

Since len is run-continuous, then $\forall \delta > 0$, there is a $\varepsilon > 0$ such that, for any interval (s,t) with $|t-s| < \varepsilon$,

$$\mathrm{len}(\gamma|_{[s,t]}) < \varepsilon, \quad \mathrm{len}(\hat{\gamma}|_{[-t,-s]}) < \varepsilon$$

where $\hat{\gamma}(t) \stackrel{\text{def}}{=} \gamma(-t)$; and then $b(\gamma(t),\gamma(s)) < \delta$, $b(\gamma(s),\gamma(t)) < \delta$, by our definition of b. We define the density $|T|$ as

$$|T| \stackrel{\text{def}}{=} \sup_i (t_{i+1} - t_i)$$

when $T = \{t_1 < t_2 < \dots t_n\}$.

Consider any T with density $|T| \le \varepsilon$; for any point $s \in [0,1]$, there $\exists t, t' \in T$, with $|t'-t| \le \varepsilon$ and $t \le s \le t'$:

$$b(\xi(s),\gamma(s)) \le b(\xi(s),\xi(t)) + b(\gamma(t),\gamma(s)) \le \mathrm{len}(\xi|_{[t,s]}) + \delta \le$$
$$\le \mathrm{len}(\xi|_{[t,t']}) + \delta \le \mathrm{len}(\gamma|_{[t,t']}) + \delta \le 2\delta$$

and similarly $b(\gamma(s),\xi(s)) \le 2\delta$

We can order the sets T by inclusion, so that the class of such sets is a direct class; moreover $T \subset T'$ implies

$$\inf_{\xi \in \Xi_T} \mathrm{len}\xi \le \inf_{\xi \in \Xi_{T'}} \mathrm{len}\xi$$

so we can write

$$\sup_T \inf_{\xi \in \Xi_T} \mathrm{len}\xi = \lim_{T \to \infty} \inf_{\xi \in \Xi_T} \mathrm{len}\xi$$

As $T \to \infty$, we have $|T| \to 0$, and then $\xi \to \gamma$ uniformly; since $\mathrm{len}\gamma$ is l.s.c by hypothesis, then

$$\mathrm{len}^b\gamma = \lim_T \inf_{\Xi_T} \mathrm{len}\xi \ge \mathrm{len}\gamma$$

and then $\mathrm{len}\gamma = \mathrm{len}^b\gamma$. $\square$

Corollary 4.43 *Suppose we are given an asymmetric space (M,b); let $\mathcal{C}$ be the class of rectifiable curves with continuous run-length, and $\mathrm{len} \stackrel{def}{=} \mathrm{len}^b$: this is a length structure satisfying all above hypotheses.* len^b *induces* b^g *which induces* $\mathrm{len}^{b^g}\gamma$. $\mathrm{len}^b\gamma$ *is lower semi continuous since it is the supremum of the continuous functionals*

$$\sum_{i=1}^n b\big(\gamma(t_{i-1}),\gamma(t_i)\big)$$

By the lemma

$$\mathrm{len}^b\gamma = \mathrm{len}^{b^g}\gamma$$

This immediately entails the result:

Corollary 4.44 *The operation $b \mapsto b^g$ is idempotent, that is, b^g is path-metric, idem est $b^g = (b^g)^g$.*

5 Applications

§5.i Asymmetric metric manifold

Suppose now that M is a differential manifold [20], with an atlas $\mathcal{A}$ and a topology τ^M.

Definition 5.1 ($\mathcal{A}$-Lip) *We will say that a curve $\xi : [\alpha, \beta] \to M$ is locally Lipschitz w.r.t. $\mathcal{A}$, if for any local chart $\varphi : U \to \mathbb{R}^n$ in the atlas $\mathcal{A}$, $\varphi \circ \xi$ is locally Lipschitz for $t \in \xi^{-1}(U)$.*

Hypotheses 5.2 *Suppose that we are given a positive function $F : TM \to \mathbb{R}^+$; suppose*

- *F is lower-semi-continuous,*
- *$v \mapsto F(x, v)$ is positively 1-homogeneous, that is,*

$$F(x, \lambda v) = F(x, v)\lambda \quad \forall \lambda \geq 0$$

- *$F(x, v) = 0$ iff $v = 0$.*

We define the *length* $\mathrm{len}^L \gamma$ of a locally Lipschitz curve $\xi : [0, 1] \to M$ as

$$\mathrm{len}^L \gamma = \int_0^1 F(\xi(s), \dot{\xi}(s))\, ds \tag{5.3}$$

Remark 5.4 (The regular Finsler case) *If we would suppose that*

- *F is continuous, and C^∞ on the slit tangent bundle $TM \setminus 0$*
- *$v \mapsto F^2(x, v)$ is strongly convex [21] for $v \neq 0$, $\forall x$,*

then this section would be exactly what is found in section 6.2 [3]; we will instead use less regular assumptions.

As in §4.iii.2, we then define the asymmetric distance $b(x, y)$ on M to be the infimum of this length $\mathrm{len}^L \gamma$ in the class of all locally Lipschitz ξ with given extrema $\xi(0) = x, \xi(1) = y$.

In the above hypotheses 5.2, b is an asymmetric distance; indeed, b is a semidistance by 4.38, and the relation $b(x, y) \implies x = y$ can be proved by the methods in the following Lemma. This metric is also called a *Finsler metric*, and is naturally associated to Hamilton-Jacobi equations; see [18] and references therein.

There may be a confusion here, since we have two topologies at hand: the topology τ^M of the differential manifold M, and the topology τ^b induced by b; but we may use this lemma, that simply restates the lemma 6.2.1 in [3] (with fewer regularity assumptions on F)

[20] more precisely, let M be a finite dimensional connected smooth differential manifold, without boundary
[21] that is, setting $f(v) = F^2(x, v)$ then in local coordinates the Hessian of f is positive definite for $v \neq 0$

Lemma 5.5 *Assume 5.2 and let F be locally bounded from above. Then for any $z \in M$ there exists a compact neighborhood U of z, associated with local coordinates $\varphi : U \to W \subset \mathbb{R}^n$, such that the push forward of b in local coordinates is bounded from above and below by the Euclidean norm: that is, $\exists c' > c > 0$*

$$c'|\hat{x} - \hat{y}| \geq b(x, y) \geq c|\hat{x} - \hat{y}| \tag{5.5.$\star$}$$

where $\hat{y} = \varphi(y)$ and $\hat{x} = \varphi(x)$.

Proof. Let $z \in M$ and Z be a neighborhood of z, with $\overline{Z}$ compact w.r.t. $\tau^{\mathcal{A}}$, associated to local coordinates $\varphi : Z \to W \subset \mathbb{R}^n$, with W convex bounded.

Let $\hat{F}$ be the push-forward of F; for any $x, y \in Z$, we may join $\hat{y} = \varphi(y)$ to $\hat{x} = \varphi(x)$ with a segment, and then $b(x, y) \leq c'|\varphi(x) - \varphi(y)|$ where c' is the upper bound of $\hat{F}(w, v)$ for $w \in W$ and $|v| \leq 1$.

In the above setting, let $U \subset\subset Z$, let c be the minimum of $\hat{F}(w, v)$ for $w \in \varphi(Z)$ and $|v| = 1$ (this minimum exists and is positive, since F is l.s.c); and let ε be the minimum of $|\varphi(x) - \varphi(z)|$ for $x \in \partial U, z \in \partial Z$; let $\hat{\xi} \stackrel{\text{def}}{=} \varphi \circ \xi$ when $\xi \in Z$.

Then, for any ξ connecting x to y,

$$\text{len}^L \xi = \int_0^1 F(\xi(s), \dot{\xi}(s))\, ds \geq \int_0^t c|\dot{\hat{\xi}}(s)|\, ds \geq c(\varepsilon \wedge |\hat{x} - \hat{y}|)$$

where $t = 1$ if ξ never exits from Z, otherwise it is the first time s such that $\xi(s) \in \partial Z$: then $b(x, y) \geq c(\varepsilon \wedge |\hat{x} - \hat{y}|)$, and we choose a smaller U. □

Corollary 5.6 *Under the same hypotheses, $\tau^b = \tau^M$, i.e. b generates the same topology that the atlas of M induces. Moreover both coincide with the topology generated by the forward balls; or respectively, by the backward balls.* (22)

This topology is locally compact: then

- *$x_n \to x$ iff $b(x_n, x) \to 0$ iff $b(x, x_n) \to 0$* (23)*; and*
- *(M, b) locally admits geodesics (by 4.24).*

By (5.5.$\star$), a curve $\xi : [0, 1] \to M$ is locally Lipschitz w.r.t $\mathcal{A}$, as defined in 5.1, iff it is locally Lipschitz w.r.t b, as defined in §2.ii.

Proposition 5.7 *i). (M, b) is path-metric, that is, $b = b^g$*

ii). suppose 5.2, and $v \mapsto F(x, v)$ is convex; then for any Lipschitz γ, $\text{len}^L \gamma$ coincides with the $\text{len}^b \gamma$ defined in (2.5).

Proof. i) is a consequence of 4.4. (24) By prop. 4.4, we need to check that $\forall x, y \in M, \forall \varepsilon > 0\ \exists z \in M$ such that $b(x, z) < b(x, y)/2 + \varepsilon$ and $b(z, y) < b(x, y)/2 + \varepsilon$. Fix x, y. Let $\varepsilon > 0$ and ξ connecting x to y such that $\text{len}^L \xi < b(x, y) + 2\varepsilon$; let

$$f(t) = \int_0^t F(\xi(s), \dot{\xi}(s))\, ds$$

(22) cf. 3.8 , or sec. 6.2C in [3] for the regular case

(23) cf. 6.2.5 in [3] for the regular case

(24) If $v \mapsto F(x, v)$ is convex, then this is an immediate consequence of the second point. Alternatively, we have a proof that does not need that $L(x, \cdot)$ is convex: indeed the operation of computing b implicitly replaces F by its relaxed l.s.c.

then f is continuous, so let t_0 be the time when $f(t_0) = f(1)/2$, and $z = \xi(t_0)$; by reparametrizing, $\xi'(s) = \xi(st_0)$, $\xi''(s) = \xi(1-(1-s)(1-t_0))$; then since $b(x,z) < \mathrm{len}^L\xi' = f(1)/2 < b(x,y)/2+\varepsilon$, and $b(z,y) < \mathrm{len}^L\xi'' = f(1)/2 < b(y,z)/2+\varepsilon$.

ii). The proof is similar to the proof of lemma 4.39. We set $\mathcal{C}$ to be the class of Lipschitz curves, and $\mathrm{len} \stackrel{\text{def}}{=} \mathrm{len}^L$. Fix γ, let $L \stackrel{\text{def}}{=} \mathrm{len}^L\gamma$. If L is small, then the curve is contained in local coordinates. Indeed, choose, as per 5.5, U a compact neighborhood of $\gamma(0)$, associated to local coordinates $\varphi : U \to W \subset \mathbb{R}^n$. We suppose for the moment that $L < R$, with $B^+(\gamma(0), R) \subset U$: then $\gamma(t) \in U\ \forall t \in [0,1]$. We fix $\eta = (R-L)/2$ in the proof of the aforementioned lemma 4.39: then any curve $\xi \in \Xi_{T,\eta}$ is wholly contained in U: indeed by (4.42), $\mathrm{len}^b\xi \le \mathrm{len}^L\xi \le \mathrm{len}^b\gamma + \varepsilon \le L+\varepsilon < R$. As $T \to \infty$, $\xi \to \gamma$ uniformly.

By (5.5.$\star$), the uniform convergence happens also in local coordinates. By well-known theory (cf. thm. 2.3.3 in [7]), $\mathrm{len}^L\gamma$ is l.s.c, and we conclude as in the Lemma 5.5.

If we don't suppose that $\gamma(t) \in U\ \forall t \in [0,1]$: we fix a very dense partition $\hat{T}$, so that we can associate to any point $t_i \in \hat{T}$ a local chart $\varphi_i : U_i \to \mathbb{R}^n$, so that $\bigcup_i U_i$ covers the image of γ, and more precisely $\gamma([t_i, t_{i+1}]) \subset U_i$. We repeat the above argument in any U_i. □

Lemma 5.8 (Reparametrization) *For any $\gamma : [\alpha,\beta] \to M$ Lipschitz, define $l : [\alpha,\beta] \to [0,L]$*

$$l(t) = \int_a^t F(\gamma(s), \dot\gamma(s))\, ds$$

(where $L = \mathrm{len}\gamma$).

Then there is a unique ξ such that $\gamma = \xi \circ l$, and

$$F(\xi(s), \dot\xi(s)) = 1 \qquad \tilde\forall s \in [0,L] \tag{5.8.$\star$}$$

Proof. We apply 4.19, and just note that $\ell^\xi(t) = t$ means that $\int_0^t F(\xi,\dot\xi) = t$ □

We can then state this version of the *Hopf-Rinow* theorem

Theorem 5.9 *Assume that the function $F : TM \to \mathbb{R}^+$*

- *is locally bounded from above*
- *$v \mapsto F(x,v)$ is positively 1-homogeneous, and convex,*

and 5.2 holds.

Then the asymmetric metric space (M,b) is path-metric, locally compact, and locally admits geodesics (that we can always reparametrize so that they satisfy (5.8.$\star$)).

If (M,b) satisfies any (=all) of the conditions i,ii,iii *in the Hopf–Rinow theorem 4.33, then for any $x,y \in M$, there is a locally Lipschitz curve ξ connecting them that minimizes (5.3), satisfying (5.8.$\star$).*

Further applications of these ideas are in [16]; in particular, in the appendix of [16], we consider a Lagrangian $L : TM \to \mathbb{R}^+$ (that is defined as $L(x,v) = F(x,v)^2$) and its Legendre-Fenchel dual Hamiltonian $H : T^*M \to \mathbb{R}^+$; we show that, if $H \in C^{1,1}$ and $p \mapsto H(x,p)$ is positively 2-homogeneous and strictly convex, we can define an exponential map and add it to the statement of the Hopf-Rinow theorem, to obtain a statement exactly as in ch. VI of [3].

Acknowledgements

The author thanks Prof. S. Mitter, who has reviewed and corrected various versions of this paper, and suggested many improvements; and Prof. L. Ambrosio for the discussions.

Contents

References

[1] L. Ambrosio and P. Tilli. *Selected topics in "analysis in metric spaces"*. appunti. edizioni Scuola Normale Superiore, Pisa, 2000.

[2] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savarè. *Gradient flows in metric spaces and in the Wasserstein space of probability measures*. Birkhüser, 2004.

[3] D. Bao, S. S. Chern, and Z. Shen. An introduction to Riemann-Finsler geometry. (October 1, 1999 version).

[4] H. Busemann. Local metric geometry. *Trans. Amer. Math. Soc.*, 56:200–274, 1944.

[5] H. Busemann. *The geometry of geodesics*, volume 6 of *Pure and applied mathematics*. Academic Press (New York), 1955.

[6] H. Busemann. *Recent syntetic differential geometry*, volume 54 of *Ergebnisse der Mathematik und ihrer Grenzgebiete*. Springer Verlag, 1970.

[7] G. Buttazzo. *Semicontinuity, relaxation and integral representation in the calculus of variation*. Number 207 in scientific & technical. Longman, 1989.

[8] A. Duci and A. Mennucci. Banach-like metrics and metrics of compact sets. in preparation, 2007.

[9] A. Fathi. Weak KAM theorem in Lagrangian dynamics. (Preliminary version, 2 Feb 2001).

[10] W. H. Fleming and H. M. Soner. *Controlled Markov processes and viscosity solutions*. Springer–Verlag, 1993.

[11] P. Fletcher and W.F.Lindgren. *Quasi-uniform spaces*, volume 77 of *Lecture notes in pure and applied mathematics*. Marcel Dekker, 1982.

[12] M. Gromov. *Metric Structures for Riemannian and Non-Riemannian Spaces*. Birkhäuser, 1999.

[13] M.K.Vamanamurthy I.L.Reilly, P.V.Subrahmanyam. Cauchy sequences in quasi-pseudo-metric spaces. *Monat.Math.*, 93:127–140, 1982.

[14] J. C. Kelly. Bitopological spaces. *Proc. London Math. Soc.*, 13(3):71–89, 1963.

[15] H. P. Künzi. complete quasi-pseudo-metric spaces. *Acta Math. Hung.*, 59(1-2):121–146, 1992.

[16] A. C. G. Mennucci. Regularity and variationality of solutions to Hamilton-Jacobi equations. part ii: variationality, existence, uniqueness. 1st (draft) version, online since 2005 at `http://cvgmt.sns.it/papers/men05/`, 2nd version ibidem, 2006.

[17] B. B. Phadke. Nonsymmetric weakly complete g-spaces. *Fundamenta Mathematicae*, 1974.

[18] Antonio Siconolfi. Metric character of Hamilton-Jacobi equations. *Transactions of the american mathematical society*, 355(5):1987–2009, 2003. electronically published on January 8, 2003.

[19] E. M. Zaustinsky. *Spaces with non-symmetric distances*. Number 34 in Mem. Amer. Math. Soc. AMS, 1959.