

A nonlinear d'Alembert comparison theorem and causal differential calculus on metric measure spacetimes

Tobias Beran,¹ Mathias Braun,² Matteo Calisti,³
Nicola Gigli,⁴ Robert J. McCann,⁵ Argam Ohanyan,⁶
Felix Rott,⁷ Clemens Sämann⁸

March 20, 2025

Abstract We introduce a variational first-order Sobolev calculus on metric measure spacetimes. The key object is the maximal weak subslope of an arbitrary causal function, which plays the role of the (Lorentzian) modulus of its differential. It is shown to satisfy certain chain and Leibniz rules, certify a locality property, and be compatible with its smooth analog. In this setup, we propose a quadraticity condition termed *infinitesimal Minkowskianity*, which singles out genuinely Lorentzian structures among Lorentz–Finsler spacetimes. Moreover, we establish a comparison theorem for a nonlinear yet elliptic p -d'Alembertian in a weak form under the timelike measure contraction property. As a particular case, this extends Eschenburg's classical estimate past the timelike cut locus.

Keywords Nonsmooth gravity, general relativity, Lorentzian length space, metric measure spacetime, d'Alembert comparison, p -Laplacian, strong energy condition, timelike Ricci curvature bounds, low regularity, Finsler spacetime.

2020 Mathematics Subject Classification Primary 51K10, 83C75; Secondary 35J92, 35Q75, 49Q22, 53C21, 53C50.

¹Faculty of Mathematics, University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria.
tobias.beran@univie.ac.at

²Institute of Mathematics, EPFL, 1015 Lausanne, Switzerland. mathias.braun@epfl.ch

³Faculty of Mathematics, University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria.
matteo.calisti@univie.ac.at

⁴SISSA, Via Bonomea 265, 34136 Trieste, Italy. ngigli@sissa.it

⁵Department of Mathematics, University of Toronto, 40 St. George Street Room 6290, Toronto, Ontario M5S 2E4, Canada. mccann@math.toronto.edu

⁶Department of Mathematics, University of Toronto, 45. St. George Street Room PG 111, Toronto, Ontario M5S 2E5, Canada. argam.ohanyan@utoronto.ca

⁷SISSA, Via Bonomea 265, 34136 Trieste, Italy. frott@sissa.it

⁸Faculty of Mathematics, University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria.
clemens.saemann@univie.ac.at

Contents

1	Introduction	2
1.1	Order-completeness; tangent notions for worldlines, curves of measures, and measures over worldlines	5
1.2	Smooth paradigm: Lagrangian-Hamiltonian duality	6
1.3	The Sobolev cone of cotangent fields; infinitesimal Minkowskianity	8
1.4	Nonsmooth Lagrangian–Hamiltonian duality	10
1.5	Timelike Ricci curvature bounds and applications	11
1.6	Two-sided perturbations, hyperbolic norms, compatibility with smooth notions	13
2	Curves in the nonsmooth Lorentzian setting	14
2.1	Infinity conventions	14
2.2	Notions of spacetimes	14
2.3	The spacetime of Borel probability measures	17
2.4	Causal speed	22
2.5	Left-continuous causal paths and their topology	28
2.6	Action and geodesics	31
2.7	Lifting curves of probability measures to measures on curves	35
2.8	Non-branching notions	41
3	Lorentzian differential and Sobolev calculus	44
3.1	Some basic properties of causal functions	44
3.2	Test plans on metric measure spacetimes	48
3.3	Weak subslopes	49
3.4	Calculus rules	53
4	Horizontal (inner) and vertical (outer) derivatives	56
4.1	Horizontal differentiation	56
4.2	Perturbation of causal functions	58
4.3	Vertical right-differentiation	62
4.4	Relation of horizontal and vertical derivatives	67
5	Effects of timelike Ricci curvature assumptions	70
5.1	Synthetic timelike curvature-dimension bounds	70
5.2	Good geodesics	72
5.3	A converse Hawking–King–McCarthy theorem	80
5.4	A metric Brenier–McCann theorem	83
5.5	Nonlinear p -d’Alembert comparison	86
5.6	Distributional p -d’Alembert operator	91
5.7	Modifications for past TMCP spaces	93
A	Appendices	96
A.1	Compatibility with the smooth setting	96
A.2	Null distance and perturbations of causal functions	100
A.3	Hyperbolic norms and polarization on cones	103
	References	106

1 Introduction

Einstein’s theory of gravity is set on a smooth Lorentzian manifold, but the singularity theorems of Hawking [67] and Penrose [103] show one cannot expect smoothness to continue to hold in the far past (and future): the Big Bang and black hole singularities are unavoidable consequences of

the theory. Furthermore, the linearization of the Einstein field equations which relate the Ricci curvature of spacetime to its matter content — being a wave equation — propagates singularities (in the form of gravitational waves) without smoothing them. For these reasons and others discussed in [75] and its references, it appears highly desirable to relax the smooth Lorentzian framework of Einstein’s theory to allow for rougher, more flexible, settings.

Developments in positive (Riemannian) signature provide a model for such relaxations. The first is the metric geometry of Alexandrov [2], Gromov [66], and others, where sectional curvature bounds are defined using triangle comparison. The second is the metric measure geometry developed following work of Fukaya [56] and Bakry–Émery [15], in which a measure is also used to define Ricci bounds, as in

- the Ricci limit spaces of Cheeger–Colding [45, 46, 47],
- the curvature-dimension spaces $\text{CD}(K, N)$ of Sturm [113, 114] and Lott–Villani [78], which satisfy a lower bound of K on the Ricci curvature and upper bound of N on the dimension expressed using refined versions of McCann’s entropic displacement convexity [82, 102, 48, 120], and
- the Riemannian $\text{RCD}(K, N)$ spaces of Ambrosio–Gigli–Savaré and Gigli, which for $N = \infty$ [11] and $N \in (1, +\infty)$ [60] single out among curvature-dimension spaces those which are infinitesimally Hilbertian.

It is in the last setting that several of the most powerful results of the theory have been achieved, such as the splitting theorem of Gigli [63] and its consequences [70, 92], a second-order calculus [61], and the non-branching result of Deng [50]. Here ideas related to calculus of variations, gradient or heat flows, Sobolev calculus and integration by parts play a central role.

Recently, in Lorentzian signature, several analogous steps have been taken. Building on earlier works such as [29, 74, 14, 3, 97, 111], Kunzinger–Sämman proposed a theory of Lorentzian (pre)length spaces — based largely on a time separation function — and timelike triangle comparison [75]. (They also required a causal order, but this can be embedded into a modified time separation function which takes both signs as in McCann [84, 85].) By adding a measure, Cavalletti–Mondino were able to develop a theory of timelike curvature-dimension spaces $\text{TCD}_q^e(K, N)$ which manifest nonsmooth timelike lower Ricci bounds [39]. These are based on entropic convexity estimates from the smooth setting, that had there been shown to be equivalent to the strong energy condition of Hawking–Penrose by McCann [84] and to the Einstein field equations by Mondino–Suhr [93]. A burst of developments have followed, e.g. Braun [20, 21], Braun–McCann [25], McCann [85], Ketterer [71], and Cavalletti–Mondino [40] with Manini [34]. Of course there is also an abundance of work beyond the scope of Ricci curvature in this setting, including, e.g., by Burtscher–García-Heveling [28], Kunzinger–Steinbauer [76], Müller [94, 95], Minguzzi–Suhr [91], Beran–Harvey–Napper–Rott [16], Beran–Ohanyan–Rott–Solis [17], etc.

Just as the $\text{CD}(0, N)$ spaces of Lott–Sturm–Villani include all Banach spaces and Finsler manifolds of dimension $n \leq N$ [99], the $\text{TCD}_q^e(0, N)$ spaces of Cavalletti–Mondino include all Finsler spacetimes of dimension $n \leq N$ [26]. One of the goals of the present manuscript is to develop a nonsmooth criterion which distinguishes those timelike curvature-dimension spaces whose time separation function behaves as if it came from a signed inner-product; we call such spaces *infinitesimally Minkowskian*.

Definition 1.1 (Lorentzian TCD spaces). *We will say a metric measure spacetime satisfies the Lorentzian (entropic) timelike curvature-dimension condition, briefly $\text{LTCD}_q^e(K, N)$, if it is an infinitesimally Minkowskian $\text{TCD}_q^e(K, N)$ metric measure spacetime; (c.f. (1.11) below).*

This terminology is in analogy with the Riemannian curvature dimension condition of [11, 60]. Just as infinitesimal Hilbertianity [60] is essential for the metric splitting theorem of Gigli [63], so we expect infinitesimal Minkowskianity to play an equally crucial role in obtaining a nonsmooth analog of the Eschenburg, Galloway and Newman splitting theorems from Lorentzian geometry [54, 57, 96]. Having this goal in mind, we also pursue the development of a nonsmooth Lorentzian

analog of a negative-homogeneity d'Alembert operator, which — like the positive-homogeneity Riemannian p -Laplace operator and in sharp contrast to the the usual 1-homogeneous wave operator — is elliptic, albeit non-uniformly, on non-decreasing (viz. causal) functions. Indeed, we establish that for $0 \neq p < 1$ the p -d'Alembert operator satisfies a comparison inequality, which estimates its value on Lorentz distance functions by a bound in terms of K and N that becomes sharp in constant curvature spaces. This is an important step toward a nonsmooth metric splitting theorem under timelike Ricci rather than sectional curvature [17] bounds. To do so requires us to develop a nonsmooth Sobolev calculus for (future-directed) causal curves and functions, which may be of independent interest. One of the challenges we face is that such functions — due to the monotonicity enforced by the arrow of time — form a convex Sobolev cone, rather than a Sobolev space. In separate work inspired by the present one, five of us [23] show our p -d'Alembert comparison bound leads to a new proof of the Lorentzian splitting theorems due to Eschenburg, Galloway and Newman in the smooth setting. Using this p -d'Alembertian approach to the Lorentzian splitting theorems, Caponio–Ohanyan–Ohta [33] are able to prove a splitting theorem for Finsler spacetimes, generalizing a previous such result due to Lu–Minguzzi–Ohta [79]. Moreover, our method can be used to prove Lorentzian splitting theorems for metrics of low regularity, see [24].

Setting and overview

Let us now give an overview of the strategy, organization and results of the present paper. As detailed in §2, we work in a setting slightly different from that of previous authors; we do not require any form of global hyperbolicity and replace it with some form of order completeness. This allows our framework to accommodate infinite dimensions, possibly facilitating its amenability to an eventual passage to a quantum theory of gravity, in whatever form that may take. We start by introducing the concept of metric spacetime, namely a set M equipped with a (signed) time separation function $\ell: M \times M \rightarrow \{-\infty\} \cup [0, +\infty]$ which (is zero on the diagonal and) satisfies the reverse triangle inequality. We use ℓ to define the transitive relations $\{\ell \geq 0\}$ and $\{\ell > 0\}$; these relations are called *causal* and *chronological*, and denoted by \leq and \ll respectively. To develop any sort of calculus we need to have a topology τ on M : rather than fixing a predetermined one, we shall leave the choice of topology independent from that of time separation, asking only for minimal relations. We shall typically assume that τ is Polish and that the set $\{\ell \geq 0\}$ is closed in M^2 and call the resulting structure (M, τ, ℓ) a Polish metric spacetime. Then, inspired by the Dedekind rather than Cauchy's strategy for constructing the real numbers, we say that (M, τ, ℓ) is *forward* if every non-decreasing sequence $(y_i)_{i \in \mathbf{Z}} \subset M$ which is order-bounded $x \leq y_i \leq y_{i+1} \leq z$ admits a future limit $y_{+\infty} = \lim_{i \rightarrow +\infty} y_i$. Similarly (M, τ, ℓ) is *backward* if every such sequence admits a past limit $y_{-\infty} = \lim_{i \rightarrow -\infty} y_i$. In principle, one might decide to work with spacetimes that are both forward and backward, but we eschew this assumption as we are skeptical about its stability (see also the discussion in [58]). Our most important existence results (such as the limit curve Theorem 2.30 or the lifting Theorem 2.43) are stated on forward spacetimes, typically under the additional relation between τ and ℓ that the topology is *locally causally convex*, i.e. every point has a neighbourhood basis made of causally convex sets. All these results could equally well be formulated in backward spacetimes. We show forward metric spacetimes lie somewhere between causally simple and globally hyperbolic metric spacetimes on the nonsmooth causal ladder (Remark 2.11).

A curve $\gamma: [0, 1] \rightarrow M$ is called *causal* if $s \leq t$ implies $\gamma_s \leq \gamma_t$; dually, a function $f: M \rightarrow [-\infty, +\infty]$ is called *causal* if $x \leq y$ implies $f(x) \leq f(y)$. Since our space M may not be a manifold, we need to develop structures to compensate for the absence of a tangent and cotangent space. We rely on causal curves for the former purpose, and causal functions for the latter. The nonsmoothness of our spaces prevents us from defining tangent or cotangent vectors pointwise; instead — inspired by [10, 60] — we represent causal tangent fields using probability measures π on causal curves, and (exact, future-directed) cotangent fields using causal functions f or more heuristically, their differentials. The operation of horizontally differentiating f along π , which corresponds to a global average of the initial derivative of f along the vector field represented by π , defines a positively bilinear pairing of these objects. These proxies for causal tangent and cotangent fields are the

subjects of §2 and §3, respectively. We shall also need to set up a convex-analytic Legendre transform to identify tangent fields with cotangent fields and vice versa; this operation depends on ℓ and a conjugate pair of exponents $p^{-1} + q^{-1} = 1$ with $0 \neq q < 1$ and is homogeneous but — as in [60] and in contrast to both smooth Lorentzian and nonsmooth infinitesimally Hilbertian geometry — is nonlinear and often multivalued. These challenges — and the fact that causal functions form a convex cone rather than a vector space — are addressed in §4, where we also compute the “vertical” (i.e. outer variational) derivative of the p -Dirichlet energy at f , and relate it to the tangent field proxies π corresponding to the nonlinear Legendre-transform of the cotangent field proxied by f . Inspired by analogous results concerning the nonlinear Laplacian which arises in the Finslerian [101], Hamiltonian [100], and nonsmooth [60] settings, in §5 we define and estimate relative to constant curvature models the p -d’Alembert operator $\square_p f$ in terms of the curvature and dimension parameters (K, N) of the forward metric measure spacetime.

Remark 1.2 (Beyond global hyperbolicity and $q < 0$). Although we allow $0 \neq q < 1$ of both signs, in the Lorentzian optimal transport literature to date [52, 115, 84, 93, 39, 38, 21, 20, 25], only the cost function ℓ^q/q with $q \in (0, 1)$ (sometimes with $q = 1$ included) has been considered. However, since the function $u_q(z) := z^q/q$ (c.f. Section 2.1) has the same monotonicity and concavity properties for $q \in (0, 1)$ as for $q < 0$, the majority of the fundamental results established in the above literature can be generalized to the case $q < 0$. Moreover, global hyperbolicity is almost universally assumed in the references cited above, but it turns out that for most purposes it suffices that the given spacetime be forward (c.f. Definition 2.1). Any result we mention in this work for $0 \neq q < 1$ on forward spacetimes while citing previous works belongs to the category of results that can be generalized in this manner; it is mainly in §5 that such results are required and each of these few is flagged by a reference to Remark 1.2. The detailed treatment of these generalizations to the setting $q < 0$ (and even more general Lorentz costs) and to forward spacetimes is the content of an ongoing work in progress [32]. ■

We now lay out the contents of each section in further detail, with an interlude between sections §1.1 and §1.3 describing more of the smooth motivation for the ensuing constructions. Before doing so, let us note that inspired by the current work, a very satisfactory and more constructive approach to the existence of the p -d’Alembert operator has been developed by Braun [22]. This enables him not only to establish many of its fundamental properties for the first time, but also to give an exact representation formula covering most cases of interest using one-dimensional localization and especially [39, 40, 25] in q -essentially timelike non-branching spaces [21].

1.1 Order-completeness; tangent notions for worldlines, curves of measures, and measures over worldlines

Much like Cauchy-completeness is central in classical metric geometry to obtain many of the key existence results in that setting, so is the order-completeness encoded in the definition of forward (or backward) spacetimes in our setting for similar purposes. The first instance where this naturally occurs is in the class of curves we consider: in a forward spacetime any causal curve bounded from above and defined on a dense subset of $[0, 1]$ including zero can uniquely be extended to a left continuous causal curve on $[0, 1]$ and if the underlying spacetime is Polish, then so is the space of left-continuous causal curves equipped with a natural L^1 -distance. Under suitable compactness assumptions on the set of curves or the underlying spacetime, this completeness of the space of curves improves to compactness (Theorem 2.30), a result that we interpret as a limit curve theorem in our setting (c.f. [86]) and is reminiscent of Helly’s selection theorem.

A leitmotiv of this work will be reliance on causal monotonicity in Lorentzian signature instead of the Lipschitz or absolute continuity typically required in Riemannian signature. After laying out the details of our setting, the remainder of §2 is largely devoted to defining the speed of a causal curve. Heuristically, one would expect this to be given pointwise by a limit such as

$$|\dot{\gamma}_t| = \lim_{h \downarrow 0} \frac{\ell(\gamma_t, \gamma_{t+h})}{h}. \quad (1.1)$$

Instead, we take this limit in the distributional sense and call the resulting measure the *causal speed* $|\dot{\gamma}|$ of γ . Moreover, we show $|\dot{\gamma}|$ coincides on the interval $(0, 1)$ with the maximum non-negative Radon measure ν satisfying $\nu((s, t)) \leq \ell(\gamma_s, \gamma_t)$ for all $0 < s < t < 1$. The latter is shown to exist and to be unique by differentiation of the time separation function restricted to the curve in §2.4; its absolutely continuous part agrees with the pointwise limit (1.1) for almost all $t \in [0, 1]$. When $\ell(\gamma_s, \gamma_t) = G(t) - G(s)$ for some non-decreasing function G , then $\nu = G'$ coincides with the distributional derivative G' of G .

Associated to the causal speed are various power-law actions (with power $0 \neq q < 1$) whose maximizers are by definition *causal geodesics*. This point of view has the virtue of being inherited by curves in the space $\mathcal{P}(M)$ of probability measures on M : using optimal transport to maximize the cost function $\ell(x, y)^q$ over couplings with prescribed marginals μ_0 and μ_1 defines a (signed) time separation function $\ell_q(\mu_0, \mu_1)$ on the space $\mathcal{P}(M)$ of Borel probability measures. In Section 2.3 we show that if (M, τ, ℓ) is a Polish (forward) metric spacetime, then so is $(\mathcal{P}(M), \ell_q)$ equipped with the narrow topology. A relevant result of the chapter, and the main one of Section 2.7 is the Lorentzified version of a lifting theorem of Lisini [77], see also [9], that asserts that any suitable causal curve $\mu: [0, 1] \rightarrow \mathcal{P}(M)$ of measures can be lifted to a probability measure π on causal curves — called a *plan* (or “dynamical plan”) — such that the expectation of the q -action of the individual curves γ under π agrees with the q -action of the path μ in $(\mathcal{P}(M), \ell_q)$. In particular, causal geodesics for ℓ_q lift to plans π concentrated on causal geodesic curves γ for ℓ . We use the plan π as a proxy for a tangent to the curve of measures: whereas the causal speed represents only the magnitude of the tangent, the plan also encodes its direction.

The final section of the chapter discusses various non-branching notions. A notable one that arises in this setting is that of ‘timelike non-branching at 0 (resp. 1)’. A spacetime has this property if any two timelike geodesics that agree on the whole $(0, 1)$ also agree at 0 (resp. at 1). Of course, in positive signature this always occurs by the continuity of geodesics, but in our framework this cannot be excluded, and in some circumstances it is necessary to ask for this property, at least in a suitable \mathfrak{m} -a.e. version, see Definition 2.56.

1.2 Smooth paradigm: Lagrangian-Hamiltonian duality

In §3 we explore a cotangent calculus in duality with the tangent concepts from §2. This must be developed using positive linear rather than linear operations: the need to rely on causality of our functions and curves means their tangent spaces should be represented by convex cones rather than vector spaces. We therefore resort to concepts from convex analysis. Let us recall them in the familiar setting of an n -dimensional time-oriented Lorentzian manifold M equipped with a smooth globally hyperbolic metric tensor g of signature $(+, -, \dots, -)$. A vector $v \in T_x M$ is called *causal* if $g(v, v) \geq 0$ and *timelike* if the previous inequality holds strictly. Moreover the analogous terminology applies also to covectors, curves and functions: a curve is said to have property Q if and only if its tangent vectors all have property Q; a smooth function $f \in C^1(M)$ is said to have property Q if and only if $df(x)$ has property Q for all $x \in M$. The set of timelike vectors has two connected components, called the *future* and *past*; the vectors in the closure of the future component are called *future-directed*. The time-orientability of M means the future can (and has) been chosen to vary continuously with $x \in M$. A causal co-vector is called *future-directed* if its action on one (and hence all) timelike future vectors is non-negative; it is called *past-directed* otherwise. Fix dual exponents $0 \neq q < 1$ and $p^{-1} + q^{-1} = 1$. On the tangent space $T_x M$ to M at x we define the Lagrangian

$$L(v) := \begin{cases} \frac{1}{q} g(v, v)^{q/2} & \text{if } v \text{ is future-directed causal} \\ -\infty & \text{else,} \end{cases}$$

with the convention that $0^q = +\infty$ if $q < 0$. Remarkably, L is a concave function [84, 88]; its strict concavity in the timelike future was fundamental for McCann [84] and Mondino–Suhr [93]. We define the q -action of a smooth curve $\gamma: [0, 1] \rightarrow M$ by integrating L over the curve. We define the

time separation $\ell(x, y)$ between two events $x, y \in M$ by declaring $\frac{1}{q}\ell(x, y)^q$ to be the supremum of the q -actions of curves joining $\gamma_0 = x$ to $\gamma_1 = y$, where it is intended — and consistent — that $\ell(x, y) = -\infty$ if no future-directed causal curve from x to y exists (as in [84]); this definition is independent of $0 \neq q < 1$. Global hyperbolicity is well-known to ensure this supremum is attained unless no future-directed causal curve joins x to y ; the maximizers have constant speed and, when timelike, satisfy the geodesic equation. We define the concave Hamiltonian dual to L on the cotangent space by

$$H(\omega) := \inf_{v \in T_x M} [\omega(v) - L(v)] = \begin{cases} \frac{1}{p} g^*(\omega, \omega)^{p/2} & \text{if } \omega \text{ is future-directed causal} \\ -\infty & \text{else,} \end{cases}$$

where g^* is the cotangent metric induced by g . Obviously, H is also concave and together with L satisfies the Fenchel–Young inequality

$$L(v) + H(\omega) \leq \omega(v); \quad (1.2)$$

moreover, equality holds if and only if $\omega = DL(v)$ — or equivalently, if and only if $v = DH(\omega)$. Thus the v -derivative DL of L and its inverse map DH set up a nonlinear correspondence between (future-directed timelike) tangent and cotangent vectors at x . Define

$$\|v\| := \lim_{q \uparrow 1} L(v), \quad \|\omega\|_* := \lim_{p \uparrow 1} H(\omega). \quad (1.3)$$

For v and ω both future-directed, this allows us to re-express (1.2) in a form

$$\frac{1}{q} \|v\|^q + \frac{1}{p} \|\omega\|_*^p \leq \omega(v) \quad (1.4)$$

better suited to the nonsmooth setting: it requires only the Lorentzian magnitudes of v and ω as well as their bilinear pairing $\omega(v)$ to be known, but not the directions of v and ω . Since equality holds if and only if $\omega = DL(v)$, (1.4) tests whether or not v and ω are identified under the nonlinear duality DL . In the nonsmooth setting of §4, our Sobolev calculus will use an integrated version of (1.4) to establish the analogous nonlinear identification of tangent and cotangent field pairs. The limits defined by (1.3) give examples of objects we refer to as *hyperbolic norms*; their properties are further explored in Appendix A.3 and [58].

Before returning to the nonsmooth setting, let us also point out that for causal functions $f \in C^\infty(M)$ it is the convex functional

$$\mathcal{E}(f) := - \int_M H(df) d\text{vol}_g \quad (1.5)$$

which plays the role of the p -Dirichlet energy. Convexity of \mathcal{E} helps to determine when the (vertical) directional derivative

$$\lim_{t \rightarrow 0} \frac{\mathcal{E}(f + t\varphi) - \mathcal{E}(f)}{t} = - \int_M d\varphi(DH(df)) d\text{vol}_g =: \int_M \varphi \frac{\delta \mathcal{E}}{\delta f} d\text{vol}_g \quad (1.6)$$

exists. It makes the variational derivative

$$\frac{\delta \mathcal{E}}{\delta f} = \text{div}_g (\|df\|_*^{p-2} \nabla f) =: -\square_p f$$

a non-uniformly elliptic operator that we call the p -*d'Alembertian*. (This is analogous to how uniform convexity of the usual 2-Dirichlet energy makes the Riemannian Laplacian elliptic.) Notice however a conceptual issue: for the two-sided derivative (1.6) to exist, it is necessary that $f + t\varphi$ be causal (i.e. have $df + t d\varphi$ future-directed) for sufficiently small t of *both signs*. Since the set of causal functions is only a convex cone, outside of the smooth setting we will frequently need

to be satisfied with a one-sided derivative and an inequality in place of (1.6). Nevertheless, this construction allows us to prove a weak nonsmooth p -d'Alembert comparison theorem in §5. In particular, the latter establishes the 2-d'Alembert comparison inequality

$$\square_2 g^o \leq \frac{\dim M - 1}{|g^o|} \quad (1.7)$$

for the *Lorentz distance function* $g^o := -\ell(\cdot, o)$ on a spacetime with non-negative timelike Ricci curvature; notice that $\square_p f$ is independent of p when $\|df\|_* = 1$. Although (1.7) was proved in the region where g^o is smooth by [54], our extension across the cut locus seems new; it can be viewed as an analog of Calabi's result for the Riemannian Laplacian [31], even though both our formulations — in the weak as opposed to viscosity sense — and our methods of proof are utterly different. For a more direct proof of this extension inspired by the present work yet using more classical techniques, see [23]. We return now to the nonsmooth setting with one caveat: in the nonsmooth setting, the notions of causal and timelike will tacitly be taken to mean future-directed unless otherwise mentioned.

1.3 The Sobolev cone of cotangent fields; infinitesimal Minkowskianity

Having established nonsmooth tangent notions in §2, we turn to the problem of developing nonsmooth cotangent notions in §3. Although the consideration of linear functions on Minkowski space makes it tempting to define the slope at $y \in M$ of a causal function $f: M \rightarrow \mathbf{R}$ using (backward and forward) limits like

$$|\partial^- f|(y) \doteq \liminf_{x \uparrow y} \frac{f(y) - f(x)}{\ell(x, y)} \quad \text{or} \quad |\partial^+ f|(y) \doteq \liminf_{z \downarrow y} \frac{f(z) - f(y)}{\ell(y, z)} \quad (1.8)$$

restricted to $x \ll y \ll z$, in nonsmooth spaces it is not obvious when these limits should agree. It is even less obvious whether such a pointwise definition will cooperate with the kind of integration by parts formulas needed to give meaning in a weak sense to the nonsmooth analog of the p -d'Alembertian $\square_p f$. We focus instead to a different strategy inspired by developments in positive signature [73, 43, 110, 10, 60], as well as by the above-mentioned Lagrangian-Hamiltonian duality.

We require our forward metric spacetime (M, ℓ) to be equipped with a Radon reference measure \mathfrak{m} to express integration by parts. The triple (M, ℓ, \mathfrak{m}) is called a metric measure spacetime. Relative to \mathfrak{m} , the plans π on causal curves which appear in integration by parts formulas should not concentrate too much mass at any particular point in time and spacetime; quantifying a version of this condition leads to the notion of a *test plan*, namely a plan π admitting $C > 0$ such that for every $t \in [0, 1]$, the push-forward $(e_t)_* \pi$ under the evaluation map $e_t(\gamma) := \gamma_t$ is \mathfrak{m} -absolutely continuous with Radon-Nikodým density bounded from above by C . Remarkably, this Lorentzian notion is even simpler than its Riemannian antecedent [10], which in turn was inspired by the regular Lagrangian flows of Ambrosio-DiPerna-Lions [51, 5] (that, in relation with lower Ricci curvature bounds, turned out to be connected to Cheeger-Colding's segment inequality [44, Theorem 2.11]).

Instead of (1.8), the slope $|df|: M \rightarrow [0, +\infty]$ that we seek to define should heuristically be characterized as the largest function G satisfying

$$f(\gamma_1) - f(\gamma_0) \geq \int_0^1 G(\gamma_t) |\dot{\gamma}_t| dt \quad (1.9)$$

for all causal curves γ . However, it is not clear that such a function exists; moreover, for rough spaces and discontinuous functions, the values of f may be ambiguous at individual points or even along individual curves. Thus we hope instead that (1.9) holds along most curves. Lacking a preferred reference measure to define a notion of 'almost every' curve [110, 10], we instead require (1.9) to hold when integrated against every test plan, and deal with equivalence classes of functions f that agree \mathfrak{m} -a.e. For this integration to make sense, we restrict our attention to \mathfrak{m} -measurable

causal functions; if a causal function is not known to be measurable we may therefore call it *rough causal*. In §3.1 we show that every rough causal function is continuous outside a countable union of achronal sets, hence rough causality of f already implies \mathfrak{m} -measurability in many spaces of interest, such as those having synthetic lower Ricci curvature bounds as in Corollary 5.11.

There are two equivalent ways of defining a “candidate” for $|df|$ rigorously. Either pathwise (as in Lemma 3.13) or more robustly, in integrated form:

Definition 1.3 (Weak subslope). *A Borel function $G: M \rightarrow [0, +\infty]$ is a weak subslope of f if for all test plans π ,*

$$\int [f(\gamma_1) - f(\gamma_0)] d\pi(\gamma) \geq \iint_0^1 G(\gamma_t) |\dot{\gamma}_t| dt d\pi(\gamma). \quad (1.10)$$

Under mild assumptions on f and ℓ , both limits (1.8) become weak subslopes (Proposition 3.18). Weak subslopes are stable under limits (Proposition 3.12). Moreover, the set of weak subslopes for f is stable under pointwise maxima (see (3.14)). Combining these two properties allows us to deduce the existence and uniqueness of a maximal G satisfying (1.10) which we call a *maximal weak subslope* and denote by $|df|$. This represents a Lorentzian analog of the notions of minimal weak upper gradient introduced and studied in positive signature by Koskela–MacManus, Cheeger, Shanmugalingam, and Ambrosio–Gigli–Savaré [73, 43, 110, 10]. In the smooth setting we show this notion coincides with $\|df\|_*$ pointwise a.e. in Theorem A.2.

Inspired by the notion of infinitesimal Hilbertianity introduced by Gigli [60], we now formulate our definition of infinitesimal Minkowskianity. It is based on requiring the usual parallelogram law to hold for $f + g$ and f .

Definition 1.4 (Infinitesimal Minkowskianity). *We call (M, ℓ, \mathfrak{m}) infinitesimally Minkowskian if for every two causal functions $f, g: M \rightarrow \mathbf{R}$, we have*

$$2|d(f + g)|^2 + 2|df|^2 = |dg|^2 + |d(2f + g)|^2 \quad \mathfrak{m}\text{-a.e.} \quad (1.11)$$

Coupled with the a.e. concavity we shall eventually prove for $|df|$, Theorem 4.16 shows this parallelogram law is akin to asserting that $|df|(x)$ can be polarized to produce an indefinite inner product with Lorentzian signature for \mathfrak{m} -a.e. $x \in M$ (see also Lemma A.12). Simple variants of the results established in Appendix A.1 — whose proofs we omit for simplicity — show that a smooth Finsler spacetime is infinitesimally Minkowskian if and only if it is a Lorentzian spacetime. A related preceding result by Braun–Ohta [26] shows Finsler spacetimes with timelike sectional curvature bounds are already Lorentzian manifolds, hence infinitesimally Minkowskian. Another open problem already expressed in [38]: for metric measure spaces, infinitesimal Hilbertianity permits curvature-dimension bounds to be reformulated equivalently in terms of a Bochner-type inequality [11, 53]; it would be interesting to know whether infinitesimal Minkowskianity permits an analogous reformulation of timelike curvature-dimension conditions for metric measure spacetimes. In the smooth setting, one direction of this equivalence can be found in [100, 23]. In the nonsmooth case, irrespective of infinitesimal Minkowskianity, the timelike curvature-dimension condition is connected to a special Bochner-type inequality for Lorentz distance functions (whose simplifying advantage is that their “Hessian” vanishes identically in the tangential direction) by Braun [22] based on the localization formalism of Cavalletti and Mondino [37].

The remainder of §3 is devoted to establishing calculus rules in the necessary generality for our applications, including concavity a.e. of the maximal weak subslope (Proposition 3.21), locality properties in Proposition 3.22, chain rule (Proposition 3.23), and Leibniz rule (Proposition 3.24). As a consequence, although the maximal weak subslope defines only the magnitude $|df|$ of the slope and not its direction, we expect our approach paves the way to defining a genuine differential df for a causal function following the strategies from positive signature by Sauvageot [108, 109], Weaver [121], and Gigli [61].

1.4 Nonsmooth Lagrangian–Hamiltonian duality

In §4 we set up the negative-homogeneity correspondence between tangent and cotangent objects induced by a pair of exponents $p^{-1}+q^{-1} = 1$ with $0 \neq p < 1$. Inspired by the smooth Fenchel–Young inequality (1.4), one expects causal curves γ and causal functions f to satisfy

$$df(\dot{\gamma}_0) \geq \frac{1}{q} |\dot{\gamma}_0|^q + \frac{1}{p} |df|^p, \quad (1.12)$$

provided we can assign an appropriate meaning to the positively bilinear pairing on the left which represents the horizontal (inner) derivative of f along the curve γ . In the nonsmooth setting, we establish this inequality not for individual curves, but only after integrating against an arbitrary test plan π over curves (Proposition 4.1). Much as in positive signature [11, 60], the positively bilinear pairing which appears on the left of (1.12) will be replaced by the so-called “horizontal derivative” to obtain

$$\liminf_{t \downarrow 0} \int \frac{f(\gamma_t) - f(\gamma_0)}{t} d\pi(\gamma) \geq \frac{1}{p} \int |df|^p(\gamma_0) d\pi(\gamma) + \liminf_{t \downarrow 0} \frac{1}{tq} \int \int_0^t |\dot{\gamma}_r|^q dr d\pi(\gamma).$$

When the opposite inequality also holds, with \limsup replacing \liminf on the left, we say π represents the initial p -gradient of f (just as the Hamiltonian gradient $\dot{\gamma}_0 = DH(df) = \|df\|_*^{p-2} \nabla f$ from the smooth case attains equality in (1.12)). Heuristically, we may think of π as a tangent field also denoted by $|df|^{p-2} \nabla f \mu_0$ waiting to act linearly on the causal proxy φ for the cotangent field $d\varphi$, where $\mu_0 = (\mathbf{e}_0)_* \pi$. This follows the strategy used by Gigli [60] in positive signature metric measure geometry, which in turn was inspired by De Giorgi’s theory of metric gradient flows [49]. Although the plan π representing the initial p -gradient of f is never unique, for a given f the set of such π is convex. Continuing the analogy with $\dot{\gamma}_0 = DH(df)$ that has been rigorously developed in positive signature [10], in the very special case $(\mathbf{e}_0)_* \pi = \mathbf{m}$ we may view π as inducing a subgradient of the nonsmoothly adapted convex p -energy

$$\mathcal{E}(f) := -\frac{1}{p} \int_M |df|^p d\mathbf{m} \quad (1.13)$$

(via horizontal differentiation, which also leads to a weak definition for the divergence). To pursue this analogy further is beyond the scope of this paper.

This sets up the desired negative-homogeneity correspondence between the tangent objects π and cotangent objects df .

The remainder of §4 is devoted to developing and using more calculus rules to nonsmoothly compare the two sides of (1.6), with (1.13) replacing (1.5) in the “vertical” derivative on the left, while the right side is replaced by horizontal differentiation of φ along π . Theorem 4.15 then shows that an analog of identity (1.6) continues to hold as an inequality for single-sided derivatives, viz.

$$\lim_{\varepsilon \downarrow 0} \int \frac{|d(f + \varepsilon\varphi)|^p(\gamma_0) - |df|^p(\gamma_0)}{p\varepsilon} d\pi(\gamma) \leq \liminf_{t \downarrow 0} \int \frac{\varphi(\gamma_t) - \varphi(\gamma_0)}{t} d\pi(\gamma). \quad (1.14)$$

By *linearity* of the right-hand side, when $f + \varepsilon\varphi$ is causal for small ε of *both* signs the complementary inequality

$$\lim_{\varepsilon \uparrow 0} \int \frac{|d(f + \varepsilon\varphi)|^p(\gamma_0) - |df|^p(\gamma_0)}{p\varepsilon} d\pi(\gamma) \geq \limsup_{t \downarrow 0} \int \frac{\varphi(\gamma_t) - \varphi(\gamma_0)}{t} d\pi(\gamma)$$

follows. In this case, if one can show the two-sided vertical derivatives on the respective left-hand sides exist and are equal — as we establish when two-sided perturbations are causal and the spacetime is infinitesimally Minkowskian, cf. Theorem 4.16 — then (1.14) improves to an equality, meaning the horizontal ($t \downarrow 0$) derivative also exists and coincides with the vertical ($\varepsilon \rightarrow 0$) derivative. Some settings in which this improvement becomes possible are identified in Appendix A.2. Here, one of our contributions is to identify (for a given function f) a specific and

more tangible class of test functions φ satisfying the previous causality obstructions on $f + \varepsilon\varphi$; cf. Proposition A.4.

Note the asymmetric roles played by f and φ in the estimates above: whereas causality implies f is analogous to a non-decreasing function on the real line, being a difference of causal functions, $\varepsilon\varphi = (f + \varepsilon\varphi) - f$ is analogous to a function of locally bounded variation on the real line; this analogy proves quite fruitful. When both f and φ are causal, if $|df|$ and $|d\varphi|$ have strictly positive real values \mathfrak{m} -a.e., equation 4.31 shows $d^+f(\nabla\varphi) = d^+\varphi(\nabla f)$ \mathfrak{m} -a.e. for infinitesimally Minkowskian forward spacetimes. This expression becomes a positively bilinear indefinite quadratic form which can be extended bilinearly to the vector space generated by causal functions. Effectively, in some sense we recover a bilinear metric tensor \mathfrak{m} -a.e. from ℓ in this case.

1.5 Timelike Ricci curvature bounds and applications

The last part of our paper is devoted to applications of our first-order Sobolev calculus (which require additional hypotheses related to continuity of ℓ_+ and of timelike geodesics). More significantly, they require a proxy for timelike lower Ricci curvature bound K and upper dimension bound N . The synthetic version of this which we adopt is a hybrid combination $\text{TMCP}_+^h(K, N)$ of the timelike measure contraction property $\text{TMCP}_+(K, N)$ of Braun [21] and the entropic variant $\text{TMCP}_+^e(K, N)$ of Cavalletti–Mondino [39] upon which it is based. Although $\text{TMCP}_+(K, N)$ is a priori *stronger* than $\text{TMCP}_+^e(K, N)$, they are conjecturally equivalent under appropriate non-branching hypotheses (as already known for their positive signature counterparts [41, 35]), while the reduced variant $\text{TMCP}_+^*(K, N)$ of $\text{TMCP}_+(K, N)$ is equivalent to $\text{TMCP}_+^e(K, N)$ under such circumstances [21]. The two conditions $\text{TMCP}_+(K, N)$ and $\text{TMCP}_+^*(K, N)$ are defined by requiring different (K, N) -dependent convexity estimates for the N -Renyi entropy functional

$$\mathcal{S}_N(\mu) := - \int_M \rho^{(N-1)/N} \, \mathfrak{m}, \quad \text{where } \mu = \rho \mathfrak{m} + \mu^\perp \quad \text{with } \mu^\perp \perp \mathfrak{m}, \quad (1.15)$$

along suitable ℓ_q -geodesics (μ_t) joining an arbitrary probability measure μ_0 to a Dirac mass in its timelike future; when $K = 0$, both follow from ordinary convexity of the assignment $t \mapsto \mathcal{S}_N(\mu_t)$. The Cavalletti–Mondino condition instead requires a growth bound for the Boltzmann-Shannon entropy

$$\mathcal{S}_\infty(\mu) := \begin{cases} \int_M \rho \log \rho \, \mathfrak{m} & \text{if } \mu = \rho \mathfrak{m} \text{ and } \mathfrak{m}[\text{spt } \mu] < \infty, \\ +\infty & \text{otherwise} \end{cases} \quad (1.16)$$

along the same curves which follows from ordinary convexity in the case $N = 0$. Our hybrid condition $\text{TMCP}_+^h(K, N)$ requires suitable ℓ_q -geodesics to satisfy both estimates. We need it only to gain compactness; whenever the compactness is provided by globally hyperbolicity, $\text{TMCP}_+(K, N)$ would suffice.

We note that qualitatively, all our results in Section 5 will hold irrespective of whether we assume $\text{TMCP}_+^h(K, N)$ or $\text{TMCP}_+^{h,*}(K, N)$ (or $\text{TMCP}_+(K, N)$ or $\text{TMCP}_+^*(K, N)$ when globally hyperbolicity holds). The benefit of the unstarred conditions is that they provide quantitative consequences in *sharp* form even in the absence of any non-branching hypothesis. Notably, an interesting setting where this comparison theory thus applies are vector spaces with hyperbolic norms, cf. Appendix A.3. As our Proposition A.11 below establishes, Euclidean space endowed with *any* appropriate hyperbolic norm has non-negative synthetic timelike Ricci curvature, yet these structures may admit many branching geodesics in general.

For compatibility with the test plans of our Sobolev calculus, the entropic displacement convexity which defines $\text{TMCP}_+^h(K, N)$ must be satisfied by geodesics of measures having density bounds. The construction of these “good” geodesics in Theorem 5.7 and Theorem 5.12 involves new subtleties beyond those from the globally hyperbolic case [20]. To enforce initial and final conditions (and one-sided continuity) the concept of non-branching at 0 or 1 we already alluded to comes into play; it relaxes concepts from [75, 85]. To handle Dirac endpoints with forward-completeness instead of

compactness, we exploit the narrow coercivity of \mathcal{S}_∞ on finite volume subsets of M and the growth bounds from $\text{TMCP}_+^e(K, N)$.

The first result that we show is a nonsmooth converse to the Hawking–King–McCarthy–Malament theorem [68, 80]. Working in the smooth context, with the time separation function built out of the metric tensor, they showed that a bijection which preserves the class of continuous timelike curves is in fact a smooth conformal isometry. If it also preserves the time separation then it preserves the metric tensor. In the nonsmooth context, we have instead used the time separation to define a maximal weak subslope $|df|: M \rightarrow [0, +\infty]$ for causal functions $f: M \rightarrow \mathbf{R}$. We begin by showing that (M, ℓ, \mathbf{m}) satisfying $\text{TMCP}_+^h(K, N)$ implies the *Sobolev-to-steepness* property (Theorem 5.17), inspired by Gigli’s Sobolev-to-Lipschitz property in positive signature [63]. This means the condition $|df| \geq 1$ \mathbf{m} -a.e. implies a “reverse Lipschitz” condition we call *1-steepness*, viz.

$$f(y) - f(x) \geq \ell(x, y)$$

for all $x, y \in \text{spt } \mathbf{m}$. This implies ℓ can be recovered on the support of \mathbf{m} from the assignment $f \mapsto |df|$. It also implies that if $T: \text{spt } \mathbf{m}_1 \rightarrow \text{spt } \mathbf{m}_2$ is a surjection between two such spaces that preserves measure, causality, and maximal weak subslopes, then it also preserves ℓ , i.e. is a global isometry (Theorem 5.16). Much like in positive signature, this result makes use of Sobolev notions viable for deriving precise metric conclusions.

Our second application is set in the same class of spaces. We connect our first-order Sobolev calculus with the optimal transport of appropriate Borel probability measures $\mu_0, \mu_1 \in \mathcal{P}(M)$ by showing a metric [10, 60] Brenier–McCann [27, 81, 83, 84] theorem. Under appropriate hypotheses [39], Cavalletti–Mondino have shown a coupling of μ_0 and μ_1 is ℓ_q -optimal if and only if it is supported in the ℓ^q/q -superdifferential $\partial_{\ell^q/q} f$ of an ℓ^q/q -concave function f . In Theorem 5.19, we show that if a test plan has marginals $(\mathbf{e}_0, \mathbf{e}_1)_* \pi$ supported in $\partial_{\ell^q/q} f$, then π represents the initial p -gradient of f . Moreover, π -a.e. γ satisfies the *equality*

$$\ell(\gamma_0, \gamma_1) = |df|^{p-1}(\gamma_0). \quad (1.17)$$

Thus μ_0 -a.e., our maximal weak subslope captures the distance transported into the future. This also opens the door for a natural notion of exponentiation on nonsmooth spacetimes with synthetic curvature bounds akin to [65].

This result plays a role in our third application, which is a *sharp* p -d’Alembert comparison theorem for $\text{TMCP}_+^h(K, N)$ spaces (Theorems 5.24 and 5.28). The theorem has the cleanest form in the case $K = 0$. For ℓ^q/q -concave functions f with $\partial_{\ell^q/q} f \subset \{\ell > 0\}$, it states the following inequality in a weak form:

$$\square_p f \leq N. \quad (1.18)$$

Applying (1.18) to the function $f^o = -\ell(\cdot, o)^q/q$, the chain rule yields a weak expression of the bound

$$\square_p(-\ell(\cdot, o)) \leq \frac{N-1}{\ell(\cdot, o)}$$

on $I^-(o)$ for the p -d’Alembertian of the Lorentz distance to $o \in \text{spt } \mathbf{m}$. On $I^+(o)$, a similar logic yields a weak reformulation of

$$\square_p \ell(o, \cdot) \geq -\frac{N-1}{\ell(o, \cdot)}$$

under the modified hypothesis $\text{TMCP}_-^h(0, N)$, in which transport involves a Dirac endpoint in the past rather than the future. Along the way, some key inequalities have to be “reversed”; this is discussed separately in Section 5.7.

Based on the analogous result in positive signature [60], the initial idea of our proof is to consider a mass distribution μ_0 given in terms of an appropriate non-negative test function φ supported in $I^-(o)$, and to transport μ_0 to $\mu_1 := \delta_o$. The corresponding ℓ^q/q -concave function is $f = f^o$ on $I^-(o)$. (To simplify the presentation in this introduction, we neglect technicalities pertaining to causality, integrability, etc., of appropriate linear combinations of φ and f for now.) More

precisely, since the entropy functional (1.15) is given by the convex integrand $s_N(r) = -r^{(N-1)/N}$, we will choose μ_0 to be \mathbf{m} -absolutely continuous with density $\rho_0 = (c_0 \varphi)^{N/(N-1)}$, where c_0 is a normalization constant. Theorems 5.7 and 5.19 plus Theorem 5.12 establish the existence of a plan π representing the initial p -gradient of f such that the assignment $t \mapsto \mathcal{S}_N(\mu_t)$ satisfies an appropriate ℓ_q -geodesic convexity estimate, where $\mu_t := (\mathbf{e}_t)_* \pi$. This convexity entails

$$\limsup_{t \downarrow 0} \frac{\mathcal{S}_N(\mu_t) - \mathcal{S}_N(\mu_0)}{t} \leq \mathcal{S}_N(\mu_1) - \mathcal{S}_N(\mu_0) = c_0 \int_M \varphi \, \mathbf{d}\mathbf{m}. \quad (1.19)$$

On the other hand, ordinary convexity of s_N gives the complementary bound relating (1.19) to a horizontal derivative, which in turn can be bounded by a vertical derivative as in (1.14):

$$\begin{aligned} \liminf_{t \downarrow 0} \frac{\mathcal{S}_N(\mu_t) - \mathcal{S}_N(\mu_0)}{t} &\geq \liminf_{t \downarrow 0} \int \frac{s'_N \circ \rho_0(\gamma_t) - s'_N \circ \rho_0(\gamma_0)}{t} \, \mathbf{d}\pi(\gamma) \\ &\geq \liminf_{\varepsilon \downarrow 0} \int_M \frac{|\mathbf{d}(f + \varepsilon s'_N \circ \rho_0)|^p - |\mathbf{d}f|^p}{p\varepsilon} \, \mathbf{d}\mu_0 \\ &= \int_M \mathbf{d}^+(s'_N \circ \rho_0)(\nabla f) |\mathbf{d}f|^{p-2} \, \mathbf{d}\mu_0 \\ &= \frac{c_0}{N} \int_M \mathbf{d}^+ \varphi(\nabla f) |\mathbf{d}f|^{p-2} \, \mathbf{d}\mathbf{m}, \end{aligned}$$

where the last identity follows from a suitable chain rule and the specific choice of ρ_0 . Canceling c_0 gives the desired weak formulation of (1.18), viz.

$$\int_M \mathbf{d}^+ \varphi(\nabla f) |\mathbf{d}f|^{p-2} \, \mathbf{d}\mathbf{m} \leq N \int_M \varphi \, \mathbf{d}\mathbf{m}.$$

For non-zero K , the initial ‘‘chord above tangent’’ bound (1.19) is complicated by the presence of additional distortion coefficients depending on $\ell(\gamma_0, \gamma_1)$, which are related to $|\mathbf{d}f|^{p-1}(\gamma_0)$ by our metric Brenier–McCann theorem (1.17).

In Section 5.6 we discuss how the p -d’Alembert comparison estimate can be coupled with infinitesimal Minkowskianity (or related notions) to develop a notion of ‘distributional’ p -d’Alembertian and start studying its properties. A basic one is that the Riesz–Daniell theorems allow us to represent such an operator as a Radon measure, suitably defined.

In the final Section 5.7 we comment on how our results are modified if we work on $\text{TMCP}_-^h(K, N)$ spaces rather than on $\text{TMCP}_+^h(K, N)$ ones (while keeping forward completeness).

1.6 Two-sided perturbations, hyperbolic norms, and compatibility with smooth notions

We include three appendices. The first establishes compatibility between many of the nonsmooth notions introduced here and their analogs in smooth spacetimes. The second uses Kunzinger–Steinbauer’s nonsmooth adaptation [76] of the Sormani–Vega null distance [112] to give conditions under which a causal function f will admit two-sided linear perturbations, both of which remain causal. The third introduces hyperbolic norms as a generalization of Lorentz–Finsler norms, and explores weak lower Ricci bounds in this setting as well as consequences of the parallelogram law.

2 Curves in the nonsmooth Lorentzian setting

2.1 Infinity conventions

Throughout this paper, we employ the conventions

$$\begin{aligned} +\infty - (+\infty) &:= +\infty, \\ -\infty - (-\infty) &:= +\infty, \\ \pm\infty + z &:= \pm\infty \quad \text{for every } z \in [-\infty, +\infty], \\ 0 \cdot (\pm\infty) &:= 0. \end{aligned} \tag{2.1}$$

The latter is standard in measure theory. The first three will enter into play, for instance, when presenting the reverse triangle inequality for the time separation function (2.4b) or when discussing calculus with causal functions as in (3.9), as we do not want to exclude from our analysis cases in which ℓ or f take the values $\pm\infty$. We accept that our conventions imply the non-commutativity $-\infty + (+\infty) = -\infty < +\infty = +\infty + (-\infty)$. In connection with the reverse triangle inequality, it is worth to point out that for $a, b, c \in \mathbf{R}$ the above yields

$$\min\{a + b, b + a\} \leq c \quad \Leftrightarrow \quad a \leq c - b \quad \text{and} \quad b \leq c - a, \tag{2.2}$$

as can be easily checked.

For $0 \neq p \leq 1$, we also define the concave non-decreasing function $u_p : \bar{\mathbf{R}} \rightarrow \bar{\mathbf{R}}$ by setting $u_p(z) := \frac{z^p}{p}$ for $z > 0$ and then extending this by monotonicity and concavity, i.e. we put

$$\begin{aligned} u_p(z) &= -\infty && \text{for } z \in [-\infty, 0) \text{ and any } p \leq 1, \\ u_p(0) &= -\infty \quad (\text{resp. } 0) && \text{for } p < 0 \text{ (resp. } p \in (0, 1]), \\ u_p(+\infty) &= 0 \quad (\text{resp. } +\infty) && \text{for } p < 0 \text{ (resp. } p \in (0, 1]). \end{aligned} \tag{2.3}$$

2.2 Notions of spacetimes

We shall work in a setting which generalizes Lorentz(-Finsler) spacetimes and their basic causality theory to possibly nonsmooth spaces. Such an abstract approach was pioneered by Kunzinger-Sämman through their Lorentzian pre-length spaces [75] (inspired by early contributions of Busemann [29] and Kronheimer-Penrose [74]). An alternative approach where many topological properties are implied by basic axioms on the time separation function was proposed by Minguzzi-Suhr [91]. Sequella of [39, 85, 25, 30] extend [75, 91]. (We point out [85, 25] use the term ‘‘metric spacetime’’ for related yet slightly different structures than we do.) An approach more tied to a distance in the classical metric sense (the so-called null distance) is given in Sormani-Vega [112].

Let M be a set and $\ell : M \times M \rightarrow \{-\infty\} \cup [0, +\infty]$. We say that ℓ is a *time separation* if

$$\ell(x, x) = 0, \tag{2.4a}$$

$$\min\{\ell(x, y) + \ell(y, z), \ell(y, z) + \ell(x, y)\} \leq \ell(x, z). \tag{2.4b}$$

hold for any $x, y, z \in M$. We define the transitive relations *chronology* \ll and *causality* \leq on M as:

$$\ll := \ell^{-1}((0, +\infty]) = \{\ell > 0\}, \quad \leq := \ell^{-1}([0, +\infty]) = \{\ell \geq 0\}.$$

If ℓ is a time separation, so is its *time-reversal* $\ell^*(y, x) := \ell(x, y)$.

We shall denote by $\ell_+ := \max\{\ell, 0\} = \ell \vee 0$ the positive part of ℓ . For $x \in M$, the sets

$$\begin{aligned} I^+(x) &:= \{y \in M : x \ll y\}, & I^-(x) &:= \{y \in M : y \ll x\}, \\ J^+(x) &:= \{y \in M : x \leq y\}, & J^-(x) &:= \{y \in M : y \leq x\}, \end{aligned}$$

are the *chronological future and past* of x and the *causal future and past* of x , respectively. Analogously, for $X, Y \subset M$ we define

$$J^\pm(X) := \bigcup_{x \in X} J^\pm(x), \quad J(X, Y) := J^+(X) \cap J^-(Y).$$

If X and Y are both compact we refer to sets of the form $J(X, Y)$ as *causal emeralds*, or simply emeralds. If $X = \{x\}$ and $Y = \{y\}$, we will refer to these sets as *diamonds* instead of emeralds and write $J(x, y)$ for short. We shall also deal with *chronological diamonds*, i.e. sets $I(x, y)$ of the form $I^+(x) \cap I^-(y)$. A set $X \subset M$ is called *causally convex* if and only if it contains the causal diamond $J(x, y) \subset X$ for all $x, y \in X$.

Definition 2.1 (Notions of spacetimes). *We shall say that:*

- i) (M, ℓ) is a metric spacetime if ℓ is a time separation on M so that \leq is a partial order, i.e. $x \leq y$ and $y \leq x$ implies $x = y$.
- ii) (M, τ, ℓ) is a Polish metric spacetime if τ is a Polish topology (i.e. induced by a complete and separable distance) on the metric spacetime (M, ℓ) so that $\{\ell \geq 0\}$ is closed in M^2 .
- iii) (M, τ, ℓ) is a forward metric spacetime if it is a Polish metric spacetime and every sequence (x_n) such that $x_n \leq x_{n+1} \leq \bar{x}$ for every $n \in \mathbf{N}$ and some $\bar{x} \in M$ admits a τ -limit.
Symmetrically, if we ask for existence of limits of decreasing sequences bounded from below we get the notion of backward metric spacetime.

Finally, a Polish spacetime is called *locally causally convex* if every point has a neighbourhood basis made of causally convex sets.

Forward/backward spacetimes give notions of completeness which obviate our need for the local compactness guaranteed by global hyperbolicity. A new compactness result in forward metric spacetimes, Theorem 2.30 below, gives a limit curve theorem for causal curves.

Remark 2.2 (The choice of the topology). We are not going to require any further relation between τ and ℓ on top of those already asked by the concepts of Polish and forward metric spacetimes. We opt for this choice both because for the moment we do not need anything more and because we are looking for notions that get inherited by the space of Borel probability measures equipped with the time separation ℓ_q coming from optimal transport: see Section 2.3 for details.

In practical applications however, and especially in relation to convergence of spacetimes, it is surely natural to look for a topology that is induced in some way by ℓ . Also, our strongest version of the limit curve theorem Theorem 2.30 holds for locally causally convex topologies and often topologies somehow induced by ℓ have this property. Examples of ℓ -induced topologies are:

- The Alexandrov topology, i.e. the one generated by chronological diamonds,
- The chronological topology [75], i.e. the one generated by sets of the form $I^-(x)$ and $I^+(x)$ as x varies in M ,
- The smallest topologies making ℓ upper semicontinuous and its positive part ℓ_+ continuous.

These are ordered from the coarsest to the finest and are all locally causally convex. ■

Remark 2.3 (Push-up property). In any metric spacetime (M, ℓ) , the *push-up principle* holds, i.e. if $x \ll y$ and $y \leq z$ (or if $x \leq y$ and $y \ll z$) then $x \ll z$. This follows directly from the reverse triangle inequality. ■

We collect a couple of basic results about the relation of ℓ and the topology.

Lemma 2.4 (Upper semicontinuous time separations attain maxima on emeralds). *Let (M, τ, ℓ) be a Polish metric spacetime. Then emeralds are closed. If moreover ℓ_+ is real valued and upper semicontinuous, then it is bounded above on each emerald.*

Proof. Let $E \subset M$ be an emerald, say $E = J(C_0, C_1)$ for $C_0, C_1 \subset M$ compact, and $(x_n) \subset E$ be a sequence converging to a limit x . Then there are $(y_n) \subset C_0$ and $(z_n) \subset C_1$ so that $y_n \leq x_n \leq z_n$ for every $n \in \mathbf{N}$. By compactness, up to passing to a non-relabelled subsequence we can assume

that $y_n \rightarrow y \in C_0$ and $z_n \rightarrow z \in C_1$. As $\{\ell \geq 0\}$ is closed, we conclude that $y \leq x \leq z$, so that $x \in E$ as well.

For the second claim, notice that by upper semicontinuity ℓ attains its maximum value $L < \infty$ on $C_0 \times C_1$. We claim $\ell \leq L$ on E^2 . Each $x \in E$ lies in $J(x_0, x_1)$ for some $x_i \in C_i$. Similarly $y \in E$ lies in $J(y_0, y_1)$ for some $y_i \in C_i$. If $\ell(x, y) > -\infty$ then $x_0 \leq x \leq y \leq y_1$ and the reverse triangle inequality yields $\ell(x, y) \leq \ell(x_0, y_1) \leq L$ as desired. \square

Somewhat conversely, the following provides a condition ensuring lower semicontinuity of $\ell_+(\cdot, o)$ for a point $o \in M$. The assumption is tailored to the setting we will work with in Section 5: in typical applications, z comes from the t -evaluation γ_t of a rough ℓ -geodesic γ from x to o (see Definition 2.32).

Lemma 2.5 (Rough timelike geodesy yields lower semicontinuous time to a point). *Let (M, τ, ℓ) be a Polish metric spacetime with τ containing the chronological topology. Assume that a point $o \in M$ has the following property: for every $x \ll o$ and every $t \in (0, 1)$ there exists $z \in M$ such that*

$$\ell(x, z) = t \ell(x, o), \quad \ell(z, o) = (1 - t) \ell(x, o).$$

Then the assignment $z \mapsto \ell_+(z, o)$ is lower semicontinuous.

Proof. It suffices to prove that for every $c > 0$ the superlevel set $\{\ell(\cdot, o) > c\}$ is open. For every $x \in \{\ell(\cdot, o) > c\}$, by our assumption there exists $z \in M$ such that $\ell(x, z) > 0$ and $\ell(z, o) > c$. In particular, x belongs to the chronological past of z , which is open by definition. We thus proved that $\{\ell(\cdot, o) > c\}$ is the union of chronological pasts of all of its elements, hence it is open. \square

Further interesting regularity properties can be proved under the (reasonable) assumption that in our spacetime any point can be approximated by points in its chronological past, i.e.

$$x \in \overline{I^-(x)} \quad \forall x \in M. \quad (2.5)$$

This holds e.g. in metric spacetimes where every point is in the interior of a continuous timelike curve, such as in globally hyperbolic regular Lorentzian length spaces.

The notion of forward-completeness which defines forward spacetimes (Definition 2.1) does not quite coincide with the usual notion of order completeness, which would require existence of a supremum rather than a limit. These two notions are linked by the following two results.

Lemma 2.6 (Chronology determines causality). *Let (M, τ, ℓ) be a Polish metric spacetime for which (2.5) holds. Then the inclusion $I^-(x) \subset I^-(y)$ for some $x, y \in M$ implies $x \leq y$.*

Proof. This can partially be compared with one direction of [1, Prop. 2.26]. Taking closures in the given inclusion and recalling that \leq is closed we get $\overline{I^-(x)} \subset \overline{I^-(y)} \subset J^-(y)$. Then the assumption (2.5) yields $x \in J^-(y)$, which is the claim. \square

The following proposition gives sufficient conditions for the limit of any \leq -non-decreasing sequence to coincide with its \leq -supremum.

Proposition 2.7 (Limit from the past vs. order supremum). *Let (M, τ, ℓ) be a Polish metric spacetime for which (2.5) holds and so that τ contains the chronological topology. Let the limit $x = \lim_{n \rightarrow \infty} x_n$ of some non-decreasing sequence $x_n \leq x_{n+1}$ in M converge. Then $x_n \leq x$ for all $n \in \mathbf{N}$ and moreover: if $x_n \leq y \in M$ for all $n \in \mathbf{N}$ then $x \leq y$.*

Proof. Letting $m \rightarrow \infty$ in $x_n \leq x_m$ (which is valid for $n \leq m$) and using that \leq is closed we deduce that x is an \leq -upper bound for the sequence. Let y be another \leq -upper bound. Since $x_n \rightarrow x$, any neighbourhood U of x contains all x_n for large enough n . In particular, this holds for $U = I^+(z)$ for some $z \ll x$ provided by (2.5); U is open by our assumption on τ . For such U we must have $z \ll x_n$ for sufficiently large n and since $y \geq x_n$ we must also have $z \ll y$ by the push-up property (Remark 2.3). We thus proved that if $z \ll x$ then $z \ll y$, so that Lemma 2.6 implies $x \leq y$, as desired. \square

We comment on how all this relates to previously investigated settings.

Remark 2.8 (Metric spacetimes versus other abstract generalizations of spacetimes). Our setup of Polish metric spacetimes is compatible with the approaches of Kunzinger–Sämman’s Lorentzian pre-length spaces [75] and Minguzzi–Suhr’s bounded Lorentzian metric spaces [91] whenever the respective standing hypotheses intersect. (We refer to Beran–Harvey–Napper–Rott [16, §4.1] for the relation between [75] and [91].) Indeed, let (M, τ, ℓ) be a Polish metric spacetime with τ being the chronological topology. Let \mathbf{d} be a fixed metric inducing τ . If ℓ_+ is lower semicontinuous, then $(M, \mathbf{d}, \ll, \leq, \ell_+)$ is a Lorentzian pre-length space. On the other hand, a Lorentzian pre-length space $(M, \mathbf{d}, \ll, \leq, \tau)$ where the reference metric \mathbf{d} is Polish and induces the chronological topology and such that the relation \leq is closed becomes a Polish metric spacetime by defining ℓ as $-\infty$ outside \leq and τ otherwise.

Lastly, assume the pair (M, τ) is a bounded Lorentzian metric space after Minguzzi–Suhr [91]. Let τ denote the reference topology, which is Polish and in fact uniquely certifies the definition of (M, τ) being a bounded Lorentzian metric space [91, Cor. 1.7, Thm. 1.10]. Define ℓ as in the previous paragraph, where \leq is replaced by the (closed) extended causal relation from [91, §5.1]. These observations make (M, ℓ) a Polish metric spacetime. ■

Definition 2.9 (Global hyperbolicity). *A Polish metric spacetime is called globally hyperbolic if all of its emeralds are compact.*

Together with our ad hoc assumption of closedness of $\{l \geq 0\}$, this corresponds to the notion of global hyperbolicity for general topological ordered spaces suggested by Minguzzi [87].

Remark 2.10 (About global hyperbolicity). In a globally hyperbolic metric spacetime, any bounded monotone sequence has a limit, as it lives in a (compact) emerald and the closedness of \leq implies any two limits $x, x' \in M$ satisfy $x \leq x'$ and $x' \leq x$, hence $x = x'$ as \leq is a partial order. In particular, such a spacetime is both forward and backward. As we do not assume a strong relation between a metric (inducing the Polish order topology) and the causal structure (i.e. \mathbf{d} -compatibility [75, Def. 3.13]), a globally hyperbolic metric spacetime is not a globally hyperbolic Lorentzian pre-length space in the sense of [75, Def. 2.35]. On the other hand if we assume \mathbf{d} -compatibility, causal path-connectedness (i.e. every pair of causally related points can be joined by a continuous causal curve) and lower semicontinuity of ℓ_+ , then by [90, Thm. 3.7] the quintuple $(M, \mathbf{d}, \ll, \leq, \ell_+)$ is a globally hyperbolic Lorentzian pre-length space. ■

Remark 2.11 (Forward metric spacetimes on the causal ladder). A spacetime is called *causally simple* if the causal relation \leq is both a partial ordering and closed; this definition makes sense in our setting as well. By Remark 2.10, any globally hyperbolic metric spacetime is forward, while by definition any Polish metric spacetime is causally simple.

In our setting, these inclusions are strict. On the one hand, Anti-de Sitter space is causally simple but not forward-complete. Alternatively, consider any causal diamond $J(x, y)$ in the Minkowski plane but remove spacelike infinity $\partial I^+(x) \cap \partial I^-(y)$.

Also, take the diamond $J((-1, 0), (1, 0))$ in the Minkowski plane and add the horizontal line $\mathbf{R}\hat{e}_2$ to the space and extend the causal relation by $(-1, 0) \leq (0, z) \leq (1, 0)$ for all $z \in \mathbf{R}$. This space is causally simple and forward, but not globally hyperbolic (and not a manifold). ■

2.3 The spacetime of Borel probability measures

Let (M, τ, ℓ) be a Polish metric spacetime. We shall denote by $\mathcal{P}(M)$ the collection of Borel probability measures on it and by $\mathcal{P}_{\text{em}}(M) \subset \mathcal{P}(M)$ the subset of those with support in some emerald. Here we are interested in the study of the *Lorentz–Wasserstein time separation* ℓ_q (originating in Eckstein–Miller [52] in the range $q \in (0, 1]$) on $\mathcal{P}(M)$. We start with the definition:

Definition 2.12 (Lorentz–Wasserstein distance). *Let (M, τ, ℓ) be a Polish metric spacetime in which ℓ is Borel. For $\mu, \nu \in \mathcal{P}(M)$ let $\Pi_{\leq}(\mu, \nu) \subset \mathcal{P}(M \times M)$ be the set of causal couplings, i.e. of measures $\pi \geq 0$ concentrated on $\{\ell \geq 0\}$ having marginals $\mu = (\text{Pr}_1)_*\pi$ and $\nu = (\text{Pr}_2)_*\pi$; here $\text{Pr}_1(x, y) := x$ and $\text{Pr}_2(x, y) := y$ on M^2 .*

For $0 \neq q \leq 1$, and u_q from (2.3), the q -Lorentz–Wasserstein distance $\ell_q(\mu, \nu)$ of $\mu, \nu \in \mathcal{P}(M)$ is defined as $-\infty$ if $\Pi_{\leq}(\mu, \nu)$ is empty (consistently with the usual convention $\sup \emptyset := -\infty$) and otherwise as that element in $[0, +\infty]$ such that

$$u_q(\ell_q(\mu, \nu)) := \sup_{\pi \in \Pi_{\leq}(\mu, \nu)} \int_{M \times M} u_q \circ \ell \, d\pi. \quad (2.6)$$

In other words, if no infinities appear, $\ell_q(\mu, \nu)$ can be defined via

$$\ell_q(\mu, \nu) := \sup_{\pi \in \Pi_{\leq}(\mu, \nu)} \left(\int_{M \times M} \ell^q \, d\pi \right)^{\frac{1}{q}}. \quad (2.7)$$

A coupling π of μ and ν with $\mu \preceq \nu$ is called ℓ_q -optimal if it attains the supremum in (2.6) (or, equivalently, in (2.7)).

We shall prove in a moment that ℓ_q is a time separation on $\mathcal{P}(M)$. For the moment, we notice that the causal relation \preceq it induces is independent of q and can be defined as: $\mu \preceq \nu$ if and only if $\Pi_{\leq}(\mu, \nu)$ is not empty. A standard gluing argument shows that \preceq is transitive, while it is less obvious that it is antisymmetric (but see Proposition 2.16 below). For the interplay of \preceq and the narrow topology we refer to [52, Thm. 8] and [25, Thm. B.5]. We caution the reader that unlike the causal relation described above, the timelike relation \prec induced by ℓ_q can depend on q when we allow $q < 0$.

To study the properties of ℓ_q it is convenient to keep in mind some basic facts about L^q ‘norms’ for $0 \neq q < 1$, defined for $[0, +\infty]$ -valued measurable functions as

$$\|f\|_q := \left(\int f^q \, d\mathbf{m} \right)^{\frac{1}{q}} \in [0, +\infty].$$

The following is simple and well known, see also [58] for more about duality properties of L^p and L^q spaces for $p, q < 1$:

Lemma 2.13 (Basic inequalities in L^q , $0 \neq q < 1$). *Let (X, \mathbf{m}) be a σ -finite measure space, $0 \neq p < 1$ and $0 \neq q < 1$ satisfy $\frac{1}{p} + \frac{1}{q} = 1$, and $f : X \rightarrow [0, +\infty]$ be measurable. Then*

$$\|f\|_q = \inf \left\{ \int fg \, d\mathbf{m} : g : X \rightarrow [0, +\infty] \text{ is measurable with } \|g\|_p = 1 \right\} \quad (2.8)$$

and

$$\|f_1 + f_2\|_q \geq \|f_1\|_q + \|f_2\|_q. \quad (2.9)$$

Moreover, if equality holds in (2.9) and $\|f_1 + f_2\|_q \in (0, +\infty)$, then for some $\alpha \geq 0$ we have $f_1 = \alpha f_2$.

Proof. We observe that u_p is the Legendre–Fenchel transform of the concave function u_q , i.e.:

$$\begin{aligned} u_q(z) + u_p(w) &\leq zw & \forall z, w \geq 0, \\ u_q(z) + u_p(w) &= zw & \Leftrightarrow z, w > 0, u'_q(z) = w, u'_p(w) = z. \end{aligned}$$

The proof of this is trivial. We now prove \leq in (2.8). This is obvious if $\|f\|_q = 0$. Assume thus — for the moment — that $\|f\|_q \in (0, +\infty)$ and notice that the above yields

$$\frac{1}{\|f\|_q} \int fg \, d\mathbf{m} \geq \int u_q\left(\frac{f}{\|f\|_q}\right) + u_p(g) \, d\mathbf{m} = \int u_q\left(\frac{f}{\|f\|_q}\right) \, d\mathbf{m} + \int u_p(g) \, d\mathbf{m} = \frac{1}{q} + \frac{1}{p} = 1, \quad (2.10)$$

and thus the desired \leq . For the general case we find a non-decreasing sequence (f_n) increasing to f so that the $\|f_n\|_q$'s are finite and converging to $+\infty$ (here the assumption of σ -finiteness matters): letting $n \uparrow \infty$ in $\|f_n\|_q \leq \int f_n g \, d\mathbf{m} \leq \int fg \, d\mathbf{m}$, by the arbitrariness of g we conclude that \leq holds.

We pass to \geq and observe that it is obvious if $\|f\|_q = +\infty$. Then assume — up to normalization — that $\|f\|_q = 1$, notice that in this case $g := f^{\frac{q}{p}}$ satisfies $\|g\|_p = 1$ and that $\int fg \, d\mathbf{m} = \int f^q \, d\mathbf{m} = 1$ yielding the claim. Finally, if $\|f\|_q = 0$ find a decreasing sequence (f_n) converging to f so that the $\|f_n\|_q$'s are positive and converging to 0 (again, here the assumption of σ -finiteness helps). By the above discussion, for each n there is g_n with $\|g_n\|_p = 1$ so that $\|f_n\|_q = \int f_n g_n \, d\mathbf{m}$, thus in particular $f_n g_n \in L^1(\mathbf{m})$ and letting $m \rightarrow \infty$ in the bound $\int f_n g_n \, d\mathbf{m} \geq \int f_m g_n \, d\mathbf{m}$ valid for any $m \geq n$, by dominated convergence we conclude that $\|f_n\|_q \geq \int f g_n \, d\mathbf{m} \geq \inf_g \int f g \, d\mathbf{m}$. Letting $n \uparrow \infty$ we conclude.

The second part of the statement is now an easy consequence of the first. Notice indeed that

$$\|f_1 + f_2\|_q = \inf_g \int (f_1 + f_2)g \, d\mathbf{m} \geq \inf_{g_1} \int f_1 g_1 \, d\mathbf{m} + \inf_{g_2} \int f_2 g_2 \, d\mathbf{m} = \|f_1\|_q + \|f_2\|_q, \quad (2.11)$$

where g, g_1, g_2 are as in (2.8). For the equality case we observe that if $\|f_1 + f_2\|_q \in (0, +\infty)$ we proved above that there is g as in (2.8) so that $\|f_1 + f_2\|_q = \int (f_1 + f_2)g \, d\mathbf{m}$ and (2.10) shows that the only such g is given by $g = (\frac{f_1 + f_2}{\|f_1 + f_2\|_q})^{\frac{q}{p}}$. If also $\|f_1\|_q, \|f_2\|_q \in (0, +\infty)$, then the same argument and the equality in (2.8) force $g = (\frac{f_1}{\|f_1\|_q})^{\frac{q}{p}} = (\frac{f_2}{\|f_2\|_q})^{\frac{q}{p}}$ and the conclusion easily follows. If instead one of f_1, f_2 , say f_1 , has q -norm equal to 0, then equality in (2.9) implies that $\int f_2^q \, d\mathbf{m} = \int (f_1 + f_2)^q \, d\mathbf{m}$ and since both sides are real (from $\|f_1 + f_2\|_q \in (0, +\infty)$) the two integrands are in $L^1(\mathbf{m})$ and since $f_2 \leq f_1 + f_2$ holds \mathbf{m} -a.e., the equality forces $f_1 = 0$ \mathbf{m} -a.e., that is the conclusion. Finally, the case $\|f_1\|_q = \|f_2\|_q = 0$ cannot occur in conjunction with the equality in (2.9) and $\|f_1 + f_2\|_q \in (0, +\infty)$. \square

Remark 2.14. For $q \in (0, 1)$ we might have $\|f_1\|_q = \|f_2\|_q = \|f_1 + f_2\|_q = +\infty$ in (2.9) without any rigidity. Similarly, if $q < 0$ we might have $\|f_1\|_q = \|f_2\|_q = \|f_1 + f_2\|_q = 0$ without rigidity. \blacksquare

In order to prove antisymmetry of \preceq , the following lemma will be useful:

Lemma 2.15 (Recurrence of self-couplings). *Let X be a Polish space and $\pi \in \mathcal{P}(\mathsf{X}^2)$ have equal first and second marginals.*

Then for every distance \mathbf{d} inducing the Polish topology on X and $r > 0$ we have: for π -a.e. (x, y) there is a sequence $(x_n) \subset \mathsf{X}$ with $x_0 = x, x_1 = y, (x_n, x_{n+1}) \in \text{spt } \pi$ for every $n \in \mathbf{N}$ and

$$\liminf_{n \rightarrow \infty} \mathbf{d}(x_0, x_n) \leq r. \quad (2.12)$$

Proof. The proof is based on a variant ‘for couplings’ of the classical Poincaré recurrence theorem (for which we will follow the presentation from Tao [116]). Let $\mu := (\text{Pr}_1)_* \pi = (\text{Pr}_2)_* \pi$ and let (π_x) be the disintegration of π w.r.t. the projection Pr_1 on the first marginal. Define $\boldsymbol{\pi} \in \mathcal{P}(\mathsf{X}^{\mathbf{N}})$ as

$$d\boldsymbol{\pi}(x_0, x_1, \dots) = \dots d\pi_{x_2}(x_3) d\pi_{x_1}(x_2) d\pi_{x_0}(x_1) d\mu(x_0).$$

In other words, $\boldsymbol{\pi}$ is the measure given by Kolmogorov’s extension theorem associated to the measures $\pi^N \in \mathcal{P}(\mathsf{X}^N)$, where $d\pi^N(x_0, \dots, x_{N-1}) = d\pi_{x_{N-2}}(x_{N-1}) \dots d\pi_{x_0}(x_1) d\mu(x_0)$. Notice that the construction grants

$$(\text{Pr}_n, \text{Pr}_{n+1})_* \boldsymbol{\pi} = \pi \quad \forall n \in \mathbf{N} \quad (2.13)$$

and thus

$$\boldsymbol{\pi} \text{ is concentrated on sequences } (x_n) \text{ such that } (x_n, x_{n+1}) \in \text{spt } \pi \text{ for every } n \in \mathbf{N}. \quad (2.14)$$

We claim that for every $E \subset \mathsf{X}^2$ Borel we have

$$\limsup_{n \rightarrow \infty} \boldsymbol{\pi} \left(\left\{ (x_i) \in \mathsf{X}^{\mathbf{N}} : (x_0, x_1), (x_n, x_{n+1}) \in E \right\} \right) \geq \pi(E)^2. \quad (2.15)$$

Indeed, from (2.13) it follows that $\int \sum_{i < n} 1_{(\text{Pr}_i, \text{Pr}_{i+1})^{-1}(E)} \, d\boldsymbol{\pi} = n\pi(E)$ and the Cauchy-Schwarz inequality yields

$$n^2 \pi(E)^2 \leq \int \left(\sum_{i < n} 1_{(\text{Pr}_i, \text{Pr}_{i+1})^{-1}(E)} \right)^2 \, d\boldsymbol{\pi} = \sum_{i, j < n} \boldsymbol{\pi} \left(\underbrace{(\text{Pr}_i, \text{Pr}_{i+1})^{-1}(E) \cap (\text{Pr}_j, \text{Pr}_{j+1})^{-1}(E)}_{=(\text{Pr}_i, \text{Pr}_{i+1}, \text{Pr}_j, \text{Pr}_{j+1})^{-1}(E \times E)} \right).$$

From the definition it also follows that $(\text{Pr}_i, \text{Pr}_{i+1}, \text{Pr}_j, \text{Pr}_{j+1})_* \pi = (\text{Pr}_0, \text{Pr}_1, \text{Pr}_{j-i}, \text{Pr}_{j-i+1})_* \pi$ whenever $i < j$, thus calling L the lim sup in (2.15), it is not hard to see that

$$\limsup_{n \rightarrow \infty} \frac{1}{n^2} \sum_{i,j < n} \pi \left((\text{Pr}_i, \text{Pr}_{i+1})^{-1}(E) \cap (\text{Pr}_j, \text{Pr}_{j+1})^{-1}(E) \right) \leq L,$$

and the claim (2.15) follows. We now claim that for every $E \subset \mathbb{X}^2$ Borel it holds that

$$\text{for } \pi\text{-a.e. } (x, y) \in E \text{ we have } \pi_{(x,y)}(\{(x_i) : (x_n, x_{n+1}) \in E \text{ for infinite many } n\}) > 0, \quad (2.16)$$

where $(\pi_{(x,y)})$ is the disintegration of π w.r.t. the projection on the first two coordinates. Indeed, if this fails for some E , for some Borel set $F \subset E$ with $\pi(F) > 0$ we have

$$0 = \pi_{(x,y)}(\{(x_n, x_{n+1}) \in E \text{ for infinite } n\}) \geq \pi_{(x,y)}(\{(x_n, x_{n+1}) \in F \text{ for infinite } n\})$$

and therefore $\pi(\{(x_i) : (x_n, x_{n+1}) \in F \text{ for infinite } n\}) = 0$, which contradicts (2.15).

Hence (2.16) holds and applying it with $E := B \times \mathbb{X}$, where $B \subset \mathbb{X}$ is a ball of diameter $< r$ and recalling (2.14) and that $\pi_{(x,y)}$ is — for π -a.e. (x, y) — concentrated on sequences with $x_0 = x$ and $x_1 = y$, we see that for π -a.e. (x, y) with $x \in B$ there is (x_n) with $(x_n, x_{n+1}) \in \text{spt } \pi$ for every $n \in \mathbb{N}$ and $x_n \in B$ for infinite n 's. In particular $\liminf_n d(x, x_n) \leq r$ and since \mathbb{X} can be covered by a countable collection of such balls B , the conclusion follows. \square

Notice that if the coupling π in the statement is induced by a map, then the claim (2.15) reduces to the classical recurrence theorem by picking $E := F \times \mathbb{X}$ for arbitrary $F \subset \mathbb{X}$ Borel. Also, the measure π built in the statement reflects the idea that the coupling π can be seen as the map sending x to the fiber π_x of the disintegration of π w.r.t. the projection onto the first coordinate, so that ‘to iterate π means to iterate this assignment’, in a sense.

With this said, we have the following general statement; c.f. [85] [25, Theorem B.2].

Proposition 2.16 (Probability measures inherit Polish structure from metric spacetime). *Let (M, τ, ℓ) be a Polish metric spacetime so that ℓ is Borel and $0 \neq q \leq 1$.*

Then $(\mathcal{P}(M), \ell_q)$ equipped with the narrow topology is a Polish metric spacetime as well.

Proof. The reverse triangle inequality is known and based on a standard gluing argument together with the reverse triangle inequality (2.9) in L^q for $0 \neq q < 1$: see e.g. [52, Thm. 13] or [39, Prop. 2.5] and notice that the proofs do not depend on global hyperbolicity and that, even though are stated in the range $q \in (0, 1]$, remain unchanged for negative q (see also Lemma 2.13 above). We now claim that

$$\text{for any } \mu \in \mathcal{P}(M) \text{ the plan } (\text{id}, \text{id})_* \mu \text{ is the only element of } \Pi_{\leq}(\mu, \mu). \quad (2.17)$$

The fact that $(\text{id}, \text{id})_* \mu \in \Pi_{\leq}(\mu, \mu)$ is obvious. To show that it is the only element we argue by contradiction, so we assume that $\pi \in \Pi_{\leq}(\mu, \mu)$ is not concentrated on the diagonal. Let $r > 0$ and apply Lemma 2.15 to find a sequence $(x_n) \subset M$ with $x_0 \neq x_1$, $(x_n, x_{n+1}) \in \text{spt } \pi \subset \{\ell \geq 0\}$ (and thus $x_n \leq x_{n+1}$) for every $n \in \mathbb{N}$ satisfying (2.12). In particular there is $x^r = x_{n(r)}$ for a suitable $n(r)$ with $x_0 \leq x_1 \leq x^r$ and $x^r \in B_r(x_0)$. Letting $r \downarrow 0$ and using the closure of $\{\ell \geq 0\}$ we deduce that $x_0 \leq x_1 \leq x_0$ and conclude that $x_0 \leq x_1 \leq x_0$, i.e. that $x_0 = x_1$, contradicting the assumption.

Thus (2.17) holds, and therefore $\ell_q(\mu, \mu) = 0$ for any $\mu \in \mathcal{P}(M)$. To conclude that $(\mathcal{P}(M), \ell_q)$ is a metric spacetime it remains to show that $\ell_q(\mu, \nu) = \ell_q(\nu, \mu) = 0$ implies $\mu = \nu$. To see this, let $\pi_1 \in \Pi_{\leq}(\mu, \nu)$, $\pi_2 \in \Pi_{\leq}(\nu, \mu)$ and $\pi \in \mathcal{P}(M^3)$ a gluing of these along the common ν -marginal. Since π_1, π_2 are causal, the construction grants that $x \leq y \leq z$ for π -a.e. (x, y, z) and since $(\text{Pr}_1, \text{Pr}_3)_* \pi \in \Pi_{\leq}(\mu, \mu)$, we deduce from (2.17) that for π -a.e. (x, y, z) we have $x = z$, and thus that $x = y = z$ for π -a.e. (x, y, z) . In particular $\mu = \nu$, as desired.

We come to the topological aspects. The fact that the narrow topology on $\mathcal{P}(M)$ is Polish is well known (see e.g. [9, Remark 7.1.7]), so to conclude we need only to show that $\{(\mu, \nu) : \mu \leq \nu\}$ is closed. This, however, is obvious: if $(\mu_n), (\nu_n)$ narrowly converge to μ, ν respectively and $\pi_n \in \Pi_{\leq}(\mu_n, \nu_n)$ for every n , then (π_n) is tight and any narrow limit is, by the closure of $\{\ell \geq 0\}$, an element of $\Pi_{\leq}(\mu, \nu)$. \square

We now want to investigate how forward completeness of \mathbb{M} is inherited by $\mathcal{P}(\mathbb{M})$. To this aim, the following simple lemma will be useful:

Lemma 2.17 ('Limit' as Borel map). *Let \mathbb{X} be a Polish space, $\hat{\mathbb{X}} := \prod_{n \in \mathbb{N}} \mathbb{X}$ with the (Polish) product topology and consider the subset $\hat{\mathbb{X}}_{\mathbb{L}} \subset \hat{\mathbb{X}}$ of sequences admitting a limit in \mathbb{X} .*

Then $\hat{\mathbb{X}}_{\mathbb{L}}$ is a Borel subset of $\hat{\mathbb{X}}$ and the map sending a sequence in $\hat{\mathbb{X}}_{\mathbb{L}}$ to its limit in \mathbb{X} is Borel.

Proof. Let $(y_k) \subset \mathbb{X}$ be countable and dense and, for $r > 0$, consider the set $\hat{\mathbb{X}}_{r,k} \subset \hat{\mathbb{X}}$ defined as

$$\hat{\mathbb{X}}_{r,k} := \bigcup_{n \in \mathbb{N}} \{(x_i) \in \hat{\mathbb{X}} : x_i \in B_r(y_k) \forall i \geq n\} = \bigcup_{n \in \mathbb{N}} \bigcap_{i \geq n} \text{Pr}_i^{-1}(B_r(y_k))$$

that is clearly Borel. Then so is $\hat{\mathbb{X}}_r := \cup_k \hat{\mathbb{X}}_{r,k}$ and so is the map $\mathbb{L}_r : \hat{\mathbb{X}}_r \rightarrow \mathbb{X}$ sending (x_i) to y_k , where $k \in \mathbb{N}$ is the least integer such that $(x_i) \in \hat{\mathbb{X}}_{r,k}$. To conclude, notice that for any $r_j \downarrow 0$ we have $\hat{\mathbb{X}}_{\mathbb{L}} = \bigcap_j \hat{\mathbb{X}}_{r_j}$ and that for $(x_i) \in \hat{\mathbb{X}}_{\mathbb{L}}$ the limit in j of $\mathbb{L}_{r_j}((x_i))$ exists and is the limit of (x_i) . \square

We then have:

Proposition 2.18 (Probability measures inherit forward structure from metric spacetime). *Let (\mathbb{M}, τ, ℓ) be a forward metric spacetime so that ℓ is Borel and $0 \neq q < 1$.*

Then $(\mathcal{P}(\mathbb{M}), \ell_q)$ equipped with the narrow topology is a forward metric spacetime as well.

Proof. By Proposition 2.16 we need only to check forward completeness, thus let $(\mu_n) \subset \mathcal{P}(\mathbb{M})$ be such that $\mu_n \preceq \mu_{n+1} \preceq \bar{\mu}$ for every $n \in \mathbb{N}$ and some $\bar{\mu} \in \mathcal{P}(\mathbb{M})$.

Let $\hat{\mathbb{M}} := \prod_{n \in \mathbb{N} \cup \{\infty\}} \mathbb{M}$ be equipped with the product (hence Polish) topology and $\mathcal{K} \subset \mathcal{P}(\hat{\mathbb{M}})$ be the collection of measures α such that $(\text{Pr}_n)_* \alpha = \mu_n$ for any $n \in \mathbb{N}$ and $(\text{Pr}_\infty)_* \alpha = \bar{\mu}$. We claim that \mathcal{K} is narrowly compact and since it clearly is narrowly closed (by the continuity of the projections), to see this it suffices to prove that it is tight. Thus fix $\varepsilon > 0$ and find compact sets $K_n \subset \mathbb{M}$, $n \in \mathbb{N} \cup \{\infty\}$, such that $\mu_n(K_n^c) < \frac{\varepsilon}{2^n}$ for every $n \in \mathbb{N}$ and $\bar{\mu}(K_\infty^c) < \varepsilon$. Then the set $K := \prod_{n \in \mathbb{N} \cup \{\infty\}} K_n$ is compact and it is easy to see that $\alpha(K^c) < 2\varepsilon$ for every $\alpha \in \mathcal{K}$, thus proving the claim.

Now for every $n \in \mathbb{N}$ let $\pi_n \in \Pi_{\leq}(\mu_n, \mu_{n+1})$ and $\tilde{\pi}_n \in \Pi_{\leq}(\mu_n, \bar{\mu})$ (the existence of such plans is ensured by our assumptions). Using a gluing argument and then Kolmogorov's theorem we can find, for every $n \in \mathbb{N}$, a measure $\alpha_n \in \mathcal{K}$ such that

$$\begin{aligned} (\text{Pr}_i, \text{Pr}_{i+1})_* \alpha_n &= \pi_i & \forall i \in \mathbb{N}, \\ (\text{Pr}_n, \text{Pr}_\infty)_* \alpha_n &= \tilde{\pi}_n. \end{aligned} \tag{2.18}$$

By the compactness of \mathcal{K} , possibly passing to a non-relabelled subsequence we can assume that (α_n) narrowly converges to a limit $\alpha \in \mathcal{K}$. We claim that $\text{spt}(\alpha)$ is contained in the set of sequences $\mathbf{x} = (\dots, x_i, \dots, x_\infty) \in \hat{\mathbb{M}}$ such that $x_1 \leq \dots \leq x_i \leq \dots \leq x_\infty$. To see this, let $\mathbf{x} \in \text{spt}(\alpha)$ and find $\mathbf{x}^n \in \text{spt}(\alpha_n)$ converging to \mathbf{x} in $\hat{\mathbb{M}}$ (i.e. $x_i^n \rightarrow x_i$ as $n \rightarrow \infty$ for every $i \in \mathbb{N} \cup \{\infty\}$). Then for every $i \in \mathbb{N}$ and every $n \geq i$, by the defining properties (2.18) we have $x_i^n \leq x_{i+1}^n$ and $x_i^n \leq x_\infty^n$. Passing to the limit in these (recall that $\{\ell \geq 0\}$ is closed), we get our claim.

It follows by forward completeness that for α -a.e. \mathbf{x} there exists the limit $\lim(\mathbf{x})$ of $i \mapsto x_i$ and since the partially defined limit map $\lim : \hat{\mathbb{M}} \rightarrow \mathbb{M}$ is Borel (by Lemma 2.17 above), we can define $\mu_\infty := \lim_* \alpha$. We check that μ_∞ is the desired narrow limit of (μ_n) : let $\varphi \in C_b(\mathbb{M})$ and notice that

$$\lim_{n \rightarrow \infty} \int \varphi d\mu_n = \lim_{n \rightarrow \infty} \int \varphi(x_n) d\alpha(\mathbf{x}) = \int \varphi(\lim(\mathbf{x})) d\alpha(\mathbf{x}) = \int \varphi d\mu_\infty,$$

having used the dominated convergence theorem in the second identity. \square

For later use we also record the following simple fact:

Lemma 2.19 (Upper semicontinuity of ℓ_q and attainment on emeralds). *Let (M, τ, ℓ) be a Polish metric spacetime such that ℓ is upper semicontinuous, does not take the value $+\infty$ and let $0 \neq q < 1$.*

Then for every $\mu, \nu \in \mathcal{P}_{\text{em}}(M)$ with $\mu \preceq \nu$ an ℓ_q -optimal coupling exists. Moreover, if $E \subset M$ is an emerald and $(\mu_n), (\nu_n)$ sequences of measures concentrated in E narrowly converging to μ, ν respectively, then μ, ν are also concentrated on E and

$$\limsup_{n \rightarrow \infty} \ell_q(\mu_n, \nu_n) \leq \ell_q(\mu, \nu).$$

Proof. We start from the second claim. The closedness of E established in Lemma 2.4 ensures that μ, ν have support in E and that $\sup_n \ell_q(\mu_n, \nu_n) < +\infty$. Now let $\pi_n \in \Pi_{\leq}(\mu_n, \nu_n)$ be such that $\int u_q \circ \ell d\pi_n \geq u_q(\ell_q(\mu_n, \nu_n)) - \frac{1}{n}$ and notice that the tightness of $(\mu_n), (\nu_n)$ yields that of (π_n) , so that passing to a non-relabelled subsequence we can assume that the sequence narrowly converges to some $\pi \in \Pi_{\leq}(\mu, \nu)$. Since ℓ is upper semicontinuous and bounded from above on E^2 , the narrow convergence yields

$$\limsup_{n \rightarrow \infty} \int u_q \circ \ell d\pi_n \leq \int u_q \circ \ell d\pi \leq u_q(\ell_q(\mu, \nu)),$$

and the conclusion follows. For the first claim just apply the above with $\mu_n = \mu$ and $\nu_n = \nu$ for every $n \in \mathbf{N}$. \square

2.4 Causal speed

Let us fix a metric spacetime (M, ℓ) . A map $\gamma : [0, 1] \rightarrow M$ is called *causal path* if it is monotone, i.e. $t, s \in [0, 1]$ with $t \leq s$ imply $\gamma_t \leq \gamma_s$. Note that the time-orientation is inherently given by the relation \leq , so that all causal paths are future directed. We do not assume causal paths are continuous (at this stage M does not even have a topology, but even when topology will be added, continuity will not be imposed).

The goal of this section is to discuss in which sense a causal path admits a *causal speed*. Heuristically, in analogy to the speed of absolutely continuous curves in metric geometry [4], its causal speed $|\dot{\gamma}_t|$ at an appropriate $t \in [0, 1]$ should be given as $\lim_{h \downarrow 0} \ell(\gamma_t, \gamma_{t+h})/h$. Making this idea precise requires a generalization of the well-known fact that the distributional derivative of a monotone function $F : \mathbf{R} \rightarrow \mathbf{R}$ is a non-negative Radon measure on the interior of its interval of finiteness. This is also related to the classical result of Lebesgue asserting that monotone functions are differentiable \mathcal{L}^1 -a.e.

Theorem 2.23 and its Corollary 2.24 establish similar differentiability results for time separation functions $T : [0, 1] \times [0, 1] \rightarrow \{-\infty\} \cup [0, +\infty]$ near the diagonal of the unit square. Choosing $T(s, t) := \ell(\gamma_s, \gamma_t)$ allows us to associate the desired causal speed and — for any suitable choice of Lagrangian — a global action with each causal path γ in M . Choosing $T(s, t) := F(t) - F(s)$ recovers the distributional and a.e. derivatives of a monotone function F as above.

Let us start by recalling a few standard facts.

Definition 2.20 (Vitali cover). *A Vitali cover of a set $E \subset \mathbf{R}$ is a collection \mathcal{V} of closed non-trivial intervals such that for every $x \in E$ and every $\varepsilon > 0$ there is $I \in \mathcal{V}$ containing x (so in particular the collection is a cover) with $\mathcal{L}^1(I) \leq \varepsilon$.*

For a proof of the following well known theorem we refer to Bogachev [18, Thm. 5.5.1].

Theorem 2.21 (Vitali's covering lemma). *Let $E \subset \mathbf{R}$ be Borel with $\mathcal{L}^1(E) < \infty$, $\varepsilon > 0$, and \mathcal{V} a Vitali cover of E . Then there are finitely many intervals $I_n \in \mathcal{V}$ that are disjoint and so that $\mathcal{L}^1(E \setminus \bigcup_n I_n) < \varepsilon$.*

The following simple lemma will be useful in defining the causal speed:

Lemma 2.22 (Non-oscillation criterion). *Let $f, \omega : (0, 1) \rightarrow (0, +\infty)$ be such that $\lim_{h \downarrow 0} \omega(h) = 0$ and so that for every $n \in \mathbf{N}$, every finite sequence $(\alpha_i) \subset [0, 1]$ with $\sum \alpha_i = 1$, and every $h \in (0, 1)$ we have*

$$\sum_{i=0}^n \alpha_i f(\alpha_i h) \leq f(h) + \omega(h).$$

Then the limit $\lim_{h \downarrow 0} f(h)$ exists.

Proof. Let $h, \alpha \in (0, 1]$ and n be the integer part of $1/\alpha$. Our assumption yields

$$n\alpha f(\alpha h) \leq n\alpha f(\alpha h) + (1 - n\alpha)f((1 - n\alpha)h) \leq f(h) + \omega(h)$$

for every $h \in (0, 1)$. Keeping h fixed and letting $\alpha \downarrow 0$ (noting that $n\alpha \rightarrow 1$ in this procedure) we obtain

$$\limsup_{h' \downarrow 0} f(h') = \limsup_{\alpha \downarrow 0} f(\alpha h) \leq f(h) + \omega(h).$$

Letting $h \downarrow 0$ we conclude $\limsup_{h' \downarrow 0} f(h') \leq \liminf_{h \downarrow 0} f(h)$, which is the claim. \square

In the sequel, consider the upper left halfsquare

$$H := \{(s, t) \in [0, 1]^2 : s \leq t\}.$$

Theorem 2.23 (Differentiation of the reverse triangle inequality). *Let $T: H \rightarrow [0, +\infty)$ be a function satisfying*

$$T(r, s) \leq T(r, t) - T(s, t) \quad (2.19)$$

for all $0 \leq r \leq s \leq t \leq 1$. Then the following hold.

(i) *There exists a unique maximal element μ among all Borel measures ν on \mathbf{R} which satisfy*

$$\nu((a, b)) \leq T(a, b), \quad \forall a, b \in \mathbf{R}, \quad a < b, \quad (2.20)$$

where here we are extending T to the whole halfplane $\{s \leq t\} \subset \mathbf{R}^2$ by putting

$$T(s, t) := T(0 \vee s \wedge 1, 0 \vee t \wedge 1). \quad (2.21)$$

(notice that this extension still satisfies (2.19)).

(ii) *The measure μ is the weak limit as $h \downarrow 0$ of the measures μ_h given by*

$$d\mu_h(t) := \frac{T(t, t+h)}{h} d\mathcal{L}^1(t).$$

Moreover, writing $\mu = \rho \mathcal{L}^1 + \mu^\perp$ for the Lebesgue decomposition of μ with respect to \mathcal{L}^1 , we have that \mathcal{L}^1 -a.e. $t \in \mathbf{R}$ satisfies

$$\rho(t) = \lim_{n \rightarrow +\infty} \frac{T(r_n, s_n)}{s_n - r_n} \quad (2.22)$$

for all sequences $(r_n), (s_n) \subset [0, 1]$ so that $r_n \uparrow t$, $s_n \downarrow t$ and $r_n < s_n$ for every $n \in \mathbf{N}$.

(iii) *For every $0 \neq q < 1$, write $u_q(z) := \frac{z^q}{q}$ (recall also the conventions in Section 2.1 for $u_q(0)$) and for a partition $P = \{0 = t_0 < t_1 < \dots < t_n = 1\}$ of $[0, 1]$ define the ‘discrete q -action’ $A_q(P)$ as*

$$A_q(P) := \sum_{i=0}^{n-1} (t_{i+1} - t_i) u_q\left(\frac{T(t_i, t_{i+1})}{t_{i+1} - t_i}\right).$$

Then

$$P \text{ finer than } Q \quad \Rightarrow \quad A_q(P) \leq A_q(Q) \quad (2.23)$$

and

$$\int_0^1 u_q(\rho) d\mathcal{L}^1 = \inf A_q(P), \quad (2.24)$$

where the infimum is taken among all partitions of $[0, 1]$.

(iv) For a partition $P = \{0 = t_0 < t_1 < \dots < t_n = 1\}$ we set $|P| := \max |t_{i+1} - t_i|$. Then for any sequence (P_k) of partitions such that $|P_k| \rightarrow 0$ we have

$$\int_0^1 u_q(\rho) d\mathcal{L}^1 = \lim_{k \rightarrow \infty} A_q(P_k). \quad (2.25)$$

Proof. (ii.a) Notice that the measures μ_h are non-negative, because T is, and that for any $[a, b] \subset \mathbf{R}$ we have

$$\mu_h([a, b]) = \frac{1}{h} \int_a^b T(t, t+h) dt = \frac{1}{h} \int_a^{a+h} \sum_{i=0}^n T(t+ih, t+(i+1)h) dt \stackrel{(2.19)}{\leq} T(a, b+2h), \quad (2.26)$$

where here n is the integer part of $\frac{b-a}{h}$. Now the existence of the weak limit follows the idea from the ‘‘subpartition lemma’’ of Korevaar–Schoen [72, Lem. 1.3.1]. Fix a non-negative function $\varphi \in C_c(\mathbf{R})$ and denote its modulus of continuity by $\omega: [0, +\infty) \rightarrow [0, +\infty)$. Pick $h > 0$ and a finite sequence $(\alpha_i) \subset [0, 1]$ with $\sum \alpha_i = 1$. Set $A_i := \sum_{j \leq i} \alpha_j$ and notice that

$$\begin{aligned} \sum_{i=1}^n \alpha_i \int \varphi d\mu_{\alpha_i h} &= \frac{1}{h} \sum_{i=1}^n \int \varphi(t) T(t, t + \alpha_i h) dt \\ &= \frac{1}{h} \int \sum_{i=1}^n \varphi(t) T(t + A_{i-1}h, t + A_i h) - (\varphi(t - A_{i-1}h) - \varphi(t)) T(t, t + \alpha_i h) dt \\ \text{(by (2.19))} \quad &\leq \int \varphi d\mu_h + \frac{1}{h} \sum_{i=1}^n \omega(h) \int T(t, t + \alpha_i h) dt \\ &= \int \varphi d\mu_h + \omega(h) \sum_{i=1}^n \alpha_i \mu_{\alpha_i h}(\mathbf{R}) \leq \int \varphi d\mu_h + \omega(h) T(0, 1), \end{aligned}$$

having used (2.26) and (2.21) in the last step. Applying Lemma 2.22 to $f(h) := \int \varphi d\mu_h$ we infer that $L(\varphi) := \lim_{h \downarrow 0} \int \varphi d\mu_h$ exists. It is then clear that this limit also exists for $\varphi \in C_c(\mathbf{R})$ possibly negative and that the resulting functional $L: C_c(\mathbf{R}) \rightarrow \mathbf{R}$ is non-negative and finite. It follows by the Riesz–Markov–Kakutani theorem that there is a (unique) non-negative Radon measure μ representing L , i.e. such that $\int \varphi d\mu = \lim_n \int \varphi d\mu_n$.

We now come to the verification of all claims for this measure μ .

(i) Let $\varepsilon > 0$ satisfy $a < b - \varepsilon$. Then lower semicontinuity on open sets under narrow convergence of finite measures and the estimate (2.26) imply

$$\mu((a, b - \varepsilon)) \leq \liminf_{h \downarrow 0} \mu_h((a, b - \varepsilon)) \leq \liminf_{h \downarrow 0} \mu_h((a, b - 2h)) \leq T(a, b).$$

Letting $\varepsilon \downarrow 0$ we see that μ satisfies (2.20).

To show maximality of μ , assume ν is any Borel measure on \mathbf{R} satisfying the claimed inequality. Since formulas (2.21) and (2.19) grant that $T(a, b) = 0$ when $a < b$ are both negative or both bigger than 1, we see that ν must be concentrated on $[0, 1]$, thus for $h > 0$ we can define $\nu_h := \nu * 1_{[0, h]}/h$. The assumption (2.20) implies $\nu_h \leq \mu_h$ for every n , and sending $h \downarrow 0$ yields $\nu \leq \mu$.

(ii.b) We now prove the identity (2.22). Given $t \in [0, 1]$ we denote by $W(t)$ the collection of closed non-trivial intervals in \mathbf{R} that contain t . Define $I, S: [0, 1] \rightarrow [0, +\infty]$ by

$$I(t) := \lim_{\varepsilon \rightarrow 0} \inf_{\substack{[r, s] \in W(t) \\ |s-r| < \varepsilon}} \frac{T(r, s)}{s-r}, \quad S(t) := \lim_{\varepsilon \rightarrow 0} \sup_{\substack{[r, s] \in W(t) \\ |s-r| < \varepsilon}} \frac{T(r, s)}{s-r}.$$

The limits exist because the relevant quantities are monotone in ε ; for the same reason, from the Lebesgue measurability of T (that follows from coordinatwise monotonicity — see [42, Thm. 4]) we get Lebesgue measurability of I and S . We claim that

1. \mathcal{L}^1 -a.e. $t \in [0, 1]$ satisfies

$$I(t) = S(t) < \infty,$$

2. the density ρ is equal to I (and hence S) \mathcal{L}^1 -a.e. on $[0, 1]$.

These will readily imply the claimed identity.

(ii.b.1) To prove the equality of I and S , it is sufficient to prove that the set $\{t \in [0, 1] : I(t) < S(t)\}$ is \mathcal{L}^1 -negligible. For $\alpha < \beta$ rational let $E := \{t \in [0, 1] : I(t) < \alpha < \beta < S(t)\}$. By the arbitrariness of α and β , to prove the claim it suffices to show that $\mathcal{L}^1(E) = 0$.

Fix $\varepsilon > 0$ and find an open superset $U \supset E$ with $\mathcal{L}^1(U \setminus E) < \varepsilon$ (notice that $I = S = 0$ outside $[0, 1]$, so $E \subset [0, 1]$ has finite measure). Let \mathcal{V} be the collection of closed nontrivial intervals $[r, s]$ contained in U such that $T(r, s) < \alpha(s - r)$. By definition of E , the family \mathcal{V} is a Vitali covering of E . We thus apply Vitali's covering lemma to find a finite disjoint family $[r_k, s_k]$ such that $\mathcal{L}^1(E \setminus \bigcup_k [r_k, s_k]) < \varepsilon$. We have

$$\sum_k T(r_k, s_k) \leq \alpha \sum_k (s_k - r_k) \leq \alpha \mathcal{L}^1(U) \leq \alpha(\mathcal{L}^1(E) + \varepsilon).$$

Let now $U' := \bigcup_k (r_k, s_k)$ and $F := E \cap U'$ and notice that $\mathcal{L}^1(F) > \mathcal{L}^1(E) - \varepsilon$. Also, let \mathcal{V}' be the collection of closed non-trivial intervals $[a, b]$ contained in U' such that $T(a, b) > \beta(b - a)$. Again, by definition of E we have that \mathcal{V}' is a Vitali covering of F , hence we can find a finite disjoint subfamily $[a_j, b_j]$ such that $\mathcal{L}^1(F \setminus \bigcup_j [a_j, b_j]) < \varepsilon$. Therefore,

$$\sum_j T(a_j, b_j) > \beta \sum_j (b_j - a_j) > \beta(\mathcal{L}^1(F) - \varepsilon) > \beta(\mathcal{L}^1(E) - 2\varepsilon).$$

Since each of the intervals $[a_j, b_j]$ is included in some of the $[r_k, s_k]$, from (2.19) and the fact that T is non-negative, non-increasing in the first entry and non-decreasing in the second one (from (2.19)) we deduce

$$\alpha(\mathcal{L}^1(E) + \varepsilon) \geq \sum_k T(r_k, s_k) \geq \sum_j T(a_j, b_j) > \beta \sum_j (b_j - a_j) > \beta(\mathcal{L}^1(E) - 2\varepsilon)$$

and the arbitrariness of ε and the choice $\alpha < \beta$ imply $\mathcal{L}^1(E) = 0$, as desired.

Next, we claim $\{t \in [0, 1] : S(t) = +\infty\}$ is \mathcal{L}^1 -negligible.

For β rational, let E be the set of t 's with $S(t) > \beta$, and let $\varepsilon > 0$. Then, arguing as before, we can find finite disjoint intervals $[a_j, b_j]$ such that $\mathcal{L}^1(E \setminus \bigcup_j [a_j, b_j]) < \varepsilon$ and $T(a_j, b_j) > \beta(b_j - a_j)$ for every j . It is not restrictive to assume that these intervals are all contained in $[0, 1]$. Then adding up, by (2.19) and the fact that T is non-negative we get

$$\infty > T(0, 1) \geq \sum_j T(a_j, b_j) > \beta \sum_j (b_j - a_j) > \beta(\mathcal{L}^1(E) - \varepsilon).$$

The arbitrariness of ε then yields $\mathcal{L}^1(E) < T(0, 1)/\beta$, and the claim follows.

(ii.b.2) To conclude we need to prove that ρ is equal to I , thus to S , \mathcal{L}^1 -a.e. on $[0, 1]$. To see this recall that by classical results about differentiation of measures (cf. e.g. [18, Thm. 5.8.8] and notice that it is not a coincidence that these are proved along the same lines we just used) we have $h^{-1} \mu([t, t+h]) \rightarrow \rho(t)$ as $h \downarrow 0$ for \mathcal{L}^1 -a.e. $t \in [0, 1]$, thus by the defining property (i) of μ we deduce \mathcal{L}^1 -a.e. $t \in [0, 1]$ satisfies

$$\rho(t) \leq I(t).$$

On the other hand, let $C \subset [0, 1]$ be compact with $\mu^\perp(C) = 0$. Then the weak convergence of (μ_h) to μ as $h \downarrow 0$ provided by the proof of (ii.a) and the fact that the set C is closed give

$$\int_C \rho d\mathcal{L}^1 = \mu(C) \geq \limsup_{h \downarrow 0} \mu_h(C),$$

while Fatou's lemma yields

$$\liminf_{h \downarrow 0} \mu_h(C) \geq \int_C \liminf_{h \downarrow 0} \frac{T(t, t+h)}{h} dt = \int_C I d\mathcal{L}^1 \geq \int_C \rho d\mathcal{L}^1.$$

Thus the inequalities must be equalities, and therefore $\rho = I$ almost everywhere on C . By the arbitrariness of C (and inner regularity of the Lebesgue measure) we conclude that $\rho = I$ \mathcal{L}^1 -a.e. on $[0, 1]$, as desired.

(iii)-(iv) Since u_q is concave and non-decreasing, for any $t < r < s$ we have

$$(r-t)u_q\left(\frac{T(t,r)}{r-t}\right) + (s-r)u_q\left(\frac{T(r,s)}{s-r}\right) \leq (s-t)u_q\left(\frac{T(t,r)}{r-t} + \frac{T(r,s)}{s-r}\right) \leq (s-t)u_q\left(\frac{T(t,s)}{s-t}\right)$$

which is (equivalent to) (2.23). Since u_q is also upper semicontinuous, for $a < b$ we have

$$\frac{1}{b-a} \int_a^b -u_q(\rho) d\mathcal{L}^1 \geq -u_q\left(\frac{1}{b-a} \int_a^b \rho d\mathcal{L}^1\right) \geq -u_q\left(\frac{\mu((a,b))}{b-a}\right) \geq -u_q\left(\frac{T(a,b)}{b-a}\right),$$

having used Jensen's inequality. From this, " \leq " in our claim (2.24) easily follows.

Let now (P_k) be a sequence of partitions as in (iv), i.e. such that $|P_k| \rightarrow 0$, say $P_k = \{0 = t_{k,0} < \dots < t_{k,n_k} = 1\}$. Define the 'piecewise constant' non-negative Radon measures ν_k on $[0, 1]$ as $\nu_k = \eta_k \mathcal{L}^1$, where $\eta_k(x) = \frac{T(t_{k,i}, t_{k,i+1})}{t_{k,i+1} - t_{k,i}}$ if $x \in [t_{k,i}, t_{k,i+1})$. Notice that $\nu_k([0, 1]) = \sum_i T(t_{k,i}, t_{k,i+1}) \leq T(0, 1)$, so passing to a non-relabelled subsequence we can assume that (ν_k) has a weak limit ν . We claim that $\nu = \mu$, and since this result does not depend on the particular subsequence chosen, this proves weak convergence of the full original sequence of measures to μ .

Let $[a', b'] \subset (a, b) \subset \mathbf{R}$ and then, for every $k \in \mathbf{N}$, let $[a_k, b_k]$ be the smallest interval containing $[a', b']$ with $a_k, b_k \in P_k$. Then we have

$$\nu_k([a', b']) \leq \nu_k([a_k, b_k]) = \sum_{i=i_1}^{i_2} T(t_{k,i}, t_{k,i+1}) \stackrel{(2.19)}{\leq} T(a_k, b_k),$$

where i_1, i_2 are so that $t_{k,i_1} = a_k$ and $t_{k,i_2} = b_k$. Since $|P_k| \rightarrow 0$ we eventually have $[a_k, b_k] \subset (a, b)$ and therefore

$$\nu((a', b')) \leq \liminf_k \nu_k((a', b')) \leq \limsup_k \nu_k([a', b']) \leq T(a, b).$$

By interior approximation of the interval (a, b) we see that ν satisfies (2.20) and thus, by what already proved, that $\nu \leq \mu$.

Conversely, let $[a, b] \subset (a', b') \subset \mathbf{R}$ and then, for every $k \in \mathbf{N}$, let $[a_k, b_k]$ be the smallest interval containing $[a, b]$ with $a_k, b_k \in P_k$. Then we have

$$\mu([a, b]) \leq \mu([a_k, b_k]) = \sum_{i=i_1}^{i_2} \mu([t_{k,i}, t_{k,i+1}]) \stackrel{(2.20)}{\leq} \sum_{i=i_1}^{i_2} T(t_{k,i}, t_{k,i+1}) = \nu_k([a_k, b_k]).$$

As before, from $|P_k| \rightarrow 0$ we deduce that $[a_k, b_k]$ is eventually contained in (a', b') and therefore

$$\mu([a, b]) \leq \limsup_k \nu_k([a_k, b_k]) \leq \limsup_k \nu_k([a', b']) \leq \nu([a', b']).$$

By exterior approximation of the interval $[a, b]$ we conclude that $\mu([a, b]) \leq \nu([a, b])$, then by interior approximation of open intervals we deduce $\mu((a, b)) \leq \nu((a, b))$ for any $(a, b) \subset \mathbf{R}$, then $\mu(U) \leq \nu(U)$ for any open set (as these are countable unions of disjoint intervals) and finally $\mu \leq \nu$ by outer regularity. Hence $\mu = \nu$, as claimed.

The conclusion now follows standard lower semicontinuity arguments. Define the functional V_q on the space of non-negative Radon measures on $[0, 1]$ as:

$$V_q(\nu) := - \int_0^1 u_q(\eta) d\mathcal{L}^1, \quad \text{where } \nu = \eta \mathcal{L}^1 + \nu^\perp \quad \text{with } \nu^\perp \perp \mathcal{L}^1.$$

Since $-u_q$ is lower semicontinuous and sublinear, V_q is lower semicontinuous on the space of non-negative finite measures on \mathbf{R} equipped with the weak convergence in duality with $C_c(\mathbf{R})$ (see e.g. [107, Proposition 7.7] — notice that in such reference the convex integrand is assumed to be real-valued, while in our case we have $-u_q(0) = +\infty$ if $q < 0$, still this is easily cured by approximating $-u_q$ from below with real-valued convex and lower semicontinuous functions). As we just proved that (ν_k) weakly converges to μ we deduce that

$$-V_q(\mu) \geq \limsup_{k \rightarrow \infty} (-V_q(\nu_k)),$$

which, expanding the definitions, is inequality \geq in (2.25). Since we already established the inequality \leq in (2.24), the proof is complete. \square

The previous Theorem 2.23 assumed finiteness of T (in which case the total mass of the measure μ therein has total mass bounded from above by $T(0, 1)$; in particular, μ is a Radon measure). For functions T taking infinity as a value (such as the ℓ_q -time separation between probability measures), we shall need a generalization of Theorem 2.23 given by the following corollary. Recall $S \subset \mathbf{R}^2$ is called *non-decreasing* if $(s' - s)(t' - t) \geq 0$ for all $(s, t), (s', t') \in S$; for example, the diagonal of \mathbf{R}^2 is non-decreasing.

Corollary 2.24 (Differentiation of the extended reverse triangle inequality). *Let $T: H \rightarrow [0, +\infty]$ satisfy*

$$T(r, s) \leq T(r, t) - T(s, t)$$

for every $r \leq s \leq t$. Then:

- i) *There is a maximal (with respect to inclusion) non-decreasing closed subset $S \subset H$ such that $S_\infty \subset \{T = \infty\}$ and $S_0 \subset \{T < \infty\}$, where S_∞ denotes the connected component of $H \setminus S$ containing the top left corner $(0, 1)$ of the square (but is empty if no such connected component exists) and $S_0 := H \setminus (S \cup S_\infty)$ is relatively open.*
- ii) *The intersection $S_0 \cap \partial H$ denotes a relatively open subset of the diagonal and Theorem 2.23 applies separately to each of its connected components in place of the unit interval.*
- iii) *The sum of the measures it yields can be extended to a maximal Borel measure μ on the diagonal satisfying the inequality*

$$\mu((a, b)) \leq T(a, b)$$

whenever $a \leq b$ by setting $\mu(A) := +\infty$ unless $A \subset S_0 \cap \partial H$.

Proof. Define $f(s) := \inf(\{t \in [0, 1] : T(s-t, s+t) = \infty\} \cup \{s, 1-s\})$. Then set $S := \{(s-t, s+t) \in H : t = f(s)\}$. Thus we have $S_\infty = \{(s-t, s+t) \in H : t > f(s)\}$ and $S_0 = \{(s-t, s+t) \in H : t < f(s)\}$. It is clear that S is maximal as a non-decreasing subset, and that $S_0 \subset \{T < \infty\}$. Note that as T is non-decreasing in the second coordinate and non-increasing in the first, we get that $T(s-t_0, s+t_0) \leq T(s-t, s+t)$ for $t_0 < t$, and in particular $S_\infty \subset \{T = \infty\}$.

By construction, the set S contains $(0, 0)$ and $(1, 1)$, thus $S_0 \cap \partial H$ is a subset of the diagonal. \square

The preceding results allow us to define the *causal speed* of a causal path, i.e. maps $\gamma: [0, 1] \rightarrow M$ such that $\gamma_t \leq \gamma_s$ whenever $t \leq s$. Indeed, for such a path, the map $T(s, t) := \ell(\gamma_s, \gamma_t)$ satisfies (2.19), thanks to the reverse triangle inequality on M and is non-negative by causality of γ . Hence Theorem 2.23 and Corollary 2.24 apply, thus we can give:

Definition 2.25 (Causal speed of curves in M). *Let $\gamma: [0, 1] \rightarrow M$ be a causal path with values in the metric spacetime (M, ℓ) .*

The maximal measure $|\dot{\gamma}| := \mu$ on $[0, 1]$ associated to γ via Theorem 2.23 and Corollary 2.24 will be called causal speed of γ . Its Lebesgue decomposition with respect to \mathcal{L}^1 is denoted by $|\dot{\gamma}| = |\dot{\gamma}|^{\mathcal{L}^1} + |\dot{\gamma}|^\perp$ and, abusing terminology, we shall also refer to $|\dot{\gamma}|$ as the path's causal speed.

To be explicit: maximality of $|\dot{\gamma}|$ and its absolutely continuous density $|\dot{\gamma}|$ means that

$$\int_t^s |\dot{\gamma}_r| dr \leq |\dot{\gamma}|([t, s]) \leq \ell(\gamma_t, \gamma_s) \quad \forall t, s \in [0, 1], t < s, \quad (2.27)$$

and that these are the maximal measure and L^1 -function bounded above by $\ell(\gamma_t, \gamma_s)$ in this sense. Also, property (2.22) implies, in particular, that

$$|\dot{\gamma}_t| = \lim_{h \downarrow 0} \frac{\ell(\gamma_t, \gamma_{t+h})}{h} = \lim_{h \downarrow 0} \frac{\ell(\gamma_{t-h}, \gamma_t)}{h} \quad a.e. t. \quad (2.28)$$

2.5 Left-continuous causal paths and their topology

In this section we deal with Polish metric spacetimes and study topology on and of causal paths.

Definition 2.26 (Left-continuous causal paths in M). *Let (M, τ, ℓ) be a Polish metric spacetime. A map $\gamma: [0, 1] \rightarrow M$ is called a left-continuous causal path if γ is left-continuous (i.e. $\gamma_t = \lim_{s \uparrow t} \gamma_s$ for every $t \in (0, 1]$) and causal. The space of all such paths is denoted $LCC([0, 1]; M)$.*

We choose to work with this class of curves for its good stability properties, see in particular Proposition 2.29 and Theorem 2.30. Left continuity is a concept closely related to that of forward completeness, which ensures that left limits exist. For this reason, in this section we shall mostly work on forward spacetimes.

Intuitively, when a test particle's worldline is known to be in $LCC([0, 1]; M)$, then its present position provides more information about its past than its future.

Lemma 2.27 (Countability of causal curve discontinuities). *Let (M, τ, ℓ) be a forward metric spacetime, $A \subset [0, 1]$ and $\gamma: A \rightarrow M$ causal (meaning $s, t \in A$ with $s \leq t$ imply $\gamma_s \leq \gamma_t$; no continuity assumptions).*

Then γ has at most a countable number of discontinuities.

Proof. Fix $c > 0$ and consider the set $B := \{t \in A : \limsup_{s \rightarrow t, s \in A} d(\gamma_s, \gamma_t) > c\}$: if we prove that B is countable for all $c > 0$ we are done. To derive a contradiction suppose B is uncountable. Then there are $t_0, T \in B$ with $t_0 < T$ such that $B \cap (t_0, T]$ is uncountable, hence there is $t_1 \in B \cap (t_0, T]$ so that $B \cap (t_1, T]$ is uncountable and by recursion we can build an increasing sequence $(t_n) \subset B$ with $B \cap (t_n, T]$ uncountable for every n . By definition of B there is also a sequence $(s_n) \subset A$ with $s_n \in (t_{n-1}, t_{n+1})$ and

$$d(\gamma_{s_n}, \gamma_{t_n}) > c \quad \text{for every } n > 0. \quad (2.29)$$

The two sequences of points $\gamma_{t_1}, \gamma_{t_2}, \gamma_{t_3}, \dots$ and $\gamma_{t_1}, \gamma_{s_2}, \gamma_{t_3}, \gamma_{s_4}, \dots$ are both non-decreasing and bounded from above by γ_T , thus by forward completeness they must both have a limit. Since the second contains a subsequence of the first, these limits must coincide, and thus we must have $d(\gamma_{t_n}, \gamma_{s_n}) \rightarrow 0$. This, however, is in contradiction with (2.29). \square

Forward-completeness will be used to deduce the following. Note the explicit exclusion of the starting point from the statement.

Proposition 2.28 (Recovering a left-continuous causal path from a dense set of values). *Let (M, τ, ℓ) be a forward metric spacetime, $\mathcal{D} \subset (0, 1]$ dense and $\eta: \mathcal{D} \rightarrow M$ a \leq -monotone map.*

Then there is a unique left-continuous causal curve $\gamma: (0, 1) \rightarrow M$ such that $\eta(r) \leq \gamma_t \leq \eta(s)$ for every $r, s \in \mathcal{D}$ with $r < t \leq s$. In particular, if η is left-continuous, then we have $\gamma|_{\mathcal{D}} = \eta|_{\mathcal{D}}$.

If η has an upper bound, then γ has a unique extension to $(0, 1]$ sharing the same properties.

Proof. For $t \in (0, 1)$ we define

$$\gamma_t := \lim_{\substack{s \uparrow t, \\ s \in \mathcal{D}}} \eta(s).$$

Forward-completeness and the existence of $r \in (t, 1) \cap \mathcal{D}$ (so that $\eta(r)$ is an upper bound for $\{\eta(s) : s \leq t\}$), ensures that the limit exists, thus the definition is well-posed. The construction

ensures continuity from the left. Causality of γ is clear from the monotonicity of η and the closedness of \leq , thus we have the claimed existence.

For uniqueness, let $\tilde{\gamma}: (0, 1) \rightarrow M$ be another such path. Given any $t \in (0, 1)$ let $s \in \mathcal{D}$ with $s < t$. As $\gamma_s \leq \eta(s) \leq \tilde{\gamma}_t$ by assumption, letting $s \uparrow t$ reveals $\gamma_t \leq \tilde{\gamma}_t$. Reversing the roles of $\tilde{\gamma}$ and γ and using antisymmetry of \leq gives $\tilde{\gamma}_t = \gamma_t$. The last claim is now obvious. \square

We now topologize the space $\text{LCC}([0, 1]; M)$ of left-continuous causal paths. To this aim, let us fix a bounded distance d inducing the Polish topology on our metric spacetime (in asking boundedness we are not losing generality, as we can always replace a given d with $1 \wedge d$). Then we define a distance D on $\text{LCC}([0, 1]; M)$ via

$$D(\gamma, \eta) := d(\gamma_0, \eta_0) + \int_0^1 d(\gamma_t, \eta_t) dt.$$

On $\text{LCC}([0, 1]; M)$ we put the topology induced by D . As the next result shows, this induced topology depends only on the topology τ of M and not on the particular choice of d (see also [85, Remarks 9 and 14] for similar considerations in more regular settings):

Proposition 2.29 (Left-continuous causal paths inherit Polishness). *Let (M, τ, ℓ) be a forward metric spacetime. Then $(\text{LCC}([0, 1]; M), D)$ is a complete and separable metric space and the following are equivalent:*

- i) $D(\gamma_n, \gamma) \rightarrow 0$ as $n \rightarrow \infty$
- ii) Any sequence $n_k \uparrow \infty$ admits a subsequence $n_{k_j} \uparrow \infty$ such that $\gamma_{n_{k_j}, t} \rightarrow \gamma_t$ as $j \uparrow \infty$ for $t = 0$ and for a.e. $t \in [0, 1]$.

If moreover the topology is locally causally convex, then these are also equivalent to:

- iii) $\gamma_{n, t} \rightarrow \gamma_t$ as $n \rightarrow \infty$ for $t = 0$ and for any continuity point t of γ ,
- iv) $\gamma_{n, t} \rightarrow \gamma_t$ as $n \rightarrow \infty$ for a set of t 's that is dense in $[0, 1]$ and contains 0.

Proof. Since M is separable, so is the space $L^1([0, 1]; M)$ of measurable maps $f: [0, 1] \rightarrow M$ with respect to the L^1 -distance $\int_0^1 d(\cdot, \cdot) dt$ (recall that d is bounded). Since $(\text{LCC}([0, 1]; M), D)$ is a subspace of $M \times L^1([0, 1]; M)$, we see that $\text{LCC}([0, 1]; M)$ is separable.

For completeness, let $(\gamma^n) \subset \text{LCC}([0, 1]; M)$ be a Cauchy sequence. First, this implies the sequence (γ_0^n) of initial points converges to some $\gamma_0 \in M$ by completeness of M . Second, if necessary we pass to a non-relabelled subsequence such that $\sum_n D(\gamma^n, \gamma^{n+1}) < +\infty$ and thus that converges for \mathcal{L}^1 -a.e. $t \in [0, 1]$. Call $\mathcal{D} \subset [0, 1]$ the set of t 's for which we have pointwise convergence and η_t the corresponding limit (with $\eta_0 = \gamma_0$). Let $\gamma \in \text{LCC}([0, 1]; M)$ be associated to η via Proposition 2.28 and notice that by Lemma 2.27 we have that $\gamma_t = \eta_t$ for every $t \in \mathcal{D}$ except at most a countable set.

We shall prove that $D(\gamma^n, \gamma) \rightarrow 0$ as $n \rightarrow +\infty$, the argument being standard. By construction, for a.e. t and for $t = 0$ we have $d(\gamma_t, \gamma_t^n) \leq \sum_{i \geq n} d(\gamma_t^{i+1}, \gamma_t^i)$, thus by monotone convergence

$$\limsup_{n \rightarrow +\infty} D(\gamma, \gamma^n) = \limsup_{n \rightarrow +\infty} \left(d(\gamma_0, \gamma_0^n) + \int_0^1 d(\gamma_t, \gamma_t^n) dt \right) \leq \limsup_{n \rightarrow +\infty} \sum_{i \geq n} D(\gamma^{i+1}, \gamma^i) = 0,$$

having used $\sum_n D(\gamma^n, \gamma^{n+1}) < +\infty$ in the last step. This argument also shows that $(i) \Rightarrow (ii)$. The converse implication follows directly from the definition of D .

We pass to the locally causally convex case.

$(i) \Rightarrow (iii)$. Since $d(\gamma_{n,0}, \gamma_0) \leq D(\gamma_n, \gamma)$, the convergence $\gamma_{n,0} \rightarrow \gamma_0$ is clear. Now let t be a continuity point of γ . We will show that any subsequence (n_k) has a further extraction (n_{k_j}) such that $\gamma_{n_{k_j}, t} \rightarrow \gamma_t$, as this suffices to conclude. Thus let $n_k \uparrow \infty$ be arbitrary and notice that, as above, there is a subsequence (n_{k_j}) such that $\gamma_{n_{k_j}, s} \rightarrow \gamma_s$ for a.e. $s \in [0, 1]$. Now let U be an

arbitrary causally convex neighbourhood of γ_t . Then there are $t_1 < t < t_2$ with $\gamma_{t_1}, \gamma_{t_2} \in U$ such that $\gamma_{n_{k_j}, t_i} \rightarrow \gamma_{t_i}$, $i = 1, 2$. Hence $\gamma_{n_{k_j}, t_i}$ is eventually in U , $i = 1, 2$, and by causal convexity and causality of the curves we get that $\gamma_{n_{k_j}, t}$ is eventually in U . This proves that $\gamma_{n_{k_j}, t} \rightarrow \gamma_t$, as desired.

(iii) \Rightarrow (i). Direct consequence of Lemma 2.27 and the dominated convergence theorem.

(iii) \Rightarrow (iv). Direct consequence of Lemma 2.27.

(iv) \Rightarrow (iii). Arguments similar to those in “(i) \Rightarrow (iii)” apply: let U be a causally convex neighbourhood of γ_t , where t is a continuity point of γ . Then for $t_1 < t < t_2$ belonging to the given dense set and sufficiently close to t we have $\gamma_{t_1}, \gamma_{t_2} \in U$, hence eventually $\gamma_{n, t_1}, \gamma_{n, t_2} \in U$ and by causal convexity eventually we get $\gamma_{n, t} \in U$. \square

The following compactness result is reminiscent of Helly’s selection theorem:

Theorem 2.30 (Limit curve theorem). *Let (M, τ, ℓ) be a forward metric spacetime. Let $\Gamma \subset \text{LCC}([0, 1]; M)$ be so that:*

i) *For some $\mathcal{D} \subset [0, 1]$ Borel with $\mathcal{L}^1(\mathcal{D}) = 1$ and $0 \in \mathcal{D}$ the set $\{\gamma_t : \gamma \in \Gamma\}$ is relatively compact for each $t \in \mathcal{D}$,*

ii) *There is a compact set $K \subset M$ such that for every $\gamma \in \Gamma$ there is $x \in K$ with $\gamma_1 \leq K$.*

Then Γ is \mathcal{D} -relatively compact.

If moreover the topology is locally causally convex, then we can weaken (i) above to

i') *For some $\tilde{\mathcal{D}} \subset [0, 1]$ dense with $0 \in \tilde{\mathcal{D}}$ the set $\{\gamma_t : \gamma \in \Gamma\}$ is relatively compact for each $t \in \tilde{\mathcal{D}}$.*

Proof. Since we are dealing with a distance, compactness is equivalent to sequential compactness. Thus let $(\gamma_n) \subset \Gamma$ and $(x_n) \subset K$ be corresponding points as in (ii); our goal is to prove that (γ_n) has a \mathcal{D} -convergent subsequence. Let $\mathcal{C} \subset \mathcal{D}$ be countable dense and containing 0. Then a diagonal argument and the assumption of relative compactness grant that up to passing to a non-relabelled subsequence we can assume that $n \mapsto \gamma_{n, t}$ converges to some limit $\eta(t)$ for every $t \in \mathcal{C}$ and similarly that $x_n \rightarrow x$. Clearly we have $\eta(t) \leq x$ for any $t \in \mathcal{C}$, so we can apply Proposition 2.28 to obtain from η a left-continuous causal path γ on $(0, 1]$ and extend it by putting $\gamma_0 := \eta(0)$ (this does not affect causality and left continuity — here we used that $0 \in \mathcal{C}$ to have a limit point $\eta(0)$).

We claim that $\mathcal{D}(\gamma_n, \gamma) \rightarrow 0$. For this it suffices, by Proposition 2.29 above, to show that for a.e. $t \in [0, 1]$ and for $t = 0$ we have $\gamma_{n, t} \rightarrow \gamma_t$. We are going to show that this holds for any continuity point $t \in \mathcal{D}$ of γ ; (by Lemma 2.27 this suffices). Fix such a t and notice that by compactness, after passing to a non-relabelled subsequence we can assume that $\gamma_{n, t} \rightarrow p$ for some $p \in M$. For any $r, s \in \mathcal{C}$ with $r < t < s$ passing to the limit in $\gamma_{n, r} \leq \gamma_{n, t} \leq \gamma_{n, s}$ we deduce that $\gamma_r \leq p \leq \gamma_s$. Letting $r \uparrow t$ and $s \downarrow t$, by the continuity of γ at t we conclude that $\gamma_t \leq p \leq \gamma_t$, and so $p = \gamma_t$. Since this conclusion holds independently of the subsequence chosen, the proof in this case is complete.

In the locally causally convex case we instead argue as follows. Let, as before, t be a continuity point of γ , then let U be a causally convex neighbourhood of γ_t and notice that the relations $\gamma_{s_1} \leq \eta(s_2) \leq \gamma_{s_3}$ valid for $s_1 < s_2 < s_3$ together with the continuity of γ at g ensure that $\eta(s) \in U$ for s sufficiently close to t . In particular this holds if s is also in $\tilde{\mathcal{D}}$, so for such s we eventually have $\gamma_s^n \in U$ and thus by causal convexity we conclude that eventually we have $\gamma_t^n \in U$. We thus proved that $\gamma_t^n \rightarrow \gamma_t$ for any continuity point t , which again suffices to conclude. \square

In what follows we shall often consider the evaluation maps $\mathbf{e}_t : \text{LCC}([0, 1]; M) \rightarrow M$ defined for $t \in [0, 1]$ as $\gamma \mapsto \gamma_t$ and the map $\mathbf{e} : \text{LCC}([0, 1]; M) \times [0, 1] \rightarrow M$ sending (γ, t) to γ_t . We shall frequently use the following basic result without explicitly mentioning it:

Proposition 2.31 (Evaluation maps are Borel). *Let (M, τ, ℓ) be a Polish metric spacetime. Then $\mathbf{e}_t : \text{LCC}([0, 1]; M) \rightarrow M$ is Borel for any $t \in (0, 1]$ and continuous for $t = 0$. Also $\mathbf{e} : \text{LCC}([0, 1]; M) \times [0, 1] \rightarrow M$ is Borel.*

Proof. The continuity of e_0 is clear, thus let $t \in (0, 1]$. Since M is Polish we can establish measurability of e_t by proving that $e_t^{-1}(B_r(p))$ is Borel for any $p \in M$ and $r > 0$. Left continuity implies

$$e_t^{-1}(B_r(p)) = \bigcup_{n \in \mathbf{N}} \bigcap_{m \geq n} \left\{ \gamma \in \text{LCC}([0, 1]; M) : \int_{0 \wedge (t - \frac{1}{m})}^t d(\gamma_s, p) ds < \frac{r}{m} \right\}.$$

Thus the claim follows from the fact that the set in parenthesis is D-open for every m . The same argument shows that $e^{-1}(B_r(p))$ is the union of $e_0^{-1}(B_r(p)) \times \{0\}$ and

$$\bigcup_{n \in \mathbf{N}} \bigcap_{m \geq n} \left\{ (\gamma, t) \in \text{LCC}([0, 1]; M) \times (0, 1] : \int_{0 \wedge (t - \frac{1}{m})}^t d(\gamma_s, p) ds < \frac{r}{m} \right\},$$

and the second claim follows as well. \square

Another map we shall need to use is the ‘restriction map’: given $s, t \in [0, 1]$ with $s < t$, we denote by $\text{restr}_s^t : \text{LCC}([0, 1]; M) \rightarrow \text{LCC}([0, 1]; M)$ the map

$$\text{restr}_s^t(\gamma)_r := \gamma_{(1-r)s+rt}.$$

Intuitively, this map first restricts the input curve to $[s, t]$ and then “stretches” it to $[0, 1]$.

From Proposition 2.31 it is not hard to see that restr_s^t is Borel for any $s, t \in [0, 1]$ with $s < t$ (and continuous if $s = 0$). Indeed, given $\eta \in \text{LCC}([0, 1]; M)$ and $r > 0$, we see that $D(\text{restr}_s^t(\gamma), \eta) < r$ if and only if there is $q \in [0, r] \cap \mathbf{Q}$ such that $d(\gamma_s, \eta_0) < q$ and $\int_0^1 d(\gamma_{(1-v)s+vt}, \eta_v) dv < r - q$. Equivalently, in symbols

$$(\text{restr}_s^t)^{-1}(B_r(\eta)) = \bigcup_{q \in [0, r] \cap \mathbf{Q}} \left(e_s^{-1}(B_q(\eta_0)) \cap \left\{ \gamma : \int_0^1 d(\gamma_{(1-v)s+vt}, \eta_v) dv < r - q \right\} \right),$$

which shows that $(\text{restr}_s^t)^{-1}(B_r(\eta))$ is Borel.

2.6 Action and geodesics

Let (M, ℓ) be a metric spacetime, $\gamma : [0, 1] \rightarrow M$ a causal curve and $0 \neq q \leq 1$. We define the q -action $\mathcal{A}_q(\gamma) \in \mathbf{R}$ as

$$\mathcal{A}_q(\gamma) := \begin{cases} \int_0^1 u_q(|\dot{\gamma}_t|) dt & \text{if } 0 \neq q < 1 \\ |\dot{\gamma}|([0, 1]) & \text{if } q = 1, \end{cases} \quad (2.30)$$

where we recall $u_q(z) = \frac{1}{q}z^q$ for $z > 0$ and from (2.3) that $u_q(0)$ is equal to 0 (resp. $-\infty$) if $q \in (0, 1)$ (resp. $q < 0$). In particular, for $q < 0$ we have $\mathcal{A}_q(\gamma) \leq 0$ and $\mathcal{A}_q(\gamma) = -\infty$ as soon as $|\dot{\gamma}| = 0$ holds on an \mathcal{L}^1 -non-negligible set.

Irrespective of the sign of q , Theorem 2.23 and Jensen’s inequality imply

$$\mathcal{A}_q(\gamma) \leq u_q(\ell(\gamma_0, \gamma_1)) \quad \left(= \frac{1}{q} \ell(\gamma_0, \gamma_1)^q \text{ if } q \vee \ell(\gamma_0, \gamma_1) > 0 \right). \quad (2.31)$$

In analogy with the definition of (constant speed) geodesics in the classical setting of metric spaces, we propose the following definition:

Definition 2.32 (Geodesics on M). *Let (M, ℓ) be a metric spacetime and $x, y \in M$ be with $x \leq y$. We call $\gamma : [0, 1] \rightarrow M$ a rough geodesic from x to y provided $x \leq \gamma_t \leq y$ for every $t \in [0, 1]$ and*

$$\ell(\gamma_t, \gamma_s) = (s - t)\ell(x, y), \quad \forall t, s \in [0, 1], t \leq s. \quad (2.32)$$

If M is also Polish and the rough geodesic γ belongs to $\text{LCC}([0, 1]; M)$ we call it a geodesic, thus dropping the term ‘rough’. The collection of all geodesics on M will be denoted $\text{CGeo}(M) \subset \text{LCC}([0, 1]; M)$, the ‘C’ standing for ‘causal’.

A causal (possibly rough) geodesic is called *timelike* if $\ell(\gamma_0, \gamma_1) \in (0, +\infty)$. The collection of all left continuous, timelike geodesics is denoted $\text{TGeo}(M) \subset \text{CGeo}(M)$. We use the term ℓ -geodesic synonymously with geodesic, particular when more than one time separation is under discussion.

Remark 2.33. The term *rough* is used analogously in [25]. Although not needed here, we reserve the label $\text{AGeo}(M)$ for the set of *affinely parameterized* causal geodesics defined as the D-closure of $\text{TGeo}(M)$; a related closure was instead denoted by $\text{CGeo}(M)$ in the final version of [85].

Notice that in order for γ to be a geodesic from x to y we are *not* insisting on $\gamma_0 = x$ or $\gamma_1 = y$, still, for any $t \in [0, 1]$ we have:

$$\ell(x, \gamma_t) = \ell(\gamma_0, \gamma_t) \quad \text{and} \quad \ell(\gamma_t, y) = \ell(\gamma_t, \gamma_1) \quad \text{in addition to} \quad \ell(x, y) = \ell(\gamma_0, \gamma_1). \quad (2.33)$$

Indeed, if $\ell(x, y) = +\infty$ the claim follows easily from (2.32) and $x \leq \gamma_t \leq y$, while if $\ell(x, y) < +\infty$, from non-negativity of ℓ on \leq we deduce $\ell(x, y) \geq \ell(x, \gamma_t) + \ell(\gamma_t, y) \geq \ell(\gamma_0, \gamma_t) + \ell(\gamma_t, \gamma_1) = \ell(x, y)$, having used (2.32) in the last step. Since $\ell(x, y) < +\infty$, the claim follows. Sufficient conditions to ensure that $\gamma_0 = x$ and/or $\gamma_1 = y$ will be explored in Section 2.8.

The link between geodesics and the q -action functionals is given by the following result:

Proposition 2.34 (Geodesics maximize the q -action). *Let (M, ℓ) be a metric spacetime, $\gamma : [0, 1] \rightarrow M$ causal and $x, y \in M$ be with $x \leq \gamma_0$, $\gamma_1 \leq y$ and $\ell(\gamma_0, \gamma_1) < +\infty$. Then the following are equivalent.*

(i) *The curve γ is a rough geodesic from x to y .*

(ii) *We have:*

$$(s - t) \ell(x, y) \leq \ell(\gamma_t, \gamma_s), \quad \forall t, s \in [0, 1], \quad t \leq s. \quad (2.34)$$

(iii) *There exists a real constant $c \geq 0$ such that $|\dot{\gamma}| = c$ \mathcal{L}^1 -a.e. and $\mathcal{A}_1(\gamma) \geq \ell(x, y)$.*

(iv) *For all $0 \neq q < 1$ we have*

$$\mathcal{A}_q(\gamma) = u_q(\ell(x, y)) \quad \left(= \frac{1}{q} \ell(x, y)^q \text{ if } q \vee \ell(\gamma_0, \gamma_1) > 0 \right).$$

(v) *For some $0 \neq q < 1$ we have $\mathcal{A}_q(\gamma) = u_q(\ell(x, y))$.*

Proof.

(i) \implies (iii). By (2.32) we see that the measure μ_h from Theorem 2.23 is equal to $\ell(x, y) \mathcal{L}^1 \llcorner [0, 1-h]$ on $[0, 1-h]$. In particular, the associated maximal measure μ is $\ell(x, y) \mathcal{L}^1$, from which the claim follows.

(iii) \implies (iv). The assumption forces $|\dot{\gamma}| = \ell(x, y)$ \mathcal{L}^1 -a.e, so the claim follows from the definition of $\mathcal{A}_q(\gamma)$.

(iv) \implies (v). Obvious.

(v) \implies (ii). Let $0 \neq q < 1$ satisfy the assumption. Then

$$\frac{1}{q} \ell(x, y)^q = \frac{1}{q} \int_0^1 |\dot{\gamma}_t|^q dt \stackrel{\text{(Jensen)}}{\leq} \frac{1}{q} \left[\int_0^1 |\dot{\gamma}_t| dt \right]^q \stackrel{(2.27)}{\leq} \frac{1}{q} \ell(\gamma_0, \gamma_1)^q \leq \frac{1}{q} \ell(x, y)^q.$$

Since by assumption we have $\ell(\gamma_0, \gamma_1) < +\infty$, we deduce that $\ell(x, y) < +\infty$ as well. If $\ell(x, y) = 0$ the conclusion is trivial, thus we can assume $\ell(x, y) \in (0, +\infty)$. In this case by strict concavity of $z \mapsto \frac{1}{q} z^q = u_q(z)$, the equality in Jensen’s inequality forces $|\dot{\gamma}| = \ell(x, y)$ \mathcal{L}^1 -a.e. The conclusion (2.34) follows from (2.27).

(ii) \implies (i). By the reverse triangle inequality for ℓ , for every $t, s \in [0, 1]$ with $t \leq s$ we have

$$\ell(x, y) \geq \ell(\gamma_0, \gamma_1) \geq \ell(\gamma_0, \gamma_t) + \ell(\gamma_t, \gamma_s) + \ell(\gamma_s, \gamma_1) \geq [t + (s - t) + (1 - s)] \ell(x, y) = \ell(x, y).$$

Since $\ell(\gamma_0, \gamma_1) < +\infty$ by assumption, this forces $\ell(x, y) < +\infty$ and thus equality throughout. \square

Remark 2.35 (Relation to customary lengths). Definition 2.32 is inspired by the length functional L_ℓ induced by ℓ [75, Def. 2.24] and its maximizers. We shortly comment on the connections between maximizers of L_ℓ and our action functionals, to relate our approach to others used in the literature. To adapt our setting to [75], we shall assume ℓ_+ is lower semicontinuous.

On the space $\text{LCC}([0, 1]; M)$ we define

$$L_\ell(\gamma) := \inf \left\{ \sum_{i=0}^{n-1} \ell(\gamma_{t_i}, \gamma_{t_{i+1}}) : n \in \mathbf{N}, 0 = t_0 < t_1 < \dots < t_n = 1 \right\}.$$

By (2.27) and the definition of L_ℓ , we have

$$\mathcal{A}_1(\gamma) \leq L_\ell(\gamma) \leq \ell(\gamma_0, \gamma_1). \quad (2.35)$$

A geodesic clearly attains the second inequality, and every $\gamma \in \text{LCC}([0, 1]; M)$ with $\mathcal{A}_1(\gamma) = \ell(\gamma_0, \gamma_1)$ maximizes L_ℓ by (2.35). On the other hand, if γ is a geodesic, Proposition 2.34 implies $\mathcal{A}_1(\gamma) = \ell(\gamma_0, \gamma_1)$, hence equality holds throughout (2.35). In order to relate general maximizers of L_ℓ (which do not need to be affinely parametrized) to those of \mathcal{A}_1 , further assumptions are needed. Suppose ℓ is upper semicontinuous and finite on \leq , and moreover we suppose that ℓ_+ is lower semicontinuous. Then ℓ_+ is continuous. Hence, if $\gamma \in \text{LCC}([0, 1]; M)$ is a *timelike* maximizer of L_ℓ , by [75, Cor. 3.35] it has a strictly increasing reparametrization σ which is a geodesic. As noted above, this implies $\mathcal{A}_1(\sigma) = \ell(\sigma_0, \sigma_1) = \ell(\gamma_0, \gamma_1)$. On the other hand, as \mathcal{A}_1 is easily seen to be invariant under strictly increasing reparametrizations, equality holds in (2.35). ■

Remark 2.36 ('Snowflakes'). A standard construction in positive signature in metric geometry is that of 'snowflaking' a metric. Say that (M, d) is a metric space. Then for any $\alpha \in (0, 1)$ the function d^α is still a distance, it is equivalent to the original one and has the property that any non-constant curve has infinite d^α -length.

A similar construction is in place for spacetimes (albeit in this case 'snowflake' does not convey the appropriate geometric intuition): given a metric spacetime (M, ℓ) and $\alpha > 1$, the function ℓ^α (where $\ell^\alpha(x, y) := -\infty$ is intended if $\ell(x, y) = -\infty$) is still a time separation inducing the same chronological and causal relations. If γ is a causal curve and $t \in [0, 1]$ such that $\lim_{h \downarrow 0} \frac{\ell(\gamma_t, \gamma_{t+h})}{h}$ exists and is finite, then clearly $\lim_{h \downarrow 0} \frac{\ell(\gamma_t, \gamma_{t+h})^\alpha}{h} \rightarrow 0$. It follows that the q -energy, $0 \neq q < 1$, of any causal curve w.r.t. ℓ^α is zero. ■

We come to topological properties of geodesics:

Proposition 2.37 (Causal and null geodesics form closed sets). *Let (M, τ, ℓ) be a forward spacetime such that ℓ_+ is continuous.*

Then the set of causal geodesics from Definition 2.32 is \mathcal{D} -closed and that of timelike geodesics is Borel.

Proof. Let $(\gamma^n) \subset \text{CGeo}(M)$ be a sequence \mathcal{D} -converging to $\gamma \in \text{LCC}([0, 1]; M)$. By (2.32) we see that

$$\ell(\gamma_s^n, \gamma_t^n) = (t - s) \ell(\gamma_0^n, \gamma_1^n), \quad \forall s, t \in [0, 1], s \leq t.$$

By Proposition 2.29 we see that for the Borel set $\mathcal{D} \subset [0, 1]$ containing 0 from the statement, passing to the limit in the above (notice that $1 \in \mathcal{D}$ may not hold) we get

$$\ell(\gamma_s, \gamma_t) = \frac{t-s}{t'} \ell(\gamma_0, \gamma_{t'}) \quad \forall s, t, t' \in \mathcal{D}, s \leq t \leq t'.$$

Then by the left continuity of γ as $t' \rightarrow 1$ and the continuity of ℓ_+ again we easily conclude that γ is a causal geodesic. The argument also shows that the collections of causal geodesics γ satisfying $\ell(\gamma_0, \gamma_1) = +\infty$ (and similarly $\ell(\gamma_0, \gamma_1) = 0$) are both closed as well, so the second claim follows. □

The q -action (2.30) inherits upper semicontinuity from ℓ via identity (2.24):

Lemma 2.38 (Upper semicontinuity of the q -action). *Let (M, τ, ℓ) be a metric spacetime with ℓ upper semicontinuous and let $0 \neq q < 1$. Assume $(\gamma_n) \subset \text{LCC}([0, 1]; M)$ D -converges to γ and that either $\sup_n \ell(\gamma_n(0), \gamma_n(1)) < +\infty$ or $\gamma_n(1) \rightarrow \gamma(1)$. Then*

$$\limsup_{n \rightarrow +\infty} \mathcal{A}_q(\gamma_n) \leq \mathcal{A}_q(\gamma).$$

Proof. We can pass to a non-relabelled subsequence realizing the lim sup and then — by Proposition 2.29 — to a further extraction so that $\gamma_{n,t} \rightarrow \gamma_t$ for every $t \in \mathcal{D}$, where $\mathcal{D} \subset [0, 1]$ is a Borel set of full measure containing 0. We can thus find partitions $P_k = \{0 = t_{k,0}, \dots, t_{k,n_k} = 1\}$ of $[0, 1]$ contained in $\mathcal{D} \cup \{1\}$ with $|P_k| \rightarrow 0$ (notation from item (iv) of Theorem 2.23). Notice that since u_q is non-decreasing, upper semicontinuity of ℓ implies that of $u_q \circ \ell$, thus if we are in the case $\gamma_{n,1} \rightarrow \gamma_1$ we have

$$\begin{aligned} \sum_i (t_{k,i+1} - t_{k,i}) u_q\left(\frac{\ell(\gamma_{t_{k,i}}, \gamma_{t_{k,i+1}})}{t_{k,i+1} - t_{k,i}}\right) &\geq \limsup_{n \rightarrow \infty} \sum_i (t_{k,i+1} - t_{k,i}) u_q\left(\frac{\ell(\gamma_{n,t_{k,i}}, \gamma_{n,t_{k,i+1}})}{t_{k,i+1} - t_{k,i}}\right) \\ &\stackrel{\text{(by (2.24))}}{\geq} \limsup_{n \rightarrow +\infty} \mathcal{A}_q(\gamma_n), \end{aligned}$$

so that letting $k \rightarrow \infty$ and recalling item (iv) of Theorem 2.23 we are done.

If we don't know whether $\gamma_{n,1} \rightarrow \gamma_1$, still the same argument shows that

$$\limsup_{n \rightarrow +\infty} \mathcal{A}_q(\gamma_n, [0, T]) \leq \mathcal{A}_q(\gamma, [0, T]) \quad \forall T \in \mathcal{D},$$

where $\mathcal{A}_q(\gamma, [0, T]) := \int_0^T u_q(|\dot{\gamma}_t|) dt$ is the q -action on the interval $[0, T]$. It is easy to see that $\lim_{T \uparrow 1} \mathcal{A}_q(\gamma, [0, T]) = \mathcal{A}_q(\gamma, [0, 1])$, so the conclusion will follow if we show that

$$\limsup_{T \uparrow 1} \limsup_n \mathcal{A}_q(\gamma_n, [T, 1]) \leq 0.$$

To see this, let $\sup_n \ell(\gamma_n(0), \gamma_n(1)) < L < \infty$ and notice that Jensen's inequality applied to the concave upper semicontinuous function u_q yields

$$\mathcal{A}_q(\gamma_n, [T, 1]) \leq (1 - T) u_q\left((1 - T)^{-1} \int_T^1 |\dot{\gamma}_{n,t}| dt\right) \leq (1 - T) u_q((1 - T)^{-1} L) = \frac{L^q}{q} (1 - T)^{1-q},$$

where the second inequality follows from monotonicity of u_q and the bound (2.27). The conclusion now follows from the fact that $q < 1$. \square

We conclude the section with a couple of simple results about the structure of (rough) ℓ_q -geodesics. In particular these will show ℓ_q -geodesics to/from Dirac masses do not depend on q . Note that the concepts of geodesic defined in Definition 2.32 apply equally well to the metric spacetime $(\mathcal{P}(M), \ell_q)$ equipped with its narrow topology.

Proposition 2.39 (On the structure of ℓ_q -geodesics). *Let (M, τ, ℓ) be a Polish metric spacetime in which ℓ is upper semicontinuous and does not take the value $+\infty$ and let $0 \neq q < 1$. Fix $\nu_0, \nu_1 \in \mathcal{P}_{\text{em}}(M)$ with $\ell_q(\nu_0, \nu_1) \in (0, +\infty)$ and (μ_t) a rough ℓ_q -geodesic from ν_0 to ν_1 .*

Fix $t, s \in [0, 1]$ with $t \leq s$ and $\pi^0, \pi^t, \pi, \pi^s, \pi^1 \in \mathcal{P}(M^2)$ ℓ_q -optimal for (ν_0, μ_0) , (μ_0, μ_t) , (μ_t, μ_s) , (μ_s, μ_1) and (μ_1, ν_1) respectively and a gluing $\hat{\pi} \in \mathcal{P}(M^6)$ of these along their common marginals. Then the coupling $(\text{Pr}_1, \text{Pr}_6)_ \hat{\pi}$ is ℓ_q -optimal for (ν_0, ν_1) and*

$$\ell(x_0, y_t) = \ell(y_0, y_t), \quad \ell(y_t, x_1) = \ell(y_t, y_1), \quad \ell(x_0, x_1) = \ell(y_0, y_1), \quad \ell(y_t, y_s) = (s - t)\ell(x_0, x_1) \quad (2.36)$$

holds for $\hat{\pi}$ -a.e. $(x_0, y_0, y_t, y_s, y_1, x_1)$.

Proof. Note $\ell_q(\nu_0, \nu_1) > 0$ forces $\hat{\pi}[\{\ell > 0\}] = 1$ when $q < 0$. For $\hat{\pi}$ as in the statement the inequality

$$\ell(x_0, x_1) \geq \ell(x_0, y_0) + \ell(y_0, y_t) + \ell(y_t, y_s) + \ell(y_s, y_1) + \ell(y_1, x_1) \quad (2.37)$$

holds for $\hat{\pi}$ -a.e. $(x_0, y_0, y_t, y_s, y_1, x_1)$ and therefore

$$\begin{aligned} \ell_q(\nu_0, \nu_1) &\geq \left(\int \ell^q(x_0, x_1) d\hat{\pi} \right)^{\frac{1}{q}} \\ \text{(by (2.37))} \quad &\geq \left(\int (\ell(x_0, y_0) + \ell(y_0, y_t) + \ell(y_t, y_s) + \ell(y_s, y_1) + \ell(y_1, x_1))^q d\hat{\pi} \right)^{\frac{1}{q}} \\ &\stackrel{*}{\geq} \ell_q(\nu_0, \mu_0) + \ell_q(\mu_0, \mu_t) + \ell_q(\mu_t, \mu_s) + \ell_q(\mu_s, \mu_1) + \ell_q(\mu_1, \nu_1) \stackrel{(2.33)}{=} \ell_q(\nu_0, \nu_1), \end{aligned}$$

where in the starred inequality we used (2.9) and optimality of the given plans. It follows that all the inequalities are equalities, and also inspecting the equality case in (2.9) the conclusion quickly follows. \square

From the above it follows that ℓ_q -geodesics to/from Dirac masses do not depend on q :

Proposition 2.40 (On ℓ_q -geodesics with Dirac endpoints). *Let (M, τ, ℓ) be a Polish metric spacetime in which ℓ is upper semicontinuous and does not take the value $+\infty$. Also, let $\nu \in \mathcal{P}_{\text{em}}(M)$ and $\bar{x} \in M$ be such that $\log(\ell(\cdot, \bar{x})) \in L^\infty(\nu)$ and let $(\mu_t)_{t \in [0,1]} \subset \mathcal{P}(M)$.*

Then (μ_t) is a rough ℓ_q -geodesic from ν to $\delta_{\bar{x}}$ for some $0 \neq q < 1$ if and only if it is so for every $0 \neq q < 1$. In this case, for $t, s \in [0, 1]$, $t \leq s$ and $\pi \in \Pi_{\leq}(\mu_t, \mu_s)$, the ℓ_q -optimality of π is also independent of q .

Similarly statements hold for rough ℓ_q -geodesics from Dirac masses to other measures.

Proof. Suppose that (μ_t) is a rough ℓ_q -geodesic from ν to $\delta_{\bar{x}}$ for a given $0 \neq q < 1$, fix $t, s \in [0, 1]$ with $t \leq s$ and let $\pi^0, \pi, \pi^1 \in \mathcal{P}(M^2)$ be ℓ_q -optimal for (ν, μ_t) , (μ_t, μ_s) and $(\mu_s, \delta_{\bar{x}})$ respectively. Let $\hat{\pi} \in \mathcal{P}(M^4)$ be a gluing of these along the common marginals and notice that Proposition 2.39 above (which applies because $\log(\ell(\cdot, \bar{x})) \in L^\infty(\nu)$ implies $\ell_q(\nu, \delta_{\bar{x}}) \in (0, +\infty)$) says that for $\hat{\pi}$ -a.e. (x_0, y_t, y_s, x_1) we have

$$\ell(x_0, y_t) = t\ell(x_0, x_1) \quad \text{and} \quad \ell(y_t, y_s) = (s-t)\ell(x_0, x_1) \quad \text{and} \quad \ell(y_s, x_1) = (1-s)\ell(x_0, x_1) \quad (2.38)$$

and logarithmically bounded. Now let $0 \neq q' < 1$ be another exponent and notice that

$$\begin{aligned} \ell_{q'}(\nu, \delta_{\bar{x}}) &\geq \ell_{q'}(\nu, \mu_t) + \ell_{q'}(\mu_t, \mu_s) + \ell_{q'}(\mu_s, \delta_{\bar{x}}) \\ &\geq \left(\int \ell^{q'}(x_0, y_t) d\hat{\pi} \right)^{\frac{1}{q'}} + \left(\int \ell^{q'}(y_t, y_s) d\hat{\pi} \right)^{\frac{1}{q'}} + \left(\int \ell^{q'}(y_s, x_1) d\hat{\pi} \right)^{\frac{1}{q'}} \\ \text{(by (2.38))} \quad &= \left(\int \ell^{q'}(x_0, x_1) d\hat{\pi} \right)^{\frac{1}{q'}} = \ell_{q'}(\nu, \delta_{\bar{x}}), \end{aligned}$$

thus again all the inequalities must be equalities, implying in particular that the plans π^0, π, π^1 are also $\ell_{q'}$ -optimal. This computation and (2.38) also show that $\ell_{q'}(\mu_t, \mu_s) = (s-t)\ell_{q'}(\nu, \delta_{\bar{x}})$, so the arbitrariness of t, s imply that (μ_t) is also a rough $\ell_{q'}$ -geodesic. \square

2.7 Lifting curves of probability measures to measures on curves

The main result of this section is Theorem 2.43, where we prove a ‘‘lifting’’ result for left-continuous causal paths on the space of Borel probability measures $\mathcal{P}(M)$ on a given Polish metric spacetime (M, τ, ℓ) .

Thanks to Proposition 2.16 we know that $(\mathcal{P}(M), \ell_q)$ is a Polish metric spacetime; we shall always endow $\mathcal{P}(M)$ with the narrow topology, so that $\text{LCC}([0, 1]; \mathcal{P}(M))$ refers to narrowly left continuous causal curves of measures. In particular, the definition of q -action we gave in (2.30) also applies here, giving the functional \mathcal{A}_q on causal curves in $\mathcal{P}(M)$ defined as

$$\mathcal{A}_q(\mu_\cdot) := \begin{cases} \frac{1}{q} \int_0^1 |\dot{\mu}_t|_q^q dt & \text{if } 0 \neq q < 1, \\ |\dot{\mu}|_1([0, 1]) & \text{if } q = 1, \end{cases}$$

where here and below $|\dot{\mu}_t|_q$ denotes the causal speed w.r.t. ℓ_q and for clarity we shall often write $\frac{1}{q}|\dot{\mu}_t|_q^q$ in place of $u_q(|\dot{\mu}_t|_q)$, with the conventions analogous to (2.3) implicitly understood. Calling a Borel probability measure π on left continuous causal curves a *plan* (or dynamical plan), we have the following general relation between actions at the level of measures and at the level of curves:

Lemma 2.41 (Relation between ℓ - and ℓ_q -causal speeds). *Let (M, τ, ℓ) be a Polish metric spacetime, $0 \neq q < 1$ and $\pi \in \mathcal{P}(\text{LCC}([0, 1]; M))$. Put $\mu_t := (\mathbf{e}_t)_* \pi$ for every $t \in [0, 1]$. Then $(\mu_t) \in \text{LCC}([0, 1]; \mathcal{P}(M))$ and*

$$\mathcal{A}_q(\mu_t) \geq \frac{1}{q} \int \int_0^1 |\dot{\gamma}_r|^q dr d\pi(\gamma). \quad (2.39)$$

Moreover, the pointwise bound

$$\frac{1}{q} |\dot{\mu}_t|_q^q \geq \frac{1}{q} \int |\dot{\gamma}_t|^q d\pi(\gamma) \quad \text{a.e. } t \quad (2.40)$$

holds if either $q \in (0, 1)$ or $\int |\dot{\gamma}_t|^q d\pi(\gamma) \in L^1([0, 1]; \mathcal{L}^1)$.

Proof. For $t, s \in [0, 1]$ with $t \leq s$ notice that since π is concentrated on causal curves the plan $(\mathbf{e}_t, \mathbf{e}_s)_* \pi$, that is admissible for (μ_t, μ_s) , is concentrated on $\{\ell \geq 0\}$, thus $t \mapsto \mu_t \in \mathcal{P}(M)$ is causal. For left continuity let $\varphi \in C_b(M)$ and notice that since π is concentrated on left continuous curves, an application of the dominated convergence theorem yields $\lim_{s \uparrow t} \int \varphi(\gamma_s) d\pi(\gamma) = \int \varphi(\gamma_t) d\pi(\gamma)$, showing narrow left continuity of (μ_t) .

We turn to (2.39). For any $t \in (0, 1)$ and any $h \in (0, 1 - t)$, $(\mathbf{e}_t, \mathbf{e}_{t+h})_* \pi$ is a causal coupling of μ_t and μ_{t+h} by construction. Thus

$$\frac{\ell_q(\mu_t, \mu_{t+h})^q}{q} \geq \int \frac{\ell(\gamma_t, \gamma_{t+h})^q}{q} d\pi(\gamma) \geq \frac{1}{q} \int \left[\int_t^{t+h} |\dot{\gamma}_r| dr \right]^q d\pi(\gamma) \geq \frac{h^{q-1}}{q} \int \int_t^{t+h} |\dot{\gamma}_r|^q dr d\pi(\gamma).$$

where the second and third inequality follow from the bound (2.27) and Jensen's inequality respectively. Now Theorem 2.23(iii) yields (2.39). For (2.40) we divide the above by h^q and let $h \downarrow 0$: the left hand side converges to $\frac{1}{q} |\dot{\mu}_t|_q^q$ for a.e. t by Theorem 2.23(ii), so we discuss the right hand side. If $q \in (0, 1)$ we notice that for $g : [0, 1] \rightarrow [0, +\infty]$ Borel we have $\liminf_h \frac{1}{h} \int_t^{t+h} g d\mathcal{L}^1 \geq g(t)$ for a.e. t (because $\liminf_h \frac{1}{h} \int_t^{t+h} g d\mathcal{L}^1 \geq \liminf_h \frac{1}{h} \int_t^{t+h} n \wedge g d\mathcal{L}^1 \geq n \wedge g(t)$ for every n and a.e. t), thus the claim follows by the non-negativity of $\frac{1}{q} |\dot{\gamma}_t|^q$. If instead $q < 0$ the conclusion follows from our integrability assumption and Lebesgue's differentiation theorem. \square

We are interested in understanding whether, given (μ_t) , we can find π for which equality holds in (2.39). To this aim, the following compactness criterion will be useful (compare with the limit curve Theorem 2.30).

Lemma 2.42 (A tightness criterion). *Let (M, τ, ℓ) be a forward spacetime and $\mathcal{G} \subset \mathcal{P}(\text{LCC}([0, 1]; M))$ a collection of measures such that:*

- i) *For some Borel $\mathcal{D} \subset [0, 1]$ of full measure with $0 \in \mathcal{D}$ we have: for any $t \in \mathcal{D}$ there is a compact set K_t such that $(\mathbf{e}_t)_* \pi$ is concentrated on K_t for every $\pi \in \mathcal{G}$ and $t \in \mathcal{D}$,*
- ii) *There is $\mathcal{K} \subset \mathcal{P}(M)$ tight so that for all $\pi \in \mathcal{G}$ there is $\mu \in \mathcal{K}$ with $(\mathbf{e}_1)_* \pi \leq \mu$.*

Then \mathcal{G} is tight.

If moreover the topology τ on M is locally causally convex, then we can weaken (i) above to

- i') *For some $\tilde{\mathcal{D}} \subset [0, 1]$ dense with $0 \in \tilde{\mathcal{D}}$ the set $\{(\mathbf{e}_t)_* \pi : \pi \in \mathcal{G}\} \subset \mathcal{P}(M)$ is tight for all $t \in \tilde{\mathcal{D}}$.*

Proof. Let $\varepsilon > 0$ and find $K \subset M$ compact such that $\nu(K^c) < \varepsilon$ for every $\nu \in \mathcal{K}$. Consider the collection \hat{K} of curves γ such that $\gamma_t \in K_t$ for every $t \in \mathcal{D}$ and $\gamma_1 \leq x$ for some $x \in K$: by Theorem 2.30, \hat{K} is compact in $\text{LCC}([0, 1]; M)$ and by our assumptions we have $\pi(\hat{K}) \geq 1 - \varepsilon$, proving the claim in this case.

Under the assumption (i') we instead argue as follows. It is not restrictive to assume that $\tilde{\mathcal{D}}$ is countable. Let $\varepsilon > 0$, find $\{\varepsilon_t > 0 : t \in \tilde{\mathcal{D}}\}$ such that $\sum_{t \in \tilde{\mathcal{D}}} \varepsilon_t < \varepsilon$ and then — by the tightness assumption — $H_t \subset M$ compact so that $(e_t)_* \pi(H_t) > 1 - \varepsilon_t$ for every $\pi \in \mathcal{G}$ and $t \in \tilde{\mathcal{D}}$. Also, let $H_1 \subset M$ be compact so that $\nu(H_1) > 1 - \varepsilon_1$ for every $\nu \in \mathcal{K}$.

For $\pi \in \mathcal{G}$ let $\nu \in \mathcal{K}$ be so that $(e_1)_* \pi \leq \nu$ and $\pi \in \Pi_{\leq}((e_1)_* \pi, \nu)$: since π is concentrated on $\{(x, y) : x \leq y\}$ we have $(e_1)_* \pi(J^-(H_1)^c) \leq \pi(M \times (H_1^c)) = \nu(H_1^c) < \varepsilon$.

Let $\hat{H} := \{\gamma \in G : \gamma_t \in H_t \text{ for any } t \in \tilde{\mathcal{D}} \text{ and } \gamma_1 \in J^-(H_1)\}$ and notice that by Theorem 2.30 we know that \hat{H} is relatively compact. To conclude observe that $\hat{H}^c \subset \cup_{t \in \tilde{\mathcal{D}}} e_t^{-1}(H_t^c) \cup e_1^{-1}(J^-(H_1)^c)$ and thus $\pi(\hat{H}^c) \leq (e_1)_* \pi(J^-(H_1)^c) + \sum_{t \in \tilde{\mathcal{D}}} (e_t)_* \pi(H_t^c) \leq \varepsilon + \sum_{t \in \tilde{\mathcal{D}}} \varepsilon_t < 2\varepsilon$ for every $\pi \in \mathcal{G}$. \square

We are now ready to state and prove the main result of the section. Notice that in order to gain tightness we shall need to assume something either at the level of the topology of M (conditions (A), (B)) or at the level of the specific curve of measures considered (condition (C)). In relation to (B), it is worth recalling that several natural topologies induced by ℓ are locally causally convex (see Remark 2.2).

Theorem 2.43 (Lifting paths of measures to measures on paths). *Let (M, τ, ℓ) be a forward metric spacetime such that ℓ is upper semicontinuous and does not take the value $+\infty$. Let $(\mu_t) \in \text{LCC}([0, 1]; \mathcal{P}(M))$ satisfy $\mu_t \in \mathcal{P}_{\text{em}}(M)$ for every $t \in [0, 1]$ and assume at least one of the following holds:*

(A) M is globally hyperbolic, i.e. emeralds are compact;

(B) the topology on M is locally causally convex;

(C) for some $D \subset [0, 1]$ dense with $0 \in D$ the set $\cup_{t \in D} \text{spt } \mu_t$ is relatively compact.

Let $0 \neq q < 1$. Then (μ_t) is induced by a plan $\pi \in \mathcal{P}(\text{LCC}([0, 1]; M))$ with the property

$$\mathcal{A}_q(\mu_\cdot) = \frac{1}{q} \iint_0^1 |\dot{\gamma}_t|^q dt d\pi(\gamma). \quad (2.41)$$

Moreover, if $\mathcal{A}_q(\mu_\cdot) > -\infty$, then any such plan also satisfies the a.e. pointwise identity

$$|\dot{\mu}_t|_q^q = \int |\dot{\gamma}_t|^q d\pi(\gamma) \quad \text{a.e. } t.$$

Proof.

STEP 0: PRELIMINARY CONSIDERATIONS. From the bound (2.31) and Lemma 2.4 we have $\mathcal{A}_q(\mu_\cdot) \leq \frac{1}{q} \ell_q^q(\mu_0, \mu_1) < \infty$, thus if $\mathcal{A}_q(\mu_\cdot) > -\infty$ the function $t \mapsto |\dot{\mu}_t|_q^q$ is in $L^1([0, 1])$. It follows that the second statement is a consequence of the first one together with Lemma 2.41: integrating (2.40) yields (2.39) whence comparison with (2.41) and the integrability just noted force equality in (2.40) for a.e. t . Here for $q < 0$ we have used the identity (2.41) and the assumption $\mathcal{A}_q(\mu) > -\infty$ to ensure that $\int |\dot{\gamma}_\cdot|^q d\pi(\gamma) \in L^1([0, 1]; \mathcal{L}^1)$ so that (2.40) holds.

STEP 1: CONSTRUCTION OF THE PLAN. The idea, originating in an analogous standard construction in different settings [9, 77, 119], is to build π as the narrow limit of piecewise constant plans which “interpolate” (μ_t) at intermediate points.

Let $n \mapsto P_n := \{0 = t_0^n < \dots < t_n^n = 1\}$ be a sequence of partitions of $[0, 1]$ containing $0, 1$, so that $P_n \subset P_{n+1}$ for every $n \in \mathbf{N}$ and with $|P_n| \rightarrow 0$ (notation from Theorem 2.23): under the assumption (C) we pick these so that $P_n \subset D \cup \{1\}$ for any n , while under assumptions (A), (B) we pick these arbitrarily.

Use the upper semicontinuity of ℓ and Lemma 2.19 to find an ℓ_q -optimal plan $\pi_i^n \in \Pi_{\leq}(\mu_{t_i^n}, \mu_{t_{i+1}^n})$. Recursively glue these plans along their common marginals to produce $\pi^n \in \mathcal{P}(M^{n+1})$, then consider the map $F^n : M^{n+1} \rightarrow \text{LCC}([0, 1]; M)$ sending (x_0, \dots, x_{n+1}) to the curve that is equal to x_i on $(\frac{t_{i-1}^n + t_i^n}{2}, \frac{t_i^n + t_{i+1}^n}{2}]$, where $t_{-1}^n := 0$ and $t_{n+1}^n := 1$. Notice that F^n is continuous and set $\pi^n := F^n_* \pi^n$.

We claim that the sequence (π^n) is tight and to this aim we shall make use of one of assumptions (A), (B), (C).

(A) Let $\varepsilon > 0$ and find a compact set $K \subset M$ with $\mu_0(K^c) + \mu_1(K^c) < \varepsilon$. Then $E := J(K, K)$ is compact by assumption and thus by Theorem 2.30 (with $\mathcal{D} = [0, 1]$ and $K_t = E$ for every $t \in [0, 1]$) the set $\hat{E} := \{\gamma \in \text{LCC}([0, 1]; M) : \gamma_t \in E, \forall t \in [0, 1]\} = e_0^{-1}(J^+(K)) \cap e_1^{-1}(J^-(K))$ is compact in $\text{LCC}([0, 1]; M)$. Since for every $n \in \mathbf{N}$ we have $(e_0)_* \pi^n = \mu_0$ and $(e_1)_* \pi^n = \mu_1$, we deduce

$$\pi^n(\hat{E}^c) \leq \pi^n(e_0^{-1}(K^c)) + \pi^n(e_1^{-1}(K^c)) = \mu_0(K^c) + \mu_1(K^c) < \varepsilon \quad \forall n \in \mathbf{N}$$

proving the desired tightness.

(B) We shall apply Lemma 2.42. We have $(e_1)_* \pi^n = \mu_1$ for every $n \in \mathbf{N}$, so assumption (ii) of Lemma 2.42 holds. Moreover, $\tilde{\mathcal{D}} := \cup_n P_n$ is dense in $[0, 1]$ and for every $t \in \tilde{\mathcal{D}}$ the measure $(e_t)_* \pi^n$ is eventually equal to μ_t (here we used that $P_n \subset P_{n+1}$), thus assumption (i') of Lemma 2.42 holds. Hence Lemma 2.42 gives the desired tightness.

(C) We shall apply again Lemma 2.42 and as before we notice that $(e_1)_* \pi^n = \mu_1$ for every $n \in \mathbf{N}$, so assumption (ii) holds. Also, for every $t \in [0, 1)$ and $n \in \mathbf{N}$ we have $(e_t)_* \pi^n \in \{\mu_t : t \in \cup_n P_n\}$, thus if we call $K \subset M$ a compact set containing $\text{spt } \mu_t$ for any $t \in \mathcal{D} \supset \cup_n P_n$, then assumption (i) of Lemma 2.42 is satisfied with $\mathcal{D} = [0, 1)$ and $K_t := K$ for any $t \in \mathcal{D}$. Thus, as before, Lemma 2.42 gives the desired tightness.

Having proved tightness, after passing to a non-relabelled subsequence we can therefore assume that (π^n) narrowly converges to some $\pi \in \mathcal{P}(\text{LCC}([0, 1]; M))$.

We claim that $(e_t)_* \pi = \mu_t$ for every $t \in [0, 1]$: since $t \mapsto (e_t)_* \pi$ is narrowly left continuous (Lemma 2.41), by Proposition 2.28 applied with $\mathcal{P}(M)$ in place of M it suffices to prove $\mu_s \preceq (e_t)_* \pi \preceq \mu_t$ for all dyadics $s < t$. Thus fix such s, t and notice that for any $\varphi \in C_b(M)$ the map $\gamma \mapsto \frac{1}{t-s} \int_s^t \varphi(\gamma_r) dr$ is continuous from $\text{LCC}([0, 1]; M)$ to \mathbf{R} (by the choice of D-topology on $\text{LCC}([0, 1]; M)$ and an application of the dominated convergence theorem), thus

$$\frac{1}{t-s} \iint_s^t \varphi(\gamma_r) dr d\pi^n(\gamma) \rightarrow \frac{1}{t-s} \iint_s^t \varphi(\gamma_r) dr d\pi(\gamma) \quad \text{as } n \rightarrow \infty.$$

In other words, the measures $\mu_{s,t}^n := \frac{1}{t-s} \int_s^t (e_r)_* \pi^n dr$ (the integral being interpreted in the weak sense) narrowly converge to $\mu_{s,t} := \frac{1}{t-s} \int_s^t (e_r)_* \pi dr$ and since $\mu_s \preceq \mu_{s,t}^n \preceq \mu_t$ for every n , by the closure of $\{\ell \geq 0\}$ we deduce that $\mu_s \preceq \mu_{s,t} \preceq \mu_t$. Finally, to establish the only remaining claim $\mu_s \preceq (e_t)_* \pi \preceq \mu_t$ of this paragraph, we use left narrow continuity and \preceq -monotonicity of $r \mapsto (e_r)_* \pi$ to deduce that $\mu_{r,t}$ converges narrowly and \preceq -monotonically to $(e_t)_* \pi$ as $r \uparrow t$, and recall from Proposition 2.16 that the causal relation \preceq on $\mathcal{P}(M)$ is narrowly closed.

STEP 2: PROOF OF THE ACTION IDENTITY. By Lemma 2.41 it suffices to prove the reverse inequality \leq to (2.39) also holds. To this aim we start imitating the notation of Theorem 2.23 and define, for $\gamma \in \text{LCC}([0, 1]; M)$ and a partition $P = \{0 = t_0 < \dots < t_n = 1\}$ of $[0, 1]$, the quantity

$$A_q(\gamma, P) := \sum_{i=0}^{n-1} (t_{i+1} - t_i) u_q \left(\frac{\ell(\gamma_{t_i}, \gamma_{t_{i+1}})}{t_{i+1} - t_i} \right).$$

Then notice that the construction and the properties (2.23) and (2.25) give

$$A_q(\mu) = \lim_{n \rightarrow +\infty} \int A_q(\gamma, P_n) d\pi^n(\gamma) \leq \liminf_{n \rightarrow \infty} \int u_q(\ell(\gamma_0, \gamma_1)) d\pi^n(\gamma).$$

Now fix $t_0, t_1 \in (0, 1)$ belonging to some of the P_n 's — and thus eventually to all of them — with $t_0 < t_1$ and notice that applying the above to the curve $t \mapsto \mu_{(1-t)t_0+tt_1}$ gives

$$\int_{t_0}^{t_1} u_q(|\dot{\mu}_t|) dt \leq \liminf_{n \rightarrow \infty} \int u_q(\ell(\gamma_{t_0}, \gamma_{t_1})) d\pi^n(\gamma) \leq \limsup_{n \rightarrow \infty} \int \int_0^{t_0} \int_{t_1}^1 u_q(\ell(\gamma_t, \gamma_s)) ds dt d\pi^n(\gamma)$$

having used also the monotonicity of u_q and the causality of the γ 's in the second inequality. Arguing as above, starting from the observation that for $\varphi \in C_b(M)$ the map $\text{LCC}([0, 1]; M) \ni \gamma \mapsto$

$\int_0^{t_0} \int_{t_1}^1 \varphi(\gamma_t, \gamma_s) ds dt \in \mathbf{R}$ is continuous we see that the sequence $\int_0^{t_0} \int_{t_1}^1 (\mathbf{e}_t, \mathbf{e}_s)_* \boldsymbol{\pi}^n ds dt$ narrowly converges to $\int_0^{t_0} \int_{t_1}^1 (\mathbf{e}_t, \mathbf{e}_s)_* \boldsymbol{\pi} ds dt$. Since ℓ is upper semicontinuous and bounded from above on E (Lemma 2.4), we can pass to the limit in the above and get

$$\int_{t_0}^{t_1} u_q(|\dot{\mu}_t|) dt \leq \int \int_0^{t_0} \int_{t_1}^1 u_q(\ell(\gamma_t, \gamma_s)) ds dt d\boldsymbol{\pi}(\gamma) \leq \int u_q(\ell(\gamma_0, \gamma_1)) d\boldsymbol{\pi}(\gamma),$$

having used again the monotonicity of u_q and causality of the γ 's. Letting $t_0 \downarrow 0$ and $t_1 \uparrow 1$ we obtain $\int_0^1 u_q(|\dot{\mu}_t|) dt \leq \int u_q(\ell(\gamma_0, \gamma_1)) d\boldsymbol{\pi}(\gamma)$. A simple scaling argument now readily implies

$$\int_0^1 u_q(|\dot{\mu}_t|) dt \leq \int A_q(\gamma, P) d\boldsymbol{\pi}(\gamma) \quad \text{for every partition } P \text{ of } [0, 1].$$

To conclude it therefore suffices to show that

$$\int \int_0^1 u_q(|\dot{\gamma}_t|) dt d\boldsymbol{\pi}(\gamma) = \lim_{k \rightarrow \infty} \int A_q(\gamma, P_k) d\boldsymbol{\pi}(\gamma), \quad \text{for given partitions } (P_k) \text{ with } |P_k| \rightarrow 0.$$

To see this, recall that $A_q(\gamma, P) \leq u_q(\ell(\gamma_0, \gamma_1)) \leq u_q(\sup_{E^2} \ell) < \infty$ for any $\gamma \in \text{spt } \boldsymbol{\pi}$, then use (2.25) in conjunction with the monotone convergence theorem. \square

It is worth encoding the property established by the previous result in a definition:

Definition 2.44 (Lifting causal curves of measures). *Let (M, τ, ℓ) be a Polish metric spacetime, $\boldsymbol{\pi} \in \mathcal{P}(\text{LCC}([0, 1]; M))$ and $(\mu_t) \in \text{LCC}([0, 1]; \mathcal{P}(M))$. We say that $\boldsymbol{\pi}$ is a lifting of (μ_t) provided $\mu_t = (\mathbf{e}_t)_* \boldsymbol{\pi}$ for each $t \in [0, 1]$ and (2.41) holds.*

To clarify the upcoming discussion, we recall that according to Definition 2.32, a rough timelike ℓ_q -geodesic is a curve $(\mu_t) \subset \mathcal{P}(M)$ such that

$$0 < \ell_q(\mu_t, \mu_s) = (s - t) \ell(\mu_0, \mu_1) < \infty \quad \forall t, s \in [0, 1], t < s.$$

This, however, does not imply that ‘the mass moves all along timelike geodesics’; more precisely, if $\boldsymbol{\pi}$ is a lifting of (μ_t) we do not necessarily have that $\boldsymbol{\pi}$ is concentrated on timelike geodesics. We thus give the following:

Definition 2.45 (Strongly timelike ℓ_q -geodesics). *Let (M, τ, ℓ) be a Polish metric spacetime, $0 \neq q < 1$ and $(\mu_t) \subset \mathcal{P}(M)$ a rough timelike ℓ_q -geodesic. We say that (μ_t) is a rough strongly timelike ℓ_q -geodesic provided furthermore for any $t, s \in [0, 1]$ with $t < s$ that any ℓ_q -optimal coupling of (μ_t, μ_s) is concentrated on $\{\ell \in (0, +\infty)\}$.*

We then have the following:

Corollary 2.46 (Lifting ℓ_q -geodesics). *Let (M, τ, ℓ) be a forward metric spacetime in which ℓ is upper semicontinuous and does not take the value $+\infty$. Let $0 \neq q < 1$ and $(\mu_t) \in \text{LCC}([0, 1]; \mathcal{P}(M))$ be a causal ℓ_q -geodesic from ν_0 to ν_1 , with $\nu_0, \nu_1 \in \mathcal{P}_{\text{em}}(M)$. Assume also that $\ell_q(\nu_0, \nu_1) \in (0, +\infty)$ and that either of (A), (B), (C) of Theorem 2.43 hold.*

Then (μ_t) is induced by a plan $\boldsymbol{\pi} \in \mathcal{P}(\text{LCC}([0, 1]; M))$ which is concentrated on causal geodesics and so that $(\mathbf{e}_0, \mathbf{e}_1)_ \boldsymbol{\pi}$ is ℓ_q -optimal. Moreover we have*

$$u_q(\ell_q(\nu_0, \nu_1)) = \int u_q(\ell(\gamma_0, \gamma_1)) d\boldsymbol{\pi}(\gamma). \quad (2.42)$$

If (μ_t) is a strongly timelike ℓ_q -geodesic, then every $\boldsymbol{\pi}$ as above is concentrated on timelike geodesics.

Conversely, let $\boldsymbol{\pi} \in \mathcal{P}(\text{LCC}([0, 1]; M))$ and $\nu_0, \nu_1 \in \mathcal{P}_{\text{em}}(M)$ be such that $\nu_0 \preceq (\mathbf{e}_0)_ \boldsymbol{\pi}$ and $(\mathbf{e}_1)_* \boldsymbol{\pi} \preceq \nu_1$. Let $\mu_t := (\mathbf{e}_t)_* \boldsymbol{\pi}$ for every $t \in [0, 1]$. Then (μ_t) is an ℓ_q -geodesic from ν_0 to ν_1 if and only if*

$$\frac{1}{q} \iint_0^1 |\dot{\gamma}_t|^q dt d\boldsymbol{\pi}(\gamma) \geq u_q(\ell_q(\nu_0, \nu_1)). \quad (2.43)$$

Proof. Let π be built from (μ_t) via Theorem 2.43. Then

$$\int u_q(\ell(\gamma_0, \gamma_1)) d\pi(\gamma) \leq u_q(\ell_q(\mu_0, \mu_1)) \leq u_q(\ell_q(\nu_0, \nu_1))$$

and since (μ_t) is a ℓ_q -geodesic from ν_0 to ν_1 we also have

$$u_q(\ell_q(\nu_0, \nu_1)) = \int_0^1 u_q(|\dot{\mu}_t|_q) \stackrel{(2.41)}{=} \iint_0^1 u_q(|\dot{\gamma}_t|) dt d\pi(\gamma) \stackrel{(2.24)}{\leq} \int u_q(\ell(\gamma_0, \gamma_1)) d\pi(\gamma).$$

The assumption $\ell_q(\nu_0, \nu_1) \in (0, +\infty)$ implies that $u_q(\ell_q(\nu_0, \nu_1)) \in \mathbf{R}$ and thus the equalities in the above force the optimality of $(\mathbf{e}_0, \mathbf{e}_1)_*\pi$ and that $\int_0^1 u_q(|\dot{\gamma}_t|) dt = \ell(\gamma_0, \gamma_1)$ for π -a.e. γ , that by Proposition 2.34 means that γ is a geodesic. The claim about strong timelike ℓ_q -geodesics is now obvious by definition and the optimality of $(\mathbf{e}_0, \mathbf{e}_1)_*\pi$.

All this also establishes the ‘only if’ in the last claim. For the ‘if’ we put $\mu_t := (\mathbf{e}_t)_*\pi$ for every $t \in [0, 1]$ and notice that

$$u_q(\ell_q(\nu_0, \nu_1)) \stackrel{(2.43)}{\leq} \int \mathcal{A}_q(\gamma) d\pi(\gamma) \stackrel{(2.39)}{\leq} \mathcal{A}_q(\mu) \stackrel{(2.31)}{\leq} u_q(\ell_q(\mu_0, \mu_1)) \leq u_q(\ell_q(\nu_0, \nu_1)),$$

having used the causality relation in the last step. The claim follows. \square

As a consequence of this discussion we can see the relation between M being geodesic and $\mathcal{P}(M)$ being so. Here and below we say that a metric spacetime is geodesic if for any two points x, y with $x \leq y$ a causal geodesic from x to y exists.

Proposition 2.47 (Heredity of geodesy on metric spacetimes). *Let (M, τ, ℓ) be a Polish metric spacetime in which ℓ is upper semicontinuous and ℓ_+ is continuous and real valued.*

If (M, τ, ℓ) is geodesic then $(\mathcal{P}_{\text{em}}(M), \ell_q)$ is geodesic for every $0 \neq q < 1$.

Conversely, assume that $(\mathcal{P}_{\text{em}}(M), \ell_q)$ is geodesic for some $0 \neq q < 1$ and that the topology of M satisfies (A) or (B) of Theorem 2.43. Then every $x, y \in M$ with $\ell(x, y) \in [0, +\infty)$ are joined by a geodesic.

Proof. Assume that M is geodesic. Notice that from Proposition 2.37 we easily get that the multivalued map taking couples $(x, y) \in \{\ell \geq 0\}$ returning the set of causal geodesics from x to y has closed graph. Hence by standard measurable selection arguments (see e.g. [19, Thm. 6.9.2]) there is a universally measurable map $\text{GeodSel} : \{\ell \geq 0\} \rightarrow \text{CGeo}(M)$ such that $\text{GeodSel}(x, y)$ is one such geodesic for every $(x, y) \in \{\ell \geq 0\}$. Now let $\mu, \nu \in \mathcal{P}_{\text{em}}(M)$ be with $\mu \preceq \nu$ and let π be a ℓ_q -optimal plan (this exists by Lemma 2.19). A direct computation shows that putting $\pi := \text{GeodSel}_*\pi$, the curve $t \mapsto (\mathbf{e}_t)_*\pi$ is a ℓ_q -geodesic from μ to ν .

Conversely, assume that $(\mathcal{P}_{\text{em}}(M), \ell_q)$ is geodesic and let $x, y \in M$ with $x \leq y$. If $\ell(x, y) = 0$ then the lightlike curve

$$\gamma_t := \begin{cases} x & \text{if } t = 1/2 \\ y & \text{if } t \in (1/2, 1] \end{cases}$$

can be verified to be geodesic (using conventions (2.3) and (2.6) if $q < 0$). Therefore, assume $\ell(x, y) \in (0, \infty)$ and let (μ_t) be an ℓ_q -geodesic from δ_x to δ_y . Since $\ell_q(\delta_x, \delta_y) = \ell(x, y) \in (0, +\infty)$, by Corollary 2.46 above (μ_t) admits a lifting π that is concentrated on $\text{CGeo}(M)$. We claim that π -a.e. γ is a geodesic from x to y (and thus in particular at least a geodesic exists). To see this, notice that since (μ_t) is a ℓ_q -geodesic from δ_x to δ_y , we have $\delta_x \preceq \mu_0 = (\mathbf{e}_0)_*\pi$ and $(\mathbf{e}_1)_*\pi = \mu_1 \preceq \delta_y$, so that π -a.e. γ satisfies $x \leq \gamma_0$ and $\gamma_1 \leq y$ and therefore $u_q(\ell(\gamma_0, \gamma_1)) \leq u_q(\ell(x, y))$. Then (2.42) and the fact that $u_q(\ell(x, y)) \in \mathbf{R}$ forces $u_q(\ell(\gamma_0, \gamma_1)) = u_q(\ell(x, y))$ for π -a.e. γ , which is the conclusion. \square

2.8 Non-branching notions

In this section we shall start working with Polish metric spacetimes equipped with a reference measure. Let us isolate the relevant definition:

Definition 2.48 (Metric measure spacetime). *A metric measure spacetime, or shortly a mm spacetime, is a quadruple $(M, \tau, \ell, \mathfrak{m})$ such that (M, τ, ℓ) is a Polish metric spacetime and \mathfrak{m} is a non-negative and non-zero Radon measure on M .*

A forward (resp. backward) metric measure spacetime, or shortly a forward mm spacetime, is a metric measure spacetime so that (M, τ, ℓ) is also forward (resp. backward).

In the classical theory of optimal transport it is well known that a non-branching condition implies strong regularity properties of geodesics of measures (see [119] for an overview on the topic). In the context of metric geometry, it has been understood in [59] that a lower Ricci bound and a non-branching assumption yield existence of optimal maps (see also [65] and [36] for further results).

In non-smooth Lorentzian signature a similar picture is — starting from [39] — emerging, see also [21] for subsequent developments. In this section we review some of the existing notions in our framework and propose the concept of ‘non-branching at 0/1’.

Notice that in our setting one should exclude geodesics with either 0 or infinite length, because these are invariant by reparametrization and thus always branch. In other words, a non-branching condition only makes sense if required for timelike geodesics. Also, as in [39], one might distinguish between ‘forward’ and ‘backward’ non-branching, as discussed in the next definition, coming from [39, Def. 1.10]:

Definition 2.49 (Timelike non-branching conditions). *Let (M, τ, ℓ) be a Polish metric spacetime and G a set of timelike geodesics. We say that G is forward (resp. backward) timelike non-branching if any two $\gamma, \eta \in G$ that agree on $[0, t]$ (resp. $[t, 1]$) for some $t \in (0, 1)$ agree on the whole $[0, 1]$.*

We then say that (M, τ, ℓ) is forward (resp. backward) timelike non-branching if $\text{TGeo}(M)$ is so.

This concept applies equally well to the spacetime $(\mathcal{P}(M), \ell_q)$, but in this case it is better to restrict the attention to strongly timelike ℓ_q -geodesics. The problem is that a lifting of general timelike ℓ_q -geodesics may give positive mass to non-timelike geodesics and reparametrizing these curves we easily see that in general circumstances $(\mathcal{P}(M), \ell_q)$ is never non-branching.

With this said, the timelike non-branching of M is related to that of $\mathcal{P}(M)$:

Proposition 2.50 (Heredity of timelike non-branchingness). *Let (M, τ, ℓ) be a forward spacetime. Let $0 \neq q < 1$ and assume that the collection of strongly timelike ℓ_q -geodesics in $\mathcal{P}_{\text{em}}(M)$ is forward (resp. backward) timelike non-branching. Then (M, τ, ℓ) is forward (resp. backward) timelike non-branching.*

Conversely, assume that (M, τ, ℓ) is forward (resp. backward) timelike non-branching, that ℓ is upper semicontinuous and does not take the value $+\infty$ and that either (A) or (B) of Theorem 2.43 hold. Then for every $0 \neq q < 1$ the collection of strongly timelike ℓ_q -geodesics in $\mathcal{P}_{\text{em}}(M)$ is forward (resp. backward) timelike non-branching.

Proof. We deal with the ‘forward’ case, the ‘backward’ one being analogous. Let $0 \neq q < 1$ and suppose that the collection of strongly timelike ℓ_q -geodesics in $\mathcal{P}_{\text{em}}(M)$ is forward timelike non-branching. Notice that for any $x, y \in M$ we have $\ell(x, y) = \ell_q(\delta_x, \delta_y)$, hence $t \mapsto \gamma_t \in (M, \ell)$ is a timelike geodesic if and only if $t \mapsto \delta_{\gamma_t} \in (\mathcal{P}(M), \ell_q)$ is strongly timelike. The conclusion follows.

Conversely, assume that (M, τ, ℓ) is forward timelike non-branching and let $(\mu_t^1), (\mu_t^2) \subset \mathcal{P}_{\text{em}}(M)$ be two strongly timelike ℓ_q -geodesics that agree on $[0, T]$ for some given $T \in (0, 1)$. Let $\pi^1, \pi^2 \in \mathcal{P}(\text{TGeo}(M))$ be liftings of $(\mu_t^1), (\mu_t^2)$ respectively as in Corollary 2.46 (that we can apply thanks to the assumptions on M). Produce a new plan $\pi \in \mathcal{P}(\text{TGeo}(M))$ by gluing $(\text{restr}_0^T)_* \pi^1$ with $(\text{restr}_T^1)_* \pi^2$ along their common marginal μ_T^1 — obtaining a measure on the space of paths on $[0, 2]$ — and then reparametrize the curves in the obvious (piecewise affine) way to obtain a measure on curves on $[0, 1]$. It is clear that π is still a lifting of (μ_t^2) , so by Corollary 2.46 above is concentrated

on $\text{TGeo}(\mathbb{M})$. Also, the construction ensures that $(\text{restr}_0^T)_* \boldsymbol{\pi} = (\text{restr}_0^T)_* \boldsymbol{\pi}^1$ and since our non-branching assumption can equivalently be stated by saying that $\text{restr}_0^T : \text{TGeo}(\mathbb{M}) \rightarrow \text{TGeo}(\mathbb{M})$ is injective, this proves that $\boldsymbol{\pi}^1 = \boldsymbol{\pi}$, and thus that $\mu_t^1 = (\mathbf{e}_t)_* \boldsymbol{\pi}^1 = (\mathbf{e}_t)_* \boldsymbol{\pi} = \mu_t^2$ for every $t \in [0, 1]$, as desired. \square

When dealing with *measured* spacetimes, it is natural to tailor the above concept in order to allow a ‘negligible set of exceptions’, so to say.

We shall use the following measure-theoretic variant of the above non-branching condition, akin to that introduced by Braun [21] for $q \in (0, 1)$ (inspired by Rajala–Sturm [106] in positive signature). Here and below a family of measures in $\mathcal{P}_{\text{em}}(\mathbb{M})$ is said to be of bounded compression if for some $C > 0$ each of the measures is $\leq C\mathbf{m}$.

Definition 2.51 (Timelike q -essential non-branching). *Let $(\mathbb{M}, \tau, \ell, \mathbf{m})$ be a forward mm spacetime and $0 \neq q < 1$. We say that \mathbb{M} is forward (resp. backward) timelike q -essentially non-branching, provided for any strongly timelike ℓ_q -geodesic $(\mu_t) \subset \mathcal{P}_{\text{em}}(\mathbb{M})$ with bounded compression and for any lifting $\boldsymbol{\pi} \in \mathcal{P}(\text{LCC}([0, 1]; \mathbb{M}))$ we have that $\boldsymbol{\pi}$ is concentrated on a forward (resp. backward) timelike non-branching subset of $\text{TGeo}(\mathbb{M})$.*

The analogue of Proposition 2.50 for this last definition is the following result:

Proposition 2.52 (When bounded compression strongly timelike ℓ_q -geodesics are non-branching). *Let $(\mathbb{M}, \ell, \mathbf{m})$ be a forward mm spacetime and $0 \neq q < 1$.*

Assume that the collection of bounded compression strongly timelike ℓ_q -geodesics in $\mathcal{P}_{\text{em}}(\mathbb{M})$ is forward (resp. backward) non-branching. Then $(\mathbb{M}, \tau, \ell, \mathbf{m})$ is forward (resp. backward) timelike q -essentially non-branching.

Conversely, assume that $(\mathbb{M}, \tau, \ell, \mathbf{m})$ is forward (resp. backward) timelike q -essentially non-branching, that ℓ is upper semicontinuous and does not take the value $+\infty$ and that and that either (A) or (B) of Theorem 2.43 hold. Then the collection of bounded compression strongly timelike ℓ_q -geodesics in $\mathcal{P}_{\text{em}}(\mathbb{M})$ is forward (resp. backward) non-branching.

Proof. Suppose that \mathbb{M} is not forward q -essentially non-branching. Then there is a bounded compression strongly timelike ℓ_q -geodesic (μ_t) and a lifting $\boldsymbol{\pi} \in \mathcal{P}(\text{TGeo}(\mathbb{M}))$ of it that is not concentrated on a timelike non-branching subset of $\text{TGeo}(\mathbb{M})$. With a bit of work (see e.g. [59] for analogous arguments in positive signature) it is not hard to see that there are $t \in (0, 1)$ and two Borel sets $\Gamma_1, \Gamma_2 \subset \text{TGeo}(\mathbb{M})$ so that $(\text{restr}_0^t)_*(\boldsymbol{\pi} \llcorner \Gamma_1) = (\text{restr}_0^t)_*(\boldsymbol{\pi} \llcorner \Gamma_2)$ and $(\mathbf{e}_s)_*(\boldsymbol{\pi} \llcorner \Gamma_1) \neq (\boldsymbol{\pi} \llcorner \Gamma_2)$ for some $s > t$ (in particular $\boldsymbol{\pi}(\Gamma_1) = \boldsymbol{\pi}(\Gamma_2) > 0$). Then define the plans $\boldsymbol{\pi}^i := \boldsymbol{\pi}(\Gamma_i)^{-1} \boldsymbol{\pi} \llcorner \Gamma_i$ and the measures $\mu_t^i := (\mathbf{e}_t)_* \boldsymbol{\pi}^i$ for $i = 1, 2, t \in [0, 1]$. The construction ensures that $(\mu_t^1), (\mu_t^2)$ are strongly timelike ℓ_q -geodesics with bounded compression, that they agree on $[0, t]$ and with $\mu_s^1 \neq \mu_s^2$, as desired.

The converse implication follows along the very same lines used to prove Proposition 2.50. \square

We turn to a sort of ‘infinitesimal variant’ of the timelike non-branching condition. Notice that this concept is trivial in positive signature, due to the continuity of geodesics:

Definition 2.53 (Non-branching at 0 or 1). *Let (\mathbb{M}, ℓ) be a metric spacetime and G a collection of rough timelike geodesics. We say that G is timelike non-branching at 0 (resp. at 1) if any two geodesics in G that agree on $(0, 1)$ also agree at 0 (resp. at 1).*

If this holds with G being the collection of all rough timelike geodesics, then we say that (\mathbb{M}, ℓ) is timelike non-branching at 0 (resp. at 1).

It is worth noticing that under suitable compactness assumptions, the uniqueness encoded in this condition precisely detects continuity of geodesics:

Proposition 2.54 (Geodesics are not rough in a forward spacetime non-branching at 1). *Let (\mathbb{M}, τ, ℓ) be a Polish metric spacetime that is timelike non-branching at 0 (resp. 1). Also, let $\gamma : [0, 1] \rightarrow \mathbb{M}$ be a rough timelike geodesic such that for any $t \in [0, 1]$ and $s_n \downarrow t$ (resp. $s_n \uparrow t$) there is $n_k \uparrow \infty$ such that $k \mapsto \gamma_{s_{n_k}}$ admits a limit.*

Then γ is right- (resp. left-) continuous.

Proof. We imitate the proof of [85, Lem. 5]. We prove right-continuity of the rough timelike geodesic γ at $t = 0$, then right-continuity at other times follows by scaling, and the arguments for left continuity are analogous. Call y a limit of a sequence $k \mapsto \gamma_{s_{n_k}}$ for $s_{n_k} \downarrow 0$ as in the statement. Let (η_t) be equal to (γ_t) on $(0, 1]$ and defined as $\eta_0 := y$ for $t = 0$. We clearly have $|\dot{\eta}_t| = |\dot{\gamma}_t|$ for a.e. t and

$$A_1(\eta) \geq \limsup_{s \downarrow 0} \int_s^1 |\dot{\eta}_t| dt = \limsup_{s \downarrow 0} \int_s^1 |\dot{\gamma}_t| dt = A_1(\gamma) \geq \ell(\gamma_0, \gamma_1),$$

having used that γ is an ℓ -geodesic. Since $y \geq \gamma_0$ (here we use that $\{\ell \geq 0\}$ is closed), by item (iii) in Proposition 2.34 this suffices to establish that η is also a rough timelike ℓ -geodesic, so that non-branchingness at 0 forces $y = \eta_0 = \gamma_0$. This suffices to conclude right-continuity of γ at 0. \square

Clearly, a sufficient condition for left-limits to exist is the forward completeness that we typically assume. More relevant for us however, will be the notion of timelike non-branching at 0, which is linked to continuity at $t = 0$ of geodesics in $\mathcal{P}(\mathbb{M})$.

The following is a variant of Proposition 2.50. Recall that timelike (and strongly timelike) ℓ_q -geodesics (μ_t) satisfy $\ell_q(\mu_0, \mu_1) \in (0, \infty)$ as part of their definition.

Proposition 2.55 (Hereditiy of non-branchingness at 0 and/or 1). *Let (\mathbb{M}, τ, ℓ) be a forward metric spacetime. Let $0 \neq q < 1$ and assume that the collection \mathcal{C}_q of strongly timelike ℓ_q -geodesics in $\mathcal{P}_{\text{em}}(\mathbb{M})$ has the following property: if $(\mu_t) \in \mathcal{C}_q$ and $\nu \in \mathcal{P}_{\text{em}}(\mathbb{M})$ satisfies $\nu \preceq \mu_t$ for all $t \in (0, 1)$ and $\ell_q(\nu, \mu_1) = \ell_q(\mu_0, \mu_1)$, then $\nu = \mu_0$ (resp. $\mu_t \preceq \nu$ for every $t \in (0, 1)$ and $\ell_q(\mu_0, \nu) = \ell_q(\mu_0, \mu_1)$, then $\nu = \mu_1$). Then (\mathbb{M}, τ, ℓ) is timelike non-branching at 0 (resp. 1).*

Conversely, assume that (\mathbb{M}, τ, ℓ) is timelike non-branching at 0 (resp. 1), that ℓ is upper semicontinuous and does not take the value $+\infty$ and that either (A) or (B) of Theorem 2.43 hold. Then for every $0 \neq q < 1$, the collection \mathcal{C}_q has the property described above (respectively).

Proof. The proof is similar to that of Proposition 2.50. We deal with the case of non-branching at 0, that of non-branching at 1 being analogous.

Assume that the collection \mathcal{C}_q of strongly timelike ℓ_q -geodesics in $\mathcal{P}_{\text{em}}(\mathbb{M})$ has the indicated property. Since $\ell(x, y) = \ell_q(\delta_x, \delta_y)$ for any $x, y \in \mathbb{M}$ and the inclusion $\mathbb{M} \ni x \mapsto \delta_x \in \mathcal{P}_{\text{em}}(\mathbb{M})$ preserves the time separations, we easily deduce that $\text{TGeo}(\mathbb{M})$ has the following property: for $\gamma \in \text{TGeo}(\mathbb{M})$ and $x \in \mathbb{M}$ with $x \leq \gamma_t$ for all $t \in (0, 1)$ and $\ell(x, \gamma_1) = \ell(\gamma_0, \gamma_1)$ we must have $x = \gamma_0$. It is easy to see that this is equivalent to the claim of (\mathbb{M}, τ, ℓ) being timelike non-branching at 0, so in this case the conclusion follows.

Conversely, assume that (\mathbb{M}, ℓ) is timelike non-branching at 0, let $(\mu_t) \in \mathcal{C}_q$ be a strongly timelike ℓ_q -geodesic and $\nu \in \mathcal{P}_{\text{em}}(\mathbb{M})$ with $\nu \preceq \mu_t$ for any $t \in (0, 1)$ and $\ell_q(\nu, \mu_1) = \ell_q(\mu_0, \mu_1)$. Let $\pi \in \mathcal{P}(\text{TGeo}(\mathbb{M}))$ be a lifting of (μ_t) as in Corollary 2.46 and, for every $t \in (0, 1)$, let $\pi_t \in \Pi_{\leq}(\nu, \mu_t)$ be ℓ_q -optimal. Use a gluing argument to find $\alpha_t \in \mathcal{P}(\mathbb{M} \times \text{TGeo}(\mathbb{M}))$ such that

$$(\text{Pr}_2)_* \alpha_t = \pi, \quad \text{and} \quad (\text{Pr}_1, \mathbf{e}_t)_* \alpha_t = \pi_t.$$

Notice that the same arguments used in the proof of Theorem 2.43 easily implies that for any $t_n \downarrow 0$ the sequence (α_{t_n}) is tight, thus up to passing to a non-relabelled subsequence it admits a narrow limit α . Since α_t is concentrated on couples (x, γ) with $\gamma \in \text{TGeo}(\mathbb{M})$ and $x \leq \gamma_t$, the closure of $\{\ell \geq 0\}$ easily implies that α is concentrated on couples (x, γ) with $\gamma \in \text{TGeo}(\mathbb{M})$ and $x \leq \gamma_t$ for any $t \in (0, 1]$. The construction also yields

$$(\text{Pr}_2)_* \alpha = \pi, \quad \text{and} \quad (\text{Pr}_1)_* \alpha = \nu,$$

thus from the assumption that (\mathbb{M}, ℓ) is timelike non-branching at 0, to conclude it suffices to prove that for α -a.e. (x, γ) we have $\ell(x, \gamma_1) = \ell(\gamma_0, \gamma_1)$ (as this forces $x = \gamma_0$ and thus $\nu = \mu_0$, as

required). To this aim, notice that

$$\begin{aligned}
u_q \circ \ell_q(\nu, \mu_1) &\geq \int u_q \circ \ell(x, \gamma_1) \, d\alpha(x, \gamma) \\
(\text{as } x \leq \gamma_t \text{ for } \alpha\text{-a.e. } (x, \gamma)) &\geq \int \sup_{t \in (0,1)} u_q \circ \ell(\gamma_t, \gamma_1) \, d\alpha(x, \gamma) \\
(\text{by (2.24) with } P = \{0, 1\}) &\geq \iint_0^1 u_q(|\dot{\gamma}_t|) \, dt \, d\pi(\gamma) = u_q \circ \ell_q(\mu_0, \mu_1),
\end{aligned}$$

having used Corollary 2.46 in the last step. Since by assumption we have $\ell_q(\nu, \mu_1) = \ell_q(\mu_0, \mu_1) \in (0, +\infty)$, all the inequalities are in fact equalities and all the values are real: this can only be possible if $\ell(x, \gamma_1) = \sup_{t \in (0,1)} \ell(\gamma_t, \gamma_1) = \ell(\gamma_0, \gamma_1)$, forcing $x = \gamma_0$, for α -a.e. (x, γ) . This implies $\nu = \mu_0$, as desired. \square

We are interested in forcing curves of measures with bounded compression to be non-branching at 0. For this, it is sufficient to introduce the following weakening of Definition 2.53 above (analogous to Definition 2.51):

Definition 2.56 (*q-essential timelike non-branching at 0 or 1*). Fix $0 \neq q < 1$.

We say that a forward mm spacetime $(M, \tau, \ell, \mathfrak{m})$ is *q-essentially timelike non-branching at 0 (resp. at 1)* if the following holds. For each bounded compression strongly timelike ℓ_q -geodesic $(\mu_t) \subset \mathcal{P}_{\text{em}}(M)$ and $\nu \in \mathcal{P}_{\text{em}}(M)$ with bounded compression such that $\nu \preceq \mu_t$ for all $t \in (0, 1)$ and $\ell_q(\nu, \mu_1) = \ell_q(\mu_0, \mu_1)$ we must have $\nu = \mu_0$ (resp. $\mu_t \preceq \nu$ for every $t \in (0, 1)$ and $\ell_q(\mu_0, \nu) = \ell_q(\mu_0, \mu_1)$ we must have $\nu = \mu_1$).

Proposition 2.54 gives the following result, which is our reason for introducing this last concept:

Corollary 2.57 (Continuity for bounded compression ℓ_q -geodesics on a fixed emerald). Let $(M, \tau, \ell, \mathfrak{m})$ be a forward mm spacetime with \mathfrak{m} giving finite mass to each emerald, $0 \neq q < 1$ and (μ_t) a rough strongly timelike ℓ_q -geodesic of bounded compression such that $\text{spt } \mu_t \subset E$ for some fixed emerald E and all $t \in [0, 1]$.

If M is *q-essentially timelike non-branching at 0 (resp. at 1)* then (μ_t) is right (resp. left) continuous.

Proof. We know that $\mu_t \leq C\mathfrak{m} \llcorner E$ for some $C > 0$ and every $t \in [0, 1]$, which readily implies that the collection $\{\mu_t : t \in [0, 1]\}$ is tight, hence relatively narrowly compact. The conclusion follows from Proposition 2.54. \square

3 Lorentzian differential and Sobolev calculus

3.1 Some basic properties of causal functions

Definition 3.1 (Causal function). Let (M, ℓ) be a metric spacetime. A function $f: M \rightarrow \bar{\mathbf{R}}$ is called a *causal function* if it is *causally monotone*, i.e. $x \leq y$ implies $f(x) \leq f(y)$.

The domain $\text{Dom}(f)$ of a causal function $f: M \rightarrow \bar{\mathbf{R}}$ is the set of all $x \in M$ such that $f(x)$ is a real number.

Notice that the domain of a causal function is always causally convex.

If f and g are causal functions and $\lambda_1, \lambda_2 \geq 0$, then $\lambda_1 f + \lambda_2 g$ is again a causal function. If f and g are non-negative, then $f g$ is a causal function. If $\varphi: \bar{\mathbf{R}} \rightarrow \bar{\mathbf{R}}$ is non-decreasing on the image of f , then $\varphi \circ f$ is again a causal function. Lastly, arbitrary (pointwise) infima and suprema of causal functions are again causal functions.

The following lemma is an analogue for causal functions of the fact that a non-decreasing map of the line to itself can have only countably many discontinuities. This notably includes the case where f attains the values $\pm\infty$.

Recall a subset of M is *achronal* if none of its elements are related by \ll .

Lemma 3.2 (Chronological almost continuity). *Let (M, ℓ) be a metric spacetime and $f: M \rightarrow \bar{\mathbf{R}}$ causal. Define $f^\pm: M \rightarrow \bar{\mathbf{R}}$ by*

$$f^+(y) = \inf_{z \in I^+(y)} f(z), \quad \text{and} \quad f^-(y) = \sup_{x \in I^-(y)} f(x), \quad (3.1)$$

where we employ the usual conventions $\inf \emptyset := +\infty$ and $\sup \emptyset := -\infty$. Then:

- (i) The functions f^+ and f^- are causal functions.
- (ii) We have $f^- \leq f \leq f^+$ everywhere on M .
- (iii) Outside a countable union of achronal sets, $f^+ = f^-$.

Let M be equipped with a topology containing the chronological one. Then we also have:

- (iv) If a sequence $(y_n) \subset I^\pm(y)$ converges to y , then $\lim_n f^\pm(y_n) = f^\pm(y)$.
- (v) Both f^+ and $-f^-$ are upper semicontinuous. Hence f is continuous at every point of the complement of the countable union provided by (iii).

Proof.

- (i) For $x \leq y \ll z$ we have $x \ll z$ by the push-up property (Remark 2.3), whence $f^+(x) \leq f^+(y)$ as desired. An analogous argument shows f^- is a causal function.
- (ii) This is immediate from the causal property of f and the definitions (3.1).
- (iii) Given a number $q \in \mathbf{Q}$ we define

$$S_q^- := \{y \in M : f^-(y) < q < f(y)\} \quad \text{and} \quad S_q^+ := \{y \in M : f(y) < q < f^+(y)\}.$$

We claim that the sets S_q^\pm are achronal. Indeed, if $y_1, y_2 \in S_q^+$ were points satisfying $y_1 \ll y_2$, we would have $f^+(y_1) \leq f(y_2)$ by (3.1). This would produce the contradiction $q < f^+(y_1) \leq f(y_2) < q$. An analogous argument applies to S_q^- .

Since $f^+(y) > f^-(y)$ implies $y \in S_q^+ \cup S_q^-$ for some $q \in \mathbf{Q}$, (iii) is proven.

- (iv) We only show the statement for the function f^+ , the argument for f^- follows analogous lines. By the transitivity of \ll , $I^+(y_n)$ is contained in $I^+(y)$ for every $n \in \mathbf{N}$. Therefore $f^+(y) \leq f^+(y_n)$ by (3.1), and consequently

$$f^+(y) \leq \liminf_{n \rightarrow +\infty} f^+(y_n).$$

On the other hand, let $z \in I^+(y)$ be arbitrary. (Our assumption on the given sequence implies such a point does exist.) Since (y_n) converges to y and the set $I^-(z)$ is open, we eventually have $y_n \in I^-(z)$. Choosing z as a competitor in the definition (3.1) of $f^+(y_n)$ implies $\limsup_{n \rightarrow +\infty} f^+(y_n) \leq f(z)$ and therefore

$$\limsup_{n \rightarrow +\infty} f^+(y_n) \leq f^+(y)$$

by the arbitrariness of z .

- (v) Since the chronological topology makes $I^-(z)$ open for each $z \in M$, the formula

$$f^+(y) = \inf_{z \in M} f(z) + \infty \mathbf{1}_{M \setminus I^-(z)}(y)$$

shows f^+ is an inf of upper semicontinuous functions, hence itself upper semicontinuous. Similarly f^- is lower semicontinuous and both f^\pm are continuous at points where they agree, as is f . \square

As a direct consequence of the above we have:

Corollary 3.3 (Measurability of causal functions). *Let (M, ℓ) be a metric spacetime equipped with a $(\sigma$ -algebra and a) measure \mathfrak{m} that assigns zero outer measure to every achronal subset.*

Then for every $f: M \rightarrow \bar{\mathbf{R}}$ causal we have $f^+ = f^-$ \mathfrak{m} -a.e., where f^\pm are from (3.1). If, in addition, open sets in the chronological topology are \mathfrak{m} -measurable, then so is f .

Proof. The first claim follows from our hypotheses on \mathfrak{m} and Lemma 3.2. For the second it suffices to notice that for any $c \in \mathbf{R}$ we have

$$\{f < c\} = \left(\bigcup_{x \in \{f < c\}} I^-(x) \right) \cup \left(\{f < c\} \setminus \bigcup_{x \in \{f < c\}} I^-(x) \right),$$

that the first set on the right is chronologically open and the second achronal. \square

A reinforcement of Definition 3.1 is the subsequent analog of Lipschitz continuity from metric geometry. Recall our infinity conventions from the beginning of Section 2.

Definition 3.4 (Steepness). *Let (M, ℓ) be a metric spacetime, $L \in [0, \infty]$ and $f: M \rightarrow \bar{\mathbf{R}}$. We say that f is L -steep on $W \subset M$ if*

$$f(y) - f(x) \geq L \ell(x, y) \quad \forall x, y \in W.$$

If, in addition, $W = M$, we simply say f is L -steep.

Notice that a function is 0-steep if and only if it is a causal function. Typical examples of 1-steep functions are $\ell(x, \cdot)$ and $-\ell(\cdot, x)$ for given $x \in M$: their 1-steepness is a direct consequence of the reverse triangle inequality (2.19) (and the infinity conventions from the beginning of Section 2).

A steep function can always be extended to all of M by the following analog of the well-known McShane extension lemma from metric geometry, e.g. [122, Thm. 1.33]. Note that the conventions (2.1) about differences do not have a role in the definitions of f_\vee and f_\wedge below.

Lemma 3.5 (McShane-type extension). *Let (M, ℓ) be a metric spacetime. Let $f: W \rightarrow \bar{\mathbf{R}}$ be a given L -steep function, where $W \subset M$ is any subset and $L \geq 0$. Define the functions $f_\wedge, f_\vee: M \rightarrow \bar{\mathbf{R}}$ by*

$$\begin{aligned} f_\wedge(y) &:= \sup\{f(x) + L \ell(x, y) : x \in W \text{ with } x \leq y \text{ and } f(x) > -\infty\}, \\ f_\vee(x) &:= \inf\{f(y) - L \ell(x, y) : y \in W \text{ with } y \geq x \text{ and } f(y) < +\infty\}. \end{aligned}$$

Then f_\wedge and f_\vee are L -steep functions which coincide with f on W . Furthermore, these extensions are extremal, in the sense that every L -steep extension $g: M \rightarrow \bar{\mathbf{R}}$ of f satisfies $f_\wedge \leq g \leq f_\vee$ everywhere on M .

Proof. We show all the claims for f_\wedge , those for f_\vee being similar.

Let us first show that f_\wedge coincides with f on W . Let $x \in W$. Choosing x as a competitor in the supremum defining f_\wedge and using $\ell(x, x) = 0$ yields $f_\wedge(x) \geq f(x)$. On the other hand, by L -steepness of f on W , $L \ell(x, y) + f(x) \leq f(y)$ for any $x, y \in W$ with $f(x) > -\infty$, taking the supremum over x implies $f_\wedge(y) \leq f(y)$, which is the claim.

Next, we show L -steepness of f_\wedge on M . Let $y_1, y_2 \in M$. We claim

$$f_\wedge(y_2) - f_\wedge(y_1) \geq L \ell(y_1, y_2). \quad (3.2)$$

We may assume that $y_1 \leq y_2$, otherwise the claim is trivial. Let $x \in W$ be a competitor in the definition of $f_\wedge(y_1)$ (we can assume such x exists, otherwise $f_\wedge(y_1) = -\infty$ and according to the convention (2.1) the inequality (3.2) holds). Notice that x is also a competitor in the definition of $f_\wedge(y_2)$ and that in this case the reverse triangle inequality can be written as $\ell(y_1, y_2) + \ell(x, y_1) \leq \ell(x, y_2)$ and thus trivially

$$f(x) + L \ell(x, y_2) \geq f(x) + L \ell(x, y_1) + L \ell(y_1, y_2).$$

Taking the sup in x we obtain that

$$f_\wedge(y_2) \geq f_\wedge(y_1) + L \ell(y_1, y_2) \quad (3.3)$$

and thus in particular that f_\wedge is causal. Then, since by the conventions (2.1) the claim (3.2) holds provided either $f_\wedge(y_2) = +\infty$ or $f_\wedge(y_1) = -\infty$, we can assume that $f_\wedge(y_1), f_\wedge(y_2) \in \mathbf{R}$. In this case, though, the claim (3.2) is a direct consequence of (3.3).

Finally, let g be any L -steep extension of f to M and $y \in M$. We need to prove that

$$g(y) \geq f_\wedge(y). \quad (3.4)$$

To this end, let $x \in W$ be a competitor in the definition of $f_\wedge(y)$ (as before: if such x does not exist we have $f_\wedge(y) = -\infty$ and (3.4) follows). By L -steepness of g we have

$$g(y) - f(x) = g(y) - g(x) \geq L \ell(x, y). \quad (3.5)$$

By assumption we know that $f(x) > -\infty$ while we see from the above that if $f(x) = +\infty$ for some competitor x , then $g(y) = +\infty$ and (3.4) follows. Thus we can assume that $f(x) \in \mathbf{R}$ for every x competitor in the definition of $f_\wedge(y)$. In turn, the relation (3.5) tells us that $g(y) \geq f(x) + L \ell(x, y)$ for any such x , thus taking the supremum in x we conclude the statement. \square

A key object of study of this manuscript is differential calculus with causal functions. To develop this a first step is to leverage the concept of causal speed introduced above to define a notion of ‘modulus of differential’ by duality. In Polish metric spacetimes, relevant proxies of what will turn out to be ‘the correct’ definition are the *forward* and *backward* slopes $|\partial^+ f|$ and $|\partial^- f|$, respectively, defined as:

$$\begin{aligned} |\partial^+ f|(x) &:= \liminf_{y \downarrow x} \frac{f^+(y) - f^-(x)}{\ell(x, y)} := \sup_{y \in U \cap J^+(x)} \inf_{y \in U \cap J^+(x)} \frac{f^+(y) - f^-(x)}{\ell(x, y)}, \\ |\partial^- f|(x) &:= \liminf_{z \uparrow x} \frac{f^+(x) - f^-(z)}{\ell(z, x)} := \sup_{z \in U \cap J^-(x)} \inf_{z \in U \cap J^-(x)} \frac{f^+(x) - f^-(z)}{\ell(z, x)}, \end{aligned} \quad (3.6)$$

where both suprema are taken over all neighborhoods U of x . In these formulas, $0/0$ is intended to be equal to $+\infty$, thus, for instance, if x has a neighborhood U for which there is no $y \in U$ different from x with $y \geq x$, then $|\partial^+ f|(x) = +\infty$. Similarly for $|\partial^- f|$. If f is L -steep, then clearly we have

$$|\partial^+ f|(x), |\partial^- f|(x) \geq L \quad \forall x \in M. \quad (3.7)$$

The functions $|\partial^\pm f|$ could be considered as Lorentzian counterparts of the ascending/descending slopes appearing in positive signature and they will have a similar role in the proof of the metric Brenier–McCann theorem (compare [10, Eq. (2.6), Thm. 10.3] with (3.6) above and Theorem 5.19 respectively). In this direction, the analog of the local Lipschitz constant as e.g. defined in [10] would be the *local steepness constant* $\text{st}(f): M \rightarrow [0, +\infty]$ defined through the assignment

$$\text{st}(f)(x) := \min\{|\partial^+ f|(x), |\partial^- f|(x)\}.$$

In the smooth setting, the reverse Cauchy-Schwarz inequality yields:

$$\partial_t(f \circ \gamma)_t = \text{d}f_{\gamma_t}(\gamma'_t) \geq \|\text{d}f\|_* \|\gamma'_t\|.$$

The following result offers a first non-smooth counterpart to this.

Proposition 3.6 (Pathwise growth exceeds integrated forward and backward slopes). *Let (M, τ, ℓ) be a Polish metric spacetime, $f: M \rightarrow \mathbf{R}$ a causal function and $\gamma: [0, 1] \rightarrow M$ a causal curve. Then there is a Borel function $g: [0, 1] \rightarrow [0, +\infty]$ with $g_t \geq \max\{|\partial^+ f|, |\partial^- f|\}(\gamma_t)$ for every $t \in [0, 1]$ and such that*

$$f(\gamma_1) - f(\gamma_0) \geq \int_0^1 g_t |\dot{\gamma}_t| dt. \quad (3.8)$$

If the topology τ contains the chronological one and ℓ_+ is lower semicontinuous, then $|\partial^+ f|$ and $|\partial^- f|$ are Borel measurable.

Proof. If either $f(\gamma_1) = +\infty$ or $f(\gamma_0) = -\infty$ the conventions (2.1) ensure that (3.8) hold with $g \equiv +\infty$. Taking causality of f, γ into account we can thus assume that f is real valued.

Lemma 3.2 asserts $f^- \leq f \leq f^+$ with equality holding outside a countable union Σ of achronal sets. Let $S := \text{spt} |\dot{\gamma}| \subset [0, 1]$ denote the smallest closed set outside of which the causal speed of γ vanishes. Its complement consists of countably many intervals (a, b) . Unless $a, b \in S$ are the endpoints of one of these intervals, if $a < b$ then Theorem 2.23(i) and the Definition 2.25 of causal speed show $0 < |\dot{\gamma}|(a, b) \leq \ell(\gamma_a, \gamma_b)$. This chronological relation shows $S \cap \gamma^{-1}(\Sigma)$ is at most countable. Thus the non-decreasing functions $f^\pm \circ \gamma$ satisfy $f^- \circ \gamma \leq f \circ \gamma \leq f^+ \circ \gamma$ on $[0, 1]$, with equality holding on S outside of the countable set $S \cap \gamma^{-1}(\Sigma)$.

By Lebesgue's theorem, see e.g. [55, Theorem 3.23], (or by Theorem 2.23 applied to $T(s, t) := f(\gamma_t) - f(\gamma_s)$) a monotone real-valued function is differentiable Lebesgue almost everywhere and letting $(f \circ \gamma)'$ be this pointwise derivative yields

$$f(\gamma_1) - f(\gamma_0) \geq \int_0^1 (f \circ \gamma)'(t) dt.$$

Now let $s_t := \liminf_{h \downarrow 0} \min\{\frac{\ell(\gamma_t, \gamma_{t+h})}{h}, \frac{\ell(\gamma_{t-h}, \gamma_t)}{h}\}$ and notice that $t \mapsto s_t$ is Lebesgue measurable (as a consequence of the monotonicity of ℓ in its entries — see the already cited [42, Thm. 4]) and equal to $|\dot{\gamma}_t|$ for a.e. t . Let $D \subset S$ be the set of t 's where $s_t > 0$ for which the differentials of $f \circ \gamma$ and $f^\pm \circ \gamma$ exist and agree. Set $g_t := \frac{1}{s_t} \lim_{h \rightarrow 0} \frac{f(\gamma_{t+h}) - f(\gamma_t)}{h}$ on D and $g_t := +\infty$ on $[0, 1] \setminus D$. It is clear that with this choice (3.8) holds, so we just have to prove that $g_t \geq \max\{|\partial^+ f|, |\partial^- f|\}(\gamma_t)$ holds for every $t \in D$. To see this notice that

$$(f \circ \gamma)'(t) = \lim_{h \downarrow 0} \frac{f^+(\gamma_{t+h}) - f^-(\gamma_t)}{h} \geq s_t \liminf_{h \downarrow 0} \frac{f^+(\gamma_{t+h}) - f^-(\gamma_t)}{\ell(\gamma_t, \gamma_{t+h})} \geq |\partial^+ f|(\gamma_t) s_t.$$

and similarly

$$(f \circ \gamma)'(t) = \lim_{h \downarrow 0} \frac{f^+(\gamma_t) - f^-(\gamma_{t-h})}{h} \geq s_t \liminf_{h \downarrow 0} \frac{f^+(\gamma_t) - f^-(\gamma_{t-h})}{\ell(\gamma_{t-h}, \gamma_t)} \geq |\partial^- f|(\gamma_t) s_t,$$

proving the claim. The g as defined is Lebesgue measurable, but modifying it in a suitable Lebesgue measurable set of measure 0 by setting it to $+\infty$ therein, we retain all of the required properties and obtain Borel measurability of g .

We turn to Borel measurability of $|\partial^+ f|$ and $|\partial^- f|$. Let $F : M^2 \rightarrow [0, +\infty]$ be defined as

$$F(x, y) := \begin{cases} \frac{f^+(y) - f^-(x)}{\ell(x, y)} & \text{if } \ell(x, y) > 0, \\ +\infty & \text{otherwise.} \end{cases}$$

Recall that Lemma 3.2(v) yields upper semicontinuity of $\pm f^\pm$, and also $f^+(y) \geq f(y) \geq f(x) \geq f^-(x)$ whenever $\ell(x, y) \geq 0$. As the set $\{x : \ell(x, y) > 0\}$ is open for each $y \in M$, the assumed lower semicontinuity of ℓ_+ implies the assignment $(x, y) \mapsto F(x, y)$ is upper semicontinuous.

Now let (U_n) be a countable base for the Polish topology of M and notice that from the monotonicity of the quantity $\inf_{y \in U \cap J^+(x)} (f^+(y) - f^-(x))/\ell(x, y)$ in U with respect to inclusion it follows that

$$|\partial^+ f|(x) = \sup_{n \in \mathbf{N}} \inf_{y \in U_n} F(x, y)$$

for every $x \in M$, hence $|\partial^+ f|$ is Borel measurable. Similarly, one shows that $|\partial^- f|$ is Borel measurable. \square

3.2 Test plans on metric measure spacetimes

The concept of ‘weak subslope’ will be given by duality with test plans. The latter are defined by:

Definition 3.7 (Test plans). *Let $(M, \tau, \ell, \mathbf{m})$ be a mm spacetime. A measure $\pi \in \mathcal{P}(\text{LCC}([0, 1]; M))$ is called a test plan if it has bounded compression, i.e. there exists a constant $C > 0$ such that $(\mathbf{e}_t)_* \pi \leq C \mathbf{m}$ for every $t \in [0, 1]$.*

Also, we will call $\pi \in \mathcal{P}(\text{LCC}([0, 1]; M))$

- a) an initial test plan if for some $T \in (0, 1)$ and $C > 0$ we have $(\mathbf{e}_t)_* \pi \leq C \mathbf{m}$ for all $t \in [0, T]$ and moreover $(\mathbf{e}_t)_* \pi \rightarrow (\mathbf{e}_0)_* \pi$ in the narrow topology as $t \downarrow 0$,*
- b) a final test plan if for some $T \in (0, 1)$ and $C > 0$ we have $(\mathbf{e}_t)_* \pi \leq C \mathbf{m}$ for all $t \in [T, 1]$ and moreover $(\mathbf{e}_s)_* \pi \rightarrow (\mathbf{e}_1)_* \pi$ in the narrow topology as $s \uparrow 1$.*

Remark 3.8 (Initial and final limits and time reversibility). In forward metric measure spacetimes, the narrow continuity at time one (stipulated in point (b) above) holds automatically. We impose it anyways, to ensure we could work equally well with backward metric measure spacetimes, at the cost of exchanging left-notions for right-notions throughout. This ensures the theory is as symmetric as possible under time-reversal. ■

We shall refer to the condition “ $(\mathbf{e}_t)_* \pi \leq C \mathbf{m}$ for every $t \in [0, 1]$ ” by saying that π has bounded compression. Among other things, it ensures \mathbf{m} -a.e. defined functions to be well-defined along test plans as follows: for $t \in [0, 1]$, it yields $f \circ \mathbf{e}_t = g \circ \mathbf{e}_t$ π -a.e. if $f = g$ \mathbf{m} -a.e.

Example 3.9 (Statics). Let $\text{const}: M \rightarrow \text{LCC}([0, 1]; M)$ denote the continuous map sending a point $x \in M$ to the curve constantly equal to x . If $\mu \in \mathcal{P}(M)$ then $\pi := \text{const}_* \mu$ satisfies $(\mathbf{e}_t)_* \pi = \mu$ for all $t \in [0, 1]$. Thus π is a test plan if and only if $\mu \leq C \mathbf{m}$ for some $C > 0$, i.e. if and only if μ itself has bounded compression. ■

We conclude this short section describing procedures to build new test plans from old. Let (M, τ, ℓ) be a Polish metric spacetime:

Lemma 3.10 (Restrictions of test plans). *Let $(M, \tau, \ell, \mathbf{m})$ be a mm spacetime and π a test plan on it. Then for every $s, t \in [0, 1]$ with $s < t$ and every Borel set $\Gamma \subset \text{LCC}([0, 1]; M)$ with $\pi[\Gamma] > 0$, the measures $(\text{restr}_s^t)_* \pi$ and $\pi[\Gamma]^{-1} \pi \llcorner \Gamma$ are again test plans.*

Proof. Both measures clearly belong to $\mathcal{P}(\text{LCC}([0, 1]; M))$. Bounded compression in either case follows easily as well. Indeed, if $(\mathbf{e}_r)_* \pi \leq C \mathbf{m}$ for any r , then we have

$$(\mathbf{e}_r)_* [(\text{restr}_s^t)_* \pi] = (\mathbf{e}_r \circ \text{restr}_s^t)_* \pi = (\mathbf{e}_{(1-r)s+rt})_* \pi \leq C \mathbf{m}$$

and similarly

$$(\mathbf{e}_t)_* [\pi[\Gamma]^{-1} \pi \llcorner \Gamma] \leq \pi[\Gamma]^{-1} (\mathbf{e}_t)_* \pi \leq \pi[\Gamma]^{-1} C \mathbf{m}. \quad \square$$

3.3 Weak subslopes

From now on all causal functions are assumed to be \mathbf{m} -measurable unless explicitly stated otherwise; (this hypothesis becomes redundant in those cases covered by Corollary 3.3 anyways). We reserve the term *rough causal* to refer to a causal function which is not necessarily \mathbf{m} -measurable.

We have already noticed that in the smooth category, given a smooth causal function f and a smooth causal curve γ , the reverse Cauchy-Schwarz inequality yields:

$$\partial_t (f \circ \gamma)_t = \text{d}f_{\gamma_t}(\dot{\gamma}_t) \geq \|\text{d}f\|_* \|\dot{\gamma}_t\| \quad \forall t.$$

Also, it is easy to see that $\|\text{d}f\|_*$ is the maximal continuous function for which the above inequality holds for any curve and t . Starting from this consideration and inspired by analogous definitions in the metric setting [10], we propose the following ‘weak’ notion:

Definition 3.11 (Weak subslope). *Let $(M, \tau, \ell, \mathbf{m})$ be a mm spacetime and $f: M \rightarrow \bar{\mathbf{R}}$ a causal function. A Borel function $G: M \rightarrow [0, +\infty]$ is called weak subslope of f if*

$$\int [f(\gamma_1) - f(\gamma_0)] \text{d}\pi(\gamma) \geq \int \int_0^1 G(\gamma_t) |\dot{\gamma}_t| \text{d}t \text{d}\pi(\gamma) \quad \text{for every test plan } \pi. \quad (3.9)$$

Some elementary comments on this notion are in order. By non-negativity of the respective integrands, both integrals in (3.9) are well-defined, possibly with value $+\infty$. Because test plans have bounded compression, Definition 3.11 is not influenced by changes of f or G on \mathfrak{m} -negligible sets. The set of weak subslopes of a given causal function is never empty, as it always contains the function identically 0 on M .

A direct and relevant consequence of working with test plans rather than with single curves is the following stability result:

Proposition 3.12 (Stability of weak subslopes). *Let $(M, \tau, \ell, \mathfrak{m})$ be a mm spacetime and (f_n) a sequence of causal functions converging pointwise \mathfrak{m} -a.e. to a causal function f . Let (G_n) be a sequence of weak subslopes G_n of f_n which converges pointwise \mathfrak{m} -a.e. to a function G .*

Then G is a weak subslope of f .

Proof. Let π be a fixed test plan. For the moment, we assume

$$\sup_{n \in \mathbf{N}} \|f_n \circ \mathbf{e}_1 - f_n \circ \mathbf{e}_0\|_{L^\infty(\text{LCC}([0,1]; M), \pi)} < +\infty. \quad (3.10)$$

Our hypothesis and bounded compression give $f_n \circ \mathbf{e}_1 - f_n \circ \mathbf{e}_0 \rightarrow f \circ \mathbf{e}_1 - f \circ \mathbf{e}_0$ π -a.e. as $n \rightarrow +\infty$. Lebesgue's dominated convergence theorem and (3.10) imply

$$\int [f(\gamma_1) - f(\gamma_0)] d\pi(\gamma) = \lim_{n \rightarrow +\infty} \int [f_n(\gamma_1) - f_n(\gamma_0)] d\pi(\gamma).$$

On the other hand, again by our hypothesis and bounded compression, the set of all pairs $(\gamma, t) \in \text{LCC}([0, 1]; M) \times [0, 1]$ with $G_n(\gamma_t) \rightarrow G(\gamma_t)$ as $n \rightarrow +\infty$ has full $\pi \otimes \mathcal{L}^1$ -measure. Fubini's theorem and Fatou's lemma then give

$$\liminf_{n \rightarrow +\infty} \iint_0^1 G_n(\gamma_t) |\dot{\gamma}_t| dt d\pi(\gamma) \geq \iint_0^1 G(\gamma_t) |\dot{\gamma}_t| dt d\pi(\gamma).$$

This shows the desired property (3.9) under the hypothesis (3.10).

To get rid of the assumption (3.10), we argue by approximation. Let π be an arbitrary test plan. If $f \circ \mathbf{e}_1 - f \circ \mathbf{e}_0 = +\infty$ on a set of positive π -measure, the conclusion is trivial, so we can assume that $f \circ \mathbf{e}_1 - f \circ \mathbf{e}_0 < +\infty$ π -a.e. Then since $f_n \circ \mathbf{e}_1 - f_n \circ \mathbf{e}_0 \rightarrow f \circ \mathbf{e}_1 - f \circ \mathbf{e}_0$ π -a.e. as $n \rightarrow +\infty$, we obtain $\limsup_{n \in \mathbf{N}} [f_n(\gamma_1) - f_n(\gamma_0)] < +\infty$ for π -a.e. $\gamma \in \text{LCC}([0, 1]; M)$. Thus, the collection $(\Gamma_{i,j})_{i,j \in \mathbf{N}}$ of π -measurable sets defined as

$$\Gamma_{i,j} := \left\{ \gamma \in \text{LCC}([0, 1]; M) : \sup_{n \geq j} [f_n(\gamma_1) - f_n(\gamma_0)] \leq i \right\}$$

satisfies $\pi(\cup_{i,j} \Gamma_{i,j}) = 1$. As in Lemma 3.10, setting $\pi_{i,j} := \pi(\Gamma_{i,j})^{-1} \pi \llcorner \Gamma_{i,j}$ gives a collection $(\pi_{i,j})_{i,j \in \mathbf{N}}$ of test plans such that, thanks to the previous part, (3.9) holds for π replaced by $\pi_{i,j}$ for every $i, j \in \mathbf{N}$. Since $\Gamma_{i,j} \subset \Gamma_{i',j'}$ for $i \leq i'$ and $j \leq j'$, letting $i, j \rightarrow +\infty$ in the resulting inequality and using Levi's monotone convergence theorem, the claim follows. \square

Our next goal is to prove that the collection of weak subslopes of a given function has a maximal element, intended in the \mathfrak{m} -a.e. sense. To show this, it is convenient to have the following pathwise characterization of weak subslopes:

Lemma 3.13 (An essentially pathwise criterion for weak subslopes). *Let $(M, \tau, \ell, \mathfrak{m})$ be a mm spacetime, $f: M \rightarrow \bar{\mathbf{R}}$ a causal function, and $G: [0, +\infty] \rightarrow \bar{\mathbf{R}}$ Borel.*

Then G is a weak subslope of f if and only if for any test plan π we have:

- i) For π -a.e. $\gamma \in \text{LCC}([0, 1]; M)$ the function $t \mapsto G(\gamma_t) |\dot{\gamma}_t|$ belongs to $L^1_{\text{loc}}(\text{Dom}(f \circ \gamma); \mathcal{L}^1)$;*
- ii) the distributional derivative of the non-decreasing and real valued restriction of $f \circ \gamma$ to $\text{Dom}(f \circ \gamma)$ is bounded from below by $G \circ \gamma |\dot{\gamma}| \mathcal{L}^1$ on $\text{Dom}(f \circ \gamma)$.*

Item (ii) could also be replaced by

ii') the density $(f \circ \gamma)'$ of the absolutely continuous part of the distributional derivative of $f \circ \gamma$ in $\text{Dom}(f \circ \gamma)$ satisfies $(f \circ \gamma)' \geq G \circ \gamma |\dot{\gamma}|$ \mathcal{L}^1 -a.e. in $\text{Dom}(f \circ \gamma)$.

In particular, if these hold for a test plan π the function $t \mapsto G(\gamma_t) |\dot{\gamma}_t|$ belongs to $L^1([0, 1]; \mathcal{L}^1)$ for π -a.e. $\gamma \in \text{LCC}([0, 1]; \mathbf{M})$ with $f(\gamma_1) - f(\gamma_0) < +\infty$.

Proof. The equivalence of (ii) and (ii') is an obvious consequence of the fact that $f \circ \gamma$ is non-decreasing and the last claim is a trivial consequence of (ii') and the bound $\int_0^1 (f \circ \gamma)'_t dt \leq f(\gamma_1) - f(\gamma_0)$, that in turn follows from the monotonicity of $f \circ \gamma$.

Now assume first that G is a weak subslope of f . We claim that, given any $s, t \in [0, 1]$ with $s < t$, π -a.e. $\gamma \in \text{LCC}([0, 1]; \mathbf{M})$ satisfies

$$f(\gamma_t) - f(\gamma_s) \geq \int_s^t G(\gamma_r) |\dot{\gamma}_r| dr. \quad (3.11)$$

Indeed, if not there are $s, t \in [0, 1]$ with $s < t$ as well as $\varepsilon > 0$ such that the set Γ_ε of all $\gamma \in \text{LCC}([0, 1]; \mathbf{M})$ such that $[s, t] \subset \text{Dom}(f \circ \gamma)$ and

$$f(\gamma_t) - f(\gamma_s) + \varepsilon \leq \int_s^t G(\gamma_r) |\dot{\gamma}_r| dr \quad (3.12)$$

satisfies $\pi[\Gamma_\varepsilon] > 0$. Up to further reducing ε and slightly reducing Γ_ε as well, as in the proof of Proposition 3.12 we may and will assume $f \circ e_t - f \circ e_s$ to be uniformly bounded on Γ_ε . Then by Lemma 3.10, $\pi' := (\text{restr}_s^t)_* [\pi[\Gamma_\varepsilon]^{-1} \pi \llcorner \Gamma_\varepsilon]$ constitutes a test plan and is admissible in (3.9), but the latter in conjunction with (3.12) as well as the finiteness of $\int \int_0^1 G(\gamma_\alpha) |\dot{\gamma}_\alpha| d\alpha d\pi'(\gamma)$, as implied by our hypothesis on $f \circ e_t - f \circ e_s$, directly leads to a contradiction.

We then deduce that for any $s, t \in [0, 1] \cap \mathbf{Q}$ with $s < t$ the bound (3.11) holds for π -a.e. γ . Then approximating arbitrary reals number t from below and s from above we conclude that for π -a.e. γ the bound (3.11) holds for any $s, t \in [0, 1]$ with $s < t$. From this, the validity of (ii) easily follows.

Conversely, we assume (ii). Let π be a test plan. The hypothesis implies that for π -a.e. $\gamma \in \text{LCC}([0, 1]; \mathbf{M})$, the distributional derivative of the function $h: \text{Dom}(f \circ \gamma) \rightarrow \mathbf{R}$ given by

$$h(t) := f(\gamma_t) - \int_a^t G(\gamma_r) |\dot{\gamma}_r| dr$$

is non-negative (here we assume $\text{Dom}(f \circ \gamma) = [a, b]$ for notational convenience). In turn, this gives $h(s) \leq h(t)$ for \mathcal{L}^1 -a.e. $s, t \in [a, b]$ with $s < t$. Since f is a causal function, for such s and t this implies

$$f(\gamma_1) - f(\gamma_0) \geq f(\gamma_t) - f(\gamma_s) \geq \int_s^t G(\gamma_r) |\dot{\gamma}_r| dr. \quad (3.13)$$

The dependence of the right-hand side on s and t is continuous if the difference $f(\gamma_1) - f(\gamma_0)$ is finite; in this case, (3.13) holds for $s = 0$ and $t = 1$. If $f(\gamma_1) - f(\gamma_0) = +\infty$, the estimate (3.13) holds vacuously for these choices of s and t , and so does (3.13) for π -a.e. $\gamma \in \text{LCC}([0, 1]; \mathbf{M})$. Integration yields the claim. \square

This pathwise characterization immediately implies that when

$$G_1, G_2 \text{ are subslopes of } f \quad \Rightarrow \quad \max\{G_1, G_2\} \text{ is a subslope of } f. \quad (3.14)$$

Indeed, if π is a test plan, by Lemma 3.13 and with the notation of item (ii'), for π -a.e. $\gamma \in \text{LCC}([0, 1]; \mathbf{M})$ we have $(f \circ \gamma)' \geq (G_1 \circ \gamma) |\dot{\gamma}|$ and $(f \circ \gamma)' \geq (G_2 \circ \gamma) |\dot{\gamma}|$ \mathcal{L}^1 -a.e. on $\text{Dom}(f \circ \gamma)$. It follows that for π -a.e. γ we have $(f \circ \gamma)' \geq \max\{G_1, G_2\} \circ \gamma |\dot{\gamma}|$ \mathcal{L}^1 -a.e. on $\text{Dom}(f \circ \gamma)$, which by Lemma 3.13 suffices to deduce (3.14).

The main result of the section is:

Theorem 3.14 (Unique existence of maximal weak subslopes). *Let $(M, \tau, \ell, \mathfrak{m})$ be a mm spacetime and $f: M \rightarrow \bar{\mathbf{R}}$ causal. Then there exists a Borel function $G: M \rightarrow [0, +\infty]$ which is a weak subslope of f and \mathfrak{m} -a.e. no smaller than any other weak subslope of f .*

In particular, G is unique up to modifications on \mathfrak{m} -negligible sets.

Proof. The conditions on G force it to be the \mathfrak{m} -essential supremum of the collection of all weak subslopes of f : what we need to prove is that this G is still a weak subslope. From general properties of essential suprema we know that there is a sequence (G_n) of weak subslopes such that \mathfrak{m} -a.e. we have $G = \sup_n G_n$. By (3.14) the functions $H_n := \max_{i \leq n} G_i$ are still weak subslopes and by Proposition 3.12 (with $f_n = f$) we conclude that G is also a weak subslope, as desired. \square

The concept isolated by this last result deserves a definition:

Definition 3.15 (Maximal weak subslope). *Let $(M, \tau, \ell, \mathfrak{m})$ be a mm spacetime and $f: M \rightarrow \bar{\mathbf{R}}$ causal. The \mathfrak{m} -a.e. unique element G from Theorem 3.14 is called maximal weak subslope of f and denoted by $|df|$.*

Notice that if f is L -steep, then we have

$$|df| \geq L \quad \mathfrak{m} - a.e., \quad (3.15)$$

as follows directly from the maximality of $|df|$, the definition of L -steepness and the bound (2.27).

A direct consequence of the definition and Proposition 3.12 is the following:

Proposition 3.16 ('Upper semicontinuous' dependence). *Let $(M, \tau, \ell, \mathfrak{m})$ be a mm spacetime and (f_n) a sequence of causal functions on it converging pointwise \mathfrak{m} -a.e. to a causal function f . Then*

$$\liminf_{n \rightarrow +\infty} |df_n| \leq |df| \quad \mathfrak{m}\text{-a.e.} \quad (3.16)$$

Proof. The function $G_n := \inf_{i \geq n} |df_i| \leq |df_n|$ is clearly a weak subslope of f_n . Since $G_n \rightarrow \liminf_n |df_n|$ \mathfrak{m} -a.e. the conclusion follows from Proposition 3.12. \square

Remark 3.17. The result above would be false if we replace \liminf with \limsup , which is why we put the quotation marks in 'upper semicontinuous'. Still, given that in many cases of interest, e.g. in studying calculus rules in the upcoming sections, the sequence $(|df_n|)$ admits a limit \mathfrak{m} -a.e., speaking of upper semicontinuity is not entirely inappropriate.

For a counterexample to (3.16) with the \limsup consider $[0, 1]$ with the Euclidean topology and $\ell(x, y) := y - x$ if $y \geq x$ and $-\infty$ otherwise. Find a sequence $(g_n) \subset L^1([0, 1])$ of non-negative functions with $\limsup_n g_n \equiv 1$ and $\|g_n\|_{L^1} \downarrow 0$, then define $f_n(x) := \int_0^x g_n$. Quite clearly we have $|df_n| = g_n$ (see also Appendix A.1), $f_n \rightarrow 0$ but $|d0| = 0 \not\geq 1$. \blacksquare

The maximality and Proposition 3.6 directly imply:

Proposition 3.18 (Forward and backward slopes are weak subslopes). *Let $(M, \tau, \ell, \mathfrak{m})$ be a mm spacetime whose topology contains the chronological one and so that ℓ_+ is lower semicontinuous. Let $f: M \rightarrow \bar{\mathbf{R}}$ be a causal function.*

Then $|\partial^+ f|$ and $|\partial^- f|$ are weak subslopes. In particular, letting L being the steepness constant of f we have

$$|df| \geq |\partial^+ f|, |\partial^- f| \geq L \quad \mathfrak{m} - a.e. \quad (3.17)$$

Proof. By Proposition 3.6 we see that $|\partial^+ f|$ and $|\partial^- f|$ are Borel and thus also that the bound $f(\gamma_1) - f(\gamma_0) \geq \int_0^1 \max\{|\partial^+ f|, |\partial^- f|\} |\dot{\gamma}_t| dt$ holds for any causal curve γ . Integrating w.r.t. an arbitrary test plan we deduce that $|\partial^+ f|$ and $|\partial^- f|$ are weak subslopes. Then (3.17) follows from the definition of $|df|$ and (3.7). \square

3.4 Calculus rules

This part establishes several calculus rules for the maximal weak subslope of causal functions, notably locality, the chain rule, and the Leibniz rule. We start with some elementary observations about regions where $|df|$ is equal to $+\infty$. A first simple comment comes from our conventions from Section 2.1 and the maximality proclaimed in Definition 3.15: they readily imply that for any causal function f we have

$$|df| = +\infty \text{ m-a.e. on the complement of } \text{Dom}(f).$$

A related result involves the parts of our spacetime that are ‘visible’ via test plans. We introduce the concept. Let π be a given test plan and call ρ_π the \mathfrak{m} -density of the push-forward of the measure $|\dot{\gamma}_t| d\pi(\gamma) dt$ on $\text{LCC}([0, 1]; M) \times [0, 1]$ under the ‘full’ evaluation map $e(\gamma, t) := \gamma_t$. Let $\text{Vis}_\pi := \{\rho_\pi > 0\} \subset M$. We think of the Borel set Vis_π — defined up to \mathfrak{m} -negligible sets — as the set that is ‘visible’ from π , whence the notation. Then we define $\text{Vis}(M)$ as the \mathfrak{m} -essential union of all the Vis_π as π varies among test plans. Recall (see for instance [64, Prop. 3.1.9] for this basic result in measure theory) that $\text{Vis}(M)$ is characterized up to \mathfrak{m} -negligible sets as the smallest Borel set that contains all the Vis_π ’s up to \mathfrak{m} -negligible sets (this notion is closely related to that of \mathfrak{m} -essential supremum of a family of Borel functions that we have already seen in Theorem 3.14). We shall frequently use the fact that a property holds \mathfrak{m} -a.e. on $\text{Vis}(M)$ if and only if it holds \mathfrak{m} -a.e. on Vis_π for every test plan π .

We then define the ‘invisible’ set $\text{Invis}(M) := M \setminus \text{Vis}(M)$ and notice that

$$|df| = +\infty \text{ m-a.e. on } \text{Invis}(M) \text{ for any causal function } f. \quad (3.18)$$

To see this, by pointwise maximality of $|df|$ it suffices to show that $G := \infty 1_{\text{Invis}(M)}$ is a weak subslope of f and to this aim it is enough to prove that $\iint_0^1 G(\gamma_t) |\dot{\gamma}_t| dt d\pi(\gamma)$ is equal to 0 for any test plan π . This, however, is obvious from

$$\iint_0^1 G(\gamma_t) |\dot{\gamma}_t| dt d\pi(\gamma) = \int G d\mathfrak{e}_*(|\dot{\gamma}_t| \pi \times \mathcal{L}^1 \llcorner [0, 1]) = \int G \rho_\pi d\mathfrak{m} = 0.$$

Because of (3.18), our calculus rules are sometimes phrased on the complement $\text{Vis}(M)$ of $\text{Invis}(M)$ as this will simplify our formulas. Also, we will often restrict our attention to regions where $|df| < +\infty$, so that automatically from (3.18) we will not care about what happens in the invisible set $\text{Invis}(M)$.

A simple property very much in this vein is:

Lemma 3.19 (Achronal sets are invisible). *Let $(M, \tau, \ell, \mathfrak{m})$ be a mm spacetime and $A \subset M$ achronal and Borel. Then $\mathfrak{m}(A \setminus \text{Invis}(M)) = 0$.*

In particular, for any causal $f : M \rightarrow \bar{\mathbf{R}}$ we have $|df| = +\infty$ \mathfrak{m} -a.e. on A .

Proof. It suffices to prove that for any test plan π we have $\iint_0^1 1_A(\gamma_t) |\dot{\gamma}_t| dt d\pi(\gamma) = 0$. For $\gamma \in \text{LCC}([0, 1]; M)$ and $t, s \in \gamma^{-1}(A)$ with $t \leq s$ the achronality of A forces $\ell(\gamma_t, \gamma_s) = 0$. Then the causality of γ implies that the same holds for $t \leq s$ in the convex hull $I \subset [0, 1]$ of $\gamma^{-1}(A)$. It follows that $|\dot{\gamma}_t| = 0$ for a.e. $t \in I$ and in particular for a.e. $t \in \gamma^{-1}(A)$. Hence $1_A(\gamma_t) |\dot{\gamma}_t| = 0$ for a.e. $t \in [0, 1]$ and the conclusion follows. \square

Example 3.20. In the one point space the whole space is invisible, because any curve must be constant and thus have 0 causal speed. \blacksquare

With this said, we can turn to the study of calculus rules. We start with the simple

$$|df| = 0 \quad \mathfrak{m} - \text{a.e. on } \text{Vis}(M) \quad \text{if } f \text{ is constant,} \quad (3.19)$$

which can be shown noticing that for any weak subslope G and any test plan π we must have $0 \geq \iint_0^1 G(\gamma_t) |\dot{\gamma}_t| dt d\pi(\gamma) = \int G \rho_\pi d\mathfrak{m}$. We also have the following concavity properties, of which (i) is essential and (ii), though not required subsequently, offers a concrete perspective which some readers may find helpful:

Proposition 3.21 (Concavity and positive 1-homogeneity). *Let $(M, \tau, \ell, \mathbf{m})$ be a mm spacetime, $f, g: \cdot M \rightarrow \bar{\mathbf{R}}$ causal. (i) Then all $\alpha, \beta \geq 0$ satisfy*

$$|d(\alpha f + \beta g)| \geq \alpha |df| + \beta |dg| \quad \mathbf{m}\text{-a.e. on } M \quad (3.20)$$

and

$$|d(\alpha f)| = \alpha |df| \quad \mathbf{m}\text{-a.e. on } \text{Vis}(M) \text{ if } \alpha = 0 \text{ and } \mathbf{m}\text{-a.e. on } M \text{ if } \alpha > 0. \quad (3.21)$$

(ii) *A concave positively 1-homogeneous function $h_x: [0, \infty)^2 \rightarrow [0, \infty]$ can be associated to each $x \in M$ so that: all $(0, 0) \neq (\alpha, \beta) \in [0, \infty)^2$ satisfy $|d(\alpha f + \beta g)|(y) = h_y(\alpha, \beta)$ for \mathbf{m} -a.e. $y \in M$.*

Proof. (i) Evidently, $\alpha |df| + \beta |dg|$ is a weak subslope of the causal function $\alpha f + \beta g$.

The first claim (3.20) follows from Definition 3.15. The second claim is obvious if $\alpha = 0$ (here having restricted to attention to $\text{Vis}(M)$ matters), while for $\alpha > 0$ we use twice the first claim with $g = 0$ to get $|df| \geq \alpha |d(\alpha^{-1}f)| \geq |df|$, which is (3.21).

(ii) For $0 \leq \alpha \leq \alpha'$ and $0 \leq \beta \leq \beta'$, since $|d(\alpha f + \beta g)|$ is a weak subslope for $\alpha' f + \beta' g$ we find (*) $|d(\alpha f + \beta g)| \leq |d(\alpha' f + \beta' g)|$ holds \mathbf{m} -a.e. For rational coefficients and all $\mathbf{Q}_+ := \mathbf{Q} \cap [0, \infty)$ combinations of f and g , the \mathbf{m} -negligible set $M \setminus X$ of (3.20)–(3.21) can be taken independent of $\alpha, \beta \in \mathbf{Q}_+$ and of the combinations of f and g . Taking X smaller yet $\mathbf{m}(X) = 1$, (*) yields $|d(\alpha f + \beta g)| \leq |d(\alpha' f + \beta' g)|$ throughout X for α', β' also rational with $\alpha \leq \alpha'$ and $\beta \leq \beta'$. Fix $x \in X$ and assume $|d(f + g)(x)| < \infty$; otherwise a similar but simpler argument will yield $|d(\alpha f + \beta g)|(x) = \infty$ for all real positive $\alpha, \beta > 0$, (and positive 1-homogeneity for $\alpha, \beta \geq 0$ real such that $\alpha\beta = 0$ but $\alpha + \beta > 0$). For $\alpha, \beta \in [0, \gamma] \cap \mathbf{Q}_+$ and γ rational,

$$|d(\alpha f + \beta g)|(x) \leq |d(\gamma f + \gamma g)|(x) = \gamma |d(f + g)|(x) < \infty.$$

Fixing $\alpha \in \mathbf{Q}_+$, it follows that $\beta \in \mathbf{Q}_+ \mapsto |d(\alpha f + \beta g)|(x)$ is concave and locally bounded — hence locally Lipschitz except possibly at $\beta = 0$. Fixing $\beta \in \mathbf{Q}_+$ instead, the same conclusions apply to $\alpha \in \mathbf{Q}_+ \mapsto |d(\alpha f + \beta g)|(x)$. Thus $h_x(\alpha, \beta) := |d(\alpha f + \beta g)|(x)$ extends from \mathbf{Q}_+^2 to a concave positively 1-homogeneous function on $[0, \infty)^2$ (which is locally Lipschitz on the interior of this quadrant and on its boundary, but may be merely lower semicontinuous on their union). Approximating real $\alpha \in (\alpha'', \alpha')$ and $\beta \in (\beta'', \beta')$ from above and below by rational $\alpha'', \alpha', \beta'', \beta'$, the aforementioned continuity of $h_x(\cdot, \cdot)$ combines with (*) to yield $|d(\alpha f + \beta g)|(y) = h_y(\alpha, \beta)$ for \mathbf{m} -a.e. $y \in M$ unless $\alpha = 0 = \beta$. \square

We turn to locality properties of the maximal weak subslope. They will closely interact with the chain rule and the Leibniz rule (for metric measure spaces these three properties of the minimal weak upper gradient [9] are equivalent to each other [12, 61]).

We recall how calculus with BV functions on \mathbf{R} works. Thus letting $I \subset \mathbf{R}$ be an interval, recall that $u \in BV_{loc}(I)$ iff it is in $L^1_{loc}(I)$ and has distributional derivative represented by a (Radon) measure Du . We write $Du = u' \mathcal{L}^1 \llcorner I + Du^\perp$ with $Du^\perp \perp \mathcal{L}^1$ for the Lebesgue decomposition of Du . Then:

- BV1) u is \mathcal{L}^1 -a.e. equal to the difference of two monotone functions (see e.g. [7, Sect. 3.12] and references therein).
- BV2) If u coincides with the representative given above, then it is \mathcal{L}^1 -a.e. differentiable and the derivative a.e. coincides with u' (see [7, Thm. 3.107]).
- BV3) For $N \subset \mathbf{R}$ Borel and negligible we have $u' = 0$ \mathcal{L}^1 -a.e. on $u^{-1}(N)$ (see [7, Prop. 3.92]).
- BV4) For $\varphi: \mathbf{R} \rightarrow \mathbf{R}$ Lipschitz we have $\varphi \circ u \in BV_{loc}(I)$ and $(\varphi \circ u)' = \varphi' \circ u u'$ \mathcal{L}^1 -a.e. (this is the absolutely continuous part of the Vol'pert chain rule, see e.g. [7, Thm. 3.99]). For φ non-decreasing but not necessarily Lipschitz the same conclusions can be verified based on the a.e. differentiability of monotone functions and BV2)–BV4).

Lemma 3.13 and the locality in item BV3) easily yields the following:

Proposition 3.22 (Locality). *Let $(M, \tau, \ell, \mathfrak{m})$ be a mm spacetime, $f_1, f_2: M \rightarrow \bar{\mathbf{R}}$ be causal functions and $E \subset \mathbf{R}$ Borel and \mathcal{L}^1 -negligible. Then*

$$|df_1| = |df_2| \quad \mathfrak{m}\text{-a.e. on } (f_1 - f_2)^{-1}(E). \quad (3.22)$$

In particular, we have

$$|df_1| = |df_2| \quad \mathfrak{m}\text{-a.e. on } \{f_1 = f_2\} \quad (3.23)$$

and

$$|df_1| = 0 \quad \mathfrak{m}\text{-a.e. on } f_1^{-1}(E) \cap \text{Vis}(M). \quad (3.24)$$

Proof. Put $F := (f_1 - f_2)^{-1}(E)$. We shall prove that $|df_1|1_F + |df_2|1_{M \setminus F} \leq |df_2|$. Reversing the roles of f_1, f_2 this suffices to obtain (3.22). Notice that in order for $f_1(x) - f_2(x)$ to belong to $E \subset \mathbf{R}$, we must have $x \in \text{Dom}(f_1) \cap \text{Dom}(f_2)$, thus $F \subset \text{Dom}(f_1) \cap \text{Dom}(f_2)$.

Now let $\gamma \in \text{LCC}([0, 1]; M)$ and notice that on the interval $I_\gamma := \gamma^{-1}(\text{Dom}(f_1) \cap \text{Dom}(f_2))$ the function $t \mapsto f_1(\gamma_t) - f_2(\gamma_t)$ is in BV_{loc} and thus, by item **BV3**) and with the same notation we see that

$$(f_1 \circ \gamma)'_t = (f_2 \circ \gamma)'_t \quad \text{for } \mathcal{L}^1\text{-a.e. } t \in \gamma^{-1}(F). \quad (3.25)$$

Now let π be a test plan. Lemma 3.13 implies that for π -a.e. γ we have

$$(f_2 \circ \gamma)'_t \geq |df_2|(\gamma_t)|\dot{\gamma}_t| \quad \text{for } \mathcal{L}^1\text{-a.e. } t \in \gamma^{-1}(\text{Dom}(f_2)) \supset (\gamma^{-1}(\text{Dom}(f_2)) \setminus \gamma^{-1}(F))$$

and taking into account also (3.25) we see that it also holds that

$$(f_2 \circ \gamma)'_t = (f_1 \circ \gamma)'_t \geq |df_1|(\gamma_t)|\dot{\gamma}_t| \quad \text{for } \mathcal{L}^1\text{-a.e. } t \in \gamma^{-1}(F).$$

These last two bounds imply that

$$(f_2 \circ \gamma)'_t \geq (|df_1|1_F + |df_2|1_{M \setminus F})(\gamma_t)|\dot{\gamma}_t| \quad \text{for } \mathcal{L}^1\text{-a.e. } t \in \gamma^{-1}(\text{Dom}(f_2))$$

holds for π -a.e. γ , so Lemma 3.13 yields that $|df_1|1_F + |df_2|1_{M \setminus F}$ is a weak subslope of f_2 , as desired. Now (3.23) follows by picking $E := \{0\}$ in (3.22) while (3.24) by taking $f_2 \equiv 0$ in (3.22) and recalling (3.19). \square

Similarly, Lemma 3.13 and the chain rule in item **BV4**) easily imply:

Proposition 3.23 (Chain rule). *Let $(M, \tau, \ell, \mathfrak{m})$ be a mm spacetime, $f: M \rightarrow \bar{\mathbf{R}}$ causal and $\varphi: \mathbf{R} \rightarrow \mathbf{R}$ non-decreasing. Then $\varphi \circ f$ is causal and*

$$|d(\varphi \circ f)| = (\varphi' \circ f) |df| \quad \mathfrak{m}\text{-a.e. on } \text{Vis}(M) \cap (\{|df| > 0\} \cup \{\varphi' \circ f < +\infty\}), \quad (3.26)$$

where φ' is the density of the distributional derivative $D\varphi$ of φ on its domain and it is intended that $\varphi(\pm\infty) = \pm\infty$ and $\varphi'(\pm\infty) = +\infty$.

Proof. The fact that $\varphi \circ f$ is causal is trivial. Then notice that even though φ' is defined up to equality \mathcal{L}^1 -a.e., an application of (3.24) shows that identity (3.26) is well-posed, i.e. its validity is unaffected by the particular representative of φ' chosen. Also, both sides of (3.26) are \mathfrak{m} -a.e. equal to $+\infty$ on $(\varphi \circ f)^{-1}(\{\pm\infty\}) \cap \{|df| > 0\} \supset f^{-1}(\{\pm\infty\})$, thus by locality (3.22) and truncation we can assume that $\varphi: \mathbf{R} \rightarrow \mathbf{R}$ is bounded.

Now let $\gamma \in \text{LCC}([0, 1]; M)$ and use item **BV4**) and the notation therein to deduce that

$$(\varphi \circ f \circ \gamma)'_t = \varphi'(f(\gamma_t))(f \circ \gamma)'_t \quad \text{for a.e. } t \in \text{Dom}(f \circ \gamma). \quad (3.27)$$

By Lemma 3.13 this suffices both to establish that \geq holds in (3.26) (because $(f \circ \gamma)'_t \geq |df|(\gamma_t)|\dot{\gamma}_t|$ for a.e. t and π -a.e. γ for π test) and that equality holds on $f^{-1}(\{\varphi' = 0\})$. It thus remains to prove that \leq holds on $f^{-1}(\{\varphi' > 0\})$. For this we use again Lemma 3.13 to get that $(\varphi \circ f \circ \gamma)'_t \geq |d(\varphi \circ f)|(\gamma_t)|\dot{\gamma}_t|$ holds for a.e. t and π -a.e. γ for any test π , then the identity (3.27) yields $(f \circ \gamma)'_t \geq \frac{|d(\varphi \circ f)|(\gamma_t)}{\varphi'(f(\gamma_t))} |\dot{\gamma}_t|$ for a.e. t and π -a.e. γ , where we interpret the ratio as 0 if the denominator is 0. Then Lemma 3.13 once more allows to conclude. \square

Proposition 3.24 (Leibniz rule). *Let $(M, \tau, \ell, \mathfrak{m})$ be a mm spacetime and $f, g : M \rightarrow \bar{\mathbf{R}}$ non-negative and causal. Then fg is causal and*

$$|d(fg)| \geq f |dg| + g |df| \quad \mathfrak{m}\text{-a.e. on } \text{Vis}(M). \quad (3.28)$$

Proof. The fact that fg is causal is obvious. Also, \mathfrak{m} -a.e. on the set $\text{Vis}(M) \cap \{f = +\infty\} \cap \{g > 0\}$ both sides of (3.28) are equal to $+\infty$, whereas \mathfrak{m} -a.e. on $\text{Vis}(M) \cap \{g = 0\}$ both sides are 0. Thus, also switching the roles of f, g , we are left to prove the conclusion on $\text{Vis}(M) \cap \{f \in (0, +\infty)\} \cap \{g \in (0, +\infty)\}$. Replacing f with $c \vee f \wedge c^{-1}$ for $c > 0$ arbitrarily small and similarly for g and using Proposition 3.22, we can — and will — assume that f, g are $(0, +\infty)$ -valued.

In this case Proposition 3.23 yields

$$\frac{1}{fg} |d(fg)| = |d(\log(fg))| = |d(\log(f) + \log(g))| \stackrel{(3.20)}{\geq} |d(\log(f))| + |d(\log(g))| = \frac{1}{f} |df| + \frac{1}{g} |dg|,$$

which is the conclusion. \square

4 Horizontal (inner) and vertical (outer) derivatives

4.1 Horizontal differentiation

Here we want to propose a notion of “derivative” of an arbitrary causal function in the “direction” of LCC curves. Intuitively, we shall refer to the horizontal derivative of a causal function f along an initial test plan π as at the limit

$$\lim_{t \rightarrow 0} \int \frac{f(\gamma_t) - f(\gamma_0)}{t} d\pi(\gamma)$$

and we are interested in knowing when such a limit exists and/or how we can estimate it. A bound from below can be quite easily given, as we are going to show in a moment. To motivate the result, notice that in the smooth category, the reverse Cauchy-Schwarz and the Fenchel–Young inequalities (for $p, q < 1$) give that

$$df(v) \geq \|df\|_* \|v\| \geq \frac{1}{p} \|df\|_*^p + \frac{1}{q} \|v\|^q$$

for any causal function f and future directed vector v (here and below for clarity we shall often write $\frac{1}{p} z^p$ rather than the more precise $u_p(z)$ as in Section 2.1). We are going to prove a nonsmooth analogue of this. In the sequel, we say that a measure $\pi \in \mathcal{P}(\text{LCC}([0, 1]; M))$ *stays initially in a given Borel set $E \subset M$* if there is $T \in (0, 1]$ such that $(\mathbf{e}_t)_* \pi$ is concentrated on E for every $t \in [0, T]$.

Proposition 4.1 (Nonsmooth Fenchel–Young inequality). *Let $(M, \tau, \ell, \mathfrak{m})$ be a forward mm spacetime, $E \subset M$ be a Borel set. Let π be an initial test plan which stays initially in E . Let $p^{-1} + q^{-1} = 1$ with $0 \neq q < 1$. Let $f : M \rightarrow \bar{\mathbf{R}}$ be a causal function and assume that*

- (i) *if $p < 0$ we have $|df|^p \in L^1(E, \mathfrak{m} \llcorner E)$ (note: this implies that $|df|^p \in L^1(E, (\mathbf{e}_0)_* \pi)$ since π is an initial test plan, but the converse implication does not hold),*
- (ii) *if $q < 0$ we have $\liminf_{t \rightarrow 0} \frac{1}{tq} \int \int_0^t |\dot{\gamma}_r|^q dr d\pi(\gamma) > -\infty$.*

Then

$$\liminf_{t \downarrow 0} \int \frac{f(\gamma_t) - f(\gamma_0)}{t} d\pi(\gamma) \geq \frac{1}{p} \int |df|^p(\gamma_0) d\pi(\gamma) + \liminf_{t \downarrow 0} \frac{1}{tq} \int \int_0^t |\dot{\gamma}_r|^q dr d\pi(\gamma). \quad (4.1)$$

Proof. Let $T > 0$ be such that $(\text{restr}_0^t)_* \pi$ is a test plan $t \in (0, T)$. Applying the Definitions 3.11 and 3.15 of (maximal) weak subslope as provided by Theorem 3.14 to the restricted test plan $(\text{restr}_0^t)_* \pi$ we obtain

$$\int f(\gamma_t) - f(\gamma_0) d\pi(\gamma) \geq \int \int_0^t |df|(\gamma_r) |\dot{\gamma}_r| dr d\pi(\gamma)$$

$$\begin{aligned}
&\geq \iint_0^t \frac{1}{p} |df|^p(\gamma_r) + \frac{1}{q} |\dot{\gamma}_r|^q dr d\pi(\gamma) \\
&= \frac{1}{p} \iint_0^t |df|^p(\gamma_r) dr d\pi(\gamma) + \frac{1}{q} \iint_0^t |\dot{\gamma}_r|^q dr d\pi(\gamma) \quad \text{for } t \text{ small enough,}
\end{aligned}$$

where the scalar Fenchel–Young inequality has been used in the second inequality. We comment on the last equality. Since the two integrands both have a sign, and thus their integrals are well defined — possibly with values $\pm\infty$ — equality holds provided at least one of the two integrands is actually integrable. This follows from our assumptions:

- if $p < 0$, then (i) and the fact that π is an initial test plan that stays initially in E ensures that for $t > 0$ sufficiently small the first integral is finite (see also formula (4.3) below),
- if $q < 0$, then (ii) ensures that for $t > 0$ sufficiently small the second integral is $> -\infty$.

To conclude it then suffices to prove

$$\liminf_{t \rightarrow 0} \frac{1}{pt} \iint_0^t |df|^p(\gamma_r) dr d\pi(\gamma) \geq \frac{1}{p} \int |df|^p(\gamma_0) d\pi(\gamma). \quad (4.2)$$

We start with the case $p < 0$, so that $|df|^p \in L^1(E, \mathbf{m} \llcorner E)$ by assumption (i). In this case, we claim in fact that equality holds in (4.2). Let ρ_r be the density of $(\mathbf{e}_r)_* \pi$ with respect to \mathbf{m} , where $r \in [0, T]$. By Fubini’s theorem we have

$$\frac{1}{t} \iint_0^t |df|^p(\gamma_r) dr d\pi(\gamma) = \int_E |df|^p \left[\frac{1}{t} \int_0^t \rho_r dr \right] d\mathbf{m}. \quad (4.3)$$

Since $(\text{restr}_0^T)_* \pi$ is a test plan, we have $\sup_{r \in [0, T]} \|\rho_r\|_{L^\infty(\mathbf{m})} < +\infty$. Since π is an initial test plan, we have the narrow convergence $(\mathbf{e}_r)_* \pi \rightarrow (\mathbf{e}_0)_* \pi$ as $r \rightarrow 0$, which together with the uniform L^∞ bound yields $\rho_r \rightarrow \rho_0$ as $r \rightarrow 0$ in duality with $L^1(\mathbf{M}, \mathbf{m})$. Since we assumed $|df|^p \in L^1(E, \mathbf{m} \llcorner E)$, this suffices to prove that equality holds — with a full lim rather than a lim inf — in (4.2).

In the case $p > 0$ we instead argue as follows. Since \mathbf{m} is a Radon measure and the topology on \mathbf{M} is Polish (and in particular has the Lindelöf property), \mathbf{m} is σ -finite. Considering an increasing sequence of sets with finite measure covering \mathbf{M} we can thus find an increasing sequence (g_n) of $L^1(\mathbf{M}, \mathbf{m})$ functions such that $g_n(x) \uparrow +\infty$ for every $x \in \mathbf{M}$. Thus $(\frac{1}{p}|df|^p) \wedge g_n$ is in $L^1(\mathbf{M}, \mathbf{m})$ for every $n \in \mathbf{N}$ and the same arguments used above give

$$\liminf_{t \rightarrow 0} \frac{1}{t} \iint_0^t \frac{|df|^p}{p}(\gamma_r) dr d\pi(\gamma) \geq \liminf_{t \rightarrow 0} \frac{1}{t} \iint_0^t g_n \wedge \left(\frac{|df|^p}{p} \right)(\gamma_r) dr d\pi(\gamma) = \int g_n \wedge \left(\frac{|df|^p}{p} \right) \rho_0 d\mathbf{m}.$$

The conclusion follows letting $n \uparrow +\infty$ in the above and using Levi’s monotone convergence theorem noticing that $g_n \wedge \left(\frac{|df|^p}{p} \right) \uparrow \frac{|df|^p}{p}$. \square

Inspired by a similar definition for metric measure spaces [60], which is itself motivated by De Giorgi’s approach to gradient flows in metric spaces (cf. e.g. [9]), the following definition is naturally based on requiring the saturation of the inequality from Proposition 4.1. In the smooth setting, the p -gradient of f refers to the Hamiltonian gradient of f given by the nonlinear identification of the cotangent and tangent spaces induced by the Hamiltonian

$$H(\omega) := \begin{cases} \frac{1}{p} \|\omega\|_*^p, & \text{if } \omega \text{ is future directed,} \\ +\infty, & \text{otherwise.} \end{cases}$$

Definition 4.2 (Representing the initial p -gradient of a causal function). *Let $(\mathbf{M}, \tau, \ell, \mathbf{m})$ be a mm spacetime, $f: \mathbf{M} \rightarrow \bar{\mathbf{R}}$ be a causal function and $p^{-1} + q^{-1} = 1$ with $0 \neq p, q < 1$. We say that an initial test plan π represents the initial p -gradient of f if:*

- i) *There is a Borel set $E \subset \mathbf{M}$ so that π stays initially in E ;*

ii) for E as above we also have $|df|^p \in L^1(E, \mathbf{m} \llcorner E)$;

iii) $\limsup_{t \downarrow 0} \frac{1}{t} \int \int_0^t |\dot{\gamma}_r|^q dr d\pi(\gamma) < +\infty$; and

iv)

$$\limsup_{t \downarrow 0} \int \frac{f(\gamma_t) - f(\gamma_0)}{t} d\pi(\gamma) \leq \frac{1}{p} \int |df|^p(\gamma_0) d\pi(\gamma) + \liminf_{t \downarrow 0} \frac{1}{tq} \int \int_0^t |\dot{\gamma}_r|^q dr d\pi(\gamma). \quad (4.4)$$

In this case we also say that π represents the initial p -gradient of f on E .

We collect some comments about the above:

1. We are asking (ii), (iii) irrespective of the signs of p and q
2. A consequence of this definition and of Proposition 4.1 is that $f \circ e_t - f \circ e_0 \in L^1(\pi)$ for $t > 0$ sufficiently small.
3. If $p < 0$, the assumptions above — in particular $|df|^p \in L^1(E)$ — do not exclude the possibility that $|df| = +\infty$ on a subset of E of positive measure. For instance, in the one point space any function is causal with $|df| = +\infty$ and the (only) test plan is concentrated in the (only) constant curve and represents the initial p -gradient of f .
4. If the initial p -gradient of a given causal function f is represented by an initial test plan π on E , then

$$\lim_{t \rightarrow 0} \int \frac{f(\gamma_t) - f(\gamma_0)}{t} d\pi(\gamma) = \frac{1}{p} \int |df|^p(\gamma_0) d\pi(\gamma) + \liminf_{t \rightarrow 0} \frac{1}{tq} \int \int_0^t |\dot{\gamma}_r|^q dr d\pi(\gamma),$$

as follows by combining (4.1) with (4.4). In this case, imitating the proof of [60, Prop. 3.11] it is not hard to see that the limit $\lim_{t \rightarrow 0} \frac{1}{t} \int \int_0^t |\dot{\gamma}_r|^q dr d\pi(\gamma)$ exists in (4.4).

5. From Proposition 4.1 we see that if π represents the initial p -gradient of f in E and $\Gamma \subset \text{LCC}([0, 1]; M)$ is Borel with $\pi(\Gamma) > 0$, then the “restricted” plan $\pi_\Gamma := \pi(\Gamma)^{-1} \pi \llcorner \Gamma$ also represents the initial p -gradient of f in E .

4.2 Perturbation of causal functions

This section serves as preparation for the subsequent one, where we will start from a causal function f , perturb it as $f + \varepsilon g$ for some g and be interested in those cases for which the new function is still \mathbf{m} -measurable and causal. This motivates the next definition, which is the main object of study here:

Definition 4.3 (Perturbation of causal functions). *Let $(M, \tau, \ell, \mathbf{m})$ be a mm spacetime and $f : M \rightarrow \bar{\mathbf{R}}$ a causal function. We define the class $\text{Pert}(f)$ of perturbations of f as the collection of all \mathbf{m} -measurable functions $g : M \rightarrow \bar{\mathbf{R}}$ such that $f + \varepsilon g$ is a causal function for some $\varepsilon > 0$.*

We think of $\text{Pert}(f)$ as the tangent space at f to the space of causal functions. Notice that according to our infinity conventions in Section 2.1, we have $\pm\infty + \varepsilon g(x) = \pm\infty$ for any $g : M \rightarrow \bar{\mathbf{R}}$ and $x \in M$. Thus the role of g is only seen in the causally convex set $\text{Dom}(f)$.

The properties of $\text{Pert}(f)$ are better understood if we introduce the following order relation on (not necessarily causal) functions on M . By slight abuse of notation, we denote it by \preceq just like the causal relation between probability measures, cf. Section 2.7, and hope no confusion ensues.

Definition 4.4 (A pre-order on the space of functions via finite differences). *Let (M, ℓ) be a metric spacetime. For $g : M \rightarrow \bar{\mathbf{R}}$ define $\delta g : M \times M \rightarrow \bar{\mathbf{R}}$ as*

$$\delta g(x, y) := g(y) - g(x).$$

We say that g_2 is at least as steep as g_1 and write $g_1 \preceq g_2$ provided

$$\delta g_1(x, y) \leq \delta g_2(x, y) \quad \text{whenever } x \leq y.$$

An analogous definition will also be used in the case $M = \bar{\mathbf{R}}$. The following are direct consequences of the definition.

- A function $f: M \rightarrow \bar{\mathbf{R}}$ is causal if and only if $\delta f(x, y) \geq 0$ whenever $x \leq y$ and if and only if $c \preceq f$, where c is any constant, real-valued function.
- Given an interval $I \subset \bar{\mathbf{R}}$ consider the function $f_I: \bar{\mathbf{R}} \rightarrow \bar{\mathbf{R}}$ defined as 0 on I , as $-\infty$ on the left of I and $+\infty$ on the right of I . Then $f_I \preceq f_J$ if and only if $I \supset J$.
- \preceq is a relation about “first order derivatives” and not about “pointwise values” of the functions. In particular, $g_1 \preceq g_2$ does not say anything about the sign of $g_1 - g_2$ outside $\pm\infty$ unless $g_1(x) = g_2(x) \in \mathbf{R}$ at some $x \in M$ (in which case $(g_2(y) - g_1(y))(y - x) \geq 0$ for all $y \in \mathbf{R}$). It can happen that $g_1 \preceq g_2$ yet $g_1(x) > g_2(x)$ for every $x \in M$. This is for instance the case for the interval $[0, 1]$ endowed with the usual order relation \leq and the causal functions $g_1(x) := x$ and $g_2(x) := 2x - 2$. Rather, $g_1 \preceq g_2$ should be thought of as a monotonicity condition on the *derivatives* of g_1 and g_2 , as illustrated in the next remark.

Remark 4.5 (Smooth differential characterization of \preceq). Assume (M, g) is a smooth spacetime. Given $x \in M$ we define an order \leq_x on the cotangent space T_x^*M by $\omega_1 \leq_x \omega_2$ if $\omega_1(v) \leq \omega_2(v)$ for every future-directed causal vector $v \in T_xM$ (see also Appendix A.1). Then for $f_1, f_2 \in C^1(M)$, we have $f_1 \preceq f_2$ if and only if their differentials satisfy $df_1(x) \leq_x df_2(x)$ for every $x \in M$.

For the “only if” direction, fix $x \in M$ and a future-directed causal vector $v \in T_xM$. Since $x \leq \exp_x(\varepsilon v)$ for every $\varepsilon > 0$, we have

$$df_1(x)(v) = \lim_{\varepsilon \downarrow 0} \frac{f_1(\exp_x(\varepsilon v)) - f_1(x)}{\varepsilon} \leq \lim_{\varepsilon \downarrow 0} \frac{f_2(\exp_x(\varepsilon v)) - f_2(x)}{\varepsilon} = df_2(x)(v).$$

For the “if” direction, let $x \leq y$ and γ be a smooth causal curve from x to y . Then

$$f_1(y) - f_1(x) = \int_0^1 df_1(\gamma_t)(\gamma'_t) dt \leq \int_0^1 df_2(\gamma_t)(\gamma'_t) dt = f_2(y) - f_2(x).$$

Notice that we shall be able to somehow formulate the implication “ $f_1 \preceq f_2$ implies $df_1 \leq df_2$ a.e.” also in the non-smooth setting: see the monotonicity stipulated in Proposition 4.11(v) below and the second part of Lemma 4.6.

To motivate the non-smooth statement Lemma 4.6 below, we point out that the smooth duality formula $\|\omega\|_* = \inf_{|v|=1} \omega(v)$ valid for all $\omega \geq_x 0$ easily yields the implication

$$\min\{df_1(v), df_2(v)\} \leq df_3(v) \quad \forall v \text{ future directed} \quad \Rightarrow \quad \min\{\|df_1\|_*, \|df_2\|_*\} \leq \|df_3\|_* \quad (4.5)$$

for every f_1, f_2, f_3 causal functions. On the other hand, it is easy to construct causal functions f_1, f_2, f_3 such that $\max\{df_1(v), df_2(v)\} \geq df_3(v)$ for every future directed v , but for which $\max\{\|df_1\|_*, \|df_2\|_*\} \not\leq \|df_3\|_*$ (we can take suitable linear functions in flat Minkowski space). ■

The rest of this part is devoted to basic properties of \preceq that will become relevant in the subsequent discussion. We start with the nonsmooth analog of implication (4.5) above:

Lemma 4.6 (Finite difference and maximal weak subslope comparison among causal functions). *Let $(M, \tau, \ell, \mathbf{m})$ be a mm spacetime, $f_1, f_2, f_3: M \rightarrow \bar{\mathbf{R}}$ causal and such that $\min\{\delta f_1(x, y), \delta f_2(x, y)\} \leq \delta f_3(x, y)$ holds for any $x \leq y$.*

Then $\min\{|df_1|, |df_2|\} \leq |df_3|$ m-a.e. In particular, if f, \tilde{f} are causal functions with $f \preceq \tilde{f}$, then $|df| \leq |d\tilde{f}|$ m-a.e.

Proof. Notice that for γ causal, the set $\text{Dom}(f_3 \circ \gamma)$ is a subinterval of $[0, 1]$ and if it is not a point (if it is, the conclusion is trivial) our assumption easily implies $\text{Dom}(f_3 \circ \gamma) \subset \text{Dom}(f_1 \circ \gamma) \cup \text{Dom}(f_2 \circ \gamma)$.

Now recall that Lebesgue’s differentiation theorem of monotone functions on \mathbf{R} implies that if both the curve γ and the function f are causal, then the density $(f \circ \gamma)'$ of the distributional

derivative of $f \circ \gamma$ coincides \mathcal{L}^1 -a.e. on $\text{Dom}(f)$ with the classically derivative $\lim_{h \downarrow 0} \frac{f(\gamma_{t+h}) - f(\gamma_t)}{h}$ of $f \circ \gamma$.

Therefore our assumption gives $(f_3 \circ \gamma)'_t \geq \min\{(f_1 \circ \gamma)'_t, (f_2 \circ \gamma)'_t\}$ for a.e. $t \in [0, 1]$, where we interpret $(f_i \circ \gamma)'_t$ to be $+\infty$ if $t \notin \text{Dom}(f_i \circ \gamma)$. In particular, this holds for π -a.e. γ for any given test plan π , so that the conclusion follows from Lemma 3.13.

The second claim now follows by picking $f_1 \equiv 0$ in the first part. \square

Lemma 4.7 (Ordering respects addition of causal functions). *Let $(M, \tau, \ell, \mathfrak{m})$ be a mm spacetime, $f: M \rightarrow \bar{\mathbf{R}}$ causal function and $g_1, g_2: M \rightarrow \bar{\mathbf{R}}$ are arbitrary functions with $g_1 \preceq g_2$, then $f + g_1 \preceq f + g_2$.*

Proof. The proof is a simple case analysis. Let x and y be fixed points with $x \leq y$. We need to prove that

$$\delta(f + g_1)(x, y) \leq \delta(f + g_2)(x, y). \quad (4.6)$$

If $f(y) = +\infty$ then $(f + g_i)(y) = +\infty$ by our infinity conventions and thus $\delta(f + g_i)(x, y) = +\infty$ by (2.1), thus in this case (4.6) holds. Similarly, if $f(x) = -\infty$ we have $(f + g_i)(x) = -\infty$ and then $\delta(f + g_i)(x, y) = +\infty$, thus also in this case (4.6) holds.

Since f is causal it remains to check the case $f(x), f(y) \in \mathbf{R}$, but in this case it is easily verified that — according to (2.1) — we have $\delta(f + g_i)(x, y) = \delta f(x, y) + \delta g_i(x, y)$, so the conclusion follows from the assumption $g_1 \preceq g_2$. \square

We also have:

Proposition 4.8 (Properties of $\text{Pert}(f)$). *Let $(M, \tau, \ell, \mathfrak{m})$ be a mm spacetime, and $f: M \rightarrow \bar{\mathbf{R}}$ causal. Then:*

- (i) Cone. *We have $\lambda g \in \text{Pert}(f)$ for every $g \in \text{Pert}(f)$ and every $\lambda \geq 0$.*
- (ii) Convex cone. *We have $g_1 + g_2 \in \text{Pert}(f)$ for every $g_1, g_2 \in \text{Pert}(f)$.*
- (iii) Monotonicity. *If $g_1 \in \text{Pert}(f)$ and $g_1 \preceq g_2$ then $g_2 \in \text{Pert}(f)$.*
- (iv) Lattice. *If $g_1, g_2 \in \text{Pert}(f)$ then $\min\{g_1, g_2\}, \max\{g_1, g_2\} \in \text{Pert}(f)$.*
- (v) Comparison. *Let $g: M \rightarrow \bar{\mathbf{R}}$ and $\varphi_1, \varphi_2, \psi: \bar{\mathbf{R}} \rightarrow \bar{\mathbf{R}}$ be functions satisfying $\varphi_1 \preceq \psi \preceq \varphi_2$ and $\varphi_1 \circ g, \varphi_2 \circ g \in \text{Pert}(f)$. (We do not necessarily assume $g \in \text{Pert}(f)$ here.) Then $\psi \circ g \in \text{Pert}(f)$ and*

$$\begin{aligned} \min\{\delta(f + \varepsilon \varphi_1 \circ g), \delta(f + \varepsilon \varphi_2 \circ g)\}(x, y) &\leq \delta(f + \varepsilon \psi \circ g)(x, y) \\ &\leq \max\{\delta(f + \varepsilon \varphi_1 \circ g), \delta(f + \varepsilon \varphi_2 \circ g)\}(x, y). \end{aligned}$$

whenever $x \leq y$ and $\varepsilon > 0$. In particular,

$$|\text{d}(f + \varepsilon \psi \circ g)| \geq \min\{|\text{d}(f + \varepsilon \varphi_1 \circ g)|, |\text{d}(f + \varepsilon \varphi_2 \circ g)|\} \quad \mathfrak{m}\text{-a.e.} \quad (4.7)$$

Proof.

(i) This is trivial.

(ii) This statement follows from the identity $f + \varepsilon(g_1 + g_2) = (f + 2\varepsilon g_1)/2 + (f + 2\varepsilon g_2)/2$ which is consistent with (2.1), observing that $f + 2\varepsilon g_1$ and $f + 2\varepsilon g_2$ are both causal functions for sufficiently small $\varepsilon > 0$ thanks to (i).

(iii) This is a consequence of Lemma 4.7 and the characterization of causal functions in terms of the finite difference operator δ mentioned after Definition 4.4.

(iv) This follows from the identity $f + \varepsilon \min\{g_1, g_2\} = \min\{f + \varepsilon g_1, f + \varepsilon g_2\}$ which is consistent with (2.1) and the fact that the minimum of causal functions is still a causal function. Analogously, we show $\max\{g_1, g_2\} \in \text{Pert}(f)$.

(v) The inequality (4.7) follows from the first statement and Lemma 4.6.

To show the first claim, we distinguish two cases.

- Assume $g(x) \leq g(y)$. Since $\varphi_1 \preceq \psi \preceq \varphi_2$ we obtain

$$\delta(\varphi_1 \circ g)(x, y) \leq \delta(\psi \circ g)(x, y) \leq \delta(\varphi_2 \circ g)(x, y).$$

By applying Lemma 4.7 to the two-point space $\{x, y\}$ we deduce for $\varepsilon > 0$

$$\delta(f + \varepsilon\varphi_1 \circ g)(x, y) \leq \delta(f + \varepsilon\psi \circ g)(x, y) \leq \delta(f + \varepsilon\varphi_2 \circ g)(x, y).$$

- Assume $g(x) \geq g(y)$. As above, this yields

$$\delta(\varphi_2 \circ g)(x, y) \leq \delta(\psi \circ g)(x, y) \leq \delta(\varphi_1 \circ g)(x, y),$$

and applying Lemma 4.7 to $\{x, y\}$ once more we deduce

$$\delta(f + \varepsilon\varphi_2 \circ g)(x, y) \leq \delta(f + \varepsilon\psi \circ g)(x, y) \leq \delta(f + \varepsilon\varphi_1 \circ g)(x, y). \quad \square$$

In view of the chain and Leibniz rules stated in Proposition 4.11 — used especially in the proof of Corollary 5.25 as well as in [22], we study the behavior of perturbations under composition and multiplication. The hypotheses are likely not optimal, but will suffice for our purposes.

In the statement below we shall always extend the given function $\varphi : \mathbf{R} \rightarrow \mathbf{R}$ by putting $\varphi(\pm\infty) = \pm\infty$.

Lemma 4.9 (Composition and multiplication). *Let $(M, \tau, \ell, \mathfrak{m})$ be a mm spacetime and $f : M \rightarrow \bar{\mathbf{R}}$ causal. Then:*

- (i) Composition I. *Let $\varphi : \mathbf{R} \rightarrow \mathbf{R}$ be Lipschitz continuous. Then $\varphi \circ f \in \text{Pert}(f)$.*
- (ii) Composition II. *Let $\varphi : \mathbf{R} \rightarrow \mathbf{R}$ be c -steep (i.e. $\varphi(w) - \varphi(z) \geq (w - z)c$ for all $z < w$) for some $c > 0$. Then $\text{Pert}(f) \subset \text{Pert}(\varphi \circ f)$.*
- (iii) Composition III. *Let $g \in \text{Pert}(f)$ and $\varphi : \mathbf{R} \rightarrow \mathbf{R}$ be non-decreasing and Lipschitz continuous. Then $\varphi \circ g \in \text{Pert}(f)$.*
- (iv) Multiplication I. *Let $g, h \in \text{Pert}(f)$ be non-negative and bounded. Then $gh \in \text{Pert}(f)$.*
- (v) Multiplication II. *Let $g \in \text{Pert}(f)$ be bounded and $\varphi : \mathbf{R} \rightarrow \mathbf{R}$ be non-negative, bounded and Lipschitz continuous. Then $g \varphi \circ f \in \text{Pert}(f)$.*

Proof.

(i) If φ is affine the statement is straightforward. For general Lipschitz continuous φ , we use the relation $-\text{Lip } \varphi \text{ id} \preceq \varphi \preceq \text{Lip } \varphi \text{ id}$ directly implied from Lipschitz continuity and Proposition 4.8(v).

(ii) Note that $\varphi \circ f$ is a causal function, let $g \in \text{Pert}(f)$ be such that $f + g$ is causal (any $g \in \text{Pert}(f)$ can be rescaled to satisfy this). We claim that $\varphi \circ f + cg$ is also causal, which clearly suffices to conclude. Pick $x \leq y$, with $x, y \in M$. We need to prove that $\varphi(f(y)) + cg(y) \geq \varphi(f(x)) + cg(x)$. If either $\varphi(f(y)) = +\infty$ or $\varphi(f(x)) = -\infty$ the claim holds by our infinity conventions. Since $\varphi(f(y)) \geq \varphi(f(x))$, we can then assume $\varphi(f(y)), \varphi(f(x)) \in \mathbf{R}$ and thus in particular $f(y), f(x) \in \mathbf{R}$. If this is the case we have $\varphi(f(y)) \geq \varphi(f(x)) + cf(y) - cf(x)$ hence

$$\varphi(f(y)) + cg(y) \geq \varphi(f(x)) + cf(y) - cf(x) + cg(y) \geq \varphi(f(x)) + cg(x),$$

by our normalization of $g \in \text{Pert}(f)$.

(iii) As in (i), the claim is clear if φ is linear. The general case follows from the relation $0 \preceq \varphi \preceq \text{Lip } \varphi \text{ id}$, the causal property of f , and Proposition 4.8(v).

(iv) Assume, by scaling, that $f + g, f + h$ are causal, let $x \leq y$ and $\varepsilon > 0$. As already noticed, to check that $(f + \varepsilon gh)(y) \geq (f + \varepsilon gh)(x)$ we can assume that $x, y \in \text{Dom}(f)$, otherwise the claim follows from our infinity conventions. Let C be an upper bound on g and h on all of M . Then

$$(f + \varepsilon gh)(y) - (f + \varepsilon gh)(x) = f(y) - f(x) + \varepsilon g(y) [h(y) - h(x)] + \varepsilon h(x) [g(y) - g(x)]$$

$$\begin{aligned} &\geq f(y) - f(x) + C\varepsilon \max\{0, \operatorname{sgn}(h(x) - h(y))\} [h(y) - h(x)] \\ &\quad + C\varepsilon \max\{0, \operatorname{sgn}(g(x) - g(y))\} [g(y) - g(x)], \end{aligned}$$

where the last inequality follows from a straightforward case distinction, using g and h are non-negative. It is now easy to check that for $\varepsilon := \frac{1}{2C}$ the right-hand side is non-negative, thus concluding the proof.

(v) It follows from (i) and (iv) that $(g - \inf g(M))\varphi \circ f \in \operatorname{Pert}(f)$. Moreover, again (i) implies $\inf g(M)\varphi \circ f \in \operatorname{Pert}(f)$. Since $\operatorname{Pert}(f)$ is closed under addition, the claim follows. \square

4.3 Vertical right-differentiation

Recall that we put $u_p(z) := \frac{1}{p}z^p$ and that u_p is suitably extended to all of $\bar{\mathbf{R}}$ in Section 2.1.

Definition 4.10 (Vertical right derivative). *Let $(M, \tau, \ell, \mathbf{m})$ be a mm spacetime, $f : M \rightarrow \bar{\mathbf{R}}$ causal and $g \in \operatorname{Pert}(f)$. We define the vertical right derivative $d^+g(\nabla f) |df|^{p-2} : M \rightarrow \bar{\mathbf{R}}$ as*

$$d^+g(\nabla f) |df|^{p-2} := \begin{cases} \lim_{\varepsilon \downarrow 0} \frac{u_p(|d(f + \varepsilon g)|) - u_p(|df|)}{\varepsilon} & \text{if } |df| < +\infty, \\ 0 & \text{otherwise.} \end{cases} \quad (4.8)$$

For the moment, the term $d^+g(\nabla f) |df|^{p-2}$ from Definition 4.10 should and will henceforth be considered as a single expression, not as the “product” of $|df|^{p-2}$ and (the not yet defined quantity) $d^+g(\nabla f)$. For comments about the definition on $\{|df| = +\infty\}$ see Remark 4.13.

Let us comment on the well-definedness of (4.8). For $\varepsilon > 0$ sufficiently small $f + \varepsilon g$ is causal, thus the difference quotient is well defined. The concavity of the maximal weak subslope $|df|$ in f (on the space of causal functions) from Proposition 3.21 and the fact that u_p is concave and non-decreasing easily yield

$$\frac{u_p(|d(f + \varepsilon_1 g)|) - u_p(|df|)}{\varepsilon_1} \geq \frac{u_p(|d(f + \varepsilon_2 g)|) - u_p(|df|)}{\varepsilon_2} \quad \mathbf{m}\text{-a.e.}$$

whenever $\varepsilon_2 > \varepsilon_1 > 0$ are sufficiently small. It follows that for every $\varepsilon_n \downarrow 0$ the functions $\frac{u_p(|d(f + \varepsilon_n g)|) - u_p(|df|)}{\varepsilon_n}$ have a pointwise \mathbf{m} -a.e. limit which is independent of the chosen sequence, and this limit $d^+g(\nabla f) |df|^{p-2}$ satisfies

$$d^+g(\nabla f) |df|^{p-2} = \mathbf{m}\text{-ess sup}_{\varepsilon > 0} \frac{u_p(|d(f + \varepsilon g)|) - u_p(|df|)}{\varepsilon} \quad \text{on } \{|df| < +\infty\}. \quad (4.9)$$

From this it is easy to see that if f, g are both causal, thus in particular $g \in \operatorname{Pert}(f)$, then the following sort of reverse Cauchy–Schwarz inequality is in place:

$$d^+g(\nabla f) |df|^{p-2} \geq |dg| |df|^{p-1} \quad \mathbf{m}\text{-a.e.}$$

Indeed, the concave homogeneity in (3.20) gives $|d(f + \varepsilon g)| \geq |df| + \varepsilon |dg|$ and since u_p is non-decreasing we have $u_p(|d(f + \varepsilon g)|) \geq u_p(|df| + \varepsilon |dg|)$. The conclusion follows from the above discussion and the fact that $u'_p(z) = z^{p-1}$. Similar considerations give

$$d^+f(\nabla f) |df|^{p-2} = |df|^p \quad \mathbf{m}\text{-a.e. on } \{|df| < +\infty\},$$

where here we need to restrict to the set $\{|df| < +\infty\}$ to avoid conflicts, in the case $p \in (0, 1)$, with our choice $d^+g(\nabla f) |df|^{p-2} = 0$ on $\{|df| = +\infty\}$, see also Remark 4.13.

We also point out that according to our infinity conventions in Section 2.1 we have

$$d^+g(\nabla f) |df|^{p-2} = +\infty, \quad \mathbf{m}\text{-a.e. on } \{|df| = 0\} \quad \text{if } p < 0 \quad (4.10)$$

and, taking also into account the concavity in Proposition 3.21,

$$d^+g(\nabla f)|df|^{p-2} = \begin{cases} +\infty, & \mathbf{m}\text{-a.e. on } \{|df| = 0\} \cap E(f, g), \\ 0, & \mathbf{m}\text{-a.e. on } \{|df| = 0\} \setminus E(f, g), \end{cases} \quad \text{if } p \in (0, 1), \quad (4.11)$$

where $E(f, g) := \cup_{\varepsilon > 0} \{|d(f + \varepsilon g)| > 0\} = \cup_{n \in \mathbf{N}} \{|d(f + \frac{1}{n}g)| > 0\}$.

For later use it is worth demonstrating also the following. If $f : M \rightarrow \bar{\mathbf{R}}$ is causal and $g \in \text{Pert}(f)$, say $f + \bar{\varepsilon}g$ is causal for $\bar{\varepsilon} > 0$, then we claim the uniform bound from below

$$\frac{u_p(|d(f + \varepsilon g)|) - u_p(|df|)}{\varepsilon} \geq \begin{cases} -\frac{1}{\bar{\varepsilon}} u_p(|df|) & \text{if } p \in (0, 1), \\ \frac{2^{1-p}}{\bar{\varepsilon}} u_p(|df|) & \text{if } p < 0, \end{cases} \quad \forall \varepsilon \in (0, \frac{1}{2}\bar{\varepsilon}). \quad (4.12)$$

Indeed, the bound for $p \in (0, 1)$ follows from the monotonicity in ε of the left hand side together with the positivity of u_p . For $p < 0$, the same monotonicity together with $u_p < 0$ give $\frac{u_p(|d(f + \varepsilon g)|) - u_p(|df|)}{\varepsilon} \geq \frac{u_p(|d(f + \bar{\varepsilon}g/2)|)}{\bar{\varepsilon}/2}$. Then, since $f + \bar{\varepsilon}g$ is causal, from (3.20) we get $|d(f + \bar{\varepsilon}g/2)| \geq \frac{1}{2}|df|$ and the claim follows from the monotonicity of u_p .

The following two propositions collect some basic calculus rules. Compare with those valid in positive signature obtained in [60, Section 3.3].

Proposition 4.11 (Calculus rules ‘for g ’). *Let $(M, \tau, \ell, \mathbf{m})$ be a mm spacetime and $f : M \rightarrow \bar{\mathbf{R}}$ causal. Then*

- (i) Positive 1-homogeneity. *For $g \in \text{Pert}(f)$ and every $\lambda \in (0, +\infty)$ we have*

$$d^+(\lambda g)(\nabla f)|df|^{p-2} = \lambda d^+g(\nabla f)|df|^{p-2} \quad \mathbf{m}\text{-a.e.} \quad (4.13)$$

For $\lambda = 0$ the same holds \mathbf{m} -a.e. on $\{|df| > 0\}$.

- (ii) Super-additivity. *For every $g_1, g_2 \in \text{Pert}(f)$ we have*

$$d^+(g_1 + g_2)(\nabla f)|df|^{p-2} \geq d^+g_1(\nabla f)|df|^{p-2} + d^+g_2(\nabla f)|df|^{p-2} \quad \mathbf{m}\text{-a.e.} \quad (4.14)$$

- (iii) Comparison from the left. *Let $g : M \rightarrow \bar{\mathbf{R}}$ be such that $\pm g \in \text{Pert}(f)$. Then*

$$d^+g(\nabla f)|df|^{p-2} \leq -d^+(-g)(\nabla f)|df|^{p-2} \quad \mathbf{m}\text{-a.e. on } \{|df| > 0\}. \quad (4.15)$$

- (iv) Locality. *Let $g_1, g_2 \in \text{Pert}(f)$ and $E \subset \mathbf{R}$ Borel and negligible. Then*

$$d^+g_1(\nabla f)|df|^{p-2} = d^+g_2(\nabla f)|df|^{p-2} \quad \mathbf{m}\text{-a.e. on } (g_1 - g_2)^{-1}(E). \quad (4.16)$$

- (v) Monotonicity. *If $g_1, g_2 \in \text{Pert}(f)$ satisfy $g_1 \leq g_2$, we have*

$$d^+g_1(\nabla f)|df|^{p-2} \leq d^+g_2(\nabla f)|df|^{p-2} \quad \mathbf{m}\text{-a.e.}$$

- (vi) Chain rule I. *Let $g \in \text{Pert}(f)$ and let $\varphi : \mathbf{R} \rightarrow \mathbf{R}$ be a non-decreasing and Lipschitz function. Then*

$$d^+(\varphi \circ g)(\nabla f)|df|^{p-2} = \varphi' \circ g d^+g(\nabla f)|df|^{p-2} \quad (4.17)$$

holds \mathbf{m} -a.e. on $\{\varphi' \circ g > 0\} \cup \{|df| > 0\}$. Note that $\varphi \circ g \in \text{Pert}(f)$ by Lemma 4.9.

- (vii) Chain rule II. *Let $\varphi : \mathbf{R} \rightarrow \mathbf{R}$ be Lipschitz continuous. Then*

$$d^+(\varphi \circ f)(\nabla f)|df|^{p-2} = \varphi' \circ f |df|^p \quad (4.18)$$

holds \mathbf{m} -a.e. on $\{|df| < +\infty\} \cap (\{\varphi' \circ f > 0\} \cup \{|df| > 0\})$. Note that $\varphi \circ f \in \text{Pert}(f)$ by Lemma 4.9.

(viii) Leibniz rule I. Let $g, h \in \text{Pert}(f)$ be non-negative and bounded. Then

$$d^+(gh)(\nabla f) |df|^{p-2} \geq g d^+h(\nabla f) |df|^{p-2} + h d^+g(\nabla f) |df|^{p-2} \quad \mathbf{m}\text{-a.e. on } \{|df| > 0\}. \quad (4.19)$$

Note that $gh \in \text{Pert}(f)$ by Lemma 4.9.

(ix) Leibniz rule II. Let $g \in \text{Pert}(f)$ be bounded and $\varphi: \mathbf{R} \rightarrow \mathbf{R}$ be non-negative, bounded and Lipschitz continuous. Then

$$d^+(g\varphi \circ f)(\nabla f) |df|^{p-2} \geq g \varphi' \circ f |df|^p + \varphi \circ f d^+g(\nabla f) |df|^{p-2} \quad \mathbf{m}\text{-a.e. on } \{|df| \in (0, +\infty)\}. \quad (4.20)$$

Note that $g\varphi \circ f \in \text{Pert}(f)$ by Lemma 4.9.

Note: the latter two Leibniz rules hold as equalities in the smooth case.

Proof.

(i) Follows directly from definition (4.8).

(ii) On $\{|df| = +\infty\}$ the conclusion is trivial, on $\{|df| = 0\}$ it follows from (4.10) and (4.11). On $\{|df| \in (0, +\infty)\}$ we argue as follows: for $\lambda \in (0, 1)$ the concavity in (3.20) gives

$$|d(f + \varepsilon((1 - \lambda)g_1 + \lambda g_2))| \geq (1 - \lambda) |d(f + \varepsilon g_1)| + \lambda |d(f + \varepsilon g_2)|,$$

thus using the concavity and non-decreasingness of u_p we get

$$u_p(|d(f + \varepsilon((1 - \lambda)g_1 + \lambda g_2))|) \geq (1 - \lambda) u_p(|d(f + \varepsilon g_1)|) + \lambda u_p(|d(f + \varepsilon g_2)|).$$

The claim follows taking also (i) into account.

(iii) On $\{|df| = +\infty\}$ both sides are 0 by definition, so we focus on $\{|df| \in (0, +\infty)\}$. For $\varepsilon > 0$ sufficiently small, the concavity in (3.20) gives $|df| \geq \frac{1}{2}|d(f + \varepsilon g)| + \frac{1}{2}|d(f - \varepsilon g)|$ thus using the concavity and non-decreasingness of u_p we get

$$u_p(|df|) \geq \frac{1}{2}u_p(|d(f + \varepsilon g)|) + \frac{1}{2}u_p(|d(f - \varepsilon g)|)$$

and therefore (notice that $u_p(|df|) \in \mathbf{R}$ on $\{|df| \in (0, +\infty)\}$ to justify cancellations) we have

$$\begin{aligned} d^+g(\nabla f) |df|^{p-2} &= \mathbf{m}\text{-ess sup}_{\varepsilon > 0} \frac{u_p(|d(f + \varepsilon g)|) - u_p(|df|)}{\varepsilon} \\ &\leq -\mathbf{m}\text{-ess sup}_{\varepsilon > 0} \frac{u_p(|d(f - \varepsilon g)|) - u_p(|df|)}{\varepsilon} = -d^+(-g)(\nabla f) |df|^{p-2}. \end{aligned}$$

(iv) Direct consequence of the locality in Proposition 3.22 and definition (4.8).

(v) On $\{|df| = +\infty\}$ the claim is obvious, so we focus on $\{|df| < +\infty\}$. The relation $g_1 \preceq g_2$ implies $f + \varepsilon g_1 \preceq f + \varepsilon g_2$ for every $\varepsilon > 0$ by Lemma 4.7. For $\varepsilon > 0$ sufficiently small both these functions are causal, thus the second claim in Lemma 4.6 yields $|d(f + \varepsilon g_1)| \leq |d(f + \varepsilon g_2)|$. Then the conclusion follows from the fact that u_p is non-decreasing.

(vi) Let us assume for the moment that, with the stated assumptions, it holds

$$d^+(\varphi \circ g)(\nabla f) |df|^{p-2} \geq \varphi' \circ g d^+g(\nabla f) |df|^{p-2} \quad (4.21)$$

and let us see how to conclude from here.

Define $\psi := (\text{Lip}(\varphi) + 1)\text{id} - \varphi$, so that ψ is also non-decreasing and Lipschitz with $\psi' > 0$ a.e. Thus from (4.13), \mathbf{m} -a.e. on $\{\varphi' \circ g > 0\} \cup \{|df| > 0\}$ we can write

$$\begin{aligned} (\text{Lip}(\varphi) + 1) d^+g(\nabla f) |df|^{p-2} &= d^+(\psi \circ g + \varphi \circ g)(\nabla f) |df|^{p-2} \\ &\quad \text{(by (4.14))} \geq d^+(\psi \circ g)(\nabla f) |df|^{p-2} + d^+(\varphi \circ g)(\nabla f) |df|^{p-2} \\ &\quad \text{(by } \geq \text{ in (4.17))} \geq \psi' \circ g d^+g(\nabla f) |df|^{p-2} + \varphi' \circ g d^+g(\nabla f) |df|^{p-2} \\ &= (\text{Lip}(\varphi) + 1) d^+g(\nabla f) |df|^{p-2}. \end{aligned}$$

Thus the inequalities are in fact equalities, and on the set $\{d^+g(\nabla f)|df|^{p-2} < +\infty\}$ this can only occur if each of the inequalities we used was in fact an equality, yielding the desired conclusion. On $\{d^+g(\nabla f)|df|^{p-2} = +\infty\}$ the conclusion is obvious on $\{\varphi' \circ g > 0\}$ (because we already know that \geq holds in (4.17)), so — inspecting the claim — to conclude it suffices to deal with the set $\{\varphi' \circ g = 0\} \cap \{|df| \in (0, +\infty)\}$ and prove that \mathfrak{m} -a.e. in there we have $|d(f + \varepsilon\varphi \circ g)| = |df|$. This, however, is clear from the characterization in Lemma 3.13, as for $\gamma \in \text{LCC}([0, 1]; M)$ and $t \in [0, 1]$ with $s \mapsto (\varphi(g(\gamma_s)), f(\gamma_s))$ differentiable at t (this holds for π -a.e. curve and \mathcal{L}^1 -a.e. t for any test plan π) and φ differentiable at $g(\gamma_t)$ with null derivative, we have $((f + \varepsilon\varphi \circ g)(\gamma))'_t = (f \circ \gamma)'_t + \varepsilon\varphi'(g(\gamma_t))(g \circ \gamma)'_t = (f \circ \gamma)'_t$.

We are thus left to proving (4.21). Fix an interval $I = [a, b] \subset (0, +\infty)$ and then a compact set $K \subset (\varphi')^{-1}(I)$. We are going to prove that

$$d^+(\varphi \circ g)(\nabla f)|df|^{p-2} \geq \min\{a d^+g(\nabla f)|df|^{p-2}, b d^+g(\nabla f)|df|^{p-2}\} \quad \mathfrak{m} - a.e. \text{ on } f^{-1}(K), \quad (4.22)$$

which by the arbitrariness of $[a, b]$ and K suffices to conclude. We can assume that K is not empty, or else the claim is obvious. Thus fix $\bar{z} \in K$ and then consider the auxiliary function $\tilde{\varphi} : \mathbf{R} \rightarrow \mathbf{R}$ defined by $\tilde{\varphi}(\bar{z}) = \varphi(\bar{z})$ and $\tilde{\varphi}' = 1_K\varphi' + a1_{\mathbf{R} \setminus K}$. Since $\mathbf{R} \setminus K$ is open, and thus a countable union of disjoint intervals (I_n) , we see that the set $E := \{\varphi(z) - \tilde{\varphi}(z) : z \in K\}$ is countable, as for any $z \in K$ the difference $\varphi(z) - \tilde{\varphi}(z)$ is equal to the sum of $\int_{I_n} \varphi' - a d\mathcal{L}^1$ over those intervals I_n that are entirely contained in either $[z, \bar{z}]$ or $[\bar{z}, z]$ (depending on whether $z > \bar{z}$ or $z < \bar{z}$). It follows from (3.22) that

$$|d(f + \varepsilon\varphi \circ g)| = |d(f + \varepsilon\tilde{\varphi} \circ g)| \quad \mathfrak{m} - a.e. \text{ on } K. \quad (4.23)$$

To conclude, let $\varphi_1(z) := az$ and $\varphi_2(z) := bz$ for any $z \in \mathbf{R}$, so that $\varphi_1 \preceq \varphi \preceq \varphi_2$. Then (4.7) ensures that

$$|d(f + \varepsilon\tilde{\varphi} \circ g)| \geq \min\{|d(f + \varepsilon a g)|, |d(f + \varepsilon b g)|\}$$

and the monotonicity of u_p yields

$$d^+(\tilde{\varphi} \circ g)(\nabla f)|df|^{p-2} \geq \min\{a d^+g(\nabla f)|df|^{p-2}, b d^+g(\nabla f)|df|^{p-2}\} \quad \mathfrak{m} - a.e.,$$

that together with (4.23) gives (4.22) and the claim. (vii) For $\varepsilon < \frac{1}{\text{Lip}(\varphi)}$, by Proposition 3.23 we have $|d(f + \varepsilon\varphi \circ f)| = |d((\text{id} + \varepsilon\varphi) \circ f)| = (1 + \varepsilon\varphi') \circ f |df|$. The conclusion follows by direct computation also noticing that $\{|df| < +\infty\} \subset \text{Vis}(M)$.

(viii) On $\{g = 0\} \cup \{h = 0\}$ the conclusion follows from (4.13), hence by a locality argument based on (4.16) to conclude it suffices to prove (4.19) in the set $\{g, h \in [c, c^{-1}]\}$ for any given $c > 0$. Fix such c and notice that by Lemma 4.9 the functions $\tilde{g} := c \vee g \wedge (c^{-1})$ and $\tilde{h} := c \vee h \wedge (c^{-1})$ still belong to $\text{Pert}(f)$. Since $z \mapsto \log(z)$ is Lipschitz on $[c^2, +\infty)$ we can apply (4.17) and get

$$\begin{aligned} \frac{1}{gh} d^+(\tilde{g}\tilde{h})(\dots) &= d^+(\log(\tilde{g}\tilde{h}))(\dots) = d^+(\log(\tilde{g}) + \log(\tilde{h}))(\dots) \\ &\stackrel{\text{(by (4.14))}}{\geq} d^+(\log(\tilde{g}))(\dots) + d^+(\log(\tilde{h}))(\dots) = \frac{1}{\tilde{g}} d^+\tilde{g}(\dots) + \frac{1}{\tilde{h}} d^+\tilde{h}(\dots), \end{aligned}$$

where (\dots) is shorthand for $(\nabla f)|df|^{p-2}$. Using again the locality property (4.16) the claim follows.

(ix) Let $c \geq 0$ be such that $g + c \geq 0$. Then we have

$$\begin{aligned} d^+(g\varphi \circ f)(\dots) &= d^+((g+c)\varphi \circ f + (-c\varphi \circ f))(\dots) \\ &\stackrel{\text{(by (4.14))}}{\geq} d^+((g+c)\varphi \circ f)(\dots) + d^+(-c\varphi \circ f)(\dots) \\ &\stackrel{\text{(by (4.19) and (4.18))}}{\geq} (g+c)d^+(\varphi \circ f)(\dots) + \varphi \circ f d^+(g+c)(\dots) - c\varphi' \circ f |df|^p \\ &\stackrel{\text{(by (4.16) and (4.18))}}{=} (g+c)\varphi' \circ f |df|^p + \varphi \circ f d^+g(\dots) - c\varphi' \circ f |df|^p \end{aligned}$$

and the conclusion follows because the term $c\varphi' \circ f |df|^p$ is finite and can be canceled out. \square

Proposition 4.12 (Calculus rules ‘for f ’). *Let $(M, \tau, \ell, \mathfrak{m})$ be a mm spacetime and $f : M \rightarrow \bar{\mathbf{R}}$ causal. Then:*

(i) Positive $p - 1$ homogeneity. Let $g \in \text{Pert}(f)$ and $\lambda \in (0, +\infty)$. Then

$$d^+g(\nabla(\lambda f))|d(\lambda f)|^{p-2} = \lambda^{p-1}d^+g(\nabla f)|df|^{p-2} \quad \mathbf{m}\text{-a.e.} \quad (4.24)$$

Notice that $g \in \text{Pert}(\lambda f)$ (trivially).

(ii) Locality. Let $\tilde{f} : M \rightarrow \bar{\mathbf{R}}$ be a causal function, $g \in \text{Pert}(f) \cap \text{Pert}(\tilde{f})$ and $E \subset \mathbf{R}$ Borel and negligible. Then

$$d^+g(\nabla f)|df|^{p-2} = d^+g(\nabla \tilde{f})|d\tilde{f}|^{p-2} \quad \mathbf{m}\text{-a.e. on } (f - \tilde{f})^{-1}(E). \quad (4.25)$$

(iii) Chain rule. Let $\varphi : \mathbf{R} \rightarrow \mathbf{R}$ be c -steep (i.e. $\varphi(w) - \varphi(z) \geq (w - z)c$ for every $z \leq w$) for some $c > 0$ and let $g \in \text{Pert}(f)$. Then

$$d^+g(\nabla(\varphi \circ f))|d(\varphi \circ f)|^{p-2} = (\varphi')^{p-1} \circ f d^+g(\nabla f)|df|^{p-2} \quad \mathbf{m}\text{-a.e.} \quad (4.26)$$

Note that g belongs to $\text{Pert}(\varphi \circ f)$ by Lemma 4.9.

Proof.

(i) For $\lambda > 0$ the claim follows \mathbf{m} -a.e. from the positive 1-homogeneity in Proposition 3.21:

$$d^+g(\nabla(\lambda f))|d(\lambda f)|^{p-2} = \lambda^{p-1} \lim_{\varepsilon \downarrow 0} \frac{|d(f + g\varepsilon/\lambda)^p - |df|^p|}{\varepsilon/\lambda} = \lambda^{p-1}d^+g(\nabla f)|df|^{p-2}.$$

(ii) Direct consequence of the definition and Proposition 3.22.

(iii) It is readily proved that for $\psi : \mathbf{R} \rightarrow \mathbf{R}$ with $\varphi \preceq \psi$ we have $\delta(\varphi \circ f + \varepsilon g) \leq \delta(\psi \circ f + \varepsilon g)$ and therefore, by Lemma 4.6, that $|d(\varphi \circ f + \varepsilon g)| \leq |d(\psi \circ f + \varepsilon g)|$. Taking into account the monotonicity of u_p we then deduce that

$$d^+g(\nabla(\varphi \circ f))|d(\varphi \circ f)|^{p-2} \leq d^+g(\nabla(\psi \circ f))|d(\psi \circ f)|^{p-2} \quad \mathbf{m} - \text{a.e.} \quad (4.27)$$

The conclusion now follows as for the proof of (4.17): we first use a locality argument to replace the original φ with a $\tilde{\varphi}$ so that $\varphi' \in [a, b] \subset (0, +\infty)$, then we use (4.27) to compare the desired quantity with that appearing with the choices $\varphi_a(z) := az$ and $\varphi_b(z) := bz$. To conclude we then observe that for linear φ 's the conclusion is given by (4.24). \square

Remark 4.13 (Limits and usefulness of conventions). The — arbitrary — choice of defining $d^+g(\nabla f)|df|^{p-2}$ to be 0 on $\{|df| = +\infty\}$ is only motivated by simplicity, as in any case we won't ever really care about what happens to such an expression where $|df| = +\infty$; setting to zero won't create integrability problems later on. A side effect of this choice is that some formulas, such as (4.18) only hold on $\{|df| < +\infty\}$: we are aware of this fact, that causes no troubles in what comes next.

In a conceptually similar direction, we have defined the homogeneity formulas (4.13) and (4.24) — and thus with the chain rules, that ultimately depend on these — with care to avoid conflicts with (4.10) or (4.11) on the set $\{|df| = 0\}$.

In contrast with the situation up to the previous chapter, we haven't found a consistent set of conventions capable of handling all possible scenarios, and in any case having these would be irrelevant for the point we want to make, which is that in all interesting cases the relevant formulas hold 'as is', without the need to resort to artificial choices. \blacksquare

Remark 4.14 (Sign choices). In line with the terminology in [60], the quantity $-d^+(-\varphi)(\nabla f)|df|^{p-2}$ appearing in (4.15) might also be called $d^- \varphi(\nabla f)|df|^{p-2}$. We haven't done so both to avoid introducing a new notation and because in our signature this would lead to the bizarre looking inequality $d^+\varphi(\dots) \leq d^-\varphi(\dots)$ where the 'minus' term is bigger than the 'plus' term.

Speaking of this, we emphasize that working with $d^+\varphi(\dots)$ rather than with the above-mentioned $d^-\varphi(\dots)$ is related to the fact that we shall work with the TMCP_+^h condition rather than with the TMCP_-^h one. In this direction it might be worth to keep in mind the following table

$\text{TMCP}_+^h(K, N)$	<i>vs</i>	$\text{TMCP}_-^h(K, N)$
$d^+\varphi(\dots)$	<i>vs</i>	$d^-\varphi(\dots)$
$\varphi \in \text{Pert}(f)$	<i>vs</i>	$-\varphi \in \text{Pert}(f)$,
initial test plans	<i>vs</i>	final test plans,
functions of the form ψ_c	<i>vs</i>	functions of the form ψ^c

(see (5.26) for the definition of ψ_c and ψ^c) and notice that the first choice dictates all the others below it. This means that if we work on $\text{TMCP}_+^h(K, N)$ spaces, as we shall do most of the time, then the d'Alembertian comparison will be obtained for functions of the form ψ_c by differentiating functions $\varphi \in \text{Pert}(f)$ along initial test plans and studying the quantity $d^+\varphi(\nabla f)|df|^{p-2}$. Symmetrically if we work on $\text{TMCP}_-^h(K, N)$ spaces, see Section 5.7 for comments.

Of totally different kind and *not* related to the above, is the choice of dealing with forward spacetimes rather than backward ones, i.e. spacetimes that are forward complete rather than backward complete (recall Definition 2.1). We shall always stick to this choice, so that in particular the results concerning TMCP_-^h spaces — even though will be proved along similar lines of thought — cannot be derived by a simple time-reversal. ■

4.4 Relation of horizontal and vertical derivatives

Here we discuss the connection between these two different notions of differentiation for arbitrary causal functions.

Our motivation is drawn once more from the case of a smooth spacetime (M, g) . Given a smooth causal function $f: M \rightarrow \mathbf{R}$ and any smooth and compactly supported function $g: M \rightarrow \mathbf{R}$, the quantity $dg(\nabla f) \|df\|_*^{p-2}$ can be computed in two ways. The first, related to our horizontal approach, is to consider a smooth curve γ with initial speed $\gamma'_0 = \|df\|_*^{p-2} \nabla f$ and differentiate at zero:

$$\left. \frac{d}{dt} g(\gamma_t) \right|_0 = dg(\nabla f) \|df\|_*^{p-2}.$$

The second, related to our vertical approach, is to differentiate the ε -dependent quantity $|d(f + \varepsilon g)|_*^p/p$ (in other words, the underlying Hamiltonian applied to the covector field $d(f + \varepsilon g)$) at zero:

$$\left. \frac{d}{d\varepsilon} \frac{\|d(f + \varepsilon g)\|_*^p}{p} \right|_0 = dg(\nabla f) \|df\|_*^{p-2}.$$

This last identity can be seen recalling that ∇f is — inspecting the definition of the musical isomorphisms — equal to the differential of $\frac{1}{2} \|\cdot\|_*^2$ applied at df , see also Appendix A.1. Since the right-hand sides of the last two formulas agree, so do the left ones.

Remarkably, this consideration has a counterpart in our context. Recall the definition of ‘plan π representing the initial p -gradient of a causal function f ’ from Definition 4.2.

Theorem 4.15 (First-order differentiation formula). *Let $(M, \tau, \ell, \mathbf{m})$ be a mm spacetime, $f: M \rightarrow \bar{\mathbf{R}}$ causal, $0 \neq p < 1$ and let π represent the initial p -gradient of f .*

Then for every $g \in \text{Pert}(f)$ we have:

- (i) *For any $t > 0$ sufficiently small the negative part of the function $g \circ \mathbf{e}_t - g \circ \mathbf{e}_0$ is in $L^1(\pi)$,*
- (ii) *and*

$$\liminf_{t \downarrow 0} \int \frac{g(\gamma_t) - g(\gamma_0)}{t} d\pi(\gamma) \geq \int d^+g(\nabla f) |df|^{p-2}(\gamma_0) d\pi(\gamma). \quad (4.28)$$

- (iii) *If $E \subset M$ Borel so that π represents the initial p -gradient of f on E , then the negative part of the function $d^+g(\nabla f) |df|^{p-2}$ is in $L^1(\mathbf{m} \llcorner E)$.*

Proof. Definition 4.2 implies, as already noticed right after it, that for $t > 0$ sufficiently small we have $f \circ e_t - f \circ e_0 \in L^1(\pi)$ (and in particular the function is a.e. finite). On the other hand, for $\bar{\varepsilon} > 0$ we have that $f + \varepsilon g$ is causal for any $\varepsilon \in [0, \bar{\varepsilon}]$, so that in particular $(f + \bar{\varepsilon}g) \circ e_t - (f + \bar{\varepsilon}g) \circ e_0$ is non-negative on $\text{LCC}([0, 1]; \mathbf{M})$. Writing $g = \frac{1}{\bar{\varepsilon}}((f + \bar{\varepsilon}g) - f)$ the claim (i) easily follows. The claim (iii) follows from the uniform bound (4.12) and the assumption $u_p(|df|) \in L^1(E, \mathbf{m} \llcorner E)$ that comes with Definition 4.2.

(ii) We now want to apply Proposition 4.1 to the function $f + \varepsilon g$ for $\varepsilon \in (0, \frac{1}{2}\bar{\varepsilon})$ and the plan π ; to this aim let us check the assumptions. If $q < 0$ the requirement for π is satisfied, as the plan satisfies Definition 4.2. If instead $p < 0$ we need to check that $u_p(|d(f + \varepsilon g)|) \in L^1(E, \mathbf{m} \llcorner E)$. As in the proof of the bound (4.12), this follows from the bound $0 \geq u_p(|d(f + \varepsilon g)|) \geq u_p(\frac{1}{2}|df|)$ and again the assumption $u_p(|df|) \in L^1(E, \mathbf{m} \llcorner E)$.

We can therefore apply Proposition 4.1 to the function $f + \varepsilon g$, thus writing the corresponding inequality (4.1) for $f + \varepsilon g$ and subtracting (4.4) for f , after the — justified — cancellations we get

$$\liminf_{t \rightarrow 0} \int \frac{g(\gamma_t) - g(\gamma_0)}{t} d\pi(\gamma) \geq \int \frac{|d(f + \varepsilon g)|^p(\gamma_0) - |df|^p(\gamma_0)}{p\varepsilon} d\pi(\gamma) \quad \forall \varepsilon \in (0, \frac{1}{2}\bar{\varepsilon}).$$

The uniform bound (4.12) and the assumption $u_p(|df|) \in L^1(E, \mathbf{m} \llcorner E)$ ensure that we can use Levi's monotone convergence theorem to pass to the limit as $\varepsilon \downarrow 0$. The conclusion follows. \square

We now analyze how the calculus developed above improves on infinitesimally Minkowskian spaces (recall Definition 1.4 from the introduction). We claim in this setting, the bound (4.28) upgrades to a genuine first order differentiation formula equating the horizontal and vertical derivatives, (4.30) below.

Theorem 4.16 (Elements of calculus on infinitesimally Minkowskian spaces). *Let $(\mathbf{M}, \tau, \ell, \mathbf{m})$ be a forward metric measure spacetime that is infinitesimally Minkowskian and $0 \neq p < 1$. Then:*

(i) *Let $f : \mathbf{M} \rightarrow \bar{\mathbf{R}}$ be causal and $\pm g \in \text{Pert}(f)$. Then*

$$-d^+(-g)(\nabla f) |df|^{p-2} = d^+g(\nabla f) |df|^{p-2} \quad \mathbf{m} - \text{a.e. on } \{|df| > 0\}; \quad (4.29)$$

(compare with inequality (4.15)).

(ii) *If, in addition, π represents the initial p -gradient of f on the Borel set E then*

$$\lim_{t \downarrow 0} \int \frac{g(\gamma_t) - g(\gamma_0)}{t} d\pi(\gamma) = \int d^+g(\nabla f) |df|^{p-2}(\gamma_0) d\pi(\gamma). \quad (4.30)$$

(iii) *For $f, g, h : \mathbf{M} \rightarrow \bar{\mathbf{R}}$ causal define $d^+g(\nabla f)$ on the set $\{|df| \in (0, +\infty)\}$ as the product of $|df|^{2-p}$ and $d^+g(\nabla f)|df|^{p-2}$. Then this quantity does not depend on p ; moreover, for $\alpha, \beta \in [0, \infty)$, both*

$$\begin{aligned} d^+g(\nabla f) &= d^+f(\nabla g) \\ \text{and } d^+(\alpha g + \beta h)(\nabla f) &= \alpha d^+g(\nabla f) + \beta d^+h(\nabla f) \end{aligned} \quad (4.31)$$

hold \mathbf{m} -a.e. on the set $\{x \in \mathbf{M} : 0 < |df| \wedge |dg| \wedge |dh| \text{ and } |d(f + g + h)| < \infty\}$.

Proof.

(i) Let $\bar{\varepsilon} > 0$ be such that $f + \varepsilon g$ is causal for every $\varepsilon \in [-\bar{\varepsilon}, \bar{\varepsilon}]$. On $\{|df| = +\infty\}$ there is nothing to prove, as both sides are 0 by convention. On the set $\{|df| \in (0, +\infty)\}$, by the smoothness of u_p on $(0, +\infty)$ it suffices to prove that

$$\lim_{\varepsilon \downarrow 0} \frac{|d(f + \varepsilon g)|^2 - |df|^2}{2\varepsilon} = \lim_{\varepsilon \downarrow 0} \frac{|df|^2 - |d(f - \varepsilon g)|^2}{2\varepsilon}, \quad (4.32)$$

where the limits are intended as in (4.8) and their existence follows from the existence of the limit in (4.8). To see the above write the defining identity (1.11) with $f - \varepsilon g, f + \varepsilon g$ in place of f, g respectively, for $\varepsilon \in (-\bar{\varepsilon}, \bar{\varepsilon})$, to get

$$2|d(f - \varepsilon g)|^2 + 2|d(2f)|^2 = |d(f + \varepsilon g)|^2 + |d(3f - \varepsilon g)|^2 \quad \mathbf{m} - a.e.$$

Rearranging and recalling the positive homogeneity in Proposition 3.21 we get

$$|d(f + \varepsilon g)|^2 - |df|^2 = 2(|d(f - \varepsilon g)|^2 - |df|^2) - 9(|d(f - \frac{\varepsilon}{3}g)|^2 - |df|^2)$$

and — since both limits in (4.32) exist — the identity (4.32) easily follows.

(ii) Applying Theorem 4.15 to $-g$ in place of g we obtain

$$\limsup_{t \rightarrow 0} \int \frac{g(\gamma_t) - g(\gamma_0)}{t} d\boldsymbol{\pi}(\gamma) \leq - \int d^+(-g)(\nabla f) |df|^{p-2}(\gamma_0) d\boldsymbol{\pi}(\gamma).$$

The conclusion follows combining Theorem 4.15 as stated with item (i) above.

(iii) For f, g causal define the auxiliary function $B(f, g)$ on $\{0 < |df| \wedge |dg|\} \cap \{|d(f + g)| < \infty\}$ as

$$2B(f, g) := |d(f + g)|^2 - |df|^2 - |dg|^2.$$

Then $B(f, g) = B(g, f)$ and is finite \mathbf{m} -a.e. by (3.20). We claim that for any f, g, h causal

$$B(f + h, g) = B(f, g) + B(h, g) \tag{4.33}$$

holds where they are all defined. From the parallelogram law (1.11) defining infinitesimal Minkowskianity we see that

$$\begin{aligned} 4|d(f + h + g)|^2 + 4|dh|^2 &= 2|d(f + 2h + g)|^2 + 2|d(f + g)|^2, \\ 2|d(f + 2h + g)|^2 + 2|df|^2 &= |d(2f + 2h + g)|^2 + |d(2h + g)|^2, \\ 2|d(h + g)|^2 + 2|dh|^2 &= |d(2h + g)|^2 + |dg|^2, \\ 2|d(f + h + g)|^2 + 2|d(f + h)|^2 &= |d(2f + 2h + g)|^2 + |dg|^2. \end{aligned}$$

The quantities on the left of each line are finite, hence the same is true of the quantities on the right. Adding the first two identities and subtracting the last two we get (4.33). It is then clear that

$$B(\alpha f, g) = \alpha B(f, g) \tag{4.34}$$

first for $\alpha \in \mathbf{N}$, then for $\alpha \in \mathbf{Q}_+$ and finally, using that $|d(\alpha f + g)| \leq |d(\beta f + g)| \leq |d(\gamma f + g)|$ holds \mathbf{m} -a.e. whenever $0 \leq \alpha \leq \beta \leq \gamma$ as a consequence of (3.20), we conclude that (4.34) holds for $\alpha \in [0, +\infty)$.

To conclude, let f, g be causal, notice that on the set $\{|df| \in (0, +\infty)\}$ the limits in (4.32) are equal to $d^+g(\nabla f)$ (from the discussions in the previous item) and therefore

$$d^+g(\nabla f) = \lim_{\varepsilon \downarrow 0} \frac{|d(f + \varepsilon g)|^2 - |df|^2}{2\varepsilon} = \lim_{\varepsilon \downarrow 0} \frac{B(f, \varepsilon g)}{\varepsilon} \stackrel{(4.34)}{=} B(f, g)$$

and the conclusion follows. \square

Remark 4.17 (Lorentzian a.e. inner product from infinitesimal Minkowskianity). In an infinitesimally Minkowskian forward spacetime, we can regard (4.31) as defining a symmetric and positively bilinear a.e. inner product on the convex cone of causal cotangent fields df and dg , at least where $0 < |df| \wedge |dg|$ and $|d(f + g)| < \infty$. \blacksquare

5 Effects of timelike Ricci curvature assumptions

5.1 Synthetic timelike curvature-dimension bounds

We introduce here the curvature conditions we are going to work with in the rest of the paper. They involve (slight modifications of) the Timelike Measure Contraction Property proposed by Braun [21], and the entropic variant of Cavalletti–Mondino [39] upon which it is based. Although these conditions are expected to become equivalent under suitable non-branching hypotheses, this is not yet established in Lorentzian signature. Braun’s formulation provides sharp constants, while Cavalletti and Mondino’s yields narrow compactness in merely forward spacetimes via Lemma 5.4 below. In the absence of global hyperbolicity we are going to need both. (When $(M, \tau, \ell, \mathbf{m})$ is globally hyperbolic on the other hand, all our results remain valid even if we replace Boltzmann’s entropy (5.1) by the trivial entropy $\mathcal{S}_\infty(\mu) := 0$; Lemma 5.4 would remain true since the compactness of emeralds $E \subset M$ which holds in this case implies tightness — hence narrow compactness — of the entire set $\mathcal{P}(E)$.)

For measures $\mu \in \mathcal{P}(E)$ supported on an emerald $E \subset M$, the Boltzmann-Shannon entropy $\mathcal{S}_\infty : \mathcal{P}_{\text{em}}(M) \rightarrow [-\log \mathbf{m}(E), \infty]$ is defined by

$$\mathcal{S}_\infty(\mu) := \begin{cases} \int_M \rho \log \rho \, d\mathbf{m} & \text{if } \mu = \rho \mathbf{m} \\ +\infty & \text{otherwise.} \end{cases} \quad (5.1)$$

It is well-defined if $\mathbf{m}(E) < \infty$ and its values lie in the indicated range by Jensen’s inequality. Cavalletti and Mondino’s $\text{TMCP}^e(K, N)$ condition asks for suitable growth bounds on $\mathcal{S}_\infty(\mu_t)$ along strongly timelike ℓ_q -geodesics starting or ending at a point mass. Braun’s $\text{TMCP}(K, N)$ asks for analogous decay estimates on the N -Rényi entropy $\mathcal{S}_N : \mathcal{P}(M) \rightarrow [-\infty, 0]$ defined by

$$\mathcal{S}_N(\mu) := - \int_M \rho^{(N-1)/N} \, d\mathbf{m} \quad \text{for } \mu = \rho \mathbf{m} + \mu^\perp, \quad \mu^\perp \perp \mathbf{m}.$$

Both are detailed in Definition 5.1 and Remark 5.2 below.

For $\kappa \in \mathbf{R}$ define the generalized sine function $\sin_\kappa : \mathbf{R} \rightarrow \mathbf{R}$ as the only smooth f solving

$$f'' + \kappa f = 0, \quad f(0) = 0, \quad f'(0) = 1,$$

so that

$$\sin_\kappa(\theta) := \begin{cases} \kappa^{-1/2} \sin(\kappa^{1/2} \theta) & \text{if } \kappa > 0, \\ \theta & \text{if } \kappa = 0, \\ |\kappa|^{-1/2} \sinh(|\kappa|^{1/2} \theta) & \text{otherwise.} \end{cases}$$

Then for $K \in \mathbf{R}$ and $N \in (1, +\infty)$ define two distortion coefficients as follows: for $\theta \geq 0$ and $t \in [0, 1]$, set

$$\sigma_{K,N}^{(t)}(\theta) := \begin{cases} \frac{\sin_{K/N}(t\theta)}{\sin_{K/N}(\theta)} & \text{if } K\theta^2 < N\pi^2, \\ +\infty & \text{otherwise} \end{cases} \quad (5.2)$$

and, by geometrically averaging,

$$\tau_{K,N}^{(t)}(\theta) := t^{1/N} \sigma_{K,N-1}^{(t)}(\theta)^{(N-1)/N}. \quad (5.3)$$

In particular, when $K = 0$ the distortion coefficients simply become $\sigma_{0,N}^{(t)}(\theta) = \tau_{0,N}^{(t)}(\theta) = t$.

The interpretation of the average (5.3) is that in the smooth setting, the term $t^{1/N}$ measures the volume distortion in all “tangential” directions (which does not see curvature), while “orthogonal” directions are influenced by curvature and quantified by $\sigma_{K,N-1}^{(t)}(\theta)^{(N-1)/N}$.

We now come to main definition of this section. Recall from Proposition 2.40 that under the stated assumption the notion of ℓ_q -geodesic to a Dirac mass does not depend on the chosen $0 \neq q < 1$.

Definition 5.1 (Hybrid timelike measure contraction property). *Let $(M, \tau, \ell, \mathbf{m})$ be a mm spacetime in which ℓ is upper semicontinuous and does not take the value $+\infty$. We say that it satisfies the hybrid future timelike measure contraction property $\text{TMCP}_+^h(K, N)$ if for every $\mu = \rho \mathbf{m} \in \mathcal{P}_{\text{em}}(M)$ and $x_1 \in \text{spt } \mathbf{m}$ with $\log(\ell(\cdot, x_1)) \in L^\infty(\mu)$ and $\mathcal{S}_\infty(\mu) < \infty$, there is an ℓ_q -geodesic (μ_t) from μ to δ_{x_1} such that*

$$\mathcal{S}_N(\mu_t) \leq - \int \tau_{K,N}^{(1-t)} \circ \ell(\cdot, x_1) \rho^{1-1/N} \text{d}\mathbf{m}, \quad \forall t \in [0, 1] \quad \text{and} \quad (5.4)$$

$$\mathcal{S}_\infty(\mu_t) \leq \mathcal{S}_\infty(\mu) - N \log \sigma_{K,N}^{(1-t)}(\|\ell(\cdot, x_1)\|_{L^2(\text{d}\mu)}), \quad \forall t \in [0, 1]. \quad (5.5)$$

Similarly, (M, ℓ, \mathbf{m}) satisfies the hybrid past timelike measure contraction property $\text{TMCP}_-^h(K, N)$ if for every $\mu = \rho \mathbf{m} \in \text{Dom}(\mathcal{S}_\infty) \subset \mathcal{P}_{\text{em}}(M)$ and $x_0 \in \text{spt } \mathbf{m}$ with $\log(\ell(x_0, \cdot)) \in L^\infty(\mu)$, there is an ℓ_q -geodesic (μ_t) from δ_{x_0} to μ such that

$$\mathcal{S}_N(\mu_t) \leq - \int \tau_{K,N}^{(t)} \circ \ell(x_0, \cdot) \rho^{1-1/N} \text{d}\mathbf{m}, \quad \forall t \in [0, 1] \quad \text{and} \quad (5.6)$$

$$\mathcal{S}_\infty(\mu_t) \leq \mathcal{S}_\infty(\mu) - N \log \sigma_{K,N}^{(t)}(\|\ell(x_0, \cdot)\|_{L^2(\text{d}\mu)}), \quad \forall t \in [0, 1]. \quad (5.7)$$

Remark 5.2 (Related timelike measure contraction properties). Requiring (5.4) but not (5.5) — or more precisely, replacing (5.1) with the trivial entropy $\mathcal{S}_\infty := 0$ — would yield a future version $\text{TMCP}_+(K, N)$ of Braun’s timelike measure contraction property. In this case, replacing the τ -distortion coefficients by the σ -distortion coefficients in (5.4) yields the *reduced future timelike measure contraction property* $\text{TMCP}_+^*(K, N)$. Similarly a future version $\text{TMCP}_+^e(K, N)$ of Cavalletti and Mondino’s timelike measure contraction property is obtained by requiring (5.5) but not (5.4) or equivalently setting $\mathcal{S}_N := -\infty$. Past versions of these various conditions are defined analogously. In all that follows, the hybrid assumption $\text{TMCP}_\pm^h(K, N)$ can be relaxed to $\text{TMCP}_\pm(K, N)$ whenever $(M, \tau, \ell, \mathbf{m})$ is globally hyperbolic.

$\text{TMCP}_+^h(K, N)$ is stronger than $\text{TMCP}_+^*(K, N)$ [21, Prop. 3.6] and allows to prove all quantitative properties we care about (especially our comparison theorems) in their sharp form. However, with appropriate modifications that we occasionally specify, our results have evident adaptations to the more general setting $\text{TMCP}_+^{h,*}(K, N)$ where the distortion coefficients τ are replaced by σ in (5.4).

Notice also that — in line with the definition in positive signature [98], [114] — the $\text{TMCP}_+^h(K, N)$ as formulated is not ‘dimensionally consistent’, meaning that a priori a space could be $\text{TMCP}_+^h(K, N)$ but not $\text{TMCP}_+^h(K, N')$ for some $N' > N$. To enforce this property one should — tautologically — work with spaces that are $\text{TMCP}_+^h(K, N')$ for all $N' \geq N$ (and similarly for $\text{TMCP}_+^{h,*}(K, N)$ and the past versions of both). In timelike q -essentially non-branching spaces, however, dimensional consistency does hold, as can be shown by localizing the inequalities defining $\text{TMCP}_+^h(K, N)$ and $\text{TMCP}_+^{h,*}(K, N)$ to a pathwise inequality following the lines of Braun [21, Thms. 4.20, 4.21] and taking appropriate powers. In particular, $\text{TMCP}_+^{h,*}(K, N)$ then recovers the entropic or reduced TMCP conditions of Cavalletti–Mondino [39] and Braun [21]. Moreover, in this situation we expect the two hybrid conditions $\text{TMCP}_+^h(K, N)$ and $\text{TMCP}_+^{h,*}(K, N)$ to be equivalent, as suggested by corresponding results of Cavalletti–Sturm [41] in positive signature. ■

Remark 5.3 (Failure to reach endpoints). Typically one assumes that the geodesic (μ_t) in the above definition satisfies $\mu_0 = \mu$ and $\mu_1 = \delta_{x_1}$, however in the current setting it seems us more natural to work with this slightly weaker variant (2.32). One of the reasons is that it seems more likely that this notion is stable by convergence of mm spacetimes (compare with the closure of $\text{CGeo}(M)$ established in Proposition 2.37). ■

Notice that by nature of the definition, the $\text{TMCP}_+^h(K, N)$ (resp. $\text{TMCP}_-^h(K, N)$) condition is only relevant for those points having non-empty chronological future (resp. past). In other words, introduce the ‘timelike final/initial sets’ M_{fin} and M_{in} of M as

$$M_{\text{fin}} := \{x \in M : I^+(x) = \emptyset\}, \quad \text{and} \quad M_{\text{in}} := \{x \in M : I^-(x) = \emptyset\}$$

and notice that these are evidently achronal but possibly empty. If the topology of M contains the chronological one, then M_{fin} and M_{in} are automatically closed; moreover, the closure of any achronal set is still achronal. It is also clear that

$$\left. \begin{array}{l} (M, \tau, \ell, \mathfrak{m}) \text{ is TMCP}_+^h(K, N) \\ M_{\text{fin}} \text{ is } \mathfrak{m}\text{-measurable} \end{array} \right\} \Rightarrow (M, \tau, \ell, \mathfrak{m} \llcorner (M_{\text{fin}})^c) \text{ is TMCP}_+^h(K, N) \quad (5.8)$$

and similarly for $\text{TMCP}_-^h(K, N)$ spaces. As we are going to see in the next section (Corollary 5.11), on $\text{TMCP}_+^h(K, N)$ spaces achronal sets are \mathfrak{m} -negligible (with the possible exception of M_{fin}), thus (5.8) is relevant in connection to \mathfrak{m} -measurability of causal functions (recall Lemma 3.2).

5.2 Good geodesics

Our Sobolev calculus is built on the concept of test plans, that in turn asks for relevant measures to have bounded compression. On the other hand, the curvature condition we are asking only imposes entropy bounds. There is therefore a regularity gap that needs to be filled to link the geometry encoded in the curvature assumption with the analysis on the space as we are developing it. In positive signature, such a link was provided by Rajala in [104, 105] in the setting of CD spaces: one of its first effects was to simplify the axiomatization of the RCD condition (making it depending only on the CD assumption plus infinitesimal Hilbertianity) along the lines suggested in [60, Remark 4.20] and put forward in [8]. Rajala's construction was extended to $\text{MCP}(K, N)$ spaces by Cavalletti-Mondino in [36]. This latter result was later 'Lorentzified' in [20]. In this section we shall show how it can be adapted to our current setting. The aim is to show existence of ℓ_q -geodesics satisfying not only the entropy bounds encoded in the TMCP^h assumption, but also suitable L^∞ -density bounds. As is well known to experts, the construction involves 3 steps:

- 1) 'ONE STEP ESTIMATES' Prove existence of an intermediate point obeying the desired bounds.
- 2) 'DISCRETE ITERATION' Construct a discretization of the desired geodesic.
- 3) 'PASSAGE TO THE LIMIT' Pass to the limit in the discretization to get the desired geodesic.

To implement this in our setting, we notice two differences w.r.t. [20]: one concerns the version of TMCP condition adopted, which has only a minor impact on the proof, and a conceptually deeper one related to our choice to give up global hyperbolicity in favor of forward completeness.

Because of this passage from 'compactness' to 'completeness', the execution of the above plan, that would be more or less standard in a globally hyperbolic setting, becomes more involved and for this reason we give some detail below on how to proceed. This will involve step (3) in particular, where we will make crucial use of the forward-narrow completeness of the space of measures established in Proposition 2.18. An additional ingredient is instead needed to adapt Rajala's construction to the forward context and get step (1) done. For this we rely on the following lemma, already known to experts:

Lemma 5.4 (Narrow coercivity of Boltzmann entropy). *Fix an emerald $E = J(C'_0, C'_1)$ with $\mathfrak{m}(E) < \infty$ in a metric measure spacetime $(M, \tau, \ell, \mathfrak{m})$. For each $c \in \mathbf{R}$, the following set is narrowly compact:*

$$S_c := \{\mu \in \mathcal{P}(E) : \mathcal{S}_\infty(\mu) \leq c\}. \quad (5.9)$$

Proof. Since $\mathfrak{m}_E := \mathfrak{m} \llcorner E$ assigns finite mass to the Polish space M , there is a sequence of compact sets $C_i \subset C_{i+1} \subset M$ which exhausts the mass of \mathfrak{m}_E . As in the discussion below [62, Eq. (3.4)], since the function $x \mapsto x \log x$ is convex and greater than -1 , Jensen's inequality can be used to estimate Boltzmann's entropy $\mathcal{S}_\infty(\mu)$ from below by $-\mathfrak{m}_E(E \setminus B) + \mu(B) \log \frac{\mu(B)}{\mathfrak{m}_E(B)}$ for any Borel set B . Hence, using (5.9), the mass of $\mu \in S_c$ outside C_i (i.e. setting $B = M \setminus C_i$) can be estimated by

$$\mu(M \setminus C_i) \leq \frac{1 + c + \mathfrak{m}_E(C_i)}{-\log \mathfrak{m}_E(M \setminus C_i)} \rightarrow 0$$

as $i \rightarrow \infty$ at a rate independent of μ . This shows S_c is tight, hence Prokhorov's theorem yields the desired narrow compactness. \square

In the discussion below we shall need the notion of λ -intermediate measure between a measure $\mu \in \mathcal{P}(M)$ and a Dirac mass $\delta_{\bar{x}}$ at a point $\bar{x} \in M$ such that $\log(\ell(\cdot, \bar{x})) \in L^\infty(\mu)$. Given such μ and \bar{x} and $\lambda \in (0, 1)$ the set $\text{Int}_\lambda(\mu, \delta_{\bar{x}})$ is defined to consist of those $\nu \in \mathcal{P}(M)$ with $\mu \preceq \nu \preceq \delta_{\bar{x}}$ for which there is $\pi \in \Pi_{\preceq}(\mu, \nu)$ such that

$$\ell(x, y) = \lambda \ell(x, \bar{x}) \quad \text{and} \quad \ell(y, \bar{x}) = (1 - \lambda) \ell(x, \bar{x}) \quad \pi - a.e. (x, y). \quad (5.10)$$

Equivalently $\nu \in \text{Int}_\lambda(\mu, \delta_{\bar{x}})$ if and only if

$$\ell_q(\mu, \nu) = \lambda \ell_q(\mu, \delta_{\bar{x}}) \quad \text{and} \quad \ell_q(\nu, \delta_{\bar{x}}) = (1 - \lambda) \ell_q(\mu, \delta_{\bar{x}}).$$

The definition of ℓ_q and reverse triangle inequality make it clear that $\text{Int}_\lambda(\mu, \delta_{\bar{x}})$ is convex. (Although not needed here, in a geodesic spacetime, Propositions 2.40 and 2.47 and their proofs can also be used to show a measure $\nu \in \text{Int}_\lambda(\mu, \delta_{\bar{x}})$ is — for any $0 \neq q < 1$ — the value at λ of a ℓ_q -geodesic from μ to $\delta_{\bar{x}}$; c.f. [84, Lemma 2.8] or [39, Theorem 2.11].)

The definition clearly implies that $\log(\ell(\cdot, \bar{x})) \in L^\infty(\nu)$ as well, so the concept can be iterated. Then notice that a simple gluing argument together with the reverse triangle inequality yield

$$\nu \in \text{Int}_{\lambda_1}(\mu, \delta_{\bar{x}}), \quad \sigma \in \text{Int}_{\lambda_2}(\nu, \delta_{\bar{x}}) \quad \Rightarrow \quad \sigma \in \text{Int}_{\lambda_3}(\mu, \delta_{\bar{x}}), \quad (5.11)$$

for $(1 - \lambda_3) = (1 - \lambda_1)(1 - \lambda_2)$.

Proposition 5.5 (Step 1 — one step estimates). *Let $(M, \tau, \ell, \mathbf{m})$ be a $\text{TMCP}_+^h(K, N)$ mm spacetime, $\mu_0 = \rho_0 \mathbf{m} \in \text{Dom}(\mathcal{S}_\infty) \subset \mathcal{P}_{\text{em}}(M)$ and $x_1 \in \text{spt } \mathbf{m}$ satisfy $\log(\ell(\cdot, x_1)) \in L^\infty(\mu_0)$. Assume $\lambda \in (0, 1)$ and $\mathbf{m}(E) < \infty$ for some emerald E containing $\{x_1\} \cup \text{spt } \mu_0$.*

Then there is $\nu_\lambda \in \text{Int}_\lambda(\mu_0, \delta_{x_1})$ so that putting $D := \text{supp } \ell(\text{spt } \mu_0 \times \{x_1\})$ we have

$$\nu_\lambda \leq \frac{1}{(1-\lambda)^N} e^{Dt\sqrt{K-(N-1)}} \|\rho_0\|_{L^\infty(\mathbf{m})} \mathbf{m}; \quad (5.12a)$$

$$\mathcal{S}_N(\nu_\lambda) \leq - \int \tau_{K,N}^{(1-\lambda)} \circ \ell(\cdot, x_1) \rho_0^{1-1/N} d\mathbf{m}; \quad (5.12b)$$

$$\mathcal{S}_\infty(\nu_\lambda) \leq \mathcal{S}_\infty(\mu_0) - N \log \sigma_{K,N}^{(1-t)}(\|\ell(\cdot, x_1)\|_{L^2(\mu_0)}). \quad (5.12c)$$

Proof. The proof is a variant of the clever manipulations in [104, 105]: we shall sketch it in the simplified case $K = 0$, where computations are more transparent. We start noticing that for μ_0, x_1 as in the assumption, the set

$$\text{Int}_\lambda^{K,N}(\mu_0, \delta_{x_1}) := \{\nu_\lambda \in \text{Int}_\lambda(\mu_0, \delta_{x_1}) \text{ satisfying (5.12b) - (5.12c)}\}$$

is not empty, as it contains the measure μ_λ from Definition 5.1; it is narrowly compact by Lemma 5.4, since each $\nu \in \text{Int}_\lambda(\mu_0, \delta_{x_1})$ vanishes outside of E . We shall show that (5.12a) can be satisfied by studying the minimum on $\text{Int}_\lambda^{K,N}(\mu_0, \delta_{x_1})$ of the modified excess functional \mathcal{F}_c defined below. If the minimizer fails to satisfy (5.12a), we derive a contradiction by constructing an admissible perturbation which lowers its excess.

We claim that there is $\nu \in \text{Int}_\lambda^{K,N}(\mu_0, \delta_{x_1})$ for which (5.12a) holds. To this aim, let $c := \frac{\|\rho_0\|_\infty}{(1-\lambda)^N}$ and $\mathcal{F}_c : \text{Dom}(\mathcal{S}_\infty) \rightarrow [0, 1]$ be defined as $\mathcal{F}_c(\mu) := \int [(1 + (\rho - c)_+^2)^{1/2} - 1] d\mathbf{m}$ for $\mu = \rho \mathbf{m}$.

Let $\mu = \rho \mathbf{m}$ be an arbitrary measure in $\text{Int}_\lambda^{K,N}(\mu_0, \delta_{x_1})$. If $\mathcal{F}_c(\mu) = 0$ we are done, otherwise $A := \{x \in M : \rho(x) > c\}$ is Borel and satisfies $\mu(A) > 0$. For $\hat{A} := M \times A$, let $\pi \in \mathcal{P}(M^2)$ be an optimal coupling between μ_0 and μ , and denote the marginals of $\pi' := \mu(A)^{-1} \pi \llcorner \hat{A}$ by $\mu'_0 := (\text{Pr}_1)_* \pi' \leq \mu(A)^{-1} \mu_0$ and $\mu' := (\text{Pr}_2)_* \pi' \leq \mu(A)^{-1} \mu$.

Since $\mathcal{S}_\infty(\mu'_0) \leq \frac{1}{\mu(A)} \mathcal{S}_\infty(\mu_0) - \log \mu(A)$ is finite and $\log(\ell(\cdot, x_1)) \in L^\infty(\mu'_0)$, the $\text{TMCP}_+^h(K, N)$ assumption yields the existence of $\mu'' \in \text{Int}_\lambda^{K,N}(\mu'_0, \delta_{x_1})$ with

$$\mathcal{S}_N(\mu'') \leq (1 - \lambda) \mathcal{S}_N(\mu'_0) \leq -(1 - \lambda) \left(\frac{\|\rho_0\|_\infty}{\mu(A)} \right)^{-\frac{1}{N}} = - \left(\frac{c}{\mu(A)} \right)^{-\frac{1}{N}}.$$

If $B \subset E$ is a Borel set where μ'' is concentrated, Jensen's inequality and the bound just proved yield $\mathfrak{m}(B) \geq \frac{\mu(A)}{c} > \mathfrak{m}(A)$. It follows that μ'' cannot concentrate entirely on A . Set $A' = B \setminus A$ and $\hat{A}' = M \times A'$. Let $\pi'' \in \mathcal{P}(M^2)$ be an optimal coupling between μ'_0 and μ'' , and set $\pi''' := \mu(A')^{-1} \pi'' \llcorner \hat{A}'$, so that $\mu''' := (\text{Pr}_2)_* \pi''' \leq \mu(A')^{-1} \mu''$ lies in $\text{Int}_\lambda(\mu'''_0, \delta_{x_1})$ where $\mu'''_0 := (\text{Pr}_1)_* \pi''' \leq \mu(A')^{-1} \mu'_0$. We have now identified a portion $\mu'''_0 \neq 0$ of μ_0 whose λ -midpoint under π (and π') lies in A , but whose λ -midpoint μ''' under π'' (and π''') lies outside A . For $\varepsilon > 0$ small enough, setting $\mu'''' = (\text{Pr}_2)_* (\frac{d\mu'''}{d\mu_0} \pi)$ makes $\mu^\varepsilon = \mu + \varepsilon(\mu''' - \mu''') \in I_\lambda(\mu_0, \delta_{x_1})$ a perturbation of μ which effectively ‘moves a bit of the mass of μ from above the threshold c to below it’. Note $\frac{d\mu'''}{d\mu_0} \leq \mu(A) \mu''(A')$ is bounded. For $\varepsilon > 0$ smaller still we claim $\mu^\varepsilon \in \text{Int}_\lambda^{K,N}(\mu_0, \delta_{x_1})$ but $\mathcal{F}_c(\mu^\varepsilon) < \mathcal{F}_c(\mu)$. Once this claim is established, choosing μ to minimize the (narrowly) lower semicontinuous excess \mathcal{F}_c on the compact set $\text{Int}_\lambda^{K,N}(\mu_0, \delta_{x_1})$ yields the desired contradiction to conclude the proposition.

The outstanding claim can be argued as follows. All three functionals of interest — $\mathcal{S}_N, \mathcal{S}_\infty$ and \mathcal{F}_c — are given by nonlinearities $s(z) \in \{-z^{1-1/N}, z \log z, (1 + (z - c)_+^2)^{1/2} - 1\}$ which are convex and differentiable on the positive reals. Formally

$$\lim_{\varepsilon \downarrow 0} \int_E \frac{s(\frac{d\mu^\varepsilon}{d\mathfrak{m}}) - s(\rho)}{\varepsilon} d\mathfrak{m} = \int s'(\rho) d(\mu''' - \mu''') < 0, \quad (5.13)$$

by the monotonicity of s' , since μ''' is concentrated on $\{\rho \leq c\}$ while μ'''' is concentrated on $\{\rho > c\}$. Since all three functionals of μ^ε are real-valued and their integrands are convex functions of ε , the dominated and monotone convergence theorems can be used to justify (5.13), arguing separately on the intersections of $\{\rho \leq c\}$ and $\{\rho > c\}$ with $\{\rho \leq 1/e\}$ and $\{\rho > 1/e\}$. For $\varepsilon > 0$ sufficiently small, we now have $\mathcal{S}_N(\mu^\varepsilon) < \mathcal{S}_N(\mu)$, $\mathcal{S}_\infty(\mu^\varepsilon) < \mathcal{S}_\infty(\mu)$ and $\mathcal{F}_c(\mu^\varepsilon) < \mathcal{F}_c(\mu)$ which establishes the desired claim and concludes the proof. \square

Proposition 5.6 (Step 2 — discrete iteration). *With the same assumptions and notation of Proposition 5.5 the following holds. Let $\lambda \in (0, 1)$, put $\nu_{\lambda,0} := \mu_0$ and then let $\nu_{\lambda,k+1} \in \text{Int}_\lambda(\nu_{\lambda,k}, \delta_{x_1})$ be given by Proposition 5.5 above with $\nu_{\lambda,k}$ in place of μ_0 .*

Then putting $t_{\lambda,k} := 1 - (1 - \lambda)^k$, for every $k \in \mathbb{N}$ we have

$$\nu_{\lambda,k} \leq \frac{1}{(1-t_{\lambda,k})^N} e^{Dt_{\lambda,k} \sqrt{K-(N-1)}} \|\rho_0\|_{L^\infty(M,\mathfrak{m})} \mathfrak{m}, \quad (5.14a)$$

$$\mathcal{S}_N(\nu_{\lambda,k}) \leq - \int \tau_{K,N}^{(1-t_{\lambda,k})} \circ \ell(\cdot, x_1) \rho_0^{1-1/N} d\mathfrak{m}; \quad (5.14b)$$

$$\mathcal{S}_\infty(\nu_{\lambda,k}) \leq \mathcal{S}_\infty(\mu_0) - N \log \sigma_{K,N}^{(1-t_{\lambda,k})} (\|\ell(\cdot, x_1)\|_{L^2(\mu_0)}); \quad (5.14c)$$

$$\nu_{\lambda,k'} \in \text{Int}_s(\nu_{\lambda,k}, \delta_{x_1}) \quad \text{for } s = \frac{t_{\lambda,k'} - t_{\lambda,k}}{1-t_{\lambda,k}} = 1 - (1-\lambda)^{k'-k} \quad \forall k' \geq k, \quad (5.14d)$$

where $D := \sup \ell(\text{spt } \mu_0 \times \{x_1\})$.

Proof. The short discussion after the definition (5.10) ensures that the measures $\nu_{\lambda,k}$ are well defined. Let us then introduce $\bar{D}_k := \sup \ell(\text{spt } \nu_{\lambda,k} \times \{x_1\})$ and $\underline{D}_k := \|\ell(\cdot, x_1)^{-1}\|_{L^\infty(\nu_{\lambda,k})}^{-1} > 0$ respectively. Then the definition (5.10) and an induction argument show that

$$\underline{D}_0(1-\lambda)^k \leq \underline{D}_k \leq \bar{D}_k \leq \bar{D}_0(1-\lambda)^k, \quad \text{and} \quad \bar{D}_0 = D. \quad (5.15)$$

Now, letting $\nu_{\lambda,k} = \eta_{\lambda,k} \mathfrak{m}$, from the bound (5.12a) we see that

$$\begin{aligned} \|\eta_{\lambda,k}\|_{L^\infty(M,\mathfrak{m})} &\leq \frac{1}{(1-\lambda)^N} e^{\bar{D}_{k-1} \lambda \sqrt{K-(N-1)}} \|\eta_{\lambda,k-1}\|_{L^\infty(M,\mathfrak{m})} \\ &\leq \dots \leq \frac{1}{(1-\lambda)^{kN}} e^{\lambda \sqrt{K-(N-1)} \sum_{i=0}^{k-1} \bar{D}_i} \|\rho_0\|_{L^\infty(M,\mathfrak{m})}. \end{aligned}$$

Together with (5.15), this gives (5.14a). For (5.14b)–(5.14c) we notice that from the monotonicity of $\theta \mapsto \tau_{K,N}^{(t)}(\theta)$ and $\theta \mapsto \sigma_{K,N}^{(t)}(\theta)$ (they are increasing if $K > 0$ and decreasing if $K < 0$), the

definition of $\nu_{\lambda,k}$ and (5.12b)–(5.12c), we get

$$\mathcal{S}_N(\nu_{\lambda,k}) \leq - \int_M \tau_{K,N}^{(1-\lambda)} \circ \ell(\cdot, x_1) (\eta_{\lambda,k-1})^{1-1/N} d\mathbf{m} \leq \tau_{K,N}^{(1-\lambda)}(\mathbf{D}_{k-1}^*) \mathcal{S}_N(\nu_{\lambda,k-1}),$$

$$\mathcal{S}_\infty(\nu_{\lambda,k}) - \mathcal{S}_\infty(\nu_{\lambda,k-1}) \leq - \log \sigma_{K,N}^{(1-\lambda)}(\|\ell(\cdot, x_1)\|_{L^2(\nu_{\lambda,k-1})}) \leq - \log \sigma_{K,N}^{(1-\lambda)}(\mathbf{D}_{k-1}^*),$$

where \mathbf{D}_{k-1}^* is equal to $\underline{\mathbf{D}}_{k-1}$ if $K > 0$ and to $\bar{\mathbf{D}}_{k-1}$ if $K \leq 0$. With this convention, from (5.15) we see that $-\tau_{K,N}^{(1-\lambda)}(\mathbf{D}_k^*) \leq -\tau_{K,N}^{(1-\lambda)}(\mathbf{D}_0^*(1-\lambda)^k)$ and similarly for $-\sigma_{K,N}^{(1-\lambda)}$. Thus taking also into account the product formulas $\prod_{i=0}^{k-1} \tau_{K,N}^{(1-\lambda)}(\theta(1-\lambda)^i) = \tau_{K,N}^{((1-\lambda)^k)}(\theta)$ and similarly for $\sigma_{K,N}^{(1-\lambda)}$ (these follow from the definitions (5.2) and (5.3) noticing that the products are telescopic), by induction from the above we conclude that

$$\mathcal{S}_N(\nu_{\lambda,k}) \leq \tau_{K,N}^{(1-t_{\lambda,k})}(\mathbf{D}_0^*) \mathcal{S}_N(\mu_0) \quad \forall k \in \mathbf{N}, \quad (5.16a)$$

$$\mathcal{S}_\infty(\nu_{\lambda,k}) \leq \mathcal{S}_\infty(\mu_0) - \log \sigma_{K,N}^{(1-t_{\lambda,k})}(\mathbf{D}_0^*) \quad \forall k \in \mathbf{N}. \quad (5.16b)$$

Now we adapt the estimates (5.16) to establish the desired entropy inequalities. The idea is to decompose the transport into separate parts where the quantity $\ell(\cdot, x_1)$ is ‘‘approximately constant’’ and combine the continuity of the distortion coefficients with a similar argument as for (5.16). Let $\varepsilon > 0$ and set $A_l := \{l\varepsilon < \ell(\cdot, x_1) \leq (l+1)\varepsilon\}$, where $l \in \mathbf{N}_0$. Clearly, these sets are mutually disjoint (and here we benefit from transporting to a Dirac mass); in particular, \mathcal{S}_N behaves additively under the decomposition of $\nu_{\lambda,k}$ into multiples of its restrictions to the A_l ’s. Since any transport toward δ_{x_1} does not mix the masses in different A_l ’s, we see that the measure $\nu_{\lambda,k+1}$ obtained applying Proposition 5.5 to $\nu_{\lambda,k}$ is equal to the suitable linear combinations of the mutually singular measures obtained applying the same statement to the restrictions of $\nu_{\lambda,k}$ to the A_l ’s. The claimed estimate thus follows from repeating the argument from the above paragraph separately for every l and finally sending $\varepsilon \rightarrow 0$. In the case of (5.14c), to obtain the desired bound we use the convexity of $r \in (0, \frac{N}{K-}\pi^2) \mapsto g(r) := \log \sigma_{K,N}^{(t)}(r^{1/2})$ for each $t \in (0, 1)$ that we now verify. When $K < 0$, setting $K = -N$ without loss of generality, $g(r) = \log \frac{\sinh(tr^{1/2})}{\sinh(r^{1/2})}$ yields

$$4r^{1/2}g''(r) = [(e^{4r^{1/2}} - 1) - t(e^{4tr^{1/2}} - 1)] + 4r^{1/2}[e^{2r^{1/2}} - t^2e^{2tr^{1/2}}]$$

with both terms in square brackets being positive. Similarly, $f(r) := \sigma_{N,N}^{(t)}(r^{1/2})$ yields

$$4r^2f''(r) = \left(\frac{r^{1/2}}{\sin r^{1/2}}\right)^2 + r^{1/2} \cot(r^{1/2}) - tr^{1/2} \cot(tr^{1/2}) - \left(\frac{tr^{1/2}}{\sin tr^{1/2}}\right)^2$$

whose positivity follows from the monotonicity of $\theta \in (0, \pi) \mapsto \theta \cot(\theta) + (\frac{\theta}{\sin \theta})^2$.

Finally, (5.14d) holds by induction from (5.11). \square

We are now ready to prove the main result of the section. In the statement below we need a few additional assumptions on the given mm spacetime: that it is forward, as in Definition 2.1 (to ensure existence of suitable limits) and that ℓ_+ is real valued (to be sure that it is bounded from above on emeralds so that suitable rigidity can be extracted from equality in the reverse triangle inequality).

Theorem 5.7 (Step 3 — passage to the limit). *With the same assumptions and notation of Proposition 5.5 and assuming furthermore that $(M, \tau, \ell, \mathbf{m})$ is forward, that $\mathbf{m}(E) < \infty$ for any emerald E , and that ℓ is upper semicontinuous and does not take the value $+\infty$.*

Then there is a curve $(\mu_t) \subset \mathcal{P}_{\text{em}}(M)$ that is a strongly timelike ℓ_q -geodesic from μ to δ_{x_1} for any $0 \neq q < 1$ such that

$$\mu_t \leq \frac{1}{(1-t)^N} e^{Dt\sqrt{K_-(N-1)}} \|\rho\|_{L^\infty(M, \mathbf{m})}, \quad (5.17a)$$

$$\mathcal{S}_N(\mu_t) \leq - \int \tau_{K,N}^{(1-t)} \circ \ell(\cdot, x_1) \rho^{1-1/N} \mathrm{d}\mathbf{m}, \quad (5.17b)$$

$$\mathcal{S}_\infty(\mu_t) \leq \mathcal{S}_\infty(\mu_0) - N \log \sigma_{K,N}^{(1-t)} (\|\ell(\cdot, x_1)\|_{L^2(\mu_0)}), \quad (5.17c)$$

for any $t \in [0, 1]$, where $D := \sup \ell(\mathrm{spt} \mu_0 \times \{x_1\})$.

Proof. We shall build upon Proposition 5.6 and borrow notation from it. Since $\mu_0 \in \mathcal{P}_{\mathrm{em}}(\mathbf{M})$, for suitable compact sets K_0, K_1 we have $\mathrm{spt} \mu_0 \subset J(K_0, K_1)$. It is then clear that $\mathrm{spt} \nu_{\lambda,k} \subset E$ for $E := J(K_0, \{x_1\})$ for any $\lambda \in (0, 1)$, $k \in \mathbf{N}$. For each $T < 1$, the assumption $\mathbf{m}(E) < \infty$ and the estimate (5.14a) yield tightness of the collection $\{\nu_{\lambda,k} : t_{\lambda,k} \leq T\}$.

Now observe that for any $t \in [0, 1]$ and $\lambda_n \downarrow 0$ there is $(k_n) \subset \mathbf{N}$ so that $t_{\lambda_n, k_n} \rightarrow t$ as $n \rightarrow \infty$. This, the tightness just mentioned and a diagonalization argument imply that there is $\mathcal{D} \subset (0, 1)$ dense and $\lambda_n \downarrow 0$ such that for any $t \in \mathcal{D}$ there is $(k_n) \subset \mathbf{N}$ such that $t_{\lambda_n, k_n} \rightarrow t$ and (ν_{λ_n, k_n}) narrowly converges to some η_t as $n \rightarrow \infty$. We also include $0, 1$ in \mathcal{D} and define $\eta_0 := \mu_0$ and $\eta_1 := \delta_{x_1}$. Since $\{\ell \geq 0\}$ is closed, we have $\eta_t \preceq \eta_{t'}$ for all $t, t' \in \mathcal{D}$ with $t \leq t'$, thus by Proposition 2.28 (recall Proposition 2.18) we see that there is a unique $(\mu_t) \in \mathrm{LCC}([0, 1]; \mathcal{P}(\mathbf{M}))$ with $\eta_r \preceq \mu_t \preceq \eta_s$ for any $r < t \leq s$, $r, s \in \mathcal{D}$.

By the stability of the bounds (5.17) w.r.t. narrow convergence it is easy — starting from (5.14a)–(5.14c) — to deduce first that (5.17) hold for η_t with $t \in \mathcal{D}$ and then that these hold for μ_t for any $t \in [0, 1]$.

Thus it only remains to prove that (μ_t) is a strongly timelike ℓ_q -geodesic for any $0 \neq q < 1$. Fix q and notice that from (5.14d) and the very definition of $\mathrm{Int}_\lambda(\mu, \delta_x)$ it easily follows that $\ell_q(\nu_{\lambda,k}, \nu_{\lambda,k'}) \geq (t_{\lambda,k'} - t_{\lambda,k}) \ell_q(\mu_0, \delta_{x_1}) \in \mathbf{R}$. Then the upper semicontinuity in Lemma 2.19 yields $\ell_q(\eta_t, \eta_s) \geq (s - t) \ell_q(\mu_0, \delta_{x_1})$ for any $t < s$, $t, s \in \mathcal{D}$. Now let $t, s \in [0, 1]$ be with $t < s$ and find $t', s' \in \mathcal{D}$ with $t < t' < s' < s$. The construction of (μ_r) ensures that $\mu_t \preceq \eta_{t'} \preceq \eta_{s'} \preceq \mu_s$ which in turn implies $\ell_q(\mu_t, \mu_s) \geq \ell_q(\eta_{t'}, \eta_{s'}) \geq (s' - t') \ell_q(\mu_0, \delta_{x_1})$. Letting $t' \downarrow t$ and $s' \uparrow s$ we conclude from Proposition 2.34 that the curve (μ_r) is a causal ℓ_q -geodesic with $\ell_q(\mu_0, \mu_1) \in \mathbf{R}$.

Also, $\mu_0 \times \delta_{x_1}$ is the only admissible coupling of (μ_0, δ_{x_1}) and by assumption it is concentrated on $\{\ell \in (0, +\infty)\}$. It follows from Proposition 2.39 that any ℓ_q -optimal coupling of (μ_0, μ_1) is also concentrated on $\{\ell \in (0, +\infty)\}$, i.e. that (μ_t) is a strongly timelike ℓ_q -geodesic, as desired. \square

Remark 5.8 (Variants). In relation to Remark 5.2, we notice that the statement of the previous Theorem 5.7 holds almost unchanged by replacing every occurrence of $\mathrm{TMCP}_+^h(K, N)$ by $\mathrm{TMCP}_+^{h,*}(K, N)$. In this case only the slightly worse density bound

$$\mu_t \leq \frac{1}{(1-t)^N} e^{Dt\sqrt{K_-(N)}} \|\rho_0\|_{L^\infty(\mathbf{M}, \mathbf{m})}$$

can be guaranteed. Similarly, if our space is $\mathrm{TMCP}_+^h(K, N')$ for every $N' \geq N$ and an ℓ_q -geodesic which satisfies each of the defining inequalities can be chosen independently of N' , then the same construction produces measures satisfying (5.17a)–(5.17c) with

$$\mathcal{S}_{N'}(\mu_t) \leq - \int \tau_{K,N'}^{(t)} \circ \ell(\cdot, x_1) \rho_0^{1-1/N'} \mathrm{d}\mathbf{m}, \quad \forall N' \geq N,$$

in place of (5.17b). Similar comments apply to Theorem 5.12 below. \blacksquare

We shall typically use Theorem 5.7 above in conjunction with the lifting of ℓ_q -geodesics given in Corollary 2.46 to obtain:

Corollary 5.9 (On existence of initial test plans with Dirac targets). *Let $(\mathbf{M}, \tau, \ell, \mathbf{m})$ be a $\mathrm{TMCP}_+^h(K, N)$ forward mm spacetime so that $\mathbf{m}(E) < \infty$ for any emerald E and ℓ is upper semicontinuous and does not take the value $+\infty$. Let $x_1 \in \mathbf{M}$ and $\mu_0 \in \mathcal{P}_{\mathrm{em}}(\mathbf{M})$ with bounded compression satisfy $\log(\ell(\cdot, x_1)) \in L^\infty(\mu_0)$. Assume also that \mathbf{M} is timelike q -essentially non-branching at 0 for some $0 \neq q < 1$ and that either (A) or (B) of Theorem 2.43 hold.*

Then there exists an initial test plan π with $(\mathbf{e}_0)_ \pi = \mu_0$ concentrated on timelike geodesics γ from γ_0 to x_1 .*

Proof. Our assumptions allow to apply first Theorem 5.7 to find (μ_t) and then Corollary 2.46 to obtain a lifting $\pi \in \mathcal{P}(\text{LCC}([0, 1]; M))$ of (μ_t) concentrated on timelike ℓ -geodesics that lifts (μ_t) and so that $(\mathbf{e}_0, \mathbf{e}_1)_* \pi$ is ℓ_q -optimal (for every $0 \neq q < 1$). Since (μ_t) is a ℓ_q -geodesic from μ_0 to δ_{x_1} , Corollary 2.46 implies that π is concentrated on geodesics γ from γ_0 to x_1 . Also, from the bound (5.17a) we see that $(\mathbf{e}_t)_* \pi \leq C \mathbf{m}$ for every $t \in [0, 1/2]$ and some $C > 0$, so that to conclude it suffices to prove that (μ_t) narrowly converges to μ_0 as $t \downarrow 0$. In turn, this follows from Corollary 2.57, the q -essential timelike non-branching at 0 and the properties of (μ_t) . \square

A useful consequence of Theorem 5.7 is the following (compare with [39, Rem. 3.10]).

Corollary 5.10 (Essential geodesy of $\text{spt } \mathbf{m}$). *Let $(M, \tau, \ell, \mathbf{m})$ be a forward $\text{TMCP}_+^h(K, N)$ mm spacetime in which ℓ is upper semicontinuous and does not take the value $+\infty$ and $\mathbf{m}(E) < +\infty$ for every emerald $E \subset M$. Assume also that either (A) or (B) of Theorem 2.43 hold.*

Then for every $y \in \text{spt } \mathbf{m}$ and \mathbf{m} -a.e. $x \in I^-(y)$ there exists a geodesic $\gamma \in \text{TGeo}(M)$ from x to y with $\gamma_t \in \text{spt } \mathbf{m}$ for every $t \in [0, 1]$.

Proof. Using that the background metric is separable and \mathbf{m} is locally finite, by the Lindelöf property we cover $I^-(y)$ with a countable number of finite \mathbf{m} -measure open sets. Each of these sets is exhausted by a sequence of compact sets of positive \mathbf{m} -measure where $\ell(\cdot, y)$ is bounded away from 0 and $+\infty$ (up to an \mathbf{m} -negligible set). Let μ be the uniform distribution of any such compact set and apply Corollary 2.46 to the geodesic (μ_t) given by Theorem 5.7 to find a lifting π of (μ_t) . The L^∞ -bounds (5.17a) ensure that $\text{spt } \mu_t \subset \text{spt } \mathbf{m}$ for every $t \in [0, 1)$ and thus that for π -a.e. γ we have $\gamma_t \in \text{spt } \mathbf{m}$ for every $t \in \mathbf{Q} \cap [0, 1)$. Since π is concentrated on $\text{TGeo}(M)$, we just proved that there is $\gamma \in \text{TGeo}(M)$ with $\gamma_t \in \text{spt } \mathbf{m}$ for every $t \in \mathbf{Q} \cap [0, 1)$. By left continuity it follows that $\gamma_t \in \text{spt } \mathbf{m}$ for every $t \in [0, 1]$, as desired. \square

Another interesting consequence of Theorem 5.7 is the following result. The argument for the proof is inspired from [59]. Surprisingly however, the non-branching assumption is only required in its weakest form, namely ‘essentially non-branching at 0’, whereas the branching at later times that must be avoided in [59] is permitted in the present context.

Corollary 5.11 (Measurability of causal functions on $\text{TMCP}_+^h(K, N)$ spacetimes). *Let $(M, \tau, \ell, \mathbf{m})$ be a $\text{TMCP}_+^h(K, N)$ forward mm spacetime with τ containing the chronological topology, $\mathbf{m}(E) < \infty$ for every emerald E and in which ℓ is upper semicontinuous and never $+\infty$. Assume also that it is q -essentially non-branching at 0 for some $0 \neq q < 1$.*

Let $A \subset M$ be achronal with closure \bar{A} disjoint from M_{fin} . Then $\mathbf{m}(\bar{A}) = 0$. In particular, if $f : M \rightarrow \bar{\mathbf{R}}$ is a rough (i.e. possibly non-measurable) causal function, then its restriction to $M \setminus M_{\text{fin}}$ is \mathbf{m} -measurable.

Proof. The second statement is a consequence of the first and of Lemma 3.2, so we focus on the first. As τ contains the chronological topology, the closure of an achronal set is still achronal. Thus we can assume $A = \bar{A}$.

We argue by contradiction and assume $\mathbf{m}(A) > 0$. From $A \cap M_{\text{fin}} = \emptyset$ we see that $A \subset \cup_{x \in M} I^-(x)$, so using the Lindelöf property of (the Polish space) A we can find a countable collection (x_n) so that $A \subset \cup_n I^-(x_n)$. Then by interior approximation we can find $x_1 \in M$ and $C \subset A$ compact with $\mathbf{m}(C) > 0$ and $\log(\ell(\cdot, x_1)) \in L^\infty(C, \mathbf{m})$. Apply Theorem 5.7 with $\mu_0 := \mathbf{m}(C)^{-1} \mathbf{m} \llcorner C$ to find (μ_t) as in the statement.

Then the uniform density bounds and the assumption of q -essentially non-branching at 0 (applied to $t \mapsto \mu_{ts}$ for some $s \in (0, 1)$) implies that (μ_t) is narrowly continuous at $t = 0$ (by Corollary 2.57).

Let $U \supset C$ be open with $\mathbf{m}(U) < \frac{3}{2} \mathbf{m}(C)$. The narrow convergence implies $\lim_{t \downarrow 0} \mu_t(U) = 1$ and letting $\mu_t = \rho_t \mathbf{m}$, the bound (5.17a) gives $\liminf_{t \downarrow 0} \mathbf{m}(\{\rho_t > 0\}) \geq \mathbf{m}(C)$. It follows that for some $t > 0$ we have $\mathbf{m}(C \cap \{\rho_t > 0\}) > 0$ and letting $\pi \in \Pi_{\leq}(\mu_0, \mu_t)$ (that exists, as $\mu_0 \preceq \mu_t$), this means that $\pi(\{(x, y) : x, y \in C\}) > 0$. Since by (2.36) we see that $\ell(x, y) = t\ell(x, x_1)$ for π -a.e. (x, y) the assumption $\log(\ell(\cdot, x_1)) \in L^\infty(C, \mathbf{m})$ implies that there are $x, y \in C$ with $\ell(x, y) > 0$, contradicting the achronality of $C \subset A$. \square

Theorem 5.7 works with the target measure being a Dirac mass. An extension of this result holds for more general target measures, provided we add a suitable non-branching assumption to our spacetime and some further assumptions on τ . Specifically, we will ask that either (A) or (B) of Theorem 2.43 hold: this is necessary as we will need to lift suitable geodesics appearing in the statement and proof. Concerning non-branching: as shown by Gigli [59] in positive signature, a non-branching assumption together with a lower Ricci bound in the form of a CD condition implies existence of optimal transport maps. This has been extended to non-branching MCP spaces by Cavalletti–Mondino [36] and then to timelike non-branching TMCP spacetimes in [39]. This latter fact will be used in the proof of the next result, that closely follows the construction in positive signature given in [36, Prop. 4.3].

Theorem 5.12 (On uniqueness of ℓ_q -geodesics and existence of optimal maps). *Let $(M, \tau, \ell, \mathbf{m})$ be a forward TMCP $_+^h(K, N)$ mm spacetime with $\mathbf{m}(E) < \infty$ for any emerald E and in which ℓ_+ is continuous and real valued. Also, let $\mu_0, \bar{\mu}_1 \in \mathcal{P}_{\text{em}}(M)$ with $\mu_0 = \rho_0 \mathbf{m}$ having bounded compression and $0 \neq q < 1$. Assume furthermore that:*

- either (A) or (B) of Theorem 2.43 hold,
- M is timelike q -essentially forward non-branching,
- M is timelike q -essentially non-branching at 0,
- $\ell_q(\mu_0, \bar{\mu}_1) \in (0, +\infty)$ and there is an ℓ_q -optimal coupling $\bar{\pi} \in \Pi_{\leq}(\mu_0, \bar{\mu}_1)$ concentrated on $\{\ell \in (c, \frac{1}{c})\}$ for some $c > 0$.

Then:

- i) There is a unique ℓ_q -geodesic (μ_t) from μ_0 to $\bar{\mu}_1$ and it is strongly ℓ_q -timelike;
- ii) (μ_t) admits a unique lifting $\pi \in \mathcal{P}(\text{LCC}([0, 1]; M))$ and is induced by a map, i.e. there is $F : M \rightarrow \text{LCC}([0, 1]; M)$ with $\mathbf{e}_0 \circ F$ being the identity μ_0 -a.e. and such that $\pi = F_* \mu_0$;
- iii) we have

$$\ell(x, F_1(x)) \in [c, \frac{1}{c}] \quad \mu_0 - \text{a.e. } x; \quad (5.18)$$

- iv) for every $t \in [0, 1)$ we have

$$\mu_t \leq \frac{1}{(1-t)^N} e^{Dt\sqrt{K_-(N-1)}} \|\rho_0\|_{L^\infty(M, \mathbf{m})} \mathbf{m}, \quad (5.19a)$$

$$\mathcal{S}_N(\mu_t) \leq - \int \tau_{K, N}^{(1-t)} \circ \ell(\gamma_0, \gamma_1) \rho_0(\gamma_0)^{-1/N} d\pi(\gamma), \quad (5.19b)$$

$$\mathcal{S}_\infty(\mu_t) \leq \mathcal{S}_\infty(\mu_0) - N \log \min\{\sigma_{K, N}^{(1-t)}(c), \sigma_{K, N}^{(1-t)}(1/c)\}. \quad (5.19c)$$

Proof. We will first consider the situation of $\bar{\mu}_1$ being a convex combination of Dirac masses, i.e. $\bar{\mu}_1 := \sum_i \lambda_i \delta_{y_i}$ for given $\lambda_1, \dots, \lambda_n \in (0, 1)$ and mutually distinct points $y_1, \dots, y_n \in \text{spt } \mathbf{m}$. For $i = 1, \dots, n$, let $\pi_i := \lambda_i^{-1} \pi \llcorner M \times \{y_i\}$ and $\mu_{0, i} := (\text{Pr}_1)_* \pi_i$. Then let $(\mu_{t, i}) \subset \mathcal{P}_{\text{em}}(M)$ be given by Theorem 5.7 applied to $\mu_{0, i}$ and δ_{y_i} and then $\pi_i \in \mathcal{P}(\text{LCC}([0, 1]; M))$ be a lifting of it as in Corollary 2.46. Also, let $\alpha_i := (\mathbf{e}_1)_* \pi_i \times \delta_{y_i}$ be the only coupling between $(\mathbf{e}_1)_* \pi_i$ and δ_{y_i} (hence ℓ_q -optimal). Then we have

$$\begin{aligned} \ell_q(\mu_{1, i}, \delta_{y_i}) &\geq \left(\int \ell^q(\gamma_0, \gamma_1) d\pi_i(\gamma) \right)^{\frac{1}{q}} + \left(\int \ell^q(x, y) d\alpha_i(x, y) \right)^{\frac{1}{q}} \\ &\geq \ell_q(\mu_{1, i}, \delta_{y_i}) + \left(\int \ell^q(x, y) d\alpha_i(x, y) \right)^{\frac{1}{q}}. \end{aligned}$$

The assumption on π ensures that $\ell_q(\mu_{1, i}, \delta_{y_i}) \in (0, +\infty)$, thus inspecting the equality case in (2.9) we deduce that α_i is concentrated on $\{(x, y) : \ell(x, y) = 0\}$.

We claim that $\mu_i \perp \mu_j$ for $i \neq j$. In positive signature, this follows from the entropy estimates coming from lower Ricci bounds, a non-branching assumption and the narrow continuity of W_2 -geodesics (see [59]). These arguments carry over also in our setting (see also [39, Thm. 3.20] and [21, Thm. 4.16]), where we remark that the assumption of timelike q -essentially non-branching at 0 is used — via Corollary 2.57 — to get the desired narrow continuity at 0. Notice that with respect to the above references, working with $q < 0$ causes no additional difficulties and that the compactness of emeralds used in the proof of this result in [39, 21] is bypassed by our requirement that they have finite \mathfrak{m} -measure; c.f. Remark 1.2.

Thus $\mu_i \perp \mu_j$ for $i \neq j$ and since $y_i \neq y_j$ as well for $i \neq j$, a further use of the non-branching assumption tells that $(\mathbf{e}_t)_* \pi_i \perp (\mathbf{e}_t)_* \pi_j$ for every $t \in [0, 1]$ and $i \neq j$. Putting $\pi := \sum_i \lambda_i \pi_i$ and $\alpha := \sum_i \lambda_i \alpha_i$, we thus have that $\alpha \in \Pi_{\leq}((\mathbf{e}_1)_* \pi, \bar{\mu}_1)$ and $(\mathbf{e}_t)_* \pi = \rho_t \mathfrak{m} = \mu_t$ with

$$\begin{aligned} \|\rho_t\|_{L^\infty(M, \mathfrak{m})} &= \max_i \lambda_i \|\rho_{i,t}\|_{L^\infty(M, \mathfrak{m})}, & \text{where } \rho_{i,t} &:= \frac{d(\mathbf{e}_t)_* \pi_i}{d\mathfrak{m}}, \\ \mathcal{S}_N(\mu_t) &= \sum_i \lambda_i^{1-1/N} \mathcal{S}_N(\mu_{i,t}), \\ \mathcal{S}_\infty(\mu_t) &\leq \sum_i \lambda_i \mathcal{S}_\infty(\mu_{i,t}), \end{aligned}$$

so that π satisfies (5.19) as a consequence of the bounds (5.17) for the $\pi_{i,t}$'s. The fact that (μ_t) is a ℓ_q -geodesic from μ_0 to $\bar{\mu}_1$ follows from the construction (notice that α is concentrated on $\{\ell = 0\}$) and (5.18) follows from Proposition 2.39.

Now assume that $\text{spt } \mu_0 \times \text{spt } \bar{\mu}_1 \subset \{\ell \in (c, \frac{1}{c})\}$. Since $\{\ell \in (c, \frac{1}{c})\}$ is open, we can find a sequence $(\bar{\mu}_1^n)$ of convex combinations of Dirac masses narrowly converging to $\bar{\mu}_1$ and so that $\text{spt } \mu_0 \times \text{spt } \bar{\mu}_1^n \subset \{\ell \in (c, \frac{1}{c})\}$ for every $n \in \mathbf{N}$ so that there are ℓ_q -optimal couplings $\bar{\pi}^n$ for $(\mu_0, \bar{\mu}_1^n)$ narrowly converging to some $\bar{\pi}'$ that, by the continuity of ℓ in $\{\ell \in (c, \frac{1}{c})\}$, is ℓ_q -optimal. It is not hard to see that we can choose the $\bar{\mu}_1^n$'s so that they all have support in a fixed emerald E , that we can assume to contain also $\text{spt } \mu_0$. The construction and the continuity of ℓ_+ ensure that $\ell_q(\mu, \bar{\mu}_1^n) \rightarrow \ell_q(\mu, \bar{\mu}_1)$ as $n \rightarrow \infty$. As before, there are $\pi^n \in \mathcal{P}(\text{LCC}([0, 1]; M))$ as above for $\mu_0, \bar{\mu}_1^n$ and notice that the construction yields $(\mathbf{e}_1)_* \pi^n \preceq \bar{\mu}_1^n$. We claim that these are tight. If (A) holds, this is a quite direct consequence of the tightness of $(\bar{\mu}_1^n)$ and the same arguments in the proof of Theorem 2.43. If (B) holds we use the uniform L^∞ estimates (5.19a) and the fact that $\mathfrak{m} \llcorner E$ is tight to conclude that $((\mathbf{e}_t)_* \pi^n)$ is tight for every $t < 1$. This suffices, as in the proof of Theorem 2.43, to obtain tightness.

Hence a non-relabelled subsequence narrowly converges to a limit π . Since the relevant couplings are concentrated on $\{\ell \in (c, \frac{1}{c})\}$ and ℓ_+ is continuous, using also Lemma 2.38 we have

$$u_q(\ell_q(\mu_0, \bar{\mu}_1)) = \lim_{n \rightarrow \infty} u_q(\ell_q(\mu_0, \bar{\mu}_1^n)) = \lim_{n \rightarrow \infty} \int \mathcal{A}_q(\gamma) d\bar{\pi}^n(\gamma) \leq \int \mathcal{A}_q(\gamma) d\pi(\gamma).$$

Arguing as in the proof of Theorem 2.43 to pass to the limit in the marginals we see that $(\mathbf{e}_1)_* \pi \preceq \bar{\mu}_1$, thus the last claim in Corollary 2.46 tells that $t \mapsto \mu_t := (\mathbf{e}_t)_* \pi$ is a ℓ_q -geodesic from μ_0 to $\bar{\mu}_1$.

Then again using the same techniques adopted in proving Theorem 2.43 when passing to the limit at the level of plans in $\mathcal{P}(\text{LCC}([0, 1]; M))$, we can pass to the limit in (5.19) stated for the π^n 's and prove that the same bounds hold for π : for (5.19a) this is trivial, passing to the limit in (5.19b) is a bit more technical, but nevertheless doable, see also the arguments in [114, Lemma 3.3]. The fact that π is induced by a map will be (briefly) mentioned in a moment; from that the bound (5.18) is trivial from the construction

Removing the assumption that $\text{spt } \mu_0 \times \text{spt } \bar{\mu}_1 \subset \{\ell \in (c, \frac{1}{c})\}$ is done by decomposing the support of the given π into countably many suitable rectangles, see e.g. the proof of [21, Prop. 3.38].

It thus remains to show the uniqueness properties in (i), (ii). These, however, follow via the same arguments used to prove $\mu_i \perp \mu_j$ above, see for instance [65] for the proof in positive signature. \square

5.3 A converse Hawking–King–McCarthy theorem

Building on the results proved in the previous section, in this one we study the interplay of our Sobolev calculus and the “metric” features of the given metric measure spacetime $(M, \tau, \ell, \mathfrak{m})$. Our motivation comes from a classical result of Hawking–King–McCarthy [68]: if (M_1, g_1) and (M_2, g_2) are two spacetimes of the same dimension, the first being strongly causal, then every map $F: M_1 \rightarrow M_2$ which preserves the respective time separations is in fact a smooth isometry, in particular $F_*g_1 = g_2$. (The converse is obvious.) Our analog of this result is stated in Theorem 5.16 below. In contradistinction to the smooth setting of Hawking–King–McCarthy, we use ℓ to define the (maximal weak sub)slope $|df|$ instead of the other way around, so it is the obvious direction from that setting which becomes challenging for us to prove.

The positive-signature predecessor of this result from metric measure geometry that inspires us is due to Gigli [63, Lem. 4.19, Prop. 4.20] in the context of the proof of the splitting theorem for RCD spaces. The so-called Sobolev-to-Lipschitz property introduced there [63, §4.1.3] was pioneered to characterize metric measure and Sobolev isometries.

We turn to the technical content. In positive signature, 1-Lipschitz functions fully characterize the underlying distance. Analogously, in our setting 1-step functions fully characterize the underlying time separations, via the duality formula

$$\ell(x, y) = \inf\{f(y) - f(x) : f : M \rightarrow \bar{\mathbf{R}} \text{ is 1-step}\}, \quad \forall x, y \in M, \quad (5.20)$$

valid on any metric spacetime (M, ℓ) . Indeed, \leq comes from the definition of 1-steepness, while for \geq we pick $f(z) := \ell(x, z)$ (in the case of smooth spacetimes, the duality formula characterizes stability of the spacetime [88, Thm. 4.6]). If we are on a metric measure spacetime, then we know from (3.15) that every 1-step function f satisfies $|df| \geq 1$ \mathfrak{m} -a.e., therefore in this case we have

$$\ell(x, y) \geq \inf\{f(y) - f(x) : f : \text{spt } \mathfrak{m} \rightarrow \bar{\mathbf{R}} \text{ causal, } |df| \geq 1 \text{ } \mathfrak{m}\text{-a.e.}\}, \quad \forall x, y \in M. \quad (5.21)$$

Understanding whether equality holds here is relevant if one wants to use the Sobolev-like calculus we just developed in order to derive precise metric information on the underlying spacetime. Still, without further assumptions, we cannot expect equality to hold in (5.21), as there could be a gap between the class of 1-step functions and that of functions with $|df| \geq 1$ \mathfrak{m} -a.e. It is then natural to propose the following:

Definition 5.13 (Sobolev-to-steepness property). *Let $(M, \tau, \ell, \mathfrak{m})$ be a mm spacetime. We say that it has the Sobolev-to-steepness property if every causal function $f: M \rightarrow \bar{\mathbf{R}}$ with $|df| \geq 1$ \mathfrak{m} -a.e. is 1-steep on $\text{spt } \mathfrak{m}$.*

Notice that unlike the positive signature case, here there is no need to pass to \mathfrak{m} -representatives, as the property of being causal already depends on the value of the function at any point.

From the previous consideration it is now easy to see that the following holds:

Proposition 5.14 (Duality formula). *A metric measure spacetime $(M, \tau, \ell, \mathfrak{m})$ satisfies the Sobolev-to-steepness property if and only if for every $x, y \in \text{spt } \mathfrak{m}$ we have*

$$\ell(x, y) = \inf\{f(y) - f(x) : f : \text{spt } \mathfrak{m} \rightarrow \bar{\mathbf{R}} \text{ causal, } |df| \geq 1 \text{ } \mathfrak{m}\text{-a.e.}\}. \quad (5.22)$$

Proof. “If”. Any causal function $f: \text{spt } \mathfrak{m} \rightarrow \bar{\mathbf{R}}$ with $|df| \geq 1$ \mathfrak{m} -a.e. is admissible in (5.22), hence all points $x, y \in \text{spt } \mathfrak{m}$ satisfy the inequality $f(y) - f(x) \geq \ell(x, y)$. This is to say f is 1-steep on $\text{spt } \mathfrak{m}$.

“Only if”. Inequality \geq holds by (5.21). For \leq we notice that every function f in the infimum is 1-steep on $\text{spt } \mathfrak{m}$ by the assumed Sobolev-to-steepness property, thus we conclude by applying (5.20) on $\text{spt } \mathfrak{m}$. \square

In the following, given a monotone map $T: \text{spt } \mathfrak{m}_1 \rightarrow \text{spt } \mathfrak{m}_2$ between two given forward mm spacetimes $(M_1, \tau_1, \ell_1, \mathfrak{m}_1)$ and $(M_2, \tau_2, \ell_2, \mathfrak{m}_2)$ we define the map $\mathcal{T}: \text{LCC}([0, 1]; \text{spt } \mathfrak{m}_1) \rightarrow$

$\text{LCC}([0, 1]; \text{spt } \mathbf{m}_2)$ as the one sending γ to the element of $\text{LCC}([0, 1]; \text{spt } \mathbf{m}_2)$ associated to the monotone map $T \circ \gamma$ via Proposition 2.28 (in other words, $\mathcal{J}(\gamma)_0 := T(\gamma_0)$ and $\mathcal{J}(\gamma)_t := \lim_{s \uparrow t} T(\gamma_s)$ for $t > 0$). Notice that by Lemma 2.27 we see that $T(\gamma_t) = \mathcal{J}(\gamma)_t$ for $t = 0$ and every $t \in (0, 1]$ except at most a countable number, thus the very definition of the (distance D inducing the) topology of $\text{LCC}([0, 1]; \text{spt } \mathbf{m}_2)$ shows that if T is Borel, then so is \mathcal{J} .

Lemma 5.15 (One-sided nonsmooth Hawking–King–McCarthy theorem). *Let $(M_1, \tau_1, \ell_1, \mathbf{m}_1)$ and $(M_2, \tau_2, \ell_2, \mathbf{m}_2)$ be two forward mm spacetimes and $T: \text{spt } \mathbf{m}_1 \rightarrow \text{spt } \mathbf{m}_2$ be a bijective Borel map. Assume that $T_* \mathbf{m}_1 \leq C \mathbf{m}_2$ for some constant $C > 0$ and let $\mathcal{J}: \text{LCC}([0, 1]; \text{spt } \mathbf{m}_1) \rightarrow \text{LCC}([0, 1]; \text{spt } \mathbf{m}_2)$ be defined as above.*

Consider the two statements:

- i) $T: \text{spt } \mathbf{m}_1 \rightarrow M_2$ is 1-steep, i.e. for every $x, y \in \text{spt } \mathbf{m}_1$ we have $\ell_2(T(x), T(y)) \geq \ell_1(x, y)$.
- ii) If $g: M_2 \rightarrow \bar{\mathbf{R}}$ is causal then so is $g \circ T: M_1 \rightarrow \bar{\mathbf{R}}$ and $|\text{d}(g \circ T)|_1 \geq |\text{d}g|_2 \circ T$ \mathbf{m}_1 -a.e.

Then (i) \Rightarrow (ii) and, conversely, if $(M_1, \ell_1, \mathbf{m}_1)$ satisfies the Sobolev-to-steepness property we also have (ii) \Rightarrow (i).

Proof. (i) \Rightarrow (ii). That $g \circ T$ is causal is obvious. To show the stated bound, let π be a test plan on M_1 . We claim that $\mathcal{J}_* \pi$ is a test plan on M_2 and since it is clearly a measure on $\text{LCC}([0, 1]; M_2)$ all we need to do is to show that $(\mathbf{e}_t)_* \mathcal{J}_* \pi \leq C' \mathbf{m}_2$ for every $t \in [0, 1]$ and some $C' > 0$. We have already noticed for every $\gamma \in \text{LCC}([0, 1]; M_1)$ that $\mathcal{J}(\gamma)_0 = T(\gamma_0)$ and $\mathcal{J}(\gamma)_t = T(\gamma_t)$ holds except at a countable number of $t \in (0, 1]$. It follows that $(\mathbf{e}_0)_* \mathcal{J}_* \pi = T_*(\mathbf{e}_0)_* \pi \leq CC'' \mathbf{m}_2$, where $C'' > 0$ is so that $(\mathbf{e}_t)_* \pi \leq C'' \mathbf{m}_1$ for every $t \in [0, 1]$. Also, for $s, t \in [0, 1]$ with $s < t$ we have $\int_s^t (\mathbf{e}_r)_* \mathcal{J}_* \pi = \int_s^t T_*(\mathbf{e}_r)_* \pi \leq (s - t)CC'' \mathbf{m}_2$, so that dividing by $s - t$ and letting $s \uparrow t$ using that $\mathcal{J}_* \pi$ is concentrated on left continuous paths we conclude that $(\mathbf{e}_t)_* \mathcal{J}_* \pi \leq CC'' \mathbf{m}_2$, establishing that $\mathcal{J}_* \pi$ is a test plan, as desired.

Now the claim follows by duality. Let $g: M_2 \rightarrow \bar{\mathbf{R}}$ be causal, note that $g \circ T$ is a causal function on $\text{spt } \mathbf{m}_1$ by (i). For a test plan π on M_1 we just proved that $\mathcal{J}_* \pi$ is a test plan on M_2 , therefore,

$$\begin{aligned} \int g \circ T(\gamma_1) - g \circ T(\gamma_0) \, \text{d}\pi(\gamma) &= \int g(\sigma_1) - g(\sigma_0) \, \text{d}\mathcal{J}_* \pi(\sigma) \\ &\geq \iint_0^1 |\text{d}g|_2(\sigma_r) |\dot{\sigma}_r| \, \text{d}r \, \text{d}\mathcal{J}_* \pi(\sigma) \geq \iint_0^1 |\text{d}g|_2(T(\gamma_r)) |\dot{\gamma}_r| \, \text{d}r \, \text{d}\pi(\gamma). \end{aligned}$$

In the last step, we have employed the hypothesized noncontractivity (i) and Theorem 2.23. This shows $|\text{d}g|_2 \circ T$ is a weak subslope of $g \circ T$. The desired inequality thus follows from the maximality asserted by Theorem 3.14.

(ii) \Rightarrow (i). Let $x, y \in M_1$ and then $g: M_2 \rightarrow \bar{\mathbf{R}}$ be defined as $g(z) := \ell_2(T(x), z)$. Then g is 1-steep, hence (by (3.15)) $|\text{d}g|_2 \geq 1$ \mathbf{m}_2 -a.e. and thus our assumptions ensure that $g \circ T$ is causal with $|\text{d}(g \circ T)|_1 \geq 1$ \mathbf{m}_1 -a.e. Since M_1 has the Sobolev-to-steepness property we deduce that $g \circ T$ is 1-steep on $\text{spt } \mathbf{m}_1$, thus in particular $g(T(y)) - g(T(x)) \geq \ell_1(x, y)$. Since the left hand side of this identity is equal to $\ell_2(T(x), T(y))$, the proof is complete. \square

Theorem 5.16 (Nonsmooth Hawking–King–McCarthy theorem). *Assume that $(M_1, \tau_1, \ell_1, \mathbf{m}_1)$ and $(M_2, \tau_2, \ell_2, \mathbf{m}_2)$ are forward mm spacetimes with the Sobolev-to-steepness property. Let $T: \text{spt } \mathbf{m}_1 \rightarrow \text{spt } \mathbf{m}_2$ be a surjective measure-preserving Borel map, i.e. $T_* \mathbf{m}_1 = \mathbf{m}_2$.*

Then the following are equivalent.

- (i) T is an isometry of metric measure spacetimes, i.e. for every $x, y \in \text{spt } \mathbf{m}$,

$$\ell_2(T(x), T(y)) = \ell_1(x, y). \quad (5.23)$$

- (ii) $f: M_2 \rightarrow \bar{\mathbf{R}}$ is causal if and only if so is $f \circ T: M_1 \rightarrow \bar{\mathbf{R}}$ and in this case $|\text{d}(f \circ T)|_1 = |\text{d}f|_2 \circ T$ holds \mathbf{m}_1 -a.e.

Proof. We shall apply Lemma 5.15 first to T and then to its inverse. To do so we must first verify that T is invertible. If (i) holds, the condition $T(x) = T(y)$ and (5.23) imply $\ell_1(x, y) = \ell_1(y, x) = 0$, which forces $x = y$. If (ii) holds a similar argument works from the causality preservation. Thus in either case T is injective, thus bijective (as we assumed surjectivity) and the inverse is also Borel (see e.g. [69, Cor. 15.2]). \square

It remains to find sufficient conditions ensuring that the Sobolev-to-steepness property holds. An answer is given by the next result:

Theorem 5.17 (Sobolev-to-steepness property from curvature conditions). *Let $(M, \tau, \ell, \mathbf{m})$ be a q -essentially timelike non-branching at 0, forward $\text{TMCP}_+^h(K, N)$ mm spacetime with τ containing the chronological topology, $\mathbf{m}(E) < \infty$ for each emerald E , and with ℓ upper semicontinuous and never $+\infty$. Assume also that either (A) or (B) of Theorem 2.43 hold.*

Let $f : M \rightarrow \mathbf{R}$ be causal with $|df| \geq 1$ \mathbf{m} -a.e. Then for every $y \in M$ we have

$$f(y) - f(x) \geq \ell(x, y) \quad \mathbf{m} - \text{a.e. } x. \quad (5.24)$$

In particular, if ℓ_+ is continuous, real valued and for all $x \ll y$ with $x, y \in \text{spt } \mathbf{m}$ we have $\mathbf{m}(U \cap J(x, y)) > 0$ for every neighbourhood U of x , then M has the Sobolev-to-steepness property.

Proof. The second claim is a trivial consequence of the first, so we focus on this one. The claimed inequality $f(y) - f(x) \geq \ell(x, y)$ is clear whenever $\ell(x, y) \in \{-\infty, 0\}$, thus we can assume $x \ll y$. In this case, by our infinity conventions (Section 2.1) the conclusion holds if either x or y do not belong to $\text{Dom}(f)$. Thus, we only have to treat the case $x, y \in \text{Dom}(f)$ with $x \ll y$.

By Corollary 5.10 for \mathbf{m} -a.e. $x \in I^-(y)$ there is a timelike geodesic γ from x to y with image contained in $\text{spt } \mathbf{m}$. Fix such x and γ : we shall prove that (5.24) holds for such x and this suffices to conclude.

By Lemma 3.2, γ crosses the set of discontinuity points of f only countably many times. Hence, fix $s, t \in [0, 1]$ with $s < t$ such that f is continuous at γ_s and γ_t and let $\varepsilon > 0$ be with $s + \varepsilon < t$. Let the metric d metrize our Polish topology of M . For sufficiently small $r > 0$, the induced ball $B_r(\gamma_s)$ has finite and positive \mathbf{m} -measure (since $\gamma_s \in \text{spt } \mathbf{m}$) and is contained in $I^+(\gamma_0) \cap I^-(\gamma_t)$, as this set is open. In particular, f is bounded on $B_r(\gamma_s)$.

Since γ is timelike, interior approximation yields a compact set $C \subset B_r(\gamma_s)$ with $\mathbf{m}(C) > 0$ and $\log(\ell(\cdot, \gamma_t)) \in L^\infty(C, \mathbf{m} \llcorner C)$. Apply Corollary 5.9 with μ_0 being the uniform distribution of C and x_1 being γ_t to get the existence of a plan π^r concentrated on timelike geodesics η from η_0 to γ_t . The L^∞ bounds in Theorem 5.7 ensure that for any $\xi \in (0, 1)$ the plan $\pi_\xi := (\text{restr}_0^{1-\xi})_* \pi^r$ is a test plan. Since f is causal we get

$$\begin{aligned} f(\gamma_t) - \int f d(\mathbf{e}_0)_* \pi^r &= f(\gamma_t) - \int f d(\mathbf{e}_0)_* \pi_\xi \geq \int [f(\sigma_1) - f(\sigma_0)] d\pi_\xi(\sigma) \\ (\pi_\xi \text{ is test}) &\geq \iint_0^1 |df|(\sigma_u) |\dot{\sigma}_u| du d\pi_\xi(\sigma) \\ (|df| \geq 1 \text{ } \mathbf{m}\text{-a.e. and } \pi_\xi[\text{TGeo}(M)] = 1) &\geq \int \ell(\sigma_0, \sigma_1) d\pi_\xi(\sigma) \\ &= (1 - \xi) \int \ell(\cdot, \gamma_t) d(\mathbf{e}_0)_* \pi^r. \end{aligned} \quad (5.25)$$

Now notice that for $r > 0$ sufficiently small the ball $\bar{B}_r(\gamma_s)$ is contained in the chronological past of $\gamma_{s+\varepsilon}$, hence the reverse triangle inequality tells that $\int \ell(\cdot, \gamma_t) d(\mathbf{e}_0)_* \pi^r \geq \ell(\gamma_{s+\varepsilon}, \gamma_t) = (t - s - \varepsilon) \ell(\gamma_0, \gamma_1)$ for r small. Thus letting first $r \downarrow 0$ then $\xi \downarrow 0$ in (5.25) using also that γ_s is a continuity point of f we get

$$f(y) - f(x) \geq f(\gamma_t) - f(\gamma_s) \geq (t - s - \varepsilon) \ell(\gamma_0, \gamma_1).$$

First letting $\varepsilon \rightarrow 0$, then $s \downarrow 0$ and $t \uparrow 1$ gives the claim. \square

5.4 A metric Brenier–McCann theorem

The purpose of this part is to link optimal transport to our Sobolev calculus. More concretely, in Theorem 5.19 we prove that optimal geodesic plans represent the initial gradients of their Kantorovich potentials after Definition 4.2. Ultimately, this will be used to prove d’Alembert comparison theorems for such potentials in the next section. Theorem 5.19 itself is a variant of Brenier’s famous polar factorization theorem [27] generalized by McCann to Riemannian manifolds [83]. Its metric counterpart is due to Ambrosio–Gigli–Savaré [10] and Gigli [60]. Versions of related results appear implicitly in the recent literature about optimal transport in Lorentzian spacetimes: see Suhr [115], McCann [84], and Remark 5.20 below.

We first recapitulate some basic notions about Kantorovich duality in the Lorentzian setting. We refer to McCann [84], Mondino–Suhr [93], and Cavalletti–Mondino [39] for details (and to Villani [119] for classical Kantorovich duality), even though our presentation differs slightly from the one in these references, as we are going to allow extended real valued cost functions and make use of the conventions in Section 2.1 in formulas such as (5.26) below.

Let us thus consider a cost function $c : M \times M \rightarrow \bar{\mathbf{R}}$ on a given Polish space M . For $\psi : M \rightarrow \bar{\mathbf{R}}$ we define the two functions $\psi_c, \psi^c : M \rightarrow \bar{\mathbf{R}}$ as

$$\psi_c(x) := \inf_{y \in M} (-c(x, y) + \psi(y)) \quad \text{and} \quad \psi^c(y) := \sup_{x \in M} (c(x, y) + \psi(x)). \quad (5.26)$$

A function $f := M \rightarrow \bar{\mathbf{R}}$ is called c -concave if it is of the form ψ_c for some ψ . If the cost function c never attains the value $+\infty$, then it is easy to see that

$$(\psi_c)^c \leq \psi \quad \text{and} \quad (\psi^c)_c \geq \psi \quad (5.27)$$

for any ψ (because $a + (-a + b) \leq b$ and $(-a) + (a + b) \geq b$ for any $a \in \mathbf{R} \cup \{-\infty\}$, $b \in \bar{\mathbf{R}}$). It follows that f is c -concave iff $f = (f^c)_c$. In particular, we have

$$f^c(y) \geq c(x, y) + f(x) \quad \forall x, y \in M. \quad (5.28)$$

The set of couples for which both sides are real numbers and equal is called the c -superdifferential of f and denoted $\partial_c f \subset M \times M$. Equivalently, we have

$$(x, y) \in \partial_c f \iff f(x), -f^c(y) > -\infty, c(x, y) \in \mathbf{R} \quad \text{and} \quad f(x) + c(x, y) \geq f^c(y). \quad (5.29)$$

We are interested in Polish metric spacetimes (M, τ, ℓ) and in the cost $c_q := u_q \circ \ell$ for $0 \neq q < 1$. It is clear that in this case functions of the form ψ_{c_q} or ψ^{c_q} are rough causal — meaning causal functions except possibly without any measurability. Indeed, if $x \leq y$, the reverse triangle inequality implies $\ell(x, z) \geq \ell(y, z)$ for every $z \in M$, thus $-u_q(\ell(x, z)) \leq -u_q(\ell(y, z))$ and therefore

$$\psi_{c_q}(x) = \inf_{z \in M} \left[(-u_q(\ell(x, z))) + \psi(z) \right] \leq \inf_{z \in M} \left[(-u_q(\ell(y, z))) + \psi(z) \right] = \psi_{c_q}(y),$$

showing that ψ_{c_q} is rough causal. One can argue similarly for ψ^{c_q} . Other than this, and the regularity this implies according to Lemma 3.2 and Corollary 5.11, it seems hard to obtain any regularity at all, the problem being that for $q \in (0, 1)$ the cost function is not lower semicontinuous (regardless of any assumed continuity of ℓ_+).

However, for $q < 0$ the rough causality of c_q -concave functions is upgraded to causality by:

Proposition 5.18 (On upper semicontinuity of c_q -concave functions for $q < 0$). *Let (M, τ, ℓ) be a Polish metric spacetime with ℓ_+ continuous and $q < 0$. Also, let $f : M \rightarrow \bar{\mathbf{R}}$ be c_q -concave.*

Then f and $-f^{c_q}$ are upper semicontinuous and $\partial_{c_q} f \cap \{\ell \in (0, +\infty)\}$ is σ -closed in $\{\ell \in (0, +\infty)\}$.

Proof. For $q < 0$ the function $-u_q : \bar{\mathbf{R}} \rightarrow \bar{\mathbf{R}}$ is continuous, non-increasing and equal to $+\infty$ in 0. Hence the continuity of ℓ_+ ensures that $(x, y) \mapsto -u_q(\ell(x, y))$ is upper semicontinuous. Since the infimum of an arbitrary family of upper semicontinuous functions is still upper semicontinuous, the claims about f and $-f^{c_q}$ follow. In particular, for every $n \in \mathbf{N}$ the sets $\{f \geq -n\}$ and $\{f^{c_q} \leq n\}$ are closed and since $u_q \circ \ell$ is continuous in $\{\ell \in [\frac{1}{n}, n]\}$, the last claim also follows. \square

The next theorem is the main result of the section. Recall the definition of the backward slope $|\partial^- f|$ from (3.6) and the notion of a plan representing the initial p -gradient of a causal function from Definition 4.2.

Theorem 5.19 (A metric Brenier–McCann theorem). *Let $(M, \tau, \ell, \mathfrak{m})$ be a forward mm spacetime satisfying either (A) or (B) in Theorem 2.43 and so that ℓ_+ is continuous and real valued. Also, let $\frac{1}{p} + \frac{1}{q} = 1$ be with $0 \neq p, q < 1$ and $f: M \rightarrow \bar{\mathbf{R}}$ a c_q -concave and \mathfrak{m} -measurable function.*

Let π be an initial test plan concentrated on $\text{TGeo}(M)$ such that $(\mathbf{e}_0, \mathbf{e}_1)_ \pi$ is concentrated on $\partial_{c_q} f$ and assume that:*

- i) for some Borel set $E \subset M$ the plan π stays initially in E and $|\text{d}f|^p \in L^1(E, \mathfrak{m} \llcorner E)$,*
- ii) $\int \ell(\gamma_0, \gamma_1)^q \text{d}\pi(\gamma) < +\infty$.*

Then π represents the initial p -gradient of f and

$$|\partial^- f|(\gamma_0) = |\text{d}f|(\gamma_0) = \ell(\gamma_0, \gamma_1)^{q-1} \quad \pi - a.e. \quad \gamma. \quad (5.30)$$

Proof. Let $(x, y) \in \partial_{c_q} f$ with $\ell(x, y) > 0$. As already mentioned, this implies $f(x) \in \mathbf{R}$ by (5.28)–(5.29). Since $f^+ \geq f \geq f^-$ from Lemma 3.2, for every $z \leq x$ we have $\ell(z, y) \geq \ell(x, y) > 0$ and thus

$$f^+(x) - f^-(z) \geq f(x) - f(z) \stackrel{(5.29)}{\geq} u_q(\ell(z, y)) - u_q(\ell(x, y)) \geq u_q((\ell(x, y) + \ell(z, x)) - u_q(\ell(x, y))).$$

Notice that for every $\varepsilon > 0$ the set $\{z : \ell(z, x) < \varepsilon\}$ is a neighbourhood of x by the assumed continuity of ℓ_+ , hence in taking the limit in the definition (3.6) of $|\partial^- f|$ we have $\ell(z, x) \rightarrow 0$. Thus, dividing the above by $\ell(z, x)$ and letting $z \uparrow x$ we have $|\partial^- f|(x) \geq \ell(x, y)^{q-1}$ and thus

$$\frac{1}{p} |\partial^- f|^p(x) \geq \frac{1}{p} \ell(x, y)^q. \quad (5.31)$$

Let now γ be a geodesic such that $\ell(\gamma_0, \gamma_1) > 0$ and $(\gamma_0, \gamma_1) \in \partial_{c_q} f$, which again implies $f(\gamma_0) \in \mathbf{R}$. By the same argument as above, we obtain the estimate

$$f(\gamma_t) - f(\gamma_0) \leq u_q(\ell(\gamma_0, \gamma_1)) - u_q(\ell(\gamma_t, \gamma_1)) = \frac{1}{q} \ell(\gamma_0, \gamma_1)^q (1 - (1-t)^q).$$

Integrating and dividing by t we get

$$\begin{aligned} \limsup_{t \downarrow 0} \int \frac{f(\gamma_t) - f(\gamma_0)}{t} \text{d}\pi(\gamma) &\leq \int \ell(\gamma_0, \gamma_1)^q \text{d}\pi(\gamma) \stackrel{(5.31)}{\leq} \int \frac{|\partial^- f|^p}{p}(\gamma_0) \text{d}\pi(\gamma) + \int \frac{\ell(\gamma_0, \gamma_1)^q}{q} \text{d}\pi(\gamma) \\ &\quad \text{(by Proposition 3.18)} \leq \int \frac{|\text{d}f|^p}{p}(\gamma_0) \text{d}\pi(\gamma) + \int \frac{\ell(\gamma_0, \gamma_1)^q}{q} \text{d}\pi(\gamma), \end{aligned}$$

where the use of Proposition 3.18 is justified by the assumed continuity of ℓ_+ . This shows π represents the initial p -gradient of f . Recalling the bound in Proposition 4.1 (which is applicable, as both the integrability requirements are satisfied), we see that the last three inequalities must be equalities. Since the integrals are finite, this easily yield the claim (5.30). \square

Remark 5.20 (Comparison with the smooth case). Theorem 5.19 should be compared to the shape of optimal maps and ℓ_q -geodesics in smooth spacetimes, cf. McCann [84, Thm. 5.8, Cor. 5.9] and Mondino–Suhr [93, Lem. 3.2]. Indeed, for simplicity assume μ is concentrated on a compact set $C \subset I^-(o)$, where o is a given point in a globally hyperbolic smooth spacetime M . Setting $\nu := \delta_o$ and

$$f^o := -u_q(\ell(\cdot, o)) \quad (= -\frac{1}{q} \ell(\cdot, o)^q \text{ on } C)$$

it is clear that the unique ℓ_q -geodesic from μ to ν is given by

$$\mu_t := \exp.(t \|\text{d}f^o\|_*^{p-2} \nabla f^o)_* \mu,$$

where p is the conjugate exponent to q . This ℓ_q -geodesic is clearly induced by the push-forward π of μ via the map $x \mapsto \exp_x(\cdot \|df^o\|_*^{q-2} \nabla f^o)$. Therefore, π -a.e. $\gamma \in \text{TGeo}(\text{M})$ satisfies $\ell(\gamma_0, \gamma_1) = \|df^o\|_*^{p-1}(\gamma_0)$, which is (5.30).

An analogous argument applies to more general situations covered by the results of [84, 93] quoted above. \blacksquare

In order to apply this last result we need to be sure that $(\mathbf{e}_0, \mathbf{e}_1)_* \pi$ is concentrated on $\partial_{c_q} f$. To this aim, the following two simple lemmas are useful.

Lemma 5.21 (A Kantorovich/Karush–Kuhn–Tucker characterization of optimal couplings). *Let (M, τ, ℓ) be a Polish metric spacetime with ℓ upper semicontinuous and never $+\infty$. Let $0 \neq q < 1$, $f : \text{M} \rightarrow \mathbf{R}$ be a c_q -concave function and $\mu, \nu \in \mathcal{P}_{\text{em}}(\text{M})$ be so that f, f^{c_q} are μ -measurable and ν -measurable, respectively, with $f \in L^1(\mu)$ and $f^{c_q} \in L^1(\nu)$.*

Assume that there is an admissible coupling π of (μ, ν) that is concentrated on $\partial_{c_q} f$. Then π is optimal and every other c_q -optimal coupling of (μ, ν) is also concentrated on $\partial_{c_q} f$.

Proof. Let $\pi' \in \Pi_{\leq}(\mu, \nu)$ be arbitrary. The marginal condition and the integrability assumption imply that $(x, y) \mapsto f^{c_q}(y) - f(x)$ is in $L^1(\pi')$ and

$$\int f^{c_q} d\nu - \int f d\mu = \int f^{c_q}(y) - f(x) d\pi'(x, y) \geq \int c_q(x, y) d\pi'(x, y), \quad (5.32)$$

the inequality being a consequence of (5.28). For π admissible and concentrated on $\partial_{c_q} f$ the above becomes

$$\int f^{c_q} d\nu - \int f d\mu = \int f^{c_q}(y) - f(x) d\pi(x, y) = \int c_q(x, y) d\pi(x, y),$$

showing that $\int c_q d\pi' \leq \int c_q d\pi$, i.e. the optimality of π . Also, if π' is optimal equality must hold in (5.32) and by the integrability assumptions — that in particular ensure that $c_q(x, y), f(x), f^{c_q}(y) \in \mathbf{R}$ for π' -a.e. (x, y) — this can only occur if equality holds in (5.28) for π' -a.e. (x, y) , proving that π' is concentrated on $\partial_{c_q} f$, as desired. \square

Lemma 5.22 (Heredity of a c_q -concave dual potential by geodesic endpoints). *Let (M, τ, ℓ) be a Polish metric spacetime with ℓ upper semicontinuous and not attaining $+\infty$. Let $0 \neq q < 1$, $f : \text{M} \rightarrow \mathbf{R}$ be a c_q -concave function and $\nu_0, \nu_1 \in \mathcal{P}_{\text{em}}(\text{M})$ be such that $\ell_q(\nu_0, \nu_1) \in (0, +\infty)$ and any c_q -optimal plan from ν_0 to ν_1 is concentrated on $\partial_{c_q} f$.*

Then for every geodesic (μ_t) from ν_0 to ν_1 , any c_q -optimal coupling of (μ_0, μ_1) is concentrated on $\partial_{c_q} f$ as well.

Proof. The assumptions on ℓ and the measures ensure that c_q -optimal couplings $\pi^1, \pi^2, \pi^3 \in \mathcal{P}(\text{M}^2)$ of (ν_0, μ_0) , (μ_0, μ_1) , (μ_1, ν_1) exist. Let $\pi \in \mathcal{P}(\text{M}^4)$ be a gluing of them along the common marginals, so that for π -a.e. (x_0, x_1, x_2, x_3) we have $x_0 \leq x_1 \leq x_2 \leq x_3$ and thus

$$\ell_q(\nu_0, \nu_1) \geq \left(\int \ell^q(x_0, x_3) d\pi \right)^{\frac{1}{q}} \geq \left(\int \ell^q(x_1, x_2) d\pi \right)^{\frac{1}{q}} = \ell_q(\mu_0, \mu_1) \stackrel{(2.33)}{=} \ell_q(\nu_0, \nu_1).$$

By our assumption on ν_0, ν_1 and f we deduce that for π -a.e. (x_0, x_1, x_2, x_3) we have $(x_0, x_3) \in \partial_{c_q} f$ and $\ell(x_0, x_3) = \ell(x_1, x_2)$ (by Proposition 2.40 for this having assumed $\ell_q(\nu_0, \nu_1) \in (0, +\infty)$ matters). Also, from (5.29) we thus see that $f(x_0) + c_q(x_0, x_3) \geq f^{c_q}(x_3)$ holds for π -a.e. quadruple, thus from the causality of f, f^{c_q} , and the causal relation $x_0 \leq x_1 \leq x_2 \leq x_3$ and identity $\ell(x_0, x_3) = \ell(x_1, x_2)$ valid for π -a.e. quadruple, we conclude that $f(x_1) + c_q(x_1, x_2) \geq f^{c_q}(x_2)$ holds for π -a.e. quadruple, meaning that π^2 is concentrated on $\partial_{c_q} f$. Since π^2 was an arbitrary c_q -optimal coupling for (μ_0, μ_1) , the proof is complete. \square

We conclude the section with a result closely related to the metric Brenier–McCann theorem and to the Sobolev-to-steepness property. It offers yet another viewpoint on the role that existence of good geodesics have in relation to Sobolev calculus: the ‘rigidity’ encoded in equality (5.34)

as opposed to inequality (5.33) should be compared to that in Proposition 5.14 as opposed to inequality (5.21). Recall that (3.15) and the 1-steepness of $\ell(o, \cdot)$ and $-\ell(\cdot, o)$ give that

$$|\mathrm{d}\ell(o, \cdot)|, |\mathrm{d}(-\ell(\cdot, o))| \geq 1, \quad \mathbf{m} - a.e., \quad (5.33)$$

on any metric measure spacetime. From the $\mathrm{TMCP}_+^h(K, N)$ condition we can — via Theorem 5.7 — get equality for the function $-\ell(\cdot, o)$; (equality for $\ell(o, \cdot)$ follows from the $\mathrm{TMCP}_-^h(K, N)$ condition, see Section 5.7).

Corollary 5.23 (Unit maximal weak subslope of time separation to a point). *Let $0 \neq q < 1$, $K \in \mathbf{R}$, $N > 1$ and a $\mathrm{TMCP}_+^h(K, N)$ forward mm spacetime $(M, \tau, \ell, \mathbf{m})$ with ℓ upper semicontinuous and never $+\infty$ be timelike q -essentially non-branching at 0 and satisfy $\mathbf{m}(E) < \infty$ for every emerald E . Assume also that either (A) or (B) of Theorem 2.43 holds.*

Let $o \in \mathrm{spt} \mathbf{m}$ and consider the functions

$$f^o := -u_q \circ \ell(\cdot, 0), \quad \text{and} \quad g^o := -\ell(\cdot, o).$$

Then

$$|\mathrm{d}f^o| = \ell(\cdot, o)^{q-1} \quad \text{and} \quad |\mathrm{d}g^o| = 1 \quad \mathbf{m}\text{-a.e. on } I^-(o). \quad (5.34)$$

Proof. Let $C \subset I^-(o)$ be a compact set with $\mathbf{m}(C) > 0$. Let $\mu_0 \in \mathcal{P}(M)$ denote the uniform distribution of C . By Corollary 5.9 there exists an initial test plan π lifting a ℓ_q -geodesic (μ_t) from μ_0 to δ_o concentrated on timelike geodesics γ from γ_0 to o , thus by Lemma 3.13 we know that given any $\varepsilon \in (0, 1)$ sufficiently small, π -a.e. $\gamma \in \mathrm{LCC}([0, 1]; M)$ satisfies

$$\varepsilon \ell(\gamma_0, o) = g^o(\gamma_\varepsilon) - g^o(\gamma_0) \geq \int_0^\varepsilon |\mathrm{d}g^o|(\gamma_r) |\dot{\gamma}_r| \mathrm{d}r = \ell(\gamma_0, o) \int_0^\varepsilon |\mathrm{d}g^o|(\gamma_r) \mathrm{d}r.$$

It follows that $\frac{1}{\varepsilon} \int_0^\varepsilon |\mathrm{d}g^o|(\gamma_r) \mathrm{d}r \leq 1$, thus an integration and the same arguments as for the proof of (4.2) give

$$\int |\mathrm{d}g^o| \mathrm{d}\mu_0 = \int |\mathrm{d}g^o|(\gamma_0) \mathrm{d}\pi(\gamma) \leq \lim_{t \rightarrow 0} \frac{1}{t} \iint_0^t |\mathrm{d}g^o|(\gamma_r) \mathrm{d}r \mathrm{d}\pi(\gamma) \leq 1.$$

By (5.33), this forces $|\mathrm{d}g^o| = 1$ $\mathbf{m} \llcorner C$ -a.e. The arbitrariness of C and inner regularity of \mathbf{m} conclude the proof of the claim for g^o . The conclusion for f^o now follows from the chain rule in Proposition 3.23. \square

5.5 Nonlinear p -d'Alembert comparison

We turn to our main result: the nonlinear p -d'Alembert comparison theorem in weak form.

To state the result it is worth to introduce the function $\tilde{\tau}_{K,N}(\theta)$ as

$$\text{for } \theta \in [0, \pi\sqrt{\frac{N-1}{K_+}}) \text{ we put } \tilde{\tau}_{K,N}(\theta) = \begin{cases} \frac{1}{N} + \frac{\theta}{N} \sqrt{K(N-1)} \cot\left(\theta \sqrt{\frac{K}{N-1}}\right) & \text{if } K > 0, \\ 1 & \text{if } K = 0, \\ \frac{1}{N} + \frac{\theta}{N} \sqrt{-K(N-1)} \coth\left(\theta \sqrt{\frac{-K}{N-1}}\right) & \text{if } K < 0, \end{cases}$$

with $\tilde{\tau}_{K,N}(\pi\sqrt{\frac{N-1}{K_+}}) := -\infty$ and $K_+ := K \vee 0$.

Notice that $\tilde{\tau}_{K,N}(\theta)$ is the derivative at 1 of the function $r \mapsto \tau_{K,N}^{(r)}(\theta)$ from (5.3). We have:

Theorem 5.24 (p -d'Alembert comparison for $-\frac{1}{q}\ell^q(\cdot, o)$). *Fix $0 \neq q < 1$, $K \in \mathbf{R}$, $N > 1$ and let a $\mathrm{TMCP}_+^h(K, N)$ forward mm spacetime $(M, \tau, \ell, \mathbf{m})$ satisfy:*

- $\mathbf{m}(E) < \infty$ for every emerald E ;

- ℓ_+ is continuous and real valued;
- timelike q -essential non-branchingness at 0;
- either (A) or (B) of Theorem 2.43, which we recall here for clarity:

- A) M is globally hyperbolic, i.e. emeralds are compact, or
- B) the topology on M is locally causally convex.

Let o be a given point in $\text{spt } \mathbf{m}$ and put $f := -u_q(\ell(\cdot, o))$. Then for every $\varphi \in \text{Pert}(f)$ non-negative, bounded and with support in an emerald contained in $I^-(o)$ we have

$$\int d^+ \varphi(\nabla f) |df|^{p-2} d\mathbf{m} \leq N \int \tilde{\tau}_{K,N} \circ \ell(\cdot, o) \varphi d\mathbf{m}, \quad (5.35)$$

where $p^{-1} + q^{-1} = 1$.

Proof. Notice that the integral in the left hand side is well-defined (a priori possibly $+\infty$) by the bound (4.12) and the identity $|df|^p = \ell(\cdot, o)^q$ in $I^-(o)$ that follows from (5.34) because on $\text{spt } \varphi$ the function $\log \ell(\cdot, o)$ is bounded. The integral on the right is well-defined (a priori possibly $-\infty$) because on $\text{spt } \varphi$ the function $\tilde{\tau}_{K,N} \circ \ell(\cdot, o)$ is bounded from above (if $K > 0$ we use the Bonnet-Myers-type estimate that ensures that the timelike diameter of a $\text{TMCP}_+(K, N)$ spacetime is $\leq \pi \sqrt{\frac{N-1}{K}}$ — see [39, Proposition 5.10][21, Cor. 3.14] and Remark 1.2).

We only discuss the details of the proof in the situation $K > 0$, as the degeneracy of the involved distortion coefficients close to the Bonnet-Myers ℓ -diameter bound $\pi \sqrt{(N-1)/K}$ requires a small extra argument. Given $K' \in (0, K)$ fixed until the end, (M, ℓ, \mathbf{m}) obeys the slightly weaker property $\text{TMCP}_+(K', N)$ [21, Prop. 4.6].

As \mathbf{m} is finite on emeralds, the assumptions on φ imply its \mathbf{m} -integrability. The conclusion of the theorem immediately follows from the locality asserted by Proposition 4.11(iv) if $\mathbf{m}(\text{spt } \varphi) = 0$, hence we assume $\text{spt } \varphi$ is not \mathbf{m} -negligible.

We start with some preparations. First, we notice that f is c_q -concave (just pick ψ to be equal to 0 on o and to $+\infty$ everywhere else in formula (5.26)) and that $\text{spt } \varphi \times \{o\} \subset \partial_{c_q} f$ (directly from (5.29), the bound $f^{c_q} \leq \psi$ that comes from (5.27) and the fact that $c_q(\cdot, o)$ is finite on $I^-(o)$).

Then, we “raise” φ to make it uniformly bounded away from zero. To this aim, let $A := J^+(\text{spt } \varphi) \cap J^-(o)$ denote the emerald spanned by $\text{spt } \varphi$ and $\{o\}$, so $\mathbf{m}(A) < +\infty$ by assumption, and notice that for every $\alpha > 0$ the function $\varphi_\alpha := \varphi + \alpha 1_A$ still belongs to $\text{Pert}(f)$ (because $f = +\infty$ outside $J^-(o)$). Now define the Borel probability measure $\mu_0 := \rho_0 \mathbf{m}$, where

$$\rho_0 := (c_\alpha \varphi_\alpha)^{\frac{N}{N-1}} \quad c_\alpha \text{ being the normalization constant.}$$

Let μ_1 and $\boldsymbol{\pi} \in \mathcal{P}(\text{LCC}([0, 1]; M))$ be given by Corollary 5.9 (with K' in place of K) and let $(\mu_t) = ((\mathbf{e}_t)_* \boldsymbol{\pi})$ be the ℓ_q -geodesic from μ_0 to δ_o induced by $\boldsymbol{\pi}$. An application of Lemma 5.22, together with the fact that $\mu_0 \times \delta_o$ is the only admissible plan for (μ_0, δ_o) show that $(\mathbf{e}_0, \mathbf{e}_1)_* \boldsymbol{\pi}$ is concentrated on $\partial_{c_q} f$. Let also $\delta := \inf_{\text{spt } \varphi} \ell(\cdot, o)$, notice that $\delta > 0$ by the continuity assumption on ℓ_+ and put $E := A \cap \{\ell(\cdot, o) > \frac{\delta}{2}\}$. Then $\boldsymbol{\pi}$ stays initially in E and $|df|^p \in L^1(E, \mathbf{m} \llcorner E)$ (as $\mathbf{m}(E) < \infty$ and recalling (5.34)). We can thus apply Theorem 5.19 and conclude that $\boldsymbol{\pi}$ represents the initial p -gradient of f .

With this said, we know that (5.4) holds with K' in place of K , thus subtracting $\mathcal{S}_N(\mu_0)$ from both sides and dividing by t leads to

$$\frac{\mathcal{S}_N(\mu_t) - \mathcal{S}_N(\mu_0)}{t} \leq \int \frac{1 - \tau_{K',N}^{(1-t)} \circ \ell(\gamma_0, \gamma_1)}{t} \rho_0(\gamma_0)^{-1/N} d\boldsymbol{\pi}(\gamma),$$

and since — by discussion made at the beginning — the functions $\frac{1 - \tau_{K',N}^{(1-t)} \circ \ell(\gamma_0, \gamma_1)}{t}$ are uniformly bounded from above on $\text{spt } \boldsymbol{\pi}$, by Fatou’s lemma for the lim sup and the fact that $\ell(\gamma_0, \gamma_1) = \ell(\gamma_0, o)$

for π -a.e. γ (by Corollary 5.9) we get

$$\limsup_{t \downarrow 0} \frac{\mathfrak{S}_N(\mu_t) - \mathfrak{S}_N(\mu_0)}{t} \leq \int \tilde{\tau}_{K',N}(\ell(\gamma_0, \gamma_1)) \rho_0(\gamma_0)^{-\frac{1}{N}} d\pi(\gamma) = c_\alpha \int \varphi_\alpha \tilde{\tau}_{K',N}(\ell(\cdot, o)) d\mathbf{m}. \quad (5.36)$$

On the other hand, given any $t \in (0, 1)$ the convexity of $z \mapsto \mathfrak{s}_N(z) := -z^{1-\frac{1}{N}}$ gives

$$\begin{aligned} \frac{\mathfrak{S}_N(\mu_t) - \mathfrak{S}_N(\mu_0)}{t} &= \int \frac{\mathfrak{s}_N \circ \rho_t - \mathfrak{s}_N \circ \rho_0}{t} d\mathbf{m} \\ &\geq \int \mathfrak{s}'_N \circ \rho_0 \frac{\rho_t - \rho_0}{t} d\mathbf{m} = \int \frac{\mathfrak{s}'_N \circ \rho_0(\gamma_t) - \mathfrak{s}'_N \circ \rho_0(\gamma_0)}{t} d\pi(\gamma) \end{aligned}$$

so that combining the metric Brenier–McCann Theorem 5.19 with the first order differentiation formula in Theorem 4.15 (notice that $\mathfrak{s}'_N \circ \rho_0 \in \text{Pert}(f)$ by item (iii) in Lemma 4.9 — here and below it matters having picked $\alpha > 0$, so that the relevant functions are Lipschitz in the range of ρ_0, φ_α and thus an argument based on locality justifies the computations) we get

$$\liminf_{t \downarrow 0} \frac{\mathfrak{S}_N(\mu_t) - \mathfrak{S}_N(\mu_0)}{t} \geq \int d^+(\mathfrak{s}'_N \circ \rho_0)(\nabla f) |df|^{p-2} \rho_0 d\mathbf{m}. \quad (5.37)$$

It is now convenient to introduce the ‘pressure’ $\mathfrak{p}_N(z) := z \mathfrak{s}'_N(z) - \mathfrak{s}_N(z) = \frac{1}{N} z^{1-\frac{1}{N}}$, so that we have $\mathfrak{p}'_N(z) = z \mathfrak{s}''_N(z)$ and $\mathfrak{p}_N(\rho_0) = \frac{1}{N} c_\alpha \varphi_\alpha$. Then the chain rule (4.17) gives

$$\begin{aligned} \int d^+(\mathfrak{s}'_N \circ \rho_0)(\nabla f) |df|^{p-2} \rho_0 d\mathbf{m} &= \int \rho_0 \mathfrak{s}''_N(\rho_0) d^+ \rho_0(\nabla f) |df|^{p-2} d\mathbf{m} \\ &= \int d^+(\mathfrak{p}_N \circ \rho_0)(\nabla f) |df|^{p-2} d\mathbf{m} = \frac{c_\alpha}{N} \int d^+ \varphi(\nabla f) |df|^{p-2} d\mathbf{m}, \end{aligned}$$

having used also (4.16) in the last step to replace φ_α with φ . Coupling all this with (5.36) we get

$$\int d^+ \varphi(\nabla f) |df|^{p-2} d\mathbf{m} \leq N \int \varphi_\alpha \tilde{\tau}_{K',N}(\ell(\cdot, o)) d\mathbf{m}$$

and letting first $\alpha \downarrow 0$ and then $K' \uparrow K$, using the monotone convergence theorem (recall the comments at the beginning of the proof to see that the integrand in the right hand side is uniformly bounded from above in $\alpha \in (0, 1)$, $K' \in (K/2, K)$) we conclude. \square

From the above result and the chain rules we established in Section 4.3 we can derive similar comparison results for other functions of the time separation. The following one is particularly relevant:

Corollary 5.25 (*p*-d’Alembert comparison for Lorentz distance). *Let p, q and $(M, \tau, \ell, \mathbf{m})$ be as in Theorem 5.24 above, o be a point in $\text{spt } \mathbf{m}$ and set $g := -\ell(\cdot, o)$.*

Then for every non-negative and bounded $\varphi \in \text{Pert}(g)$ supported in an emerald inside $I^-(o)$ we have

$$\int d^+ \varphi(\nabla g) |dg|^{p-2} d\mathbf{m} \leq \int \frac{N \tilde{\tau}_{K,N} \circ \ell(\cdot, o) - 1}{\ell(\cdot, o)} \varphi d\mathbf{m}. \quad (5.38)$$

Proof. The fact that the integrals are well defined follows as in the proof of Theorem 5.24 above. Let f be as in Theorem 5.24, notice that φ is emerald supported in $I^-(o)$ so that the continuity of ℓ_+ ensures that there is a c -steep function $\psi_q : \mathbf{R} \rightarrow \mathbf{R}$ for some $c > 0$ such that $g = \psi_q \circ f$ holds on $\text{spt}(\varphi)$. By direct computation we see that $(\psi'_q)^{p-1} \circ f = \ell(\cdot, o)^{-1}$, thus from the chain rule (4.26) and the locality property (4.25) we see that

$$d^+ \varphi(\nabla g) |dg|^{p-2} = \ell(\cdot, o)^{-1} d^+ \varphi(\nabla f) |df|^{p-2} \quad \mathbf{m} - \text{a.e. on } \text{spt}(\varphi).$$

Similar considerations justify the use of the Leibniz rule (4.20) and the chain rules (4.17), (4.26) to get that

$$\begin{aligned} \ell(\cdot, o)^{-1} d^+ \varphi(\nabla f) |df|^{p-2} &\leq d^+(\ell(\cdot, o)^{-1} \varphi)(\nabla f) |df|^{p-2} - \varphi d^+(\ell(\cdot, o)^{-1})(\nabla f) |df|^{p-2} \\ \text{(using also (5.34))} \quad &= d^+(\ell(\cdot, o)^{-1} \varphi)(\nabla f) |df|^{p-2} - \varphi \ell(\cdot, o)^{-1} \end{aligned}$$

holds \mathbf{m} -a.e. on $\text{spt}(\varphi)$. Now we use Lemma 4.9 to see that the assumption $\varphi \in \text{Pert}(g)$ implies that $\frac{\varphi}{\ell(\cdot, o)} \in \text{Pert}(f)$, so that the conclusion follows from Theorem 5.24. \square

Remark 5.26 (Eschenburg's d'Alembert comparison). The bound (5.38) is a weak form of the heuristic inequality

$$\square_p(-\ell(\cdot, o)) \leq \frac{N \tilde{\tau}_{K,N} \circ \ell(\cdot, o) - 1}{\ell(\cdot, o)} \quad \text{on } I^-(o),$$

where $\square_p := -\text{div}(|d\cdot|^{p-2} \nabla \cdot)$ is the p -d'Alembert operator. Notice also that for $K = 0$ the estimate (5.38) reads as

$$\int d^+ \varphi(\nabla g) |dg|^{p-2} \, \text{d}\mathbf{m} \leq \int \frac{N-1}{\ell(\cdot, o)} \varphi \, \text{d}\mathbf{m},$$

which is already sharp [117, 118]. Since dg has unit magnitude \mathbf{m} -a.e. on $I^-(o)$ thanks to Corollary 5.23, this extends Eschenburg's d'Alembert comparison theorem from [54, §5] across the past timelike cut locus of o . For smooth spacetimes, a simple direct proof of this extension using more classical techniques can be found in [23], where the range of validity of the theorem is widened to include timelike geodesically complete spacetimes that need not be forward. \blacksquare

Theorem 5.24 formally establishes an upper bound for $\square_p f^o$ and since \square_p is elliptic with $\square_p f = \square_p(f + c)$ for any $c \in \mathbf{R}$, we expect from the maximum principle to be able to establish upper bounds for $\square_p f$ for any c_q -concave function f . This expectation turns out to be correct, but rather than being based on the maximum principle, our argument is a variant of the one just used to prove Theorem 5.24. To carry it out we shall need a non-branching assumption to apply Theorem 5.12 and get suitable good geodesics which, as before, can be used to link lower Ricci bounds and Sobolev calculus. Before stating the result we require a preliminary lemma.

Lemma 5.27 (Timelike diameter bounds the steepness of c_q -concave functions). *Let (M, ℓ) be a metric spacetime with $L := \sup_{x \leq y} \ell(x, y) < +\infty$ and $0 \neq p, q < 1$ satisfy $\frac{1}{p} + \frac{1}{q} = 1$.*

Then any c_q -concave function is L^{q-1} -steep. In particular, if $(M, \tau, \ell, \mathbf{m})$ is a mm spacetime and f is a \mathbf{m} -measurable c_q -concave function, then

$$|df|^{p-1} \leq L \quad \mathbf{m} - \text{a.e.}, \quad \text{where} \quad L := \sup_{x \leq y} \ell(x, y).$$

Proof. The second claim follows from the first one and inequality (3.15). For the first we observe that, directly from the definition, we see that the infimum of an arbitrary family of L^{q-1} -steep functions is L^{q-1} -steep. Hence to conclude it suffices to show that for any $y \in M$ and $c \in \mathbf{R}$ the function $x \mapsto -u_q(\ell(x, y)) + c$ is L^{q-1} -steep. This, however, is obvious from the fact that $x \mapsto -\ell(x, y)$ is 1-steep and $z \mapsto -u_q(-z)$ has derivative bounded from below by L^{q-1} (and thus is L^{q-1} -steep) on $[-L, 0]$. \square

For $q < 0$ the measurability hypotheses imposed on f^{c_q} and \bar{F} in Theorem 5.28(i)-(ii) are easily satisfied by combining Proposition 5.18 with a measurable selection theorem. To focus on the new additional difficulties, we will ask in point (iii) below an integrability assumption stronger than what is actually necessary; in particular, the result below does not extend Theorem 5.24 (compare with the definition of A in the proof of Theorem 5.24).

Theorem 5.28 (*p*-d'Alembert comparison for c_q -concave functions). *Let p, q and $(M, \tau, \ell, \mathbf{m})$ be as in Theorem 5.24 and assume also that M is timelike q -essentially non-branching.*

Let $f : M \rightarrow \bar{\mathbf{R}}$ be c_q -concave and $\varphi \in \text{Pert}(f)$ non-negative, bounded, and emerald supported. Assume that:

- i) f^{c_q} is Borel and f is bounded on $\text{spt } \varphi$;*
- ii) there is a Borel $\bar{F} : \text{spt } \varphi \rightarrow M$ and an emerald E with $\bar{F}(\text{spt } \varphi) \cup \text{spt } \varphi \subset E$. Moreover, for some $c > 0$ we have $(x, \bar{F}(x)) \in \partial_{c_q} f$ and $\ell(x, \bar{F}(x)) \in (c, \frac{1}{c})$ for \mathbf{m} -a.e. $x \in \text{spt } \varphi$;*
- iii) we have $|df|^p \in L^1(E, \mathbf{m} \llcorner E)$.*

Then

$$\int d^+ \varphi (\nabla f) |df|^{p-2} d\mathbf{m} \leq N \int \tilde{\tau}_{K,N} \circ |df|^{p-1} \varphi d\mathbf{m}. \quad (5.39)$$

Proof. The backbone of the proof is the same of that of Theorem 5.24, thus we shall focus on the differences. Recall that c_q -concave functions are rough causal, as discussed right after (5.29), and then notice the assumption $\ell(x, \bar{F}(x)) \geq c > 0$ for \mathbf{m} -a.e. $x \in \text{spt } \varphi$ ensures $\varphi \mathbf{m}$ is concentrated on $M \setminus M_{\text{fin}}$. Corollary 5.11 therefore implies that f and the integrands in (5.39) are well defined and \mathbf{m} -measurable. The integral on the left is well defined (a priori with value $+\infty$) by the bound (4.12) and the integrability assumption $|df|^p \in L^1(E)$, that in particular yields integrability of $|df|^p$ on $\text{spt } \varphi \subset E$. The integral on the right hand side is also well-defined because $\tilde{\tau}_{K,N} \circ |df|^{p-1}$ is (essentially) uniformly bounded from above on $\{\varphi > 0\}$. Indeed, in the course of the proof we will show that

$$|df|^{p-1}(x) \in [c, \frac{1}{c}] \quad \mathbf{m} - \text{a.e. on } \text{spt } \varphi. \quad (5.40)$$

If $K \leq 0$ this suffices to conclude, as $\tilde{\tau}_{K,N} : \mathbf{R}^+ \rightarrow \mathbf{R}$ is continuous. If $K > 0$ we notice on one side that $\tilde{\tau}_{K,N}$ is bounded from above on $[0, \pi \sqrt{\frac{N-1}{K}}]$ and on the other that $|df|^{p-1} \leq \pi \sqrt{\frac{N-1}{K}}$ \mathbf{m} -a.e. on $\text{spt } \varphi$, this latter bound being a consequence of the Bonnet-Myers-type estimate that ensures that the timelike diameter of a $\text{TMCP}_+(K, N)$ spacetime is $\leq \pi \sqrt{\frac{N-1}{K}}$ (see [39, Proposition 5.10] [21, Cor. 3.14]) and Remark 1.2) and Lemma 5.27 above.

With this said, as in the previous case we deal with the technically slightly more involved case $K > 0$ and fix $K' \in (0, K)$. Also as before, we know that $\varphi \in L^1(\mathbf{m})$ and we can assume $\mathbf{m}(\{\varphi > 0\}) > 0$.

Let $\tilde{f} : M \rightarrow \bar{\mathbf{R}}$ be equal to f on $J^-(E)$ and $+\infty$ otherwise and notice that since f is causal so is \tilde{f} . It is then clear that for any $\alpha > 0$ we have $\varphi_\alpha := 1_E(\varphi + \alpha) \in \text{Pert}(\tilde{f})$.

We can then define the Borel probability measure $\mu_0 := \rho_0 \mathbf{m}$, where

$$\rho_0 := (c_\alpha \varphi_\alpha)^{\frac{N}{N-1}}, \quad c_\alpha \text{ being the normalization constant.}$$

Let $\bar{\mu}_1 := \bar{F}_* \mu_0$ and notice that it is emerald supported because $\bar{F}(\text{spt } \varphi)$ is contained in some emerald by assumption. The assumptions also assert $f \in L^\infty(\mu_0) \subset L^1(\mu_0)$, and that the function $x \mapsto f^{c_q}(\bar{F}(x)) = c_q(x, \bar{F}(x)) + f(x)$ is bounded (here we used that $\ell(x, \bar{F}(x)) \in [c, \frac{1}{c}]$ μ_0 -a.e.) and thus $f^{c_q} \in L^\infty(\bar{\mu}_1) \subset L^1(\bar{\mu}_1)$. Recalling the $\mu_0 \ll \mathbf{m} \llcorner (M \setminus M_{\text{fin}})$ measurability of f from above, the assumption that f^{c_q} is Borel yields the $\bar{\mu}_1$ -measurability of f^{c_q} , so Lemma 5.21 ensures that the coupling $\bar{\pi} := (\text{id}, \bar{F})_* \mu_0$ is optimal. Also, it is clearly concentrated on $\{\ell > 0\}$ and we have $\ell_q(\mu_0, \bar{\mu}_1) \in (0, +\infty)$. We can therefore apply Theorem 5.12 (with K' in place of K) and obtain an ℓ_q -geodesic (μ_t) from μ_0 to $\bar{\mu}_1$ and a lifting $\pi \in \mathcal{P}(\text{TGeo}(M))$ of it satisfying (5.19). We claim that π represents the initial p -gradient of f on the set E given in the statement and to this aim we shall verify that it satisfies the assumptions on the metric Brenier-McCann Theorem 5.19. We know from Lemma 5.21 that every ℓ_q -optimal coupling of $(\mu_0, \bar{\mu}_1)$ is concentrated on $\partial_{c_q} f$ and from Lemma 5.22 that the same holds for $(\mathbf{e}_0, \mathbf{e}_1)_* \pi$ (which is ℓ_q -optimal for (μ_0, μ_1) by construction). The identity (5.18) together with the definition of $\bar{\pi}$ yield $\log(\ell(\gamma_0, \gamma_1)) \in L^\infty(\pi)$ and since π is concentrated on timelike geodesics, this suffices to prove that the integrability assumption (ii) in

Theorem 5.19 holds. Since assumption (i) of Theorem 5.19 holds by our hypotheses (iii) and since π does not leave E we can indeed apply Theorem 5.19 to conclude that π represents the initial p -gradient of f . We notice also that for $F : M \rightarrow \text{LCC}([0, 1]; M)$ inducing π as in Theorem 5.12 we have

$$|df|^{p-1}(x) \stackrel{(5.30)}{=} \ell(x, F_1(x)) \stackrel{(5.18)}{\in} [c, \frac{1}{c}] \quad \mu_0 - a.e. \ x \in M,$$

so that (5.40) holds. With this said, the same computations leading to (5.36) yield

$$\limsup_{t \downarrow 0} \frac{\mathcal{S}_N(\mu_t) - \mathcal{S}_N(\mu_0)}{t} \leq c_\alpha \int \varphi_\alpha \tilde{\tau}_{K', N}(|df|^{p-1}) \, d\mathbf{m} = c_\alpha \int_E (\varphi + \alpha) \tilde{\tau}_{K', N}(|df|^{p-1}) \, d\mathbf{m},$$

where the use of the reverse Fatou's lemma to bring the lim sup inside the integral is justified as in the proof of Theorem 5.24.

Since π does not leave the set E where we have $f = \tilde{f}$, by Theorem 5.19 we deduce also that π represents the initial p -gradient of \tilde{f} . Since, again as in the proof of Theorem 5.24, we have $\mathcal{S}'_N \circ \rho_0 \in \text{Pert}(\tilde{f})$, the same computations as in (5.37) and the identity below it give

$$\liminf_{t \downarrow 0} \frac{\mathcal{S}_N(\mu_t) - \mathcal{S}_N(\mu_0)}{t} \geq \frac{c_\alpha}{N} \int d^+ \varphi_\alpha(\nabla \tilde{f}) |d\tilde{f}|^{p-2} \, d\mathbf{m} = \frac{c_\alpha}{N} \int d^+ \varphi(\nabla f) |df|^{p-2} \, d\mathbf{m},$$

where for the equality we used the locality properties (4.16) and (4.25). Coupling these last two bounds we get

$$\int d^+ \varphi(\nabla f) |df|^{p-2} \, d\mathbf{m} \leq N \int_E (\varphi + \alpha) \tilde{\tau}_{K', N}(|df|^{p-1}) \, d\mathbf{m}.$$

To conclude by letting $\alpha \downarrow 0$ and then $K' \uparrow K$ it suffices to prove that $\int_E \tilde{\tau}_{K', N}(|df|^{p-1}) \, d\mathbf{m} < +\infty$. For $K = 0$ this follows from $\mathbf{m}(E) < +\infty$ and for $K < 0$ from the fact that $\theta \mapsto \tilde{\tau}_{K, N}(\theta)$ has linear growth together with the extra integrability assumption made in item (iii). For $K > 0$ we use the Bonnet-Myers-type of timelike diameter bound already recalled together with the fact that on $[0, \pi \sqrt{\frac{N-1}{K}}]$ the functions $\tilde{\tau}_{K', N}$ are uniformly bounded from above (in $K \in (0, K)$) and Lemma 5.27 below. \square

Remark 5.29 (Nonsharp variants). Assuming $\text{TMCP}_+^{h,*}(K, N)$ instead of $\text{TMCP}_+^h(K, N)$ in Theorem 5.28, Theorem 5.24, and Corollary 5.25, the resulting estimates hold — with the same proof — with $\tilde{\sigma}_{K, N}$ in place of $\tilde{\tau}_{K, N}$, where

$$\tilde{\sigma}_{K, N}(\theta) = \begin{cases} \theta \sqrt{\frac{K}{N}} \cot\left(\theta \sqrt{\frac{K}{N}}\right) & \text{if } K > 0, \\ 1 & \text{if } K = 0, \\ \theta \sqrt{\frac{-K}{N}} \cot\left(\theta \sqrt{\frac{-K}{N}}\right) & \text{if } K < 0, \end{cases}$$

is the derivative at 1 of the function $r \mapsto \sigma_{K, N}^{(r)}$ from (5.2). \blacksquare

5.6 Distributional p -d'Alembert operator

The aim of this section is to propose a first answer to the question: what is the p -d'Alembertian? By no means is the presentation here exhaustive: on the contrary, our aim is more to show viability of this research direction in the nonsmooth setting. For some further results see [22].

Taking inspiration from Gigli's theory in positive signature [60], we see that the calculus developed above hints at the possibility of defining the p -d'Alembert operator 'distributionally'. Notice indeed that in the smooth category from the definition $\square_p f = -\text{div}(|df|^{p-2} \nabla f)$ we see that

$$\int d\varphi(\nabla f) |df|^{p-2} \, d\mathbf{m} = \int \varphi \square_p f \, d\mathbf{m},$$

so that $\square_p f$ can be defined weakly as the operator sending suitable φ 's to $\int d\varphi(\nabla f)|df|^{p-2} dm$. Now, in our context a priori we do not have a definition for $d\varphi(\nabla f)|df|^{p-2}$, but only of its ‘proxies’ $d^+\varphi(\nabla f)|df|^{p-2}$ and $-d^+(-\varphi)(\nabla f)|df|^{p-2}$. Still, in infinitesimally Minkowskian spaces these two agree on $\{|df| > 0\}$ (recall Theorem 4.16) and denoting their common value by $d\varphi(\nabla f)|df|^{p-2}$, from (4.13) and (4.14) we see that the linearity relation

$$d(\alpha_1\varphi_1 + \alpha_2\varphi_2)(\nabla f)|df|^{p-2} = \alpha_1 d\varphi_1(\nabla f)|df|^{p-2} + \alpha_2 d\varphi_2(\nabla f)|df|^{p-2} \quad \mathbf{m} - a.e. \text{ on } \{|df| > 0\} \quad (5.41)$$

holds for any $\alpha_1, \alpha_2 \in \mathbf{R}$ whenever $\varphi_1, \varphi_2 : M \rightarrow \mathbf{R}$ are so that $\pm\varphi_1, \pm\varphi_2 \in \text{Pert}(f)$ (notice that in this case by Proposition 4.8 we have $\alpha_1\varphi_1 + \alpha_2\varphi_2 \in \text{Pert}(f)$ and that we are insisting that the functions be real valued). Let us give a name to this space of functions: for $f : M \rightarrow \bar{\mathbf{R}}$ causal we put

$$\text{Pert}^{\text{sym}}(f) := \{\varphi : M \rightarrow \mathbf{R} : \varphi, -\varphi \in \text{Pert}(f)\}. \quad (5.42)$$

For $U \subset M$ open we also define

$$\text{Pert}_{\text{be}}^{\text{sym}}(f, U) := \{\varphi \in \text{Pert}^{\text{sym}}(f) : \text{spt } \varphi \subset U \text{ is contained in an emerald and } \varphi \text{ is bounded}\}.$$

The following is now natural:

Definition 5.30 (Distributional p -d’Alembertian). *Let $(M, \tau, \ell, \mathbf{m})$ be an infinitesimally Minkowskian mm spacetime, $U \subset M$ open, $0 \neq p < 1$ and $f : M \rightarrow \bar{\mathbf{R}}$ causal. We shall say that f has a distributional p -d’Alembertian on U provided $d^+\varphi(\nabla f)|df|^{p-2} \in L^1(U, \mathbf{m} \llcorner U)$ for every $\varphi \in \text{Pert}_{\text{be}}^{\text{sym}}(f, U)$. In this case the map*

$$\text{Pert}_{\text{be}}^{\text{sym}}(f, U) \ni \varphi \quad \mapsto \quad \int d^+\varphi(\nabla f)|df|^{p-2} dm \in \mathbf{R} \quad (5.43)$$

will be called the distributional p -d’Alembertian of f in U .

Notice that the integrability assumption $d^+(\pm\varphi)(\nabla f)|df|^{p-2} \in L^1(U, \mathbf{m} \llcorner U)$ together with (4.10) and (4.11) imply that (4.29) for $g = \varphi$ extends to the whole of M (the common value of both sides still denoted by $d\varphi(\nabla f)|df|^{p-2}$) and we conclude that the distributional p -d’Alembertian, when it exists, is a linear operator. Notice also that if $|df|^p \in L^1(E, \mathbf{m} \llcorner E)$ for every emerald $E \subset U$, then the bound (4.12) easily implies that f has a distributional p -d’Alembertian in U .

We can then interpret a one-sided bound like the one provided by Theorem 5.24 as structural information on the distributional p -d’Alembertian, which can be used to begin a regularity theory for it. For instance:

Proposition 5.31 (Existence and bounds for the distributional p -d’Alembert operator). *Using the same assumptions and notation of Theorem 5.24 and assuming the spacetime to also be infinitesimally Minkowskian, the following holds. Let $U \subset M$ be the open set $I^-(o)$ if $K \leq 0$ and $I^-(o) \cap \{\ell(\cdot, o) < \pi\sqrt{(N-1)/K}\}$ if $K > 0$.*

Then f has a distributional p -d’Alembertian on U and for every $\varphi \in \text{Pert}_{\text{be}}^{\text{sym}}(f, U)$ non-negative we have

$$\int d\varphi(\nabla f)|df|^{p-2} dm - \int \varphi d\nu \leq 0, \quad \text{where} \quad \nu := N \tilde{\tau}_{K,N} \circ \ell(o, \cdot) \mathbf{m} \llcorner U \quad (5.44)$$

(notice that ν is a Radon measure on U). Moreover, assume that for some emerald $E \subset U$ there exists a non-negative ‘cut-off’ function $\eta \in \text{Pert}_{\text{be}}^{\text{sym}}(f, U)$ that is ≥ 1 on E . Then for every $\varphi \in \text{Pert}_{\text{be}}^{\text{sym}}(f, U)$ with support in E we have

$$\left| \int d\varphi(\nabla f)|df|^{p-2} dm \right| \leq \sup_{x \in U} |\varphi(x)| \left(2|\nu|(E) \sup_{x \in M} \eta(x) + \int d\eta(\nabla f)|df|^{p-2} dm \right). \quad (5.45)$$

Proof. From (5.34) we see that $|df|^p = \ell(\cdot, o)^q$, so that $|df|^p \in L^1(E, \mathbf{m} \llcorner E)$ for every emerald $E \subset U$ and the previous discussion ensures that f has distributional p -d’Alembertian on U . The bound (5.44) is a restatement of (5.35), so we turn to the last claim.

Notice that for $\varphi \in \text{Pert}_{\text{be}}^{\text{sym}}(f, U)$ item (v) in Proposition 4.8 ensures that its positive and negative parts φ^+ and φ^- , whose support is contained in that of φ , are also in $\text{Pert}_{\text{be}}^{\text{sym}}(f, U)$. Thus writing $s := \sup_{x \in U} |\varphi(x)|$ for brevity, the functions $s\eta - \varphi^\pm$ are non-negative and in $\text{Pert}_{\text{be}}^{\text{sym}}(f, U)$. Thus putting, again for brevity, $L(\varphi) := \int d\varphi(\nabla f)|df|^{p-2} \, \text{dm}$, the bound (5.44) tells us that

$$\begin{aligned} L(\varphi^\pm) &\leq \int \varphi^\pm \, d\nu \leq s|\nu|(E), \\ L(\varphi^\pm) &= -L(s\eta - \varphi^\pm) + sL(\eta) \geq - \int s\eta - \varphi^\pm \, d\nu + sL(\eta) \geq -s(|\nu|(E) \sup_{x \in M} \eta(x) + L(\eta)) \end{aligned}$$

and the claim (5.45) follows. \square

We collect a few informal comments on this result:

- a) There is nothing special about the function f in the above beside the bound (5.44) established in Theorem 5.24: analogous statements can be made for f 's as in Corollary 5.25 or Theorem 5.28, in the latter case provided we also assume the spacetime to be non-branching as in that theorem. Similar results are also in place in ‘past’ $\text{TMCP}_-^h(K, N)$ spacetimes as discussed in the next section.
- b) The definition we gave for the distributional p -d'Alembertian depends on the space $\text{Pert}_{\text{be}}^{\text{sym}}(f, U)$, thus it is important to know that this space is rich. We do not have general results in this direction unless the spacetime is smooth, where they are provided by Appendix A.2.
- c) From bounds like (5.44) and (5.45) and results like that of Riesz and Daniell one can surely hope that the distributional p -d'Alembertian can be represented by a unique measure. See Corollary A.6 and [22] for positive results in this direction (in the latter case obtained by more constructive arguments inspired by those from [37] in positive signature) and Remark A.7 for a caveat.
- d) In order for the linearity in (5.41) to be in place, and thus for the map in (5.43) to be linear, it is not necessary to assume infinitesimal Minkowskianity. As the arguments leading to (5.41) show, it suffices that $d^+\varphi(\nabla f)|df|^{p-2}$ coincides with $-d^+(-\varphi)(\nabla f)|df|^{p-2}$ whenever $\pm\varphi \in \text{Pert}(f)$. Spaces with this property are called *infinitesimally strictly concave* in [22] (see also the concept of *infinitesimal strict convexity* defined in [60] for the positive signature case).
- e) Without assuming infinitesimal strict concavity, after noticing that the concavity (4.14) yields convexity of $\varphi \mapsto -d^+(-\varphi)(\nabla f)|df|^{p-2}$, one might say that a distributional p -d'Alembertian of f in U is any linear functional $L : \text{Pert}_{\text{be}}^{\text{sym}}(f, U) \rightarrow \mathbf{R}$ such that

$$\int d^+\varphi(\nabla f)|df|^{p-2} \, \text{dm} \leq L(\varphi) \leq \int -d^+(-\varphi)(\nabla f)|df|^{p-2} \, \text{dm} \quad \forall \varphi \in \text{Pert}_{\text{be}}^{\text{sym}}(f, U).$$

While this is doable (see also [60] for the positive signature case and [22] for more in the spacetime setting), solely from a bound like this one cannot deduce any relation between L and the right hand side in (5.35). Thus if one looks for a link like that in Proposition 5.31 above, some sort of selection procedure is necessary.

5.7 Modifications for past TMCP spaces

All results from the previous parts of Section 5 have counterparts for the past timelike measure contraction property $\text{TMCP}_-^h(K, N)$ from Definition 5.1 and these lead to p -d'Alembertian estimates for functions of the kind ψ^{c_q} , rather than ψ_{c_q} , defined in (5.27).

Roughly speaking, $\text{TMCP}_-^h(K, N)$ ‘reverses’ the transport direction of $\text{TMCP}_+^h(K, N)$ by imposing convexity properties of the involved entropies along chronological transports *from* Dirac masses *to* \mathfrak{m} -absolutely continuous distributions. Because of this, most of the results follow via

minor technical modifications — see also Remark 4.14 — so that below we only give a brief outline of the main arguments.

We stress that we will still work with the original causal structure set up by ℓ (and not its time-reversal). In particular this means that we will still work with forward spacetimes and with left continuous causal curves, rather than on backward spacetimes and right continuous curves, hence strictly speaking the results collected here cannot be derived from those already proved by time-reversal. Still, the ‘sign choice’ in our completeness assumption is almost invisible through our constructions, and reduces to the (inessential) possible discontinuity at $t = 0$ discussed in Remark 5.32 and to the (also inessential) fact that test plans are automatically continuous at $t = 1$, so that the continuity requirement in the definition of final test plan could be removed (see also Remark 3.8).

The following calculus rules are valid in any forward metric measure spacetime $(M, \tau, \ell, \mathbf{m})$. Here $0 \neq p, q < 1$ are conjugate, i.e. $\frac{1}{p} + \frac{1}{q} = 1$.

(A) *(Final version of nonsmooth Fenchel-Young inequality — compare with Proposition 4.1)*

Let $f : M \rightarrow \bar{\mathbf{R}}$ be causal, π a final test plan that stays finally in the Borel set $E \subset M$. Then

$$\liminf_{s \uparrow 1} \int \frac{f(\gamma_1) - f(\gamma_s)}{1 - s} d\pi(\gamma) \geq \frac{1}{p} \int |df|^p(\gamma_1) d\pi(\gamma) + \liminf_{s \uparrow 1} \frac{1}{(1-s)q} \iint_s^1 |\dot{\gamma}_r|^q dr d\pi(\gamma)$$

holds whenever

- (i) $|df|^p \in L^1(E, \mathbf{m} \llcorner E)$ in the case $p < 0$,
- (ii) the latter limit inferior is not equal to $-\infty$ in the case $q < 0$.

(B) *(Plans representing final p -gradients — compare with Definition 4.2)*

With f, π, E as above, we say that π represents the final p -gradient of f provided

- (i) $|df|^p \in L^1(E, \mathbf{m} \llcorner E)$
- (ii) $\limsup_{s \uparrow 1} \frac{1}{(1-s)q} \iint_s^1 |\dot{\gamma}_r|^q dr d\pi(\gamma) < +\infty$
- (iii) we have

$$\limsup_{s \uparrow 1} \int \frac{f(\gamma_1) - f(\gamma_s)}{1 - s} d\pi(\gamma) \leq \frac{1}{p} \int |df|^p(\gamma_1) d\pi(\gamma) + \liminf_{s \uparrow 1} \frac{1}{(1-s)q} \iint_s^1 |\dot{\gamma}_r|^q dr d\pi(\gamma).$$

(C) *(First order differentiation formula — compare with Theorem 4.15)*

Let π represent the final p -gradient of f on E . Then for every $-g \in \text{Pert}(f)$ we have:

- (i) for any $s < 1$ sufficiently big the positive part of the function $g \circ e_1 - g \circ e_s$ is in $L^1(\pi)$;
- (ii) the positive part of the function $-d^+(-g)(\nabla f) |df|^{p-2}$ is in $L^1(\mathbf{m} \llcorner E)$;
- (iii) we have

$$\limsup_{s \uparrow 1} \int \frac{g(\gamma_1) - g(\gamma_s)}{1 - s} d\pi(\gamma) \leq \int -d^+(-g)(\nabla f) |df|^{p-2}(\gamma_1) d\pi(\gamma). \quad (5.46)$$

(D) *(Metric Brenier-McCann theorem — compare with Theorem 5.19)*

Assume also the spacetime is forward complete, that either (A) or (B) of Theorem 2.43 hold, that ℓ_+ is continuous and real valued and let $f : M \rightarrow \bar{\mathbf{R}}$ be \mathbf{m} -measurable and of the form $f = \psi^c$ for some $\psi : M \rightarrow \bar{\mathbf{R}}$ (recall (5.27)). Let π be a final test plan concentrated on $\text{TGeo}(M)$ such that $(e_0, e_1)_* \pi$ is concentrated on the set $\partial^{c_q} f$ of couples (x, y) such that $f_c(x), c_q(x, y), f(y) \in \mathbf{R}$ and $f_c(x) + c_q(x, y) \geq f(y)$. Assume also that:

- i) for some \mathbf{m} -measurable set $E \subset M$ the plan π stays finally in E and $|df|^p \in L^1(E, \mathbf{m} \llcorner E)$,

$$\text{ii) } \int \ell(\gamma_0, \gamma_1)^q d\boldsymbol{\pi}(\gamma) < +\infty.$$

Then $\boldsymbol{\pi}$ represents the final p -gradient of f on E and

$$|\partial^+ f|(\gamma_1) = |df|(\gamma_1) = \ell(\gamma_0, \gamma_1)^{q-1} \quad \boldsymbol{\pi} - a.e. \gamma.$$

The above calculus rules are now coupled with a curvature condition: in the results below we shall assume that $(M, \tau, \ell, \mathbf{m})$ is:

- A forward $\text{TMCP}_-^h(K, N)$ mm spacetime,
- q -essentially non-branching at 1,
- such that $\mathbf{m}(E) < +\infty$ for every emerald E ,
- such that ℓ_+ is continuous and does not take the value $+\infty$,
- such that either (A) or (B) of Theorem 2.43 hold.

Then:

(E) *(Existence of good final test plans with Dirac source — compare with Thm. 5.7 and Cor. 5.9)*

Let $\mu_1 = \rho_1 \mathbf{m} \in \mathcal{P}(M)$ be supported in an emerald and have bounded compression. Let $x_0 \in \text{spt } \mathbf{m}$ satisfy $\log(\ell(x_0, \cdot)) \in L^\infty(\mu_1)$. Then there is a final test plan $\boldsymbol{\pi}$ concentrated on timelike geodesics γ from x_0 to γ_1 so that for $\mu_t := (\mathbf{e}_t)_* \boldsymbol{\pi}$ we have:

$$\begin{aligned} \mu_t &\leq \frac{1}{t^N} e^{Dt\sqrt{K_-(N-1)}} \|\rho_1\|_{L^\infty(M, \mathbf{m})}, \\ \mathcal{S}_N(\mu_t) &\leq - \int \tau_{K, N}^{(t)} \circ \ell(\cdot, x_1) \rho_1^{1-1/N} d\mathbf{m}, \\ \mathcal{S}_\infty(\mu_t) &\leq \mathcal{S}_\infty(\mu) - N \log \sigma_{K, N}^{(t)}(\|\ell(x_0, \cdot)\|_{L^2(d\mu)}), \end{aligned}$$

for any $t \in (0, 1]$, where $D := \sup \ell(\{x_0\} \times \text{spt } \mu_1)$.

Remark 5.32 (The role of forward completeness). In the proof of the above, and the analogous results Theorem 5.7 and Corollary 5.9, forward completeness is used to perform suitable limiting procedures (e.g. in the proof of the lifting theorem) but is not at all related to curvature conditions nor to the fact that here we start from a Dirac mass and end up in a diffused measure (contrary to what happens in the case of TMCP_+^h spaces). The fact that the plan $\boldsymbol{\pi}$ in the above has marginal at time 1 equal to given measure μ_1 comes from the assumption of q -essentially non-branching at 1, the density bounds and finite mass of emeralds together with Corollary 2.57.

On the other hand, we notice that μ_0 is precisely δ_{x_0} , which is a non-essential difference with Theorem 5.7 and Corollary 5.9, where we could have $\mu_1 \neq \delta_{x_1}$. This happens because the discrete procedure used to build the final geodesic keeps μ_0 always fixed, and eventually re-defines μ_1 as narrow forward limit of measures defined on a suitable dense set of times (and the previous considerations ensure that such new definition coincides with the initially given measure μ_1). Notice that in all this it might be that (μ_t) does not weakly converge to μ_0 as $t \downarrow 0$ (not even if we further assume that a narrow limit exists). \blacksquare

(F) *(p -d'Alembertian comparison for $\frac{1}{q}\ell^q(o, \cdot)$ — compare with Theorem 5.24)*

Assume also that ℓ_+ is continuous, let $o \in \text{spt } \mathbf{m}$ and define $f := \frac{1}{q}\ell(o, \cdot)^q$.

Then for every $\varphi : M \rightarrow \mathbf{R}^+$ bounded, with $\text{spt } \varphi \subset I^+(o)$ and $-\varphi \in \text{Pert}(f)$ we have

$$\int -d^+(-\varphi)(\nabla f) |df|^{p-2} d\mathbf{m} \geq - \int N \tilde{\tau}_{K, N} \circ \ell(o, \cdot) \varphi d\mathbf{m}. \quad (5.47)$$

An analogous comparison result is in place for $\ell(o, \cdot)$ and, under appropriate non-branching assumptions, also for more general functions f of the form $f = \psi^c$ in analogy with Corollary 5.25 and Theorem 5.28. All the above results can be proved following rather pedantically the analogous proofs in the $\text{TMCP}_+^h(K, N)$ case. Let us, for added clarity, quickly and formally discuss how the last statement (F) follows from the previous ones.

Let $\rho_1 := (c\varphi)^{N/(N-1)}$, $c > 0$ being the normalization constant, and let π be given by item (E) above. Then the (time-reversal of the) bound (5.4) and the identity $\tau_{K,N}^{(1)}(\theta) = 1$ give

$$\frac{\mathfrak{S}_N(\mu_1) - \mathfrak{S}_N(\mu_s)}{1-s} \geq \int -\frac{\tau_{K,N}^1 - \tau_{K,N}^{(s)}}{1-s} \circ \ell(x_0, \cdot) \rho_1^{1-\frac{1}{N}} \, \mathbf{d}\mathbf{m} \quad \xrightarrow{s \uparrow 1} \quad -c \int \tilde{\tau}_{K,N} \circ \ell(o, \cdot) \varphi \, \mathbf{d}\mathbf{m}.$$

On the other hand the bound $\mathfrak{s}_N(b) - \mathfrak{s}_N(a) \leq \mathfrak{s}'_N(b)(b-a)$ for the convex function $\mathfrak{s}_N(z) := -z^{1-1/N}$ gives

$$\frac{\mathfrak{S}_N(\mu_1) - \mathfrak{S}_N(\mu_s)}{1-s} \leq \int \frac{\mathfrak{s}'_N(\rho_1(\gamma_1)) - \mathfrak{s}'_N(\rho_1(\gamma_s))}{1-s} \, \mathbf{d}\pi(\gamma),$$

thus recalling (5.46) and then the chain rules established in Section 4.3 we deduce

$$\begin{aligned} \limsup_{s \uparrow 0} \frac{\mathfrak{S}_N(\mu_1) - \mathfrak{S}_N(\mu_s)}{1-s} &\leq \int -\mathbf{d}^+(-\mathfrak{s}'_N(\rho_1))(\nabla f) |\mathbf{d}f|^{p-2} \rho_1 \, \mathbf{d}\mathbf{m} \\ &= \frac{c}{N} \int -\mathbf{d}^+(-\varphi)(\nabla f) |\mathbf{d}f|^{p-2} \, \mathbf{d}\mathbf{m}. \end{aligned}$$

Coupling this with the previously found lower bound gives the desired conclusion (5.47).

We conclude pointing out that, again in analogy with the $\text{TMCP}_+^h(K, N)$ case, combining the metric Brenier–McCann theorem with the existence of good final test plans give the following two results, valid in any mm spacetime as in (F) above (possibly relaxing the continuity assumption on ℓ_+ to upper semicontinuity):

(G) *(Unit maximal weak subslope of time separation from a point — compare with Corollary 5.23)*

For every $o \in \text{spt } \mathbf{m}$, the functions $f := u_q \circ \ell(o, \cdot)$ and $g := \ell(o, \cdot)$ satisfy

$$|\mathbf{d}f| = \ell(o, \cdot)^{q-1} \quad \text{and} \quad |\mathbf{d}g| = 1 \quad \mathbf{m}\text{-a.e. on } I^+(o).$$

(H) *(Sobolev-to-steepness property — compare with Theorem 5.17)*

Assume also that the spacetime is q -essentially timelike non-branching at 1, that ℓ_+ is continuous and that for any $x \ll y$ with $x, y \in \text{spt } \mathbf{m}$ we have $\mathbf{m}(U \cap J(x, y)) > 0$ for any U neighbourhood of y . Then $(M, \tau, \ell, \mathbf{m})$ has the Sobolev-to-steepness property.

A Appendices

A.1 Compatibility with the smooth setting

In this section we show how our calculus notions from the non-smooth setting reduce to well known ones if the underlying space (but not the objects we are differentiating) is smooth. We shall work with strongly causal spacetimes, i.e. smooth spacetimes M so that for every $p \in M$ and neighbourhood U of p there is another neighbourhood $V \subset U$ such that any causal curve with endpoints in V remains in U .

We start with differentiation of curves:

Theorem A.1 (Speed of causal curves: smooth vs non-smooth). *Let (M, g) be a smooth strongly causal spacetime and $\gamma : [0, 1] \rightarrow M$ a causal curve.*

Then:

i) γ is differentiable at a.e. t ;

ii) for a.e. t the Lorentzian norm $\|\dot{\gamma}'_t\|$ of its derivative coincides with the Lebesgue density $|\dot{\gamma}'_t|$ of the causal speed from Definition 2.25.

Proof. Consider first Minkowski space M . Then in a basis e_1, \dots, e_n such that the cone $\{\sum \alpha_i e_i : \alpha_i \geq 0\}$ contains the future cone, the curve γ is increasing in each component, hence a.e. differentiable. In the general case we can cover M with charts (U_i, φ_i) such that $d\varphi_i$ sends the future cone at any $x \in U$ inside the future cone in flat Minkowski space. Thus the post-composition $\varphi_i \circ \gamma$, where defined, is a causal curve in Minkowski and thus a.e. differentiable. At any differentiability (and thus also continuity) point $t \in [0, 1]$ we have $\dot{\gamma}'_t = \lim_{h \rightarrow 0} v_{t,h}$ where the future vectors $v_{t,h}$ are defined as $\frac{\exp_{\gamma_t}^{-1}(\gamma_{t+h})}{h}$ (because the differential of \exp_x at x is — by construction — the identity). Thus if $t \in [0, 1]$ is such that also (2.28) holds we have

$$|\dot{\gamma}'_t| = \lim_{h \downarrow 0} \frac{\ell(\gamma_t, \gamma_{t+h})}{h} = \lim_{h \downarrow 0} \frac{\|\exp_{\gamma_t}^{-1}(\gamma_{t+h})\|_{g_{\gamma_t}}}{h} = \|\dot{\gamma}'_t\|_{g_{\gamma_t}},$$

concluding the proof. \square

We now want to study the ‘dual’ problem of differentiability of causal functions on smooth spacetimes. To do so it is worth recalling few basic facts about Lorentzian scalar products.

Thus let V be a finite dimensional real vector space and g a scalar product on it with signature $(+, -, \dots, -)$. We shall always think such g as coming with a fixed half $F \subset V$ of the cone $\{v : g(v, v) \geq 0\}$ to be called the future; g induces on F a hyperbolic norm $\|\cdot\|$ (see also the terminology in Appendix A.3) via the formula

$$\|v\| := \sqrt{g(v, v)}. \quad (\text{A.1})$$

It is well known that on F we have the reverse Cauchy–Schwarz inequality:

$$g(v, w) \geq \|v\| \|w\| \quad \forall v, w \in F. \quad (\text{A.2})$$

A direct way of proving this is to identify (V, F) with the standard Minkowski space $\mathbf{R}^{1,n}$, $n \geq 0$, and the standard future cone, with coordinates (t, x) , so that (t, x) is a future vector iff $t \geq \|x\|_{\mathbf{R}^n}$ and $g((t, x), (s, y)) = ts - \langle x, y \rangle$, where $\|\cdot\|_{\mathbf{R}^n}, \langle \cdot, \cdot \rangle$ are the Euclidean norm and scalar product respectively. Then the standard Cauchy–Schwarz inequality yields $g((t, x), (s, y)) \geq ts - \|x\|_{\mathbf{R}^n} \|y\|_{\mathbf{R}^n}$ and (A.2) follows easily. This choice of coordinates also helps check that

$$v \in F \quad \Leftrightarrow \quad g(v, v') \geq 0 \quad \forall v' \in F. \quad (\text{A.3})$$

Indeed, \Rightarrow follows from (A.2). For \Leftarrow we notice that if $v = (t, x)$ is not in F , then $t < \|x\|_{\mathbf{R}^n}$, which easily implies the existence of $s \geq \|x\|_{\mathbf{R}^n}$ such that $ts < \|x\|_{\mathbf{R}^n}^2$. Thus the choice $v' := (s, x) \in F$ proves the claim.

As g is non-singular, it induces a musical (or Riesz) isomorphism $V \ni v \mapsto v^b \in V^*$ via $v^b := g(v, \cdot)$. Letting $V^* \ni \omega \mapsto \omega^\sharp \in V$ be the inverse isomorphism, g induces a bilinear form g^* on V^* by the formula

$$g^*(\omega_1, \omega_2) := g(\omega_1^\sharp, \omega_2^\sharp) \quad \forall \omega_1, \omega_2 \in V^*$$

and it is clear that g^* still has signature $(+, -, \dots, -)$. The dual cone $F^* \subset V^*$ of F is the collection of ω 's such that $\omega(v) \geq 0$ for any $v \in F$. Then from (A.3) it directly follows that

$$\omega \in F^* \quad \Leftrightarrow \quad \omega^\sharp \in F \quad \text{and} \quad v \in F \quad \Leftrightarrow \quad v^b \in F^*.$$

Calling $\|\cdot\|_*$ the norm induced by g^* on F^* , we also point out the duality formulas

$$\|\omega\|_* = \inf_{\substack{v \in F \\ \|v\| \geq 1}} \omega(v) \quad \text{and} \quad \|v\| = \inf_{\substack{\omega \in F^* \\ \|\omega\|_* \geq 1}} \omega(v) \quad (\text{A.4})$$

valid for any $v \in F$ and $\omega \in F^*$. Indeed the reverse Cauchy–Schwarz inequality gives $\omega(v) = g(\omega^\sharp, v) \geq \|\omega^\sharp\| \|v\| = \|\omega\|_* \|v\|$, while the choice $v := \frac{\omega^\sharp}{\|\omega\|_*}$ for $\|\omega\|_* > 0$ provides equality in the first. If $\|\omega\|_* = 0$ we perturb ω^\sharp as $\omega_\varepsilon^\sharp$ with $\|\omega_\varepsilon^\sharp\| = 1$ and $\omega(\omega_\varepsilon^\sharp) \rightarrow 0$ as $\varepsilon \searrow 0$. The second claim follows in an analogous way.

Now let g_1, g_2 be two Lorentzian scalar products on the same vector space V , with future cones F_1, F_2 , respectively. Then the definition of dual cone immediately gives

$$F_1 \subset F_2 \quad \Leftrightarrow \quad F_1^* \supset F_2^*$$

and if this occurs, then (A.4) also gives

$$g_1 \leq g_2 \quad \text{on } F_1 \quad \Leftrightarrow \quad g_1^* \geq g_2^* \quad \text{on } F_2^*. \quad (\text{A.5})$$

Let us say that $g_1 \prec g_2$ if $F_1 \setminus \{0\} \subset \text{int}(F_2)$ and $g_1(v, v) < g_2(v, v)$ for any $v \in F_1 \setminus \{0\}$. With this definition it is easy to see, given g_1 , that

$$\{g_2 \text{ Lorentzian scalar products on } V : g_1 \prec g_2\} \text{ is open,}$$

w.r.t. the topology of uniform convergence on a compact subset of V^2 with non-empty interior (this is independent of the chosen subset and coincides with the topology of uniform convergence of the coefficients of the tensor in any given basis of V). Our nonstandard definition of $g_1 \prec g_2$ yields

$$g_1 = \inf\{g_2 : g_1 \prec g_2\}, \quad \text{meaning that} \quad \begin{cases} F_1 = \bigcap_{g_2 \succ g_1} F_2, \\ g_1(v, v) = \inf_{g_1 \prec g_2} g_2(v, v) \quad \forall v \in F_1. \end{cases} \quad (\text{A.6})$$

This can be seen by identifying (V, g_1) with the standard Minkowski space and then ‘slightly widening’ the light cone.

We are ready to prove the second, and last, main result of this section:

Theorem A.2 (Differential of causal functions: smooth vs non-smooth). *Let (M, g) be a smooth strongly causal spacetime and $f : M \rightarrow \mathbf{R}$ a causal function.*

Then f is vol_g -measurable and:

- i) f is differentiable at vol_g -a.e. point with $\text{D}f(x)$ in $F_x^*M \subset T_x^*M$ for vol_g -a.e. $x \in M$,*
- ii) the Lorentzian (dual) norm $\|\text{D}f\|_*$ coincides with the maximal weak subslope $|\text{d}f|$ vol_g -a.e.*

Furthermore, f is locally a BV function and $\text{D}f$ is (a representative of) the density w.r.t. vol_g of the absolutely continuous part of its distributional derivative.

Proof.

STEP 1: SMOOTH FUNCTIONS ON MINKOWSKI. We claim that for $f : \mathbf{R}^{1,n} \rightarrow \mathbf{R}$ smooth and causal we have $|\text{d}f| = \|\text{D}f\|_*$ a.e. Indeed, for $\gamma : [0, 1] \rightarrow \mathbf{R}^{1,n}$ causal, the function $f \circ \gamma$ is monotone, hence differentiable a.e. and denoting by $(f \circ \gamma)'$ its a.e. defined derivative we have $f(\gamma_1) - f(\gamma_0) \geq \int_0^1 (f \circ \gamma)'_t dt$. Since f is smooth, using Theorem A.1 we see that

$$(f \circ \gamma)'_t = \text{D}f_{\gamma_t}(\gamma'_t) \geq \|\text{D}f\|_*(\gamma_t) \|\gamma'_t\| = \|\text{D}f\|_*(\gamma_t) |\dot{\gamma}_t| \quad \text{a.e. } t \in [0, 1].$$

It follows from Lemma 3.13 that $\|\text{D}f\|_*$ is a weak subslope of f and thus by maximality that $|\text{d}f| \geq \|\text{D}f\|_*$ holds vol_g -a.e. For the converse inequality let $\mu \in \mathcal{P}(\mathbf{R}^{1,n})$ be with $\text{vol}_g \ll \mu \leq \text{vol}_g$ and $v \in T\mathbf{R}^{1,n} \sim \mathbf{R}^{1,n}$ be a future vector with $\|v\| = 1$. Let T be the map sending $x \in \mathbf{R}^{1,n}$ to the causal curve $t \mapsto x + tv$ and notice that $\pi := T_*\mu$ is a test plan so that by Lemma 3.13, since $t \mapsto f(x + tv)$ is smooth we obtain

$$\text{D}f(x + tv)(v) \geq |\text{d}f|(x + tv) \quad \mathcal{L}^{n+1} - \text{a.e. } (x, t). \quad (\text{A.7})$$

By Fubini's theorem and the translation invariance of \mathcal{L}^{n+1} this easily implies $Df(v) \geq |df|$ \mathcal{L}^{n+1} -a.e. Taking the inf on a countable dense set of v 's, by (A.4) we conclude that $\|Df\|_* \geq |df|$, as desired.

STEP 2: CAUSAL FUNCTIONS ON MINKOWSKI SPACE. Let $f : \mathbf{R}^{1,n} \rightarrow \mathbf{R}$ be causal. Pick a basis e_0, \dots, e_n of $\mathbf{R}^{1,n}$ made of vectors in F , so that the cone generated by the e_i 's is contained in F . In this coordinate system f is coordinate-wise monotone. By [42, Thm. 4] this suffices to show measurability. Also, for any $a_0, \dots, a_n \in [0, 1]$ and $i = 0, \dots, n$, the distributional derivative of $t \mapsto f(\dots, a_{i-1}, t, a_{i+1}, \dots)$ is a non-negative measure of mass bounded by $f(1, \dots, 1) - f(0, \dots, 0)$. It follows from the characterization in [7, Rem. 3.104] that f is locally a BV function, as claimed. In particular, it is approximately a.e. differentiable by [7, Thm. 3.83] and the approximate differentiability improves to actual differentiability thanks to the coordinate-wise monotonicity, see [42, Thm. 14] (see also the presentation in [89, Thm. 1.19]). Thus calling $\mathcal{D}f$ the co-vector valued measure (σ -additive only on uniformly bounded subsets — see also Remark A.7) representing the distributional derivative of f , the Lebesgue decomposition

$$\mathcal{D}f = Df \mathcal{L}^{1+n} + \mathcal{D}f^\perp \quad \text{with} \quad \mathcal{D}f^\perp \perp \mathcal{L}^{1+n} \quad (\text{A.8})$$

follows again from [7, Thm. 3.83].

We now claim that for any $E \subset \mathbf{R}^{1,n}$ Borel and bounded we have $\mathcal{D}f(E) \in F^*$. By definition of F^* and an approximation argument this will follow if we show that for any $v \in F$ and $\varphi \in C_c^\infty(\mathbf{R}^{1,n})$ non-negative we have $\int \varphi d\mathcal{D}f(v) \geq 0$ (here and below $\mathcal{D}f(v)$ is defined on bounded Borel sets as $\mathcal{D}f(v)(E) := \mathcal{D}f(E)(v)$ and — trivially — coincides with the distributional derivative of f in the direction v). We thus have

$$\int \varphi d\mathcal{D}f(v) = \int -D\varphi(v) f d\mathcal{L}^{1+n} = \lim_{h \rightarrow 0} \int \frac{\varphi(\cdot) - \varphi(\cdot + hv)}{h} f d\mathcal{L}^{1+n} = \lim_{h \rightarrow 0} \int \varphi \frac{f(\cdot + hv) - f(\cdot)}{h} d\mathcal{L}^{1+n}$$

and the latter quantity is non-negative as f is causal, $v \in F$ and $\varphi \geq 0$. Hence our claim is proved and from the decomposition (A.8) it follows that

$$Df \in F^* \quad \mathcal{L}^{1+n} - \text{a.e.} \quad \text{and} \quad \mathcal{D}f^\perp(E) \in F^* \quad \forall E \subset \mathbf{R}^{1,n} \text{ Borel and bounded.} \quad (\text{A.9})$$

It remains to prove that $\|Df\|_* = |df|$. Start by observing that for any future $v \in F$, Fubini's theorem and the translation invariance of \mathcal{L}^{1+n} yield for a.e. $x \in \mathbf{R}^{1,n}$ and a.e. $t \in [0, 1]$ that the point $x + tv$ is a differentiability point of f . Thus the arguments leading to (A.7) are still valid, hence so is (A.7), and as above we conclude that $\|Df\|_* \geq |df|$ holds \mathcal{L}^{n+1} -a.e.

For the converse inequality let (ρ_ε) be a family of mollifiers and notice that $f * \rho_\varepsilon$ is still a causal function. Thus from the previous step and Proposition 3.16 we have $|df| \geq \liminf_{\varepsilon \downarrow 0} |d(f * \rho_\varepsilon)| = \liminf_{\varepsilon \downarrow 0} \|D(f * \rho_\varepsilon)\|_*$ (here the lim inf are intended as \mathcal{L}^{1+n} -essential-lim inf). Since 'convolution' and 'distributional differentiation' commute, recalling (A.8) we have

$$\|D(f * \rho_\varepsilon)\|_* = \|(Df) * \rho_\varepsilon\|_* = \|(Df) * \rho_\varepsilon + (\mathcal{D}f^\perp) * \rho_\varepsilon\|_* \quad \text{on} \quad \mathbf{R}^{1,n}$$

and since (A.9) implies that $(Df) * \rho_\varepsilon, (\mathcal{D}f^\perp) * \rho_\varepsilon \in F^*$, the above and the reverse triangle inequality give $\|D(f * \rho_\varepsilon)\|_* \geq \|(Df) * \rho_\varepsilon\|_*$ on $\mathbf{R}^{1,n}$. Now we observe — the computations being justified because Df is locally integrable — that

$$\|(Df) * \rho_\varepsilon\|_* = \left\| \int (Df)(\cdot - y) \rho_\varepsilon(y) d\mathcal{L}^{1+n}(y) \right\|_* \geq \int \|Df\|_*(\cdot - y) \rho_\varepsilon(y) d\mathcal{L}^{1+n}(y) = \|Df\|_* * \rho_\varepsilon,$$

where the inequality follows from the concavity and continuity of $\|\cdot\|_*$ on F^* . Since $\|Df\|_* * \rho_\varepsilon \rightarrow \|Df\|_*$ \mathcal{L}^{1+n} -a.e., collecting what we proved we conclude that $|df| \geq \|Df\|_*$ \mathcal{L}^{1+n} -a.e., as desired.

STEP 3: GENERAL CASE. We argue by comparison. Say that a coordinate chart (U, φ) with $U \subset M$ open and $\varphi : U \rightarrow \mathbf{R}^{1,n}$ is 'good' provided:

- 1) U and $\varphi(U)$ are causally convex.

- 2) For any $x \in U$ the pullback $\varphi^* g_{1,n}$ of the Minkowskian tensor at $\varphi(x) \in \mathbf{R}^{1,n}$ is such that $(\varphi^* g_{1,n})_x^* \prec g_x^*$ (recall the discussion after (A.5)).

Clearly good charts cover M , and recalling the dual relation (A.5) we see that if (U, φ) is a good chart, then $f \circ \varphi^{-1}$ is a causal function on $\varphi(U) \subset \mathbf{R}^{1,n}$, hence as already argued it is measurable and a.e. differentiable with differential in $F^* \subset (\mathbf{R}^{1,n})^*$. Thus f is vol_g -measurable, vol_g -a.e. differentiable and for every good chart (U, φ) we have $D(f \circ \varphi^{-1}) \in F^* \subset (\mathbf{R}^{1,n})^*$. Then an approximation argument based on (A.6) and the fact that \prec is open shows that $Df(x) \in F_x^* M \subset T_x^* M$, thus (i) is settled.

We now prove that $|df| \geq \|Df\|_*$ holds vol_g -a.e. To see this, let π be a test plan on M and notice that for π -a.e. γ the function $f \circ \gamma$ is monotone (trivially) and differentiable at t with derivative given by $Df_{\gamma_t}(\gamma'_t) \geq \|Df_{\gamma_t}\|_* \|\gamma'_t\|$ (with an argument based on the a.e. differentiability of f , Fubini's theorem and (A.4)). It follows from the maximality of $|df|$ and Lemma 3.13 that $|df| \geq \|Df\|_*$ holds vol_g -a.e., as desired.

For the opposite inequality notice that a simple argument based on (A.6) (applied for $g_1 := g^*$) and on the fact that the relation \prec is open shows that we can find a countable collection (U_i, φ_i) of good charts such that

$$g_x^* = \inf_{i: x \in U_i} (\varphi_i^* g_{1,n})_x^* \quad \text{this being interpreted as in (A.6).}$$

Thus if we adopt the convention that $\|D(f \circ \varphi^{-1})\|_* \circ \varphi$ is equal to $+\infty$ outside U , from the above we see that

$$\|Df\|_* = \inf_i \|D(f \circ \varphi_i^{-1})\|_* \circ \varphi_i = \inf_i |d(f \circ \varphi_i^{-1})| \circ \varphi_i \quad \text{vol}_g\text{-a.e.}, \quad (\text{A.10})$$

having used — in the second identity — that we know the conclusion on flat Minkowski space. Now notice that since, for every good chart (U, φ) , the sets U and $\varphi(U)$ are causally convex and the ambient spacetimes are globally hyperbolic, we have that both U and $\varphi(U)$ are forward mm spacetimes when equipped with the structure induced by the ambient spaces. Also, from the dual relation (A.5) we see that $\ell(\varphi_i^{-1}(p), \varphi_i^{-1}(q)) \geq \|q - p\|$ for any $p, q \in \varphi_i(\bar{U}_i) \subset \mathbf{R}^{1,n}$, thus the implication (i) \Rightarrow (ii) in Lemma 5.15 allows to deduce that $|d(f \circ \varphi^{-1})| \geq |df| \circ \varphi^{-1}$. Recalling (A.10) we can conclude that $\|Df\|_* \geq |df|$, as desired. \square

An easy consequence of this last theorem, of $\lim_{\varepsilon \downarrow 0} \frac{g^*(\omega + \varepsilon \eta, \omega + \varepsilon \eta) - g^*(\omega, \omega)}{2\varepsilon} = g^*(\omega, \eta) = \omega^\sharp(\eta)$ valid for any cotangent vectors ω, η , and of the chain rule is that the value of

$$d\varphi(\nabla f) |df|^{p-2}$$

on $\{|df| \in (0, +\infty)\}$ for $f : M \rightarrow \mathbf{R}$ causal and $\varphi \in \text{Pert}(f)$ is independent of the fact that we are considering it as in Definition 4.10 or directly as the coupling of the cotangent vector field $d\varphi$ (well defined as φ is the difference of two causal functions) with the vector field ∇f multiplied by $|df|^{p-2}$. This observation is relevant in interpreting identity (A.11) in the next section.

A.2 Null distance and perturbations of causal functions

This section deals with elementary properties of causal functions and the null distance. Our goal is to show that the latter can be used to produce ‘many’ perturbations of the time separation and related functions. The technique is flexible and likely applicable to more general causal functions: we chose to focus on powers of time separation to a point to present a concrete case study, which can also be used to build a measure valued p -d’Alembertian following the presentation in Section 5.6.

The concept of null distance was originally introduced in the smooth setting by Sormani and Vega [112] as an attempt at bridging the gap between topology and causality on Lorentzian manifolds. It has subsequently been studied in the synthetic setting of Lorentzian pre-length spaces in Kunzinger and Steinbauer [76]. Let us tailor the discussion on this latter reference to our setting:

Definition A.3 (Null distance). *Let (M, ℓ) be metric spacetime and $f: M \rightarrow \bar{\mathbf{R}}$ a causal function. The null distance \hat{d}_f induced by f is defined as*

$$\hat{d}_f(x, y) := \inf \sum_{i=0}^{n-1} |f(x_i) - f(x_{i+1})|,$$

where the inf is taken among all $n \in \mathbf{N}$ and sequences $x = x_0, \dots, x_n = y$ such that for any $i = 0, \dots, n-1$ we either have $x_i \leq x_{i+1}$ or $x_{i+1} \leq x_i$.

It is clear that $\hat{d}_f: M^2 \rightarrow [0, +\infty]$ is symmetric, satisfies the triangle inequality and $\hat{d}_f(x, x) = 0$ for any x . In general, nothing more can be said, as it could both be that $\hat{d}_f \equiv 0$ (e.g. if f is constant) and that $\hat{d}_f(x, y) = +\infty$ for any $x \neq y$ (if for no $x \neq y \in M$ we have $x \leq y$).

We record the following simple general statement:

Proposition A.4 (\hat{d}_f -Lipschitz functions are perturbations). *Let (M, ℓ) be a metric spacetime, $f: M \rightarrow \mathbf{R}$ causal and \hat{d}_f the corresponding null distance. Let $g: M \rightarrow \mathbf{R}$ be \hat{d}_f -Lipschitz, i.e. $|g(x) - g(y)| \leq L\hat{d}_f(x, y)$ for some $L \geq 0$ and any $x, y \in M$.*

Then $f + \frac{1}{L}g: M \rightarrow \mathbf{R}$ is causal.

Proof. By scaling, we may and will assume that g is 1-Lipschitz with respect to \hat{d}_f . Then

$$-(g(y) - g(x)) \leq |g(y) - g(x)| \leq \hat{d}_f(x, y) \leq f(y) - f(x), \quad \forall x \leq y,$$

proving that $f + g$ is causal and thus giving the claim. \square

Thanks to this result the question is now to find sufficient conditions that ensure that \hat{d}_f -Lipschitz functions are continuous and suitably dense in the space of continuous functions, so that the distributional p -d'Alembertian as studied in Proposition 5.31 can be represented by a measure via Riesz's theorem.

We do not have general conditions on non-smooth spacetimes, but the results in [112] allow to quickly prove the following, showing that our theory is compatible with this basic case.

Proposition A.5 (Metrizing on manifolds by the null distance of time separation to a point). *Let (M, g) be a smooth, globally hyperbolic spacetime, let $o \in M$ and consider the causal function $f := \ell(o, \cdot)$. Then \hat{d}_f is a distance on $I^+(o)$ that induces the manifold topology.*

The same holds for the function $f := -\ell(\cdot, o)$ on $I^-(o)$ and, more generally, for $\frac{1}{q}\ell(o, \cdot)^q$ and $-\frac{1}{q}\ell(\cdot, o)^q$ on $I^+(o)$ and $I^-(o)$ respectively for any $0 \neq q \leq 1$.

Proof. We shall consider the spacetime $\tilde{M} := I^+(o)$. According to [112, Proposition 3.15, Theorem 5.4] the conclusion for $f = \ell(o, \cdot)$ will follow if we show that f is the cosmological time function of \tilde{M} and that is regular in the terminology of [13] recalled below.

For f to be the cosmological time function means that it satisfies $f(x) = \sup_{y \in \tilde{M}, y \leq x} \ell(y, x)$: this follows from the continuity of f on $J^+(o) \supset I^+(o)$ (see [84, Theorem 3.6]).

For f to be regular means that it is finite valued on \tilde{M} (which it clearly is) and that if $\gamma: [0, 1) \rightarrow \tilde{M}$ is any inextendible past curve, then $f(\gamma_t) \rightarrow 0$ as $t \rightarrow 1$. We verify this latter property and notice that for such γ we have $\gamma_t \in J^-(\gamma_0) \cap J^+(o) \subset M$ for any $t \in [0, 1)$. Thus using first the global hyperbolicity of M and then the causal character of γ we see that for some $p \in M$, $p \geq o$ we have $\gamma_t \rightarrow p$ as $t \rightarrow 1$. Since γ is past inextendible on \tilde{M} , we must have $p \notin \tilde{M}$, that together with $p \geq o$ yields $f(p) = 0$. It then follows from the continuity of f on $J^+(o)$ that $f(\gamma_t) \rightarrow 0$ as $t \rightarrow 1$.

The case of $f = -\ell(\cdot, o)$ is handled analogously, then the rest of the claim follows noticing that $u_q(z) = \frac{1}{q}z^q$ is smooth on $(0, \infty)$ and its inverse is smooth on $u_q((0, \infty))$. \square

We illustrate how the above can be combined with the p -d'Alembertian estimates we proved in Section 5.5 and the concept of distributional p -d'Alembertian discussed in Section 5.6. Recall also the discussion at the end of Appendix A.1 that ensure that the meaning of $d\varphi(|\nabla f|)^{p-2}$ appearing below is unambiguous.

Corollary A.6 (Distributional p -d'Alembertian in the smooth setting). *Let (M, g) be a smooth, globally hyperbolic spacetime of dimension n whose timelike Ricci curvature is no smaller than $K \in \mathbf{R}$, endowed with its canonical volume measure \mathbf{m} .*

Moreover, let $0 \neq q < 1$ and $p^{-1} + q^{-1} = 1$. Given a point $o \in M$, consider the open set $U \subset M$ defined as $I^-(o)$ if $K \leq 0$ and $I^-(o) \cap \{\ell(\cdot, o) < \pi\sqrt{(N-1)/K}\}$ if $K > 0$. Also, consider the functions $f := -\frac{1}{q}\ell(\cdot, o)^q$ and $h := -\ell(\cdot, o)$ and the Radon measures ν_f, ν_h on U given by

$$\nu_f := N \tilde{\tau}_{K,N} \circ \ell(o, \cdot) \mathbf{m}, \quad \nu_h := \frac{N \tilde{\tau}_{K,N} \circ \ell(\cdot, o) - 1}{\ell(\cdot, o)} \mathbf{m}.$$

Then f and h both have distributional p -d'Alembertian in U according to Definition 5.30 and there are unique non-positive Radon measures μ_f, μ_h on U such that

$$\int \varphi d\mu_f = \int d\varphi(\nabla f) |df|^{p-2} d\mathbf{m} - \int \varphi d\nu_f \quad \forall \varphi \in \text{Pert}_{\text{be}}^{\text{sym}}(f, U) \cap C(U) \quad (\text{A.11})$$

and similarly for h in place of f .

Proof. The collection $\text{Pert}_{\text{be}}^{\text{sym}}(f, U) \cap C(U)$ defined after (5.42) is a vector space and the map L sending $\varphi \in \text{Pert}_{\text{be}}^{\text{sym}}(f, U) \cap C(U)$ to the right-hand side of (A.11) is linear and sends non-negative functions to non-positive reals.

We claim that $L(\varphi_n) \uparrow 0$ if $(\varphi_n) \subset \text{Pert}_{\text{be}}^{\text{sym}}(f, U) \cap C(U)$ is such that $\varphi_n(x) \downarrow 0$ as $n \rightarrow \infty$ for every x . By Dini's theorem and the fact that $\text{spt } \varphi_n \subset \text{spt } \varphi_0$ (the latter being compact as the spacetime is globally hyperbolic) we get that (φ_n) converges to 0 uniformly. Also, by Proposition A.5 for every $C \subset U$ compact, there is $\eta \in \text{Pert}_{\text{be}}^{\text{sym}}(f, U) \cap C(U)$ identically 1 on C (pick $\eta := (1 - n\hat{d}_f(\cdot, C))^+$ for $n \gg 1$). Picking $C := \text{spt } \varphi_0$ and using the bound (5.45) in conjunction with the uniform convergence just proved shows our claim.

We can thus apply [19, Thm. 7.8.1] to $-L$ and deduce that $L(\varphi) = \int \varphi d\mu_f$ for a unique non-positive measure μ_f on U defined on the smallest σ -algebra \mathcal{A} making all the functions in $\text{Pert}_{\text{be}}^{\text{sym}}(f, U) \cap C(U)$ measurable. Clearly \mathcal{A} is contained in the Borel σ -algebra of U and to show that they agree it suffices to prove that any function in $C_c(U)$ can be uniformly approximated by functions in $\text{Pert}_{\text{be}}^{\text{sym}}(f, U) \cap C(U)$. This, however, is obvious from the fact that the null distance induced by f induces, in turn, the manifold topology. The fact that μ_f is Radon now follows noticing that — by (5.45) — it is finite on compact sets.

The argument for h is analogous. □

Remark A.7 (Differences of positive measures need not be measures). Since $\mu_f \leq 0$, by (A.11) it is surely tempting to say that the p -d'Alembertian of f is represented by a measure $\leq \nu_f$, this 'measure' being $\mu_f + \nu_f$. The problem with this is that it might be that $\mu_f(U) = -\infty$ and $\nu_f(U) = +\infty$, so that their sum is not a measure on U as classically intended. This happens because, more broadly speaking, the collection of Radon measures on a given topological space is not a vector space in general.

To get around this technical issue, one possibility is to keep in mind Riesz's representation theorem and to deal with the space of so-called Radon functionals, as in [37]. Another possibility is to 'ignore' this issue and to notice that objects like $\mu_f + \nu_f$ are still well-defined and σ -additive as far as one works with uniformly relatively compact subsets of U . In particular, the 'measure' of any relatively compact set is well defined and so is the integral of compactly supported functions. This sort of consideration also allows to perform operations typically done on measures (e.g. push forward, Radon-Nikodym derivatives...) whose interpretation at the level of Radon functionals would be cumbersome. Because of this some authors (see e.g. [6, Remark 2.12]) adopt this point of view and still call Radon measures objects such as $\mu_f + \nu_f$. Notice that this is what we also did in the course of the proof of Theorem A.2, when we represented by the 'measure' $\mathcal{D}f$ the distributional differential of the BV_{loc} function f . ■

A.3 Hyperbolic norms and polarization on cones

This section deals with a class of Lorentzian structures which arises on vector spaces from hyperbolic norms, such as Lorentz–Finsler norms. We discuss polarization and weak curvature properties.

Definition A.8 (Hyperbolic norm). *A hyperbolic norm on a real vector space V is a function $\mathfrak{n} : V \rightarrow [0, +\infty) \cup \{-\infty\}$ such that*

$$\mathfrak{n}(\alpha v + \beta w) \geq \alpha \mathfrak{n}(v) + \beta \mathfrak{n}(w) \quad \forall v, w \in V, \alpha, \beta \geq 0, \quad (\text{A.12a})$$

$$\mathfrak{n}(v), \mathfrak{n}(-v) \geq 0 \quad \Leftrightarrow \quad v = 0. \quad (\text{A.12b})$$

Example A.9. Let $p \in [1, +\infty)$, $d \in \mathbf{N}$, $d > 0$ and define \mathfrak{n}_p on \mathbf{R}^d as

$$\mathfrak{n}_p(v) := \left(|v_0|^p - \sum_{i=1}^{d-1} |v_i|^p \right)^{1/p}$$

if $|v_0|^p \geq \sum_{i=1}^{d-1} |v_i|^p$ and $\mathfrak{n}_p(v) = -\infty$ otherwise. Then this is a hyperbolic norm.

For $p = 2$ this coincides with the norm induced by the Lorentzian scalar product $g(v, w) = v_0 w_0 - \sum_{i>0} v_i w_i$ as in formula (A.1) and in this case (and only in this if $d > 1$) the norm \mathfrak{n}_p is positively polarizable in the sense discussed below. For more about these norms we refer to [58]. ■

Hyperbolic norms naturally induce time separation functions via the formula

$$\ell(v, w) := \mathfrak{n}(w - v) \quad \forall v, w \in V, \quad (\text{A.13})$$

and it is clear from (A.12a) that ℓ satisfies the reverse triangle inequality and from (A.12b) that the causal order induced by ℓ is a partial order. Thus, much like in the non-smooth setting, \mathfrak{n} induces the causal and chronological relations on V and the latter one induces a topology. The following is worth noting:

Proposition A.10 (Hyperbolic normed vector spaces as globally hyperbolic metric spacetimes). *Let \mathfrak{n} be an hyperbolic norm on \mathbf{R}^N , where $N \in \mathbf{N}$. Assume that there exists $v \in \mathbf{R}^N$ with $\mathfrak{n}(v) > 0$ and that the positive part \mathfrak{n}_+ of \mathfrak{n} is continuous (w.r.t. the Euclidean topology).*

Then (\mathbf{R}^N, ℓ) is a globally hyperbolic metric spacetime, where ℓ is as in (A.13), and the chronological topology coincides with the Euclidean one.

Proof. The continuity of \mathfrak{n}_+ implies that the chronological topology is weaker than the Euclidean one. For the converse inclusion it suffices to prove that there is a sufficient widening of the usual Minkowskian cones in some coordinate system that contain the future cone $\{v : \mathfrak{n}(v) \geq 0\}$ (because the chronological relation induced by the Minkowskian norm coincides with the Euclidean topology). To see this, let $Y := \{\mathfrak{n} \geq 0\} \cap \mathbf{S}^{N-1}$. Then by (A.12a), Y is a compact, convex subset of \mathbf{S}^{N-1} (considered as a Riemannian manifold) that, by (A.12b), does not contain antipodal points. Thus it is contained in a half-sphere and, by compactness, in an \mathbf{S}^{N-1} -ball of radius $R < \pi/2$. This readily implies our claim.

For global hyperbolicity we now observe that for any $C_0, C_1 \subset \mathbf{R}^N$ compact, what was already proved implies that $J(C_0, C_1) \subset \tilde{J}(C_0, C_1)$, where $\tilde{J}(C_0, C_1)$ is the emerald spanned by C_0, C_1 under this extra Minkowskian structure. Since the latter is compact and the former closed, the conclusion follows. □

We come to weak lower Ricci bounds, referring to [21] for the concept of $\text{TCD}_q(K, N)$ space. The proof below strongly imitates that of the analogous result in positive signature provided in [119, p. 926], thus we shall not give all the details.

Proposition A.11 (TCD condition for hyperbolic normed vector spaces). *Let \mathfrak{n} be as in Proposition A.10. Then the induced globally hyperbolic metric measure spacetime structure on \mathbf{R}^N (with \mathfrak{m} being the Lebesgue measure \mathcal{L}^N) satisfies $\text{TCD}_q(0, N)$ for every $0 \neq q < 1$.*

Proof. We only sketch the argument freely using terminology and notation from [26], to which the reader is referred for details. We first consider the case $n|_{\{n>0\}}$ is smooth and n is locally uniformly concave. Then the map $L: \mathbf{R}^N \times \mathbf{R}^N \rightarrow \mathbf{R}$ defined by

$$L(x, v) := \begin{cases} n^2(v) & \text{if } n(v) > 0, \\ 0 & \text{otherwise} \end{cases}$$

defines a smooth Lorentz–Finsler structure on \mathbf{R}^N . Given any pair (μ, ν) of compactly supported elements of $\mathcal{P}(M)$ with μ being \mathbf{m} -absolutely continuous which are q -separated by (π, u, v) [26, Def. 4.8], $\pi = (\text{Id}, \mathcal{F}_1)_* \mu$ is the unique maximizer of $\ell_q(\mu, \nu)$, where $\mathcal{F}_t: \mathbf{R}^N \rightarrow \mathbf{R}^N$ is the Lorentz–Finslerian displacement interpolation and is of simple linear form (both in its argument and in t ; recall that we have unique geodesics given by straight lines) by [26, Thm. 4.17]. Moreover, the unique ℓ_q -geodesic from μ to ν is given by $\mu_t := (\mathcal{F}_t)_* \mu$ [26, Cor. 4.18]. Keeping this in mind, the fact that this smooth Lorentz-Finsler spacetime satisfies $\text{TCD}_q(0, N)$ follows by the same (linear algebraic) arguments as the $\text{CD}(0, N)$ -condition for smooth, uniformly convex norms on \mathbf{R}^N , cf. [119, p. 926] for details.

For general n we first ‘widen a bit the future cone’ and then regularize the norm, thus obtaining approximating norms n_n whose future cones $F_n := \{n_n(v) \geq 0\}$ contain the limit one $F := \{n \geq 0\}$ in their interior. Given $\mu, \nu \in \mathcal{P}(\mathbf{R}^n)$ with compact support and causally ordered $\mu \preceq \nu$, we have $\mu \preceq_n \nu$ for every $n \in \mathbf{N}$, thus by the previous point we know there is a geodesic $(\mu_{n,t})$ (w.r.t. the q cost induced by the ℓ_n coming from n_n via (A.13)) from μ to ν along which the N -Renyi entropy is convex. The measures $\{\mu_{n,t} : n \in \mathbf{N}, t \in [0, 1]\}$ all have support in the same compact set, thus by diagonalization we can assume that $\mu_{n,t} \rightarrow \mu_t$ for every $t \in \mathbf{Q} \cap [0, 1]$ and some measures μ_t . The lower semicontinuity of the entropy grants that μ_t obeys the correct entropy bounds, while passing to the limit in

$$\ell_{n,q}(\mu_{n,t}, \mu_{n,s}) = (s - t)\ell_{n,q}(\mu, \nu) \quad \forall t, s \in [0, 1], t < s$$

we easily see that the μ_t ’s lie along an ℓ_q -geodesic from μ to ν . By narrow forward completeness we can extend the curve (μ_t) to all $t \in [0, 1]$ resulting in a ℓ_q -geodesic and then narrow lower semicontinuity of the entropy yields the desired conclusion. \square

Let $n: V \rightarrow \{-\infty\} \cup [0, +\infty)$ be a hyperbolic norm on a vector space V with domain (or future) $F := \{x \in V \mid n(x) \geq 0\}$. We say $n(\cdot)$ is *positively polarizable* provided

$$(x, y) := \frac{1}{2}(n(x+y)^2 - n(x)^2 - n(y)^2) \tag{A.14}$$

acts positive bilinearly on F , or equivalently if

$$(\alpha_1 x_1 + \alpha_2 x_2, y) = \alpha_1(x_1, y) + \alpha_2(x_2, y) \quad \forall x_1, x_2, y \in F, \alpha_1, \alpha_2 \geq 0. \tag{A.15}$$

Notice that by symmetry, if the above holds, a similar statement is in place for y . It is quite clear that if $n: V \rightarrow \{-\infty\} \cup [0, +\infty)$ is positively polarizable, then the form (x, y) defined by (A.14) for $x, y \in F$ extends uniquely to a indefinite inner product $\langle \cdot, \cdot \rangle$ on $\text{span}(F)$. Indeed, uniqueness is clear as well as the fact that we must have

$$\langle x_1 - x_2, y \rangle = \langle x_1, y \rangle - \langle x_2, y \rangle \quad \forall x_1, x_2, y \in F.$$

The conclusion follows observing that this formula is well posed, i.e. that if $x_1 - x_2 = z_1 - z_2$, then the corresponding right hand sides in the above agree: this is a trivial consequence of (A.15).

The next lemma reinforces and explains our Definition 1.4 of infinitesimal Minkowskianity.

Lemma A.12 (Parallelogram law and polarization of hyperbolic norms). *Let $n: V \rightarrow \{-\infty\} \cup [0, +\infty)$ be a hyperbolic norm on a real vector space V .*

Then n is positively polarizable if and only if

$$n(x + 2y)^2 + n(x)^2 = 2n(x + y)^2 + 2n(y)^2, \quad \forall x, y \in V, \text{ with } n(x), n(y) \geq 0. \tag{A.16}$$

Proof. The *only if* is a trivial consequence of the identity $\mathfrak{n}(v) = (v, v)$ and of (A.15). For the *if* we argue as in the proof of Theorem 4.16. Thus let $x, y, z \in V$ be with $\mathfrak{n}(x), \mathfrak{n}(y), \mathfrak{n}(z) \geq 0$ and notice that (A.16) yields

$$\begin{aligned} 2\mathfrak{n}(x + 2y + z)^2 + 2\mathfrak{n}(x + z)^2 &= 4\mathfrak{n}(x + y + z)^2 + 4\mathfrak{n}(y)^2, \\ \mathfrak{n}(2x + 2y + z)^2 + \mathfrak{n}(2y + z)^2 &= 2\mathfrak{n}(x + 2y + z)^2 + 2\mathfrak{n}(x)^2, \\ \mathfrak{n}(2y + z)^2 + \mathfrak{n}(z)^2 &= 2\mathfrak{n}(y + z)^2 + 2\mathfrak{n}(y)^2, \\ \mathfrak{n}(2x + 2y + z)^2 + \mathfrak{n}(z)^2 &= 2\mathfrak{n}(x + y + z)^2 + 2\mathfrak{n}(x + y)^2. \end{aligned}$$

Adding the first two identities and then subtracting the second two we conclude that $(x + y, z) = (x, z) + (y, z)$. It then follows that $(\alpha x, y) = \alpha(x, y)$ holds for $\alpha \in \mathbf{N}$ and the for $\alpha \in \mathbf{Q}^+$. The monotonicity of $\mathbf{R}^+ \ni \alpha \mapsto \mathfrak{n}(\alpha x + y)$ that comes from (A.12a) now allows to conclude that $(\alpha x, y) = \alpha(x, y)$ holds for $\alpha \in \mathbf{R}^+$, as desired. \square

Finally, let us demonstrate that if a symmetric bilinear form induces a “norm” that satisfies either the triangle inequality or the reverse triangle inequality, then the form is necessarily positive definite or of Lorentzian signature, respectively.

Lemma A.13 (Triangle inequalities indicate signature). *Let V denote a finite-dimensional real vector space and $g: V \times V \rightarrow \mathbf{R}$ a non-degenerate symmetric bilinear form. We denote by C a cone with non-empty interior contained in $\{v : g(v, v) > 0\}$ and define $\|v\| := \sqrt{g(v, v)}$ for $v \in C$.*

Then the following statements hold.

- (i) *Assume that $\|\cdot\|$ satisfies the triangle inequality, i.e. $\|v + w\| \leq \|v\| + \|w\|$ for all $v, w \in C$. Then g is positive definite.*
- (ii) *Assume that $\|\cdot\|$ satisfies the reverse triangle inequality, i.e. $\|v + w\| \geq \|v\| + \|w\|$ for all $v, w \in C$. Then g has Lorentzian signature $(+, -, \dots, -)$.*

Proof. Pick v in the interior of C and recall that, regardless of the signature of g , there is an orthonormal basis e_1, \dots, e_n of V such that $g(e_i, e_j) = 0$ for $i \neq j$ and $g(e_i, e_i) = \pm 1$ and so that $e_1 := \frac{v}{\|v\|}$. This can be proved by induction by noticing that the orthogonal complement of e_1 is a subspace of dimension exactly $n - 1$. Unless $n = 1$, in such subspace there must be a vector v' with $g(v', v') \neq 0$, so that an iteration proves the claim.

The signature of g is dictated by the signs of $g(e_i, e_i)$ for $i = 2, \dots, n$. Fix such i and notice that for $|t| \ll 1$ we have $e_1 + te_i \in C$, thus $g(e_1 + te_i, e_1 + te_i) > 0$, thus the function

$$f(t) := \|e_1 + te_i\| = \sqrt{1 + t^2 g(e_i, e_i)},$$

is well defined in a neighbourhood of 0. Then f is smooth near $t = 0$ with $f''(0) = g(e_i, e_i)$. Now, the triangle inequality for $\|\cdot\|$ implies convexity of f , hence $g(e_i, e_i) = 1$. Likewise, the reverse triangle inequality implies concavity of f , so $g(e_i, e_i) = -1$. This finishes the proof. \square

Acknowledgments

The authors are grateful to Simone Vincini for comments on an early version of this work.

The authors are grateful for the support of the Fields Institute for Research in Mathematical Sciences 2022 Thematic Program on Nonsmooth Riemannian and Lorentzian Geometry, where this collaboration began. This research was also funded in part by the Austrian Science Fund (FWF) [Grants DOI [10.55776/PAT1996423](https://doi.org/10.55776/PAT1996423), [10.55776/P33594](https://doi.org/10.55776/P33594), [10.55776/STA32](https://doi.org/10.55776/STA32), [10.55776/EFP6](https://doi.org/10.55776/EFP6) and [10.55776/J4913](https://doi.org/10.55776/J4913)]. For open access purposes, the authors have applied a CC BY public copyright license to any author accepted manuscript version arising from this submission.

MB was supported in part by the Fields Institute for the Mathematical Sciences. He gratefully acknowledges financial support by the EPFL through a Bernoulli Instructorship. A large part of

this work was carried out while he held a Postdoctoral Fellowship at the University of Toronto. NG was supported in part by the MUR PRIN-2022JJ8KER grant “Contemporary perspectives on geometry and gravity” (code 2022JJ8KER – CUP G53D23001810006) and the Fields-McMaster Dean’s Distinguished Visitorship. RM was supported in part by the Canada Research Chairs program CRC-2020-00289, Natural Sciences and Engineering Research Council of Canada Discovery Grant RGPIN- 2020–04162, and a grant from the Simons Foundation (923125, McCann). AO was supported in part by the ÖAW scholarship of the Austrian Academy of Sciences. CS was also supported by the European Research Council (ERC), under the European’s Union Horizon 2020 research and innovation programme, via the ERC Starting Grant “CURVATURE”, grant agreement No. 802689.

References

- [1] L. Aké Hau, A. J. Cabrera Pacheco, and D. A. Solis. On the causal hierarchy of Lorentzian length spaces. *Classical Quantum Gravity*, 37(21):215013, 22, 2020.
- [2] A. D. Aleksandrov. A theorem on triangles in a metric space and some of its applications. *Trudy Mat. Inst. Steklov.*, 38:5–23, 1951.
- [3] S. B. Alexander and R. L. Bishop. Lorentz and semi-Riemannian spaces with Alexandrov curvature bounds. *Comm. Anal. Geom.*, 16(2):251–282, 2008.
- [4] L. Ambrosio. Metric space valued functions of bounded variation. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 17(3):439–478, 1990.
- [5] L. Ambrosio. Transport equation and Cauchy problem for BV vector fields. *Invent. Math.*, 158(2):227–260, 2004.
- [6] L. Ambrosio and J. Bertrand. DC calculus. *Math. Z.*, 288(3-4):1037–1080, 2018.
- [7] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 2000.
- [8] L. Ambrosio, N. Gigli, A. Mondino, and T. Rajala. Riemannian Ricci curvature lower bounds in metric measure spaces with σ -finite measure. *Trans. Amer. Math. Soc.*, 367(7):4661–4701, 2012. arXiv:1207.4924.
- [9] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second edition, 2008.
- [10] L. Ambrosio, N. Gigli, and G. Savaré. Calculus and heat flow in metric measure spaces and applications to spaces with Ricci bounds from below. *Invent. Math.*, 195(2):289–391, 2014.
- [11] L. Ambrosio, N. Gigli, and G. Savaré. Metric measure spaces with Riemannian Ricci curvature bounded from below. *Duke Math. J.*, 163(7):1405–1490, 2014.
- [12] L. Ambrosio and B. Kirchheim. Currents in metric spaces. *Acta Math.*, 185(1):1–80, 2000.
- [13] L. Andersson, G. J. Galloway, and R. Howard. The cosmological time function. *Classical Quantum Gravity*, 15(2):309–322, 1998.
- [14] L. Andersson and R. Howard. Comparison and rigidity theorems in semi-Riemannian geometry. *Comm. Anal. Geom.*, 6(4):819–877, 1998.
- [15] D. Bakry and M. Émery. Diffusions hypercontractives. In *Séminaire de probabilités, XIX, 1983/84*, volume 1123 of *Lecture Notes in Math.*, pages 177–206. Springer, Berlin, 1985.

- [16] T. Beran, J. Harvey, L. Napper, and F. Rott. A Toponogov globalisation result for Lorentzian length spaces. *arXiv:2309.12733*, 2023.
- [17] T. Beran, A. Ohanyan, F. Rott, and D. A. Solis. The splitting theorem for globally hyperbolic Lorentzian length spaces with non-negative timelike curvature. *Lett. Math. Phys.*, 113(2):Paper No. 48, 47, 2023.
- [18] V. I. Bogachev. *Measure theory. Vol. I*. Springer-Verlag, Berlin, 2007.
- [19] V. I. Bogachev. *Measure theory. Vol. II*. Springer-Verlag, Berlin, 2007.
- [20] M. Braun. Good geodesics satisfying the timelike curvature-dimension condition. *Nonlinear Anal.*, 229:Paper No. 113205, 30, 2023.
- [21] M. Braun. Rényi’s entropy on Lorentzian spaces. Timelike curvature-dimension conditions. *J. Math. Pures Appl. (9)*, 177:46–128, 2023.
- [22] M. Braun. Nonsmooth d’Alembertian for Lorentz distance functions. *arXiv:2408.16525*, 2024.
- [23] M. Braun, N. Gigli, R. J. McCann, A. Ohanyan, and C. Sämann. An elliptic proof of the splitting theorems from Lorentzian geometry. *arXiv:2410.12632*.
- [24] M. Braun, N. Gigli, R. J. McCann, A. Ohanyan, and C. Sämann. Lorentzian splitting theorems for metrics and weights of regularity C^1 . In preparation.
- [25] M. Braun and R. J. McCann. Causal convergence conditions through variable timelike Ricci curvature bounds. *arXiv:2312.17158*, 2023.
- [26] M. Braun and S. Ohta. Optimal transport and timelike lower Ricci curvature bounds on Finsler spacetimes. *Trans. Amer. Math. Soc.*, 377(5):3529–3576, 2024.
- [27] Y. Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Comm. Pure Appl. Math.*, 44(4):375–417, 1991.
- [28] A. Burtscher and L. García-Heveling. Time functions on Lorentzian length spaces. *Annales Henri Poincaré*, 2024. to appear.
- [29] H. Busemann. Timelike spaces. *Dissertationes Math. (Rozprawy Mat.)*, 53:52, 1967.
- [30] A. Bykov, E. Minguzzi, and S. Suhr. Lorentzian metric spaces and GH-convergence: the unbounded case. *arXiv:2412.04311*, 2024.
- [31] E. Calabi. An extension of E. Hopf’s maximum principle with an application to Riemannian geometry. *Duke Math. J.*, 25:45–56, 1958.
- [32] M. Calisti, A. Ohanyan, and M. Sálamo Candal. Optimal transport for general Lorentz costs. *in preparation*.
- [33] E. Caponio, A. Ohanyan, and S. Ohta. Splitting theorems for weighted Finsler spacetimes via the p -d’Alembertian: beyond the Berwald case. *arXiv:2412.20783*, 2024.
- [34] F. Cavalletti, D. Manini, and A. Mondino. Optimal transport on null hypersurfaces and the null energy condition. *arxiv:2408.08986*, 2024.
- [35] F. Cavalletti and E. Milman. The globalization theorem for the curvature-dimension condition. *Invent. Math.*, 226(1):1–137, 2021.
- [36] F. Cavalletti and A. Mondino. Optimal maps in essentially non-branching spaces. *Commun. Contemp. Math.*, 19(6):1750007, 27, 2017.

- [37] F. Cavalletti and A. Mondino. New formulas for the Laplacian of distance functions and applications. *Anal. PDE*, 13(7):2091–2147, 2020.
- [38] F. Cavalletti and A. Mondino. A review of Lorentzian synthetic theory of timelike Ricci curvature bounds. *Gen. Relativity Gravitation*, 54(11):Paper No. 137, 39, 2022.
- [39] F. Cavalletti and A. Mondino. Optimal transport in Lorentzian synthetic spaces, synthetic timelike Ricci curvature lower bounds and applications. *Camb. J. Math.*, 12(2):417–534, 2024.
- [40] F. Cavalletti and A. Mondino. A sharp isoperimetric-type inequality for Lorentzian spaces satisfying timelike Ricci lower bounds. *arXiv:2401.03949*, 2024.
- [41] F. Cavalletti and K.-T. Sturm. Local curvature-dimension condition implies measure-contraction property. *J. Funct. Anal.*, 262(12):5110–5127, 2012.
- [42] Y. Chabrillac and J.-P. Crouzeix. Continuity and differentiability properties of monotone real functions of several real variables. *Math. Programming Stud.*, 30(30):1–16, 07 1987. Nonlinear analysis and optimization (Louvain-la-Neuve, 1983).
- [43] J. Cheeger. Differentiability of Lipschitz functions on metric measure spaces. *Geom. Funct. Anal.*, 9(3):428–517, 1999.
- [44] J. Cheeger and T. H. Colding. Lower bounds on Ricci curvature and the almost rigidity of warped products. *Ann. of Math. (2)*, 144(1):189–237, 1996.
- [45] J. Cheeger and T. H. Colding. On the structure of spaces with Ricci curvature bounded below. I. *J. Differential Geom.*, 46(3):406–480, 1997.
- [46] J. Cheeger and T. H. Colding. On the structure of spaces with Ricci curvature bounded below. II. *J. Differential Geom.*, 54(1):13–35, 2000.
- [47] J. Cheeger and T. H. Colding. On the structure of spaces with Ricci curvature bounded below. III. *J. Differential Geom.*, 54(1):37–74, 2000.
- [48] D. Cordero-Erausquin, R. J. McCann, and M. Schmuckenschläger. A Riemannian interpolation inequality à la Borell, Brascamp and Lieb. *Invent. Math.*, 146(2):219–257, 2001.
- [49] E. De Giorgi. New problems on minimizing movements. In *Boundary value problems for partial differential equations and applications*, volume 29 of *RMA Res. Notes Appl. Math.*, pages 81–98. Masson, Paris, 1993.
- [50] Q. Deng. *Hölder Continuity of Tangent Cones and Non-Branching in $RCD(K,N)$ Spaces*. ProQuest LLC, Ann Arbor, MI, 2021. Thesis (Ph.D.)—University of Toronto (Canada).
- [51] R. J. DiPerna and P.-L. Lions. Ordinary differential equations, transport theory and Sobolev spaces. *Invent. Math.*, 98(3):511–547, 1989.
- [52] M. Eckstein and T. Miller. Causality for nonlocal phenomena. *Ann. Henri Poincaré*, 18(9):3049–3096, 2017.
- [53] M. Erbar, K. Kuwada, and K.-T. Sturm. On the equivalence of the entropic curvature-dimension condition and Bochner’s inequality on metric measure spaces. *Invent. Math.*, 201(3):993–1071, 2015.
- [54] J.-H. Eschenburg. The splitting theorem for space-times with strong energy condition. *J. Differential Geom.*, 27(3):477–491, 1988.
- [55] G. Folland. *Real Analysis: Modern Techniques and Their Applications*. Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts. Wiley, 1999.

- [56] K. Fukaya. Collapsing of Riemannian manifolds and eigenvalues of Laplace operator. *Invent. Math.*, 87(3):517–547, 1987.
- [57] G. J. Galloway. The Lorentzian splitting theorem without the completeness assumption. *J. Differential Geom.*, 29(2):373–387, 1989.
- [58] N. Gigli. Hyperbolic Banach spaces. In preparation.
- [59] N. Gigli. Optimal maps in non branching spaces with Ricci curvature bounded from below. *Geom. Funct. Anal.*, 22(4):990–999, 2012.
- [60] N. Gigli. On the differential structure of metric measure spaces and applications. *Mem. Amer. Math. Soc.*, 236(1113):vi+91, 2015.
- [61] N. Gigli. Nonsmooth differential geometry—an approach tailored for spaces with Ricci curvature bounded from below. *Mem. Amer. Math. Soc.*, 251(1196):v+161, 2018.
- [62] N. Gigli. De Giorgi and Gromov working together. *arXiv:2306.14604*, 2023.
- [63] N. Gigli. The splitting theorem in non-smooth context, *Mem. Amer. Math. Soc.*, to appear.
- [64] N. Gigli and E. Pasqualetto. *Lectures on nonsmooth differential geometry*, volume 2 of *SISSA Springer Series*. Springer, Cham, [2020] ©2020.
- [65] N. Gigli, T. Rajala, and K.-T. Sturm. Optimal maps and exponentiation on finite-dimensional spaces with Ricci curvature bounded from below. *J. Geom. Anal.*, 26(4):2914–2929, 2016.
- [66] M. Gromov. *Metric structures for Riemannian and non-Riemannian spaces*, volume 152 of *Progress in Mathematics*. Birkhäuser Boston, Inc., Boston, MA, 1999. Based on the 1981 French original [MR0682063 (85e:53051)], With appendices by M. Katz, P. Pansu and S. Semmes, Translated from the French by Sean Michael Bates.
- [67] S. W. Hawking. The occurrence of singularities in cosmology. I. *Proc. Roy. Soc. London Ser. A*, 294:511–521, 1966.
- [68] S. W. Hawking, A. R. King, and P. J. McCarthy. A new topology for curved space-time which incorporates the causal, differential, and conformal structures. *J. Math. Phys.*, 17(2):174–181, 1976.
- [69] A. S. Kechris. *Classical descriptive set theory*, volume 156 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1995.
- [70] C. Ketterer. Cones over metric measure spaces and the maximal diameter theorem. *J. Math. Pures Appl. (9)*, 103(5):1228–1275, 2015.
- [71] C. Ketterer. Characterization of the null energy condition via displacement convexity of entropy. *J. Lond. Math. Soc. (2)*, 109(1):Paper No. e12846, 24, 2024.
- [72] N. J. Korevaar and R. M. Schoen. Sobolev spaces and harmonic maps for metric space targets. *Comm. Anal. Geom.*, 1(3-4):561–659, 1993.
- [73] P. Koskela and P. MacManus. Quasiconformal mappings and Sobolev spaces. *Studia Math.*, 131(1):1–17, 1998.
- [74] E. H. Kronheimer and R. Penrose. On the structure of causal spaces. *Proc. Cambridge Philos. Soc.*, 63:481–501, 1967.
- [75] M. Kunzinger and C. Sämann. Lorentzian length spaces. *Ann. Global Anal. Geom.*, 54(3):399–447, 2018.

- [76] M. Kunzinger and R. Steinbauer. Null distance and convergence of Lorentzian length spaces. *Ann. Henri Poincaré*, 23(12):4319–4342, 2022.
- [77] S. Lisini. Characterization of absolutely continuous curves in Wasserstein spaces. *Calc. Var. Partial Differential Equations*, 28(1):85–120, 2007.
- [78] J. Lott and C. Villani. Ricci curvature for metric-measure spaces via optimal transport. *Ann. of Math. (2)*, 169(3):903–991, 2009.
- [79] Y. Lu, E. Minguzzi, and S.-i. Ohta. Geometry of weighted Lorentz–Finsler manifolds ii: A splitting theorem. *International Journal of Mathematics*, 34(01):2350002, 2023.
- [80] D. B. Malament. The class of continuous timelike curves determines the topology of spacetime. *J. Math. Phys.*, 18(7):1399–1404, 1977.
- [81] R. J. McCann. Existence and uniqueness of monotone measure-preserving maps. *Duke Math. J.*, 80(2):309–323, 1995.
- [82] R. J. McCann. A convexity principle for interacting gases. *Adv. Math.*, 128(1):153–179, 1997.
- [83] R. J. McCann. Polar factorization of maps on Riemannian manifolds. *Geom. Funct. Anal.*, 11(3):589–608, 2001.
- [84] R. J. McCann. Displacement convexity of Boltzmann’s entropy characterizes the strong energy condition from general relativity. *Camb. J. Math.*, 8(3):609–681, 2020.
- [85] R. J. McCann. A synthetic null energy condition. *Comm. Math. Phys.*, 405(2):Paper No. 38, 24, 2024.
- [86] E. Minguzzi. Limit curve theorems in Lorentzian geometry. *J. Math. Phys.*, 49(9):092501, 18, 2008.
- [87] E. Minguzzi. Convexity and quasi-uniformizability of closed preordered spaces. *Topology Appl.*, 160(8):965–978, 2013.
- [88] E. Minguzzi. Causality theory for closed cone structures with applications. *Rev. Math. Phys.*, 31(5):1930001, 139, 2019.
- [89] E. Minguzzi. Lorentzian causality theory. *Living reviews in relativity*, 22(1):3, 2019.
- [90] E. Minguzzi. Further observations on the definition of global hyperbolicity under low regularity. *Classical Quantum Gravity*, 40(18):Paper No. 185001, 9, 2023.
- [91] E. Minguzzi and S. Suhr. Lorentzian metric spaces and their Gromov–Hausdorff convergence. *Lett. Math. Phys.*, 114(3):Paper No. 73, 2024.
- [92] A. Mondino and A. Naber. Structure theory of metric measure spaces with lower Ricci curvature bounds. *J. Eur. Math. Soc. (JEMS)*, 21(6):1809–1854, 2019.
- [93] A. Mondino and S. Suhr. An optimal transport formulation of the Einstein equations of general relativity. *J. Eur. Math. Soc. (JEMS)*, 25(3):933–994, 2023.
- [94] O. Müller. Gromov-Hausdorff metrics and dimensions of Lorentzian length spaces. *arXiv:2209.12736*, 2022.
- [95] O. Müller. Maximality and Cauchy developments of Lorentzian length spaces. *arXiv:2404.06428*, 2024.
- [96] R. P. A. C. Newman. A proof of the splitting conjecture of S.-T. Yau. *J. Differential Geom.*, 31(1):163–184, 1990.

- [97] J. Noldus. A Lorentzian Gromov-Hausdorff notion of distance. *Classical Quantum Gravity*, 21(4):839–850, 2004.
- [98] S. Ohta. On the measure contraction property of metric measure spaces. *Comment. Math. Helv.*, 82(4):805–828, 2007.
- [99] S. Ohta. Finsler interpolation inequalities. *Calc. Var. Partial Differential Equations*, 36(2):211–249, 2009.
- [100] S. Ohta. On the curvature and heat flow on Hamiltonian systems. *Anal. Geom. Metr. Spaces*, 2(1):81–114, 2014.
- [101] S. Ohta and K.-T. Sturm. Heat flow on Finsler manifolds. *Comm. Pure Appl. Math.*, 62(10):1386–1433, 2009.
- [102] F. Otto and C. Villani. Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *J. Funct. Anal.*, 173(2):361–400, 2000.
- [103] R. Penrose. Gravitational collapse and space-time singularities. *Phys. Rev. Lett.*, 14:57–59, 1965.
- [104] T. Rajala. Interpolated measures with bounded density in metric spaces satisfying the curvature-dimension conditions of Sturm. *J. Funct. Anal.*, 263(4):896–924, 2012.
- [105] T. Rajala. Improved geodesics for the reduced curvature-dimension condition in branching metric spaces. *Discrete Contin. Dyn. Syst.*, 33(7):3043–3056, 2013.
- [106] T. Rajala and K.-T. Sturm. Non-branching geodesics and optimal maps in strong $CD(K, \infty)$ -spaces. *Calc. Var. Partial Differential Equations*, 50(3-4):831–846, 2014.
- [107] F. Santambrogio. *Optimal transport for applied mathematicians*, volume 87 of *Progress in Nonlinear Differential Equations and their Applications*. Birkhäuser/Springer, Cham, 2015. Calculus of variations, PDEs, and modeling.
- [108] J.-L. Sauvageot. Tangent bimodule and locality for dissipative operators on C^* -algebras. In *Quantum probability and applications, IV (Rome, 1987)*, volume 1396 of *Lecture Notes in Math.*, pages 322–338. Springer, Berlin, 1989.
- [109] J.-L. Sauvageot. Quantum Dirichlet forms, differential calculus and semigroups. In *Quantum probability and applications, V (Heidelberg, 1988)*, volume 1442 of *Lecture Notes in Math.*, pages 334–346. Springer, Berlin, 1990.
- [110] N. Shanmugalingam. Newtonian spaces: an extension of Sobolev spaces to metric measure spaces. *Rev. Mat. Iberoamericana*, 16(2):243–279, 2000.
- [111] R. D. Sorkin and E. Woolgar. A causal order for spacetimes with C^0 Lorentzian metrics: proof of compactness of the space of causal curves. *Classical Quantum Gravity*, 13(7):1971–1993, 1996.
- [112] C. Sormani and C. Vega. Null distance on a spacetime. *Classical Quantum Gravity*, 33(8):085001, 29, 2016.
- [113] K.-T. Sturm. On the geometry of metric measure spaces. I. *Acta Math.*, 196(1):65–131, 2006.
- [114] K.-T. Sturm. On the geometry of metric measure spaces. II. *Acta Math.*, 196(1):133–177, 2006.
- [115] S. Suhr. Theory of optimal transport for Lorentzian cost functions. *Münster J. Math.*, 11(1):13–47, 2018.

- [116] T. Tao. 254a, lecture 8: The mean ergodic theorem. <https://terrytao.wordpress.com/2008/01/30/254a-lecture-8-the-mean-ergodic-theorem/>, 2008.
- [117] J.-H. Treude. Ricci curvature comparison in Riemannian and Lorentzian geometry. Master's thesis, Universität Freiburg, 2011.
- [118] J.-H. Treude and J. D. E. Grant. Volume comparison for hypersurfaces in Lorentzian manifolds and singularity theorems. *Ann. Global Anal. Geom.*, 43(3):233–251, 2013.
- [119] C. Villani. *Optimal transport*, volume 338 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009. Old and new.
- [120] M.-K. von Renesse and K.-T. Sturm. Transport inequalities, gradient estimates, entropy, and Ricci curvature. *Comm. Pure Appl. Math.*, 58(7):923–940, 2005.
- [121] N. Weaver. Lipschitz algebras and derivations. II. Exterior differentiation. *J. Funct. Anal.*, 178(1):64–112, 2000.
- [122] N. Weaver. *Lipschitz algebras*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, second edition, 2018.