CrossMark

Convex neighborhoods for Lipschitz connections and sprays

E. Minguzzi

Abstract We establish that over a $C^{2,1}$ manifold the exponential map of any Lipschitz connection or spray determines a local Lipeomorphism and that, furthermore, reversible convex normal neighborhoods do exist. To that end we use the method of Picard-Lindelöf approximation to prove the strong differentiability of the exponential map at the origin and hence a version of Gauss' Lemma which does not require the differentiability of the exponential map. Contrary to naive differential degree counting, the distance functions are shown to gain one degree and hence to be $C^{1,1}$. As an application to mathematical relativity, it is argued that the mentioned differentiability conditions can be considered the optimal ones to preserve most results of causality theory. This theory is also shown to be generalizable to the Finsler spacetime case. In particular, we prove that the local Lorentzian(-Finsler) length maximization property of causal geodesics in the class of absolutely continuous causal curves holds already for $C^{1,1}$ spacetime metrics. Finally, we study the local existence of convex functions and show that arbitrarily small globally hyperbolic convex normal neighborhoods do exist.

Keywords Lipschitz connections · Exponential map · Convex neighborhood · Distance function · Low differentiability

Mathematics Subject Classification 53B15 · 26A16 · 53B40 · 83Cxx

E. Minguzzi (✉)
Dipartimento di Matematica e Informatica "U. Dini", Università degli Studi di Firenze,
Via S. Marta 3, 50139 Florence, Italy
e-mail: ettore.minguzzi@unifi.it

Contents

1 Introduction and results

Let M be an n-dimensional paracompact connected $C^{2,1}$ manifold and let $x^\mu : U \to \mathbb{R}^n$ be a local chart where U is an open subset. Every chart induces a chart $(x^\mu, v^\mu) : \pi^{-1}(U) \to \mathbb{R}^n \times \mathbb{R}^n$, $\mu = 0, 1, \ldots, n-1$, on the tangent bundle $\pi : TM \to U$.

For the moment let us consider a second order ODE defined just over U by

$$\frac{\mathrm{d}x^\mu}{\mathrm{d}t} = v^\mu, \tag{1}$$

$$\frac{\mathrm{d}v^\mu}{\mathrm{d}t} = H^\mu(x, v), \tag{2}$$

where H^μ is locally Lipschitz. Under a coordinate change $\tilde{x}^\mu = \tilde{x}^\mu(x^\alpha)$ the system becomes $\dot{\tilde{x}}^\mu = \tilde{v}^\mu$, $\dot{\tilde{v}}^\mu = \tilde{H}^\mu(\tilde{x}, \tilde{v})$, where

$$\tilde{H}^\mu(\tilde{x}, \tilde{v}) = \frac{\partial \tilde{x}^\mu}{\partial x^\alpha \partial x^\beta} \frac{\partial x^\alpha}{\partial \tilde{x}^\gamma} \frac{\partial x^\beta}{\partial \tilde{x}^\delta} \tilde{v}^\gamma \tilde{v}^\delta + \frac{\partial \tilde{x}^\mu}{\partial x^\nu} H^\nu(x(\tilde{x}), v(\tilde{x}, \tilde{v})), \tag{3}$$

and $v^\mu(\tilde{x}, \tilde{v}) = \frac{\partial x^\mu}{\partial \tilde{x}^\nu} \tilde{v}^\nu$. The transformation shows that the following notions are well-defined as independent of the coordinate chart:

1. H^μ is positive homogeneous of second degree on the velocities, that is, for every $s > 0$, $H^\mu(x, sv) = s^2 H^\mu(x, v)$,

2. H^μ is a homogeneous quadratic form in the velocities.

In the former case we say that the second order ODE defines a *spray* over M [2,39] while in the latter more restrictive case we say that it defines a (torsionless) *connection*. In the latter case we set $H^\mu(x, v) = -\Gamma^\mu_{\alpha\beta}(x)v^\alpha v^\beta$ and we recognize in Eq. (3) the transformation rule for the Christoffel symbols $\Gamma^\mu_{\alpha\beta}$. The transformation rule (3) clarifies that M must be $C^{2,1}$ to make sense of the Lipschitz condition on sprays. The non-stationary solutions to Eqs. (1, 2) will be called *geodesics*.

Remark 1 Actually, the above notion of spray is somewhat more general than that introduced in [2,39] since these authors drop the condition $s > 0$ on (a). Our definition, consistent with current usage [4], allows us to include non-reversible pseudo-Finsler manifolds in our analysis of sprays. Clearly, if $H^\mu(x, v)$ defines a spray then $\tilde{H}^\mu(x, v) := H^\mu(x, -v)$ defines also a spray called the *reverse spray*. If $H^\mu(x, v) = H^\mu(x, -v)$ the spray is called *reversible*. If $x(t)$ is a geodesic then $x(-t)$ is a geodesic for the reverse spray but not necessarily for the original spray. Thus the direction of the parametrization of geodesics is important. If we say that two points are connected by a unique geodesic we tacitly assume that the spray under consideration is $H^\mu(x, v)$; that claim does not exclude the possible presence of a connecting geodesic, with different image, for the reverse spray.

From now on we shall consider just locally Lipschitz sprays and we shall clearly speak of connection whenever H^μ is a quadratic form. If H^μ is twice continuously differentiable with respect to the velocities on the zero section of TM, then differentiating twice $H^\mu(x, sv) = s^2 H^\mu(x, v)$ with respect to s and letting $s \to 0$ we obtain $H^\mu(x, v) = \frac{1}{2}(\partial^2 H^\mu/\partial v^\alpha \partial v^\beta(x, 0))v^\alpha v^\beta$, that is the spray is a connection [39]. Whenever the connection comes from a pseudo-Riemannian metric g we shall assume g to be $C^{1,1}_{loc}$. For notational convenience, we shall write just $C^{k,1}$ for $C^{k,1}_{loc}$. It is customary [9] to call *non-regular* the geometrical theory for which the differentiability condition on g is weaker than C^2 (or that on the connection or spray is weaker than C^1). As an example of manifold which has a $C^{1,1}$ metric g, consider a cylinder closed by two semispherical cups, where the submanifold is endowed with the metric induced from the Euclidean space.

1.1 Example: Pseudo-Finsler geometry

Pseudo-Finsler geometry [1,6,7] is a generalization of Finsler geometry [5,42] in which the fundamental tensor g is required to be non-degenerate rather than positive definite. Geodesics in pseudo-Finsler geometry are described by sprays.

Since the definitions of pseudo-Finsler manifold which can be found in the literature impose too strong differentiability conditions we provide a different definition.

Definition 1 A *pseudo-Finsler manifold* (M, g) is a paracompact connected $C^{2,1}$ manifold endowed with a $C^{1,1}$ symmetric tensor

$$g : TM \backslash 0 \to T^*M \otimes_M T^*M, \quad (x, v) \mapsto g_{(x,v)},$$

defined on the non vanishing vectors, which is non-singular and satisfies (in one and hence every chart induced from a chart on M)

$$\frac{\partial g_{(x,v)\,\mu\nu}}{\partial v^\alpha}\, v^\nu = 0, \qquad \frac{\partial g_{(x,v)\,\mu\nu}}{\partial v^\alpha}\, v^\alpha = 0. \tag{4}$$

It is called reversible if $g_{(x,v)} = g_{(x,-v)}$.

Sometimes we shall write g or g_v for $g_{(x,v)}$ either in order to shorten the notation or because we regard v as an element of $TM \backslash 0$. In the pseudo-Riemannian case g is independent of v and written without the v index.

Remark 2 The latter equation in display is equivalent to the homogeneity condition: for every $s > 0$, $g_{(x,sv)} = g_{(x,v)}$. Thus it implies that $L : TM \to \mathbb{R}$ defined by

$$L(x, v) := \frac{1}{2}\, g_{(x,v)}(v, v), \text{ for } v \neq 0, \tag{5}$$

$$L(x, 0) := 0 \tag{6}$$

is positively homogeneous of second degree, namely for every $s > 0$, and $v \neq 0$, $L(x, sv) = s^2 L(x, v)$. The former equation implies

$$\frac{\partial L}{\partial v^\mu}(x, v) = g_{(x,v)\,\mu\nu} v^\nu, \tag{7}$$

$$\frac{\partial^2 L}{\partial v^\mu \partial v^\nu} = g_{(x,v)\,\mu\nu}. \tag{8}$$

We could have defined the pseudo-Finsler manifold as a pair (M, L) in which L is positive homogeneous of second degree, and where g is defined through Eq. (8). This is the definition adopted by most authors. Indeed, differentiating twice with respect to s, $L(x, sv) = s^2 L(x, v)$, and setting $s = 1$ gives Eq. (5). Equation (7) is obtained from Eq. (8) observing that $\frac{\partial L}{\partial v^\mu}$ is positively homogeneous of first degree.

Our definition dispenses from additional differentiability conditions that would have to be imposed on L in order to define g. Furthermore, it has the advantage of making clear the connection with pseudo-Riemannian geometry. Also it clarifies that not every tensor $g_{(x,v)}$, positive homogeneous of zero degree in v, is a pseudo-Finsler metric as the first equation in (4) has to be satisfied.

A geodesic is a stationary point of the functional (with a prime we denote differentiation, typically with respect to a parameter s, if the parameter is t we often use a dot)

$$S[x] = \int_{s_0}^{s_1} L(x, x')\, ds, \qquad x : [s_0, s_1] \to M, \ x(s_0) = x_0, \ x(s_1) = x_1,$$

where $x \in C^1([s_0, s_1])$. By the same argument used above for H^μ we cannot demand g to exist and to be continuous on the zero section unless L is quadratic in the velocities, which corresponds to the case of pseudo-Riemannian geometry. In our terminology

the Finsler (Riemannian) structures are special cases of the pseudo-Finsler (pseudo-Riemannian) ones. In the former case $\sqrt{2L}$ is often denoted F.

We observe that it is always possible to introduce an auxiliary Riemannian metric h over M and to consider the unit sphere subbundle of TM. If x is kept fixed then $g_{(x,v)\,\mu\nu}$ depends only on the direction and orientation of v namely on $\hat{v}$, and since the unit sphere bundle is compact over compact subsets, $g_{(x,v)\,\mu\nu}$ is bounded in relatively compact neighborhoods of points of TM belonging to the zero section. The same observation holds for the partial derivatives with respect to x of g, and hence combinations such as $g_{v\,\alpha\beta}v^\delta$ or $\frac{\partial g_{v\,v\beta}}{\partial x^\alpha}v^\delta$ are locally Lipschitz also at the zero section once they are defined to vanish there. In particular, Eqs. (5) and (7) make sense also for $v = 0$, and for fixed x, L is $C^{1,1}$ on the zero section and $C^{3,1}$ outside it.

The Lagrangian L is constant over the geodesics because, using the Euler–Lagrange equations (we cannot invoke the Hamiltonian to obtain this result since we have not proved the convexity of L in the velocities)

$$\frac{\mathrm{d}L}{\mathrm{d}t} = \frac{\partial L}{\partial x^\mu}v^\mu + \frac{\partial L}{\partial v^\mu}\frac{\mathrm{d}v^\mu}{\mathrm{d}t} = \left(\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial v^\mu}\right)v^\mu + \frac{\partial L}{\partial v^\mu}\frac{\mathrm{d}v^\mu}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial v^\mu}v^\mu\right) = 2\frac{\mathrm{d}L}{\mathrm{d}t}$$

where in the last step we used the homogeneity of L. The spray reads

$$H^\mu(x,v) = -\frac{1}{2}g_v^{\mu\nu}\left(\frac{\partial g_{v\,v\alpha}}{\partial x^\beta} + \frac{\partial g_{v\,v\beta}}{\partial x^\alpha} - \frac{\partial g_{v\,\alpha\beta}}{\partial x^\nu}\right)v^\alpha v^\beta,$$

where it is understood that $H^\mu(x,0) = 0$. It is Lipschitz as required because $g_{(x,v)}^{-1} = g_{(x,\hat{v})}^{-1}$ depends continuously on the unit sphere bundle which is compact over compact subsets of M, thus the inverse $g_{(x,v)}^{\mu\nu}$ stays bounded in relatively compact neighborhoods of points belonging to the zero section, and combinations of the form $g_{(x,v)}^{\mu\nu}v^\beta$ are locally Lipschitz everywhere once they are defined to vanish on the zero section.

We shall return to the geometry of pseudo-Finsler spaces when we shall discuss Gauss' Lemma. Any mention to the various connections that can be introduced in this theory will be avoided in both results and proofs.

1.2 The exponential map for sprays

As we mentioned, the non-stationary solutions to Eqs. (1, 2) will be called *geodesics*. As this is a system of first order ODE over TM, according to the Picard-Lindelöf theorem, the existence and uniqueness of its solutions are guaranteed by the locally Lipschitz condition on H^μ.

Let $\gamma_v(t)$ be the unique geodesic which starts from $\pi(v)$ with velocity v. The set Ω is given by those v for which the geodesic exists at least for $t \in [0, 1]$. The exponential map $\exp\colon \Omega \to M \times M$ is given by

$$v \mapsto (\pi(v), \gamma_v(1)),$$

while the pointed exponential map at $p \in M$, is $\exp_p \colon \Omega_p \to M, \Omega_p = \Omega \cap \pi^{-1}(p)$, $\exp_p v := \gamma_v(1) = \pi_2(\exp v)$. By the homogeneity of H^μ on velocities we have

$$\gamma_{sv}(t) = \gamma_v(st), \tag{9}$$

thus the set Ω (and Ω_p) is star-shaped in the sense that if $v \in \Omega$ then $sv \in \Omega$ for every $s \in [0, 1]$. Equation (9) clarifies that it make sense to call *affine* the geodesic parameter, for any affine reparametrization of a geodesic gives a curve which solves the geodesic equation.

Remark 3 The exponential map of the reverse spray, denoted $e\tilde{x}p$ is

$$e\tilde{x}p\, v := (\pi(v), \gamma_v(-1)), \qquad e\tilde{x}p_p\, v := \gamma_v(-1),$$

and since in general $\gamma_v(-1) \neq \gamma_{-v}(1)$ this map cannot be simply expressed through the exponential map exp. Of course, if the spray is reversible it coincides with $v \mapsto \exp(-v)$.

Hartman [29] proved that for connections the uniqueness of the geodesic equation is lost if the Lipschitz condition is weakened to continuity [29]. This result was improved by Hartman and Wintner [33] [31, Exercise 6.2, Chap. 5] who considered the metric

$$ds^2 = \left(1 + |v|^{1+\alpha}\right) \left(du^2 + dv^2\right)$$

for $0 < \alpha < 1$. Its connection satisfies an Hölder condition of exponent α, and on any neighborhood of $p = (0, 0)$ one can find infinite geodesics which start from p with velocity $(1, 0)$.

These examples suggest the Lipschitz condition as the best differentiability condition that can be placed on a spray.

Remark 4 Actually, if the connection is that of a Riemannian C^2 surface of Euclidean 3-space then uniqueness of geodesics is guaranteed, and one can even build C^1 normal coordinates though the connection is just continuous [29]. Moreover, still in the 2-dimensional case under Lipschitzness of the connection one can prove results which are stronger than those considered in this work[1] [29,30,33]. Nevertheless, we shall work in the general n-dimensional case since the 2-dimensional one appears too special and less relevant for applications (it suffices to recall that in 2-dimensions any metric is locally conformally flat).

In this work we shall prove that the exponential map of every spray is a local Lipeomorphism (local bi-Lipschitz homeomorphism) and that on M any point admits arbitrarily small convex neighborhoods (Theor. 4).

In a Riemannian framework, this result can be improved in some directions. For instance, it is well known that Riemannian spaces with sectional curvature bounded from below or above find a remarkable generalization in the notion of Alexandrov

---

[1] Please notice that according to [32] claim III in [33] is incorrect.

space. In this quite general setting there are indeed results on the existence of convex neighborhoods [9, Prop. 5.5] [54,55].

We were finishing this work when we learned that Kunzinger, Steinbauer and Stojkovíc, in a recent preprint [36], have also provided a proof of the bi-Lipschitzness of the pointed exponential map and of the existence of convex neighborhoods. As us they were motivated by a recent work by Chrusciel and Grant on causality theory under low differentiability conditions [16]. Their approach is complementary to our own and deserves some comments. They consider a net of smooth Riemannian metrics g_ϵ obtained from g through convolution with a mollifier, and use methods from comparison geometry to obtain sufficiently strong estimates on the exponential maps of the regularized metrics, so as to be able to carry over the bi-Lipschitz property through the limit. In order to perform this last step in the general pseudo-Riemannian case they use some results on comparison geometry for indefinite metrics recently obtained by Chen and LeFloch [14]. They also show that the Riemannian case can be dealt with using the Rauch comparison theorem.

Our approach has several advantages among which that of being tailored to the results on convexity that we wish to prove. The differences between our strategy and more classical approaches based on the smooth category and the inverse function theorem are minimal; there is no use of comparison geometry, nor regularization is required. No prior knowledge of Riemannian geometry is actually needed, for we never use the concept of (sectional) curvature or Jacobi field. Our results are therefore obtained improving some local analytical results, without introducing advanced topics in differential geometry or touching the very foundations of the theory under consideration. This is desirable since we are actually obtaining basic results on local convexity which could be placed at the very beginning of treatments on differential geometry under low regularity. Our study may be useful, for instance, to understand the limits of pseudo-Finsler geometry and particularly Lorentzian geometry, for which a theory of the same generality of Alexandrov's is missing.[2]

We have also tried to be as a complete as possible. In this way the reader will be able to refer to the results of this work without the need of making involute arguments in the attempt of extending the results herein obtained. For instance, we prove that the *non-pointed* exponential map exp is a Lipeomorphism from a neighborhood of the zero section to a neighborhood of the diagonal on $M \times M$. This result is quite useful in applications, for instance in causality theory it is used in the proof that the causal relation over convex normal sets is closed (Theor. 12).

Unless otherwise specified, $\| \|$ will denote the Euclidean norm on $\mathbb{R}^n$. Let us recall that a function $f : O \to \mathbb{R}^k$ defined on an open set $O \subset \mathbb{R}^n$ is Lipschitz if for every $p, q \in O$,

$$\|g(p) - g(q)\| < K\|p - q\|$$

---

[2] There are well known difficulties in this generalization. They are related to the fact that sectional curvature bounds imply constant curvature [28,50]. These problems could be sidestepped imposing only bounds on the sectional curvature of timelike planes [28].

for some $K > 0$. It is *locally Lipschitz* if this inequality holds over every compact subset of O, with K dependent on the compact subset. If f depends on another variable z, then f is uniformly Lipschitz if the Lipschitz constant does not depend on z. It is *locally uniformly Lipschitz* if, chosen any compact set on the domain of z, the Lipschitz constant can be chosen to be dependent on just the compact set rather than z. As a consequence, a function of, say, two variables $f(x, y)$ which is Lipschitz is also Lipschitz in one variable uniformly in the other. A *Lipeomorphism* f is a homeomorphism for which both f and f^{-1} are locally Lipschitz. In the cases that will interest us f will be defined in an open subset $O \subset \mathbb{R}^n$, and $f(O)$ will also be an open subset of $\mathbb{R}^n$.

It is well known that the ODE $\dot{x} = f(x)$ for which f is Lipschitz admits unique solutions which have a Lipschitz dependence on the initial conditions [31, Ex. 1.2, Chap. 2] [13, Prop. 1.10.1] [38, Cor. 1.6]. As a result, the exponential map exp and its pointed version $\exp_p$ are locally Lipschitz.

We shall improve this result as in the next theorem. This refinement will be used in the proof that in a Riemannian space the geodesics are locally length minimizing in the family of absolutely continuous curves (and to prove an analogous result in the Lorentzian case). To increase readability we postpone most proofs to the next sections.

Theorem 1 *Let us consider a Lipschitz spray (Lipschitz H^μ) on a $C^{2,1}$ manifold M, and let $\varphi \colon W \to TM$, $W \subset [0, +\infty) \times TM$, $\varphi(t, v) := \gamma_v'(t)$ be the geodesic flow map defined at those $(t, v) \in \mathbb{R} \times TM$ for which the geodesic γ_v extends up to time t (so that the expression on the right-hand side makes sense). Then W is an open subset of $[0, +\infty) \times TM$ such that $[0, 1] \times \Omega \subset W$ and such that for every $s \in [0, 1]$ if $(t, v) \in W$ then $(st, v), (t, sv) \in W$. Analogously, $\Omega \subset TM$ is open and star-shaped in the sense that for every $s \in [0, 1]$ if $v \in \Omega$ then $sv \in \Omega$.*

Moreover, $\varphi(\cdot, v)$ is $C^{1,1}$, φ is locally Lipschitz and there is a star-shaped subset $\tilde{\Omega} \subset \Omega$ such that, $\Omega \backslash \tilde{\Omega}$ has zero Lebesgue measure, and for every $v \in \tilde{\Omega}$ and every $t \in [0, 1]$, $\varphi(t, \cdot)$ is differentiable at (t, v) and the differential (Jacobian) $\partial_2 \varphi(\cdot, v)$ is locally Lipschitz in t, locally uniformly with respect to those v belonging to $\tilde{\Omega}$ (that is the local Lipschitz constant can be chosen so that it does not vary in a small neighborhood of v as long as the independent variable stays in $\tilde{\Omega}$). Finally, for any $v \in \tilde{\Omega}$ we have that for almost every $t \in [0, 1]$ the following mixed differentials exist, are locally bounded and coincide

$$\partial_1 \partial_2 \varphi = \partial_2 \partial_1 \varphi.$$

These conclusions do not change if we restrict φ to some Lipschitz m-dimensional submanifold N of TM. In this case the differential ∂_2 refers to the variables of the Lipschitz chart on N and the almost everywhere existence of $\partial_2 \varphi$ must be understood in the m-dimensional Lebesgue measure of N.

One would like to prove that the (pointed) exponential map is invertible, say a local Lipeomorphism or a local diffeomorphism. Hartman [31, Exercise 6.2, Chap. 5] [32] showed in the connection case that the exponential map is actually C^1 and hence a local diffeomorphism provided the connection admits a continuous exterior derivative. We shall not impose these additional conditions.

The local injectivity of the exponential map for Lipschitz connections was proved by Whitehead in his paper on the existence of convex normal neighborhoods [62]. In [63] he observed that the result could be generalized to sprays. Whitehead uses a theorem by Picard which applies to boundary value problems of second order ODEs [31, Cor. 4.1, Chap. 12]. Using the theorem on the invariance of domain one could infer that the exponential map provides a local homeomorphism, though from here it does not seems easy to obtain that it admits a Lipschitz inverse.

Our approach relies instead on an improved version of the inverse function theorem due to Leach [40] (see also [13]). This theorem depends on Peano's definition of strong derivative [52] and on the natural corresponding notion of strong differential studied by Severi. Unfortunately, Peano's contributions in this direction are, together with many other accomplishments by the Italian mathematician, little known [20]. Nijenhuis' attempt [49] to popularize Peano's choice and Leach's inversion theorem passed essentially unnoticed. Peano's choice provides a better definition of differential, so good, in fact, that having to choose one should probably adopt it in place of the usual differential in analysis textbooks. Indeed, the strong differential leads to stronger and more elegant results, and seems to corresponds better with the intuition. We hope that this study, showing the usefulness of Peano's strong derivative for the exponential map will serve to motivate its mention in University courses.

Let us recall Peano's definition of strong differential and its basic properties. We give a general definition for Banach spaces although we shall work on $\mathbb{R}^k$ for some $k \geq 1$. We denote with $B(p, r) := \{q : |q - p| < r\}$ the open ball of radius r centered at p, and with $\bar{B}(p, r) := \{q : |q - p| \leq r\}$ the closed ball.

Definition 2 Let E and F be Banach spaces, and let $f : O \to F$, be a function defined on an open set $O \subset E$. The strong differential of f at $p \in O$ is a bounded linear transformation $L : E \to F$ which approximates changes of f in the sense that for every $\epsilon > 0$, there is a $\delta > 0$ such that if $|q_1 - p| \leq \delta$ and $|q_2 - p| \leq \delta$, then:

$$|f(q_1) - f(q_2) - L(q_1 - q_2)| \leq \epsilon |q_1 - q_2|. \tag{10}$$

Clearly, if the strong differential at p exists then it is unique. If f is strongly differentiable at p then taking $q_2 = p$ shows that it is also Fréchet differentiable and that the differentials so defined coincide. In the finite dimensional case which will interest us all norms are equivalent thus the notions of strong differentiation and strong differential are independent of the norm used.

We list some properties of strong differentiation which are easy to prove [21,40, 41,49,52].

  (i) If f is strongly differentiable at p then it satisfies a Lipschitz condition in a neighborhood of p.[3]
 (ii) If f is differentiable in a neighborhood of p and the differential is continuous at p then it is strongly differentiable at p.

---

[3] One could ask whether every Lipschitz function is strongly differentiable almost everywhere. The answer is negative already for functions defined on the real line [27, Sect. 14.4.1].

(iii) If f is strongly differentiable over a subset $A \subset E$ then the strong differential is continuous over A with respect to the induced topology.
(iv) If a continuous function $f : U_1 \times U_2 \to V$ on a product Banach space admits strong partial differentials at p (obtained keeping constant the other variable) then it is strongly differentiable at p and the total differential has the usual gradient expression in terms of partial differentials.
(v) The composition of strongly differentiable functions is strongly differentiable.
(vi) The mixed partial strong derivatives coincide wherever they exist [46].
(vii) If $f : \mathbb{R} \to \mathbb{R}$ has positive strong differential at p, then f is continuous and increasing in a neighborhood of p.

Conditions (ii) and (iii) clarify that a function is C^1 over an open set if and only if it is strongly differentiable over it. Some key theorems in analysis that require a C^1 condition on a neighborhood of p can be proved demanding the weaker condition of strong differentiability at p. An example is Leach's inversion theorem [13,35,40,49] which generalizes Dini's and which we state in a form suitable for our purposes:

Theorem 2 (Leach) *Let $f : O \to \mathbb{R}^n$, be a function defined on an open subset $O \subset \mathbb{R}^n$, such that f has strong differential $L : \mathbb{R}^n \to \mathbb{R}^n$ at $p \in O$. If L is invertible then there are an open neighborhood N_1 of p, an open neighborhood N_2 of $f(p)$, and a function $g : N_2 \to \mathbb{R}^n$ such that, $f(N_1) = N_2$, $g(N_2) = N_1$, $f|_{N_1}$ and g are one the inverse of the other, they are both Lipschitz and g has strong differential L^{-1} at $f(p)$.*

Moreover, in this case f is differentiable at $q \in N_1$ if and only if g is differentiable at $f(q)$, in which case the differentials are invertible. This last statement holds also with differentiable *replaced by* strongly differentiable.[4]

In order to clarify the connection between this inversion theorem and Clarke's [17] it is convenient to recall the notion of Clarke's generalized differential for locally Lipschitz functions:

Definition 3 The generalized Jacobian of a locally Lipschitz function $f : O \to \mathbb{R}^n$, $O \subset \mathbb{R}^k$ at p, denoted $\partial f(p)$, is the convex hull of all matrices M of the form

$$M = \lim_{p_i \to p} df(p_i)$$

where p_i converges to p, f is differentiable at p_i for each i and df denotes the usual Jacobian.

By Rademacher's theorem the generalized differential is non-empty at p and we have (see also [17,21])

Proposition 1 *If $f : O \to \mathbb{R}^n$, $O \subset \mathbb{R}^k$, is strongly differentiable at p then $\partial f(p) = \{df(p)\}$.*

---

[4] The fact that N_1 and N_2 can be chosen to be open sets such that both $f|_{N_1}$ and g are Lipschitz follows from Leach's original formulation plus (i). The statement in the last paragraph is not contained in Leach's original formulation but can be found in its proof.

Proof Indeed, take any $\epsilon > 0$ and let $\delta > 0$ be such that Eq. (10) holds. Let J be the limit of a sequence $J_i = df(p_i)$ for $p_i \to p$. We can assume $\|p_i - p\| < \delta/2$ for each i. Let e be any normalized unit vector. For each i we can find some $0 < \delta_i \leq \delta/2$ such that

$$\|f(q_i) - f(p_i) - df(p_i)(q_i - p_i)\| \leq \epsilon\|q_i - p_i\|$$

for every q_i such that $\|q_i - p_i\| \leq \delta_i$. Let $q_i = p_i + \delta_i e$, then

$$\delta_i\|(L - J_i)e\| = \|(L - J_i)(q_i - p_i)\| \leq \|f(q_i) - f(p_i) - J_i(q_i - p_i)\|$$
$$+ \|f(q_i) - f(p_i) - L(q_i - p_i)\| \leq 2\epsilon\|q_i - p_i\| = 2\epsilon\delta_i.$$

Simplifying δ_i, taking the limit $i \to \infty$ and using the arbitrariness of ϵ and e proves that $J = df(p)$ and hence $\partial f(p) = \{df(p)\}$. $\qquad\square$

Clarke proved that if $k = n$ and ∂f admits only invertible elements then f is a local Lipeomorphism. Leach's version states something more for f strongly differentiable at p, for it establishes that the inverse is strongly differentiable at $f(p)$.

Our strategy is then clear: we are going to prove the strong differentiability of the exponential map in order to deduce the Lipschitzness of the inverse by means of Leach's (or Clarke's) inversion theorem. The proof of the strong differentiability of the exponential map will pass through a local analysis based on the Picard-Lindelöf approximation method.

In the end we shall prove:

Theorem 3 *Let M be a $C^{2,1}$-manifold endowed with a locally Lipschitz spray.*

(exp) *The set Ω is open in the topology of TM. The exponential map* $\exp\colon \Omega \to M \times M$, $\Omega \subset TM$, *is locally Lipschitz.*
Moreover, $\exp$ *is strongly differentiable over the zero section, namely over the image of $p \mapsto 0_p$. The map* $\exp$ *provides a Lipeomorphism between an open star-shaped neighborhood of the zero section and an open neighborhood of the diagonal of $M \times M$.*

(exp$_p$) *For every $p \in M$ the set Ω_p is open in the topology of T_pM. The pointed exponential map* $\exp_p\colon \Omega_p \to M$, $\Omega \subset T_pM$, *is locally Lipschitz and strongly differentiable at the origin. The map* $\exp_p$ *provides a local Liperomorphism from a star-shaped open subset of Ω_p and an open neighborhood of p (for more see Theorems 1 and 4).*

In the pointed case this result can be refined. We shall need some definitions.

Definition 4 An open neighborhood N of $p \in M$ will be called *normal* if there is an open star-shaped subset $N_p \subset \Omega_p$ such that $\exp_p\colon N_p \to N$ is a Lipeomorphism.

Definition 5 An open set $C \subset M$ will be called *convex normal* if it is a normal neighborhood of each of its points. We shall say that $\bar{C}$ is *strictly convex normal* if C is convex normal and any two points of $\bar{C}$ are connected by a unique geodesic contained in C but for the endpoints. A (strictly) convex normal set is caller *reversible* if it is so also for the reverse spray.

Remark 5 In a reversible convex normal subset C for any two points $p, q \in C$ there is a geodesic $\gamma_{pq} : [0, 1] \to C$ connecting p to q and a geodesic $\gamma_{qp} : [0, 1] \to C$ connecting q to p. They coincide if the spray is reversible. Furthermore, there are geodesics for the reverse spray $\tilde{\gamma}_{pq}(t) = \gamma_{qp}(1-t)$ connecting p to q and $\tilde{\gamma}_{qp}(t) = \gamma_{pq}(1-t)$ connecting q to p. The last identities follow from the uniqueness of the connecting geodesic for the spray. Observe that while the convexity of C with respect to the spray implies the convexity of C with respect to the reverse spray, a *reversible convex normal set* has a stronger property which cannot be deduced from the corresponding property for the spray, that is, in a *reversible convex normal set C* we have that $\tilde{\exp}_p^{-1}|_C$ is a *Lipeomorphism* for every $p \in C$.

The following concept will be useful in the next section.

Definition 6 Let C be a convex normal set, let $p, q \in C$ and let $x : [0, 1] \to C$, $x(0) = p$, $x(1) = q$, be the unique geodesic connecting them. The vector $\dot{x}(1)$ is denoted $P(p, q)$ and called *position vector*.

We shall prove:

Theorem 4 *Let M be a $C^{2,1}$-manifold endowed with a locally Lipschitz spray. Let O be an open neighborhood of $p \in M$. Then there is a reversible strictly convex normal neighborhood C of p contained in O, such that* exp *establishes a Lipeomorphism between an open star-shaped subset of TC and $C \times C$. Analogously, $\tilde{\exp}$ establishes a Lipeomorphism between an open star-shaped subset of TC and $C \times C$.*

Moreover, for every chart $\{x^\mu\}$ defined in a neighborhood of p, C can be chosen equal to the open ball $B(p, \delta)$ for any sufficiently small δ (the ball is defined through the Euclidean norm induced by the coordinates).

If the spray is compatible with a pseudo-Finsler structure then this result can be further refined.

Remark 6 The previous theorems can be formulated in more generality for $C^{k,1}$ sprays over $C^{k+2,1}$ manifolds with $k \geq 0$ or for $C^{k+1,\alpha}$ sprays over $C^{k+3,\alpha}$, $\alpha \in [0, 1)$, manifolds. The exponential map and its inverse have the degree of differentiability of the spray.

These cases follow more or less straightforwardly from the Lipschitz $k = 0$ case treated here or from what is already known for C^1 sprays. For a direct proof that generalizes that for the Lipschitz case see Remark 15.

1.3 Gauss' Lemma for pseudo-Finsler sprays

Let a consider again a pseudo-Finsler geometry in which the fundamental tensor g_v is $C^{1,1}$ and the spray is Lipschitz. This means that in the pseudo-Riemannian case g is $C^{1,1}$ and the connection is Lipschitz.

The local length minimization property of geodesics in Riemannian spaces, or the local Lorentzian length maximization property of causal geodesics in Lorentzian manifolds, are proved passing through Gauss' Lemma (in the Lorentzian case see [15,51]). This Lemma is known to hold in Finsler geometry [5]

(Gauss' Lemma under sufficient differentiability conditions) Let $p \in M$, let N be a normal neighborhood of p and let $v \in \exp_p^{-1} N \backslash 0$. Let $w \in T_p M \sim T_v(T_p M)$. Then

$$g_{(d \exp_p)_v v}((d \exp_p)_v v, (d \exp_p)_v w) = g_v(v, w). \tag{11}$$

This lemma is usually expressed as above using the push forward of the exponential map [26, Theor. 3.70] [12]. As the exponential map is C^1 for C^2 metrics, one expects this lemma to be valid for C^2 metrics.

The just mentioned classical proofs of Gauss' lemma work indeed in the C^2 case. Without entering in too many details the reader should just keep in mind that for what concerns differentiability with respect to the initial conditions the exponential map, by Peano's theorem [31, Theor. 3.1], behaves better in the radial direction than in the transverse directions. As a consequence, the mixed derivative of the expression $f(t, s) = \exp_p(t v(s))$ is continuous if $v(s)$ is C^1, cf. [31, Cor. 3.2]. Thus one can use Schwarz theorem and fully justify the proof of [12].

Other proofs seem less convincing [59, Cor. 2.2]. It is very easy to forget that one cannot work directly with, say, derivatives of vector or tensor fields expressed in a normal coordinate chart or the Christoffel symbols in normal coordinates, indeed, since this chart is just C^1, vector and tensor fields on them can be at most C^0 and hence are not differentiable and similarly the Christoffel symbols are not defined [18].

The situation is worse for Lipschitz connection or sprays and $C^{1,1}$ metrics. Under this differentiability hypothesis the exponential map is just Lipschitz, thus Gauss' Lemma is not even expected to hold. However, we shall show that Gauss' Lemma still holds true in a formulation which does not require the differentiability of the exponential map.

We shall need some technical result either on (a) the differentiation under the integral sign, or on (b) Schwarz's theorem on the equality of mixed partial derivatives. We shall discuss them in Sect. 1.8. In the end we shall prove (in order to obtain a more restrictive pseudo-Riemannian version just remove the vector index from the fundamental tensor g):

Theorem 5 *Let (M, g) be a $C^{2,1}$ pseudo-Finsler manifold for which g is $C^{1,1}$ (Sect. 1.1). Let N be a normal neighborhood of $p \in M$. The function $D_p^2 \colon N \to \mathbb{R}$ defined by*

$$D_p^2(q) := 2L(p, exp_p^{-1}(q)) = g_{exp_p^{-1}(q)}(exp_p^{-1}(q), exp_p^{-1}(q)) \tag{12}$$

is $C^{1,1}$ in q and

$$dD_p^2(q) = 2g_{P(p,q)}(P(p, q), \cdot), \tag{13}$$

where $P(p, q) = \gamma'_{\exp_p^{-1} q}(1)$ is the position vector of q with respect to p. Thus the level sets of D_p^2 are orthogonal to the geodesics issued from p, and for $t, s > 0$ the $(-t)$-time flow map of $P(p, \cdot)$ is a Lipeomorphism between $(D_p^2)^{-1}(s)$ and its image on $(D_p^2)^{-1}(s \, e^{-2t})$.

Finally, Eq. (11) holds wherever $\exp_p \colon \exp_p^{-1} N \to N$ is differentiable, hence almost everywhere. Thus the usual Gauss' Lemma holds under C^2 differentiability of the metric g.

Geometrically Eq. (13) states the geodesic connecting p to q is perpendicular to a level set $D_p^2 = cnst$ (in pseudo-Finsler geometry v is perpendicular to w if $g_v(v, w) = 0$).

Observe that D_p^2 can be negative if g is not positive definite. If g is positive definite, as we shall prove in a moment in Theorem 6, $D_p(q)$ coincides with the Finsler distance between p and q in the space $(N, g|_N)$ and hence coincides with the Finsler distance on M provided the ball of radius $D_p(q)$ centered at p is contained in N.

Remark 7 It is somewhat surprising that D_p^2 is $C^{1,1}$. One would expect it to be Lipschitz because from its definition, that is Eq. (12), we see that it is built from the the inverse of the exponential map which is Lipschitz. The additional degree of differentiability comes from the fact that we can check that the differential is almost everywhere as in Eq. (13). This result can then be extended everywhere thanks to the Lipschitzness of D_p^2 and the continuity of $P(p, q)$ (see Theorem 16).

Remark 8 The proof of this result can be easily generalized to show that for (p, q) belonging to a reversible convex normal set, $D_p(q)$ is $C^{1,1}$ in (p, q) and its differential is

$$dD_p^2(q)(v_p, v_q) = 2g_{P(p,q)}(P(p, q), v_q) + 2g_{\tilde{P}(q,p)}(\tilde{P}(q, p), v_p),$$

where $v_p \in T_p M$, $v_q \in T_q M$, and $\tilde{P}$ is the position vector map according to the reverse spray. Furthermore, with Prop. 3 we shall prove that $P(p, q)$ is strongly differentiable on the diagonal of $M \times M$ hence $D_p^2(q)$ has first differential which is strongly differentiable at the origin.

We shall state the next result for pseudo-Finsler manifolds for which the fundamental tensor is either positive definite (Finsler geometry) or of signature $(-, +, \cdots, +)$ (Lorentzian-Finsler geometry). It is necessary to elaborate the last structure in the notion of *Finsler spacetime* which extends the usual notion of spacetime met in mathematical relativity.

Let us start from a Lorentzian-Finsler manifold (M, g), and let us keep in mind that if the fundamental tensor g_v does not depend on the velocity then we are back to a Lorentzian manifold [8]. Non vanishing vectors are called *spacelike*, *lightlike* or *timelike* depending on the sign of $g_v(v, v)$, namely positive, null or negative, and the terminology extends to C^1 curves provided the causality type of the tangent vector is consistent throughout the curve (which is assured for geodesics since in that case $g_v(v, v)$ is constant over the curve). A vector is *null* if it is lightlike or zero, and *non-spacelike* if it is causal or zero. At any $x \in M$ let us denote with $I_x \subset T_x M$ the subset of timelike vectors, and with J_x the subset of non-spacelike vectors and with E_x the subset of null vectors.

Beem and Perlick [6,56] have shown that each component of I_x is convex, and hence, by continuity, that each component of $J_x \backslash \{0\}$ is convex. Since $\partial L(x, v) / \partial v \neq 0$ for $v \neq 0$, the hypersurfaces $g_v(v, v) = cnst$ are imbedded submanifolds and hence each component of $E_x \backslash \{0\}$ plus $\{0\}$ is the boundary of some component of I_x. Analogously, each component of $J_x \backslash \{0\}$ plus $\{0\}$ is the closure of some component of

I_x. Furthermore, again because $\partial L(x, v)/\partial v \neq 0$ for $v \neq 0$, distinct components of $J_x \backslash \{0\}$ do not intersect.

For simplicity we shall restrict our analysis to

Definition 7 *Finsler spacetimes* are pseudo-Finsler manifolds (M, g) for which (a) g has signature $(-, +, \cdots, +)$, (b) for one (and hence every) $x \in M$, I_x has just 2 components, (c) there exists a global continuous timelike vector field which defines a notion of *future cone*.

Condition (c), can be accomplished passing to a double covering while (b) is assured under reversibility if the spacetime has dimension larger than two [47].

Given the time orientation, causal vectors are either future or past, and so the regular C^1 causal curves are either past directed or future directed . The C^1 causal curves which we shall consider will be future directed.

Remark 9 Observe that we do not assume that g is reversible. Thus we have essentially two different distributions of light cones on spacetime and hence two causality theories. In what follows, by mentioning only future directed timelike curves we restrict ourselves to one of these theories.

Two points $x, y \in M$ are said to be chronologically related in a set $S \subset M$, this being denoted $y \in I_S^+(x)$, $(x, y) \in I_S^+$ or $x \ll_S y$, if there is a future directed C^1 timelike curve from x to y contained in S. Two points are said to be causally related, this being denoted $y \in J_S^+(x)$, $(x, y) \in J_S^+$ or $x \leq_S y$, if there is a future directed C^1 causal curve from x to y contained in S or $x = y \in S$. We write $x <_S y$ if $x \leq_S y$ but $x \neq y$. If $S = M$ then we write simply $\ll, \leq$ and $<$.

A curve $\sigma : [a, b] \to M$ will be *absolutely continuous* (an AC-curve for short) if its components in one (and hence every) local chart are locally absolutely continuous. Equivalently, introduced a complete Riemannian metric on M, and denoting with ρ the corresponding distance, σ is absolutely continuous if it satisfies locally the the usual definition of absolute continuity between (topological) metric spaces. Since every pair of Riemannian metrics over a compact set is Lipschitz equivalent, and M is locally compact, this definition does not depend on the metric chosen. Analogously, we can define the concept of Lipschitz curve.

We shall say that an AC-curve $\sigma : [a, b] \to M, t \mapsto \sigma(t)$, is future directed *causal* if $\dot\sigma$ is is future directed causal almost everywhere. We do not need to define a notion of absolutely continuous timelike curve.

The Lorentzian–Finsler length of a causal AC-curve is

$$l[\sigma] = \int_a^b \sqrt{-g_{\dot\sigma}(\dot\sigma, \dot\sigma)} \, dt,$$

and it is finite because the integrand belongs to $L^1([a, b])$ as in coordinates we have $\dot\sigma^\mu \in L^1([a, b])$, $\sqrt{|\dot\sigma^\mu|} \in L^2([a, b])$.

In the Finsler case the concept of Finsler length is defined analogously but with a plus sign inside the square root. The Finsler distance from p to q is the infimum of the Finsler lengths of the C^1 curves connecting p to q. It is symmetric and hence a true

distance whenever the Finsler structure is reversible. We can also define a Lorentzian-Finsler distance between p and q with $p \leq q$ as the supremum of the Lorentzian-Finsler lengths of the causal AC-curves connecting p to q. As in Lorentzian geometry, it satisfies a reverse triangle inequality [8].

Theorem 6 *Let (M, g) be a $C^{2,1}$ pseudo-Finsler manifold for which g is $C^{1,1}$ (Sect. 1.1). Let N be a normal neighborhood of $p \in M$ and suppose that g is*

> *Finsler:*
> *Let $\sigma \colon [0, 1] \to N, s \mapsto \sigma(s)$, be any AC-curve starting from p, then its length is larger than that of the (unique) geodesic connecting its endpoints, unless its image coincides with that of that geodesic. In this last case the Finsler distance from p provides an affine parameter $r(s)$ where the dependence on s is absolutely continuous and increasing.*

> *Lorentzian–Finsler:*
> *Let $\sigma \colon [0, 1] \to N$ be any future directed causal AC-curve starting from p, then $\exp^{-1}(\sigma(s))$ is future directed causal for every $s > 0$, and if $\exp^{-1}(\sigma(\hat{s}))$ is lightlike then $\sigma|_{[0,\hat{s}]}$ coincides with a future directed lightlike geodesic segment up to parametrizations. Finally, the Lorentzian–Finsler length of σ is smaller than that of the (unique) future directed casual geodesic connecting its endpoints, unless its image coincides with that of that geodesic. In this last case the affine parameter of the geodesic is absolutely continuous and increasing with s.*

Remark 10 Physically the Lorentzian–Finsler version proves that a motion which is almost everywhere slower than light is also locally slower than light. This is the main result which allows us to develop causality theory for Lipschitz connections in Finsler spacetimes. In particular, $I_N^+(p)$ (or $J_N^+(p)$) coincides with the exponential map-image in N of the future directed timelike (resp. causal) cone at p.

Remark 11 It is natural to ask whether locally a spacelike geodesic segment minimizes the functional $\int \sqrt{g_{\dot{\sigma}}(\dot{\sigma}, \dot{\sigma})} \, dt$ over the C^1 spacelike curves connecting the same endpoints. The answer is negative already in 1+1 Minkowski spacetime, just consider almost lightlike zig-zag curves which approximate the geodesic. Their presence shows that the infimum of the functional vanishes.

1.4 Some applications to mathematical relativity

We recall that according to Hawking and Ellis [34] a future directed continuous causal curve $x : [a, b] \to M$, is a continuous curve such that for every open convex normal set C intersecting x, whenever $x([t_1, t_2]) \subset C$, $t_1 < t_2$, the points $x(t_1)$ and $x(t_2)$ are connected by a future directed causal geodesic contained in C. This definition can be imported word by word to the realm of Finsler spacetimes.

An interesting consequence of Theorem 6 is Theorem 7 which will provide a kind of converse of the well known fact that continuous causal curves are Lipschitz when parametrized with respect to the arc-length of a Riemannian metric [53] (due to the

light cones that bound the curve). Known proofs in Lorentzian geometry [11] work under stronger differentiability assumption which guarantee the validity of the usual Gauss' Lemma.

For its proof we shall need a lemma (a Lorentzian version is [48, Lemma 2.13]). Given two Lorentzian-Finsler metrics g_1 and g_2 we write $g_1 < g_2$ if the causal vectors for g_1 are timelike for g_2.

Lemma 1 *Let (M, g) be a Finsler spacetime. Let $p \in M$ then we can find in a neighborhood O of p a flat Lorentzian metric (hence independent of v) g^+, such that $g < g^+$.*

Proof Let $\{x^\mu\}$ be a coordinate chart in a neighborhood O of $p \in M$. Let $h = (\mathrm{d}x^0)^2 + (\mathrm{d}x^1)^2 + \cdots + (\mathrm{d}x^{n-1})^2$ be the usual Euclidean metric and let us consider the corresponding unit sphere subbundle of TM. Let $\hat{J}_x$, $x \in O$, be the intersection of J_x with the unit sphere at x. Since the components of $J_x \backslash \{0\}$ are convex in the linear structure of $T_x M$, $\hat{J}_x$ is made of two closed disjoined convex sets of the unit sphere at $T_x M$. We can always find a great circle separating the two convex sets (note that the sphere has dimension $n - 1$, the circle has dimension $n - 2$, thus the terminology used is not accurate for $n \neq 3$) (to prove this use for instance the stereographic projection from a point not belonging to the convex sets, use the Hahn-Banach theorem, and then project back to the sphere) and we can also rotate the coordinate system so that the hyperplane on $T_p M$ spanned by that great circle is orthogonal to ∂_0. By continuity the Lorentzian metric at p, $g^+ = -N(\mathrm{d}x^0)^2 + (\mathrm{d}x^1)^2 + \cdots + (\mathrm{d}x^{n-1})^2$, satisfies $g < g^+$ for sufficiently large N. Still by continuity of J_x on x (this continuity can be rigorously expressed in the Hausdorff metric on sets, but the details will not be needed) the relation $g < g^+$ holds in a neighborhood of p which we can redefine to be O. □

Theorem 7 *Let (M, g) be a Finsler spacetime where M is a $C^{2,1}$-manifold endowed with a $C^{1,1}$ fundamental field g. Let I be an interval of the real line. Every future directed causal AC-curve $x : I \to M$ is a future directed continuous causal curve. Every future directed continuous causal curve $x : I \to M$ once suitably parametrized (e.g. with respect to the arc-length of a Riemannian metric) becomes a future directed causal locally Lipschitz curve.*

Remark 12 It is not true that every continuous causal curve is a causal AC-curve. For instance, consider the timelike geodesic of Minkowski spacetime which satisfies $\mathbf{x} = 0$ and which is parametrized by x^0. Consider the parametrization $t = f_s^{-1}(x^0)$ where f_s is a singular monotone continuous function [57, Ex. 8.20], so that $\dot{f}_s = 0$ almost everywhere.

Proof It is sufficient to prove it for $I = [a, b]$. Suppose that x is a future directed causal AC-curve, let C be a convex normal set intersecting x, and let $t_1 < t_2$ be such that $x([t_1, t_2]) \subset C$. The set C is a normal neighborhood for $p := x(t_1)$ thus by Theorem 6 the geodesic connecting $x(t_1)$ and $x(t_2)$ is (future directed) causal.

Conversely, suppose that $x : I \to M$ is a future directed continuous causal curve and let $\bar{t} \in I$. By Lemma 1 we can find $C^{2,1}$ coordinates x^μ in a convex neighborhood

C of $p := x(\bar{t})$ such that for some $N > 0$, the Lorentzian metric $g^+ = -N(\mathrm{d}x^0)^2 + (\mathrm{d}x^1)^2 + \cdots (\mathrm{d}x^{n-1})^2$ satisfies $g < g^+$ throughout C.

The function $x^0(t)$ must be increasing in a neighborhood of $\bar{t}$. Indeed, for t_1, t_2 belonging to a sufficiently small neighborhood of $\bar{t}$, $x(t_1), x(t_2) \in C$. By Theorem 6 there is a future directed causal g-geodesic connecting $x(t_1)$ with $x(t_2)$, which is in particular a future directed g^+-causal C^1 curve. But x^0 is increasing over this type of curve since x^0 is the usual time coordinate for the subset (C, g^+) of Minkowski spacetime, which proves the claim.

Once parametrized with respect to x^0 the curve becomes Lipschitz because of the condition of g-causality which implies g^+-causality which implies $\|\mathbf{x}(t_2) - \mathbf{x}(t_1)\| \leq N|x^0(t_2) - x^0(t_1)|$. Clearly, if l is an arc-length parameter induced by the Euclidean coordinate metric $(\mathrm{d}x^0)^2 + (\mathrm{d}\mathbf{x})^2$ then $x^0(l)$ is 1-Lipschitz, and so $x(l)$ is locally Lipschitz. As all Riemannian metrics are Lipschitz equivalent over compact sets, x is locally Lipschitz whenever parametrized with respect to Riemannian arc-length. □

Continuous causal curves in Lorentzian geometry enjoy nice properties under various notions of limit [45]. The proofs presented in reference [45] hold as well under our present weaker differentiability assumptions and in the Finsler spacetime case. Analogously, as should be expected from the above equivalence, the family of absolutely continuous curves is closed under uniform convergence, a fact quite well known since the work of Tonelli on one-dimensional variational principles [10].

The theorems so far proved in the Lorentzian–Finsler case are sufficient to establish the validity of most results of mathematical relativity and especially of causality theory for $C^{1,1}$ metrics on even Finsler rather than just Lorentzian spacetimes. In particular, the notions of chronological and causal relations do not require modifications from the standard ones [34]. The chronological relation is still open, the boundaries of the causal and chronological futures of a point coincide, the achronal boundaries are still Lipschitz hypersurfaces and so on.

We wish to include a result of this type to show that most proofs can be extended word by word from the Lorentzian C^3 metric case to the Lorentzian–Finslerian $C^{1,1}$ metric case. Let us recall that a set is achronal if there is no timelike curve starting and ending at the set.

Lemma 2 *Let (M, g) be a Finsler spacetime where M is a $C^{2,1}$-manifold endowed with a $C^{1,1}$ fundamental field g. Let $p < q$ then $(p, q) \in I^+$ or every continuous causal curve connecting p to q is an achronal future directed lightlike geodesic (up to parametrizations).*

Proof Assume $(p, q) \notin I^+$ and let $\gamma : [0, 1] \to M$ be a future directed continuous causal curve such that $\gamma(0) = p$ and $\gamma(1) = q$. Since the image of γ is compact there is a finite covering with convex normal neighborhoods U_i, $i = 1, \ldots, n$. We can assume $U_i \cap U_{i+1} \neq \emptyset$ and that there are $p_i \in \gamma \cap U_i \cap U_{i+1}$, $i = 1, \ldots, n-1$, $p_0 \equiv p \in U_1$ and $p_n \equiv q \in U_n$. Since γ is a continuous causal curve by Theorem 6 $(p_i, p_{i+1}) \in J^+_{U_{i+1}}$ thus p_i and p_{i+1} are joined by a geodesic η_i in U_{i+1} and this geodesic coincides with γ between the same points or it is timelike.

Let us show that the presence of one timelike segment η_i implies $(p, q) \in I^+$. This is so because from the curve made of geodesic segments η_i one can construct a

piecewise curve made of timelike geodesic segments. Indeed, one start from η_i and translate slightly the final point of η_{i-1} along η_i so that the new connecting η'_{i-1} becomes timelike (as the Lorentzian–Finsler distance between the new endpoints is necessarily positive). Analogously, one translates slightly the starting point of η_{i+1} along η_i so that the new connecting η'_{i+1} becomes timelike. Then one continues in this way by taking as reference the timelike geodesic segments η'_{i-1} or η'_{i+1}. The corners of the so obtained piecewise timelike curve can be finally smoothed out.

Also note that if all the segments η_i are lightlike but do not join smoothly then one can, arguing as above, replace one lightlike segment with one timelike segment by moving slightly the starting endpoint along the previous segment.

In conclusion if $(p, q) \notin I^+$ the continuous causal curve must be coincident with a lightlike geodesic connecting p to q. This geodesic must be achronal, otherwise there is a timelike curve σ connecting $p', q' \in \gamma$. The continuous causal curve connecting p to p' following γ and p' to q' following σ and q' to q following γ is, by the just proved result, a lightlike geodesic which is impossible since σ is timelike. The contradiction proves that γ is achronal. □

Under the assumption of the previous theorem we have:

Corollary 1 *If $p \ll r$ and $r \leq q$ then $p \ll q$. If $p \leq r$ and $r \ll q$ then $p \ll q$.*

Proof It follows from the fact that the composition of a timelike and a causal curve, in whatever order, gives a causal curve which is not a lightlike geodesic as at some points it is timelike. □

The question as to whether causality theory could be mindlessly generalized to $C^{1,1}$ spacetime metrics was considered by Chruściel and Grant [16] among others. They developed a generalization of causality theory to continuous metrics and found that classical results involving the existence of normal neighborhoods cannot be proved in their framework. They observed:

> We note that several statements in (causality theory) concerning geodesics remain true for $C^{1,1}$ metrics; it is conceivable that all of them remain true, but justifications would be needed.

We provided arguments which prove the correctness of this expectation. Indeed, the existence of convex normal neighborhoods is the central technical tool which allows one to complete many local arguments, such as that on the openness of the chronology relation, over which causality theory is based. Once the existence of convex normal neighborhoods has been established, and the local maximization property of causal geodesics has been obtained, most (if not all) results of causality theory follow without any substantial alteration to their classical proofs. Of course those result that can be expressed only through the use of the second derivative of the metric, for instance because they use the curvature tensor, would require further discussion (especially in the Finsler case).

Working with continuous metrics as in [16] expands very much causality theory though there is a price to be paid. Some desirable results do not hold anymore, for

instance lightlike geodesics are not necessarily locally achronal (for Lipschitz connections this result is guaranteed by Theorem 6). For other differences the reader is referred to [16].

1.5 Distance balls are convex

If the spray is the Levi-Civita connection of a Riemannian metric g it is natural to ask whether the convex neighborhood can be chosen to be a distance ball, that is, if for sufficiently small δ, $D_p^{-1}([0, \delta))$ is convex normal. Whitehead gave a positive answer to this problem through a proof which demands quite strong differentiability properties on D_p and hence on the metric. He reasons that it is sufficient to introduce normal coordinates $(y^1, \ldots, y^n)$ on N, because in this way the distance balls become coincident with the coordinate balls (we proved the existence of convex normal neighborhood constructing them as coordinate balls). Unfortunately, we can apply our argument which shows the convexity of coordinate balls only for charts which are $C^{2,1}$, thus the exponential map would have to be $C^{2,1}$ and hence the metric would have to be $C^{3,1}$ (and the manifold $C^{4,1}$). A different approach [58,61] demands just the twice continuous differentiability of D_p^2, but under our assumptions D_p^2 turns out to be just $C^{1,1}$. Finally, one could invoke Morse's Lemma so as to find a coordinate system in which the level sets of D_p^2 are coordinate spheres [43]. Unfortunately, this lemma applies only if D_p^2 is C^2.

Nevertheless, we shall prove that D_p^2 is strongly convex using a Picard–Lindelöf analysis.

Let us recall [23] that a real function defined on an open convex set C of an affine space A is strongly convex with constant λ if there is a $\lambda > 0$ such that for every $x, y \in C, \alpha \in [0, 1]$,

$$ f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y) - \frac{\lambda}{2}\alpha(1 - \alpha)\|x - y\|^2. $$

A C^1 real function is strongly convex on C with constant $\lambda > 0$ if and only if its differential df is strongly monotone with constant λ, that is, for every $x, y \in C$,

$$ [df(y) - df(x)] \cdot (y - x) \geq \lambda \|y - x\|^2. $$

In a Riemannian space we have analogous definitions and results [60, Prop. 16.2]. A real function defined on an open geodesically convex set C is geodesically strongly convex with constant λ if there is $\lambda > 0$ such that for every geodesic $x : J \rightarrow C$, $\alpha \in [0, 1]$,

$$ f(x((1 - \alpha)a + \alpha b)) \leq (1 - \alpha)f(x(a)) + \alpha f(x(b)) - \frac{\lambda}{2}\alpha(1 - \alpha)D(x(a), x(b))^2. $$

A C^1 real function is geodesically strongly convex on C with constant $\lambda > 0$ if and only if its differential df is geodesically strongly monotone with constant λ, that is, for every arc-length parametrized geodesic $x : J \rightarrow C$,

$$\mathrm{d} f(\dot{x}(b)) - \mathrm{d} f(\dot{x}(a)) \geq \lambda D(x(a), x(b)).$$

The proof of the equivalence is obtained applying the result for the affine space case to the composition $f(x(t))$.

We shall prove

Theorem 8 *Let M be a $C^{2,1}$-manifold endowed with a $C^{1,1}$ Riemannian metric g and corresponding locally Lipschitz Levi-Civita connection. Let $p \in M$ and let $\epsilon > 0$. Let $x^\mu : U \to \mathbb{R}^n$ be a chart in a neighborhood of p such that, $g_{\alpha\beta}(p) = \delta_{\alpha\beta}$, $\Gamma^\mu_{\alpha\beta}(p) = 0$. Let $C \subset U$ be a coordinate ball around p which we already know to be convex normal for sufficiently small radius. We also have that for sufficiently small radius for every $q, q_1, q_2 \in C$, interpreting the minus sign and scalar product through the Euclidean structure induced by the coordinate system on C*

$$|[d D_q^2(q_2) - d D_q^2(q_1)](q_2 - q_1) - 2(q_2 - q_1)^2| \leq \epsilon(q_2 - q_1)^2, \qquad (14)$$

and for every $q \in C$ and arc-length parametrized geodesic $x : J \to C$

$$|\nabla_{\dot{x}(b)} D_q^2(x(b)) - \nabla_{\dot{x}(a)} D_q^2(x(a)) - 2D(x(a), x(b))| \leq \epsilon D(x(a), x(b)). \qquad (15)$$

In particular, $D_q^2 : C \to [0, +\infty)$ is strongly convex with parameter $\lambda = 2 - \epsilon$ with respect to both the Euclidean

$$D(q, (1-\alpha)q_1 + \alpha q_2)^2 \leq (1-\alpha)D(q, q_1)^2 + \alpha D(q, q_2)^2$$
$$- \left(1 - \frac{\epsilon}{2}\right) \alpha(1-\alpha)\|q_2 - q_1\|^2,$$

and the Riemannian structures of C

$$D(q, x((1-\alpha)a + \alpha b))^2 \leq (1-\alpha)D(q, x(a))^2 + \alpha D(q, x(b))^2$$
$$- \left(1 - \frac{\epsilon}{2}\right) \alpha(1-\alpha)D(x(0), x(1))^2.$$

Thus for any sufficiently small r the open balls $D_p^{-1}([0, r))$ are contained in C and are (strictly) convex normal neighborhoods.

As the balls $D_p^{-1}([0, r))$ are convex we can infer a number of equivalent convexity properties thanks to the equivalences for metric spaces proved in [25], for instance D_p is itself convex.

We have the following improvement of our previous formulation of Gauss' Lemma (Theorem 5) which shows that the direct product sum for the metric can be accomplished on a sphere of a chosen radius $\bar{r}$ and almost everywhere outside it.

Theorem 9 *With the notations of the previous theorem, the levels sets $D_p^{-1}(r)$ are $C^{1,1}$ hypersurfaces diffeomorphic to S^{n-1} with an induced Lipschitz metric h_r. Locally on*

C we can always find $C^{2,1}$ functions θ_i $i = 1, \ldots, n - 1$, such that in the $C^{1,1}$ chart $(r, \theta_1, \ldots, \theta_{n-1})$ the metric takes the form

$$g = dr^2 + (h_r)_{ij}(d\theta_i - A_i dr)(d\theta_j - A_j dr), \qquad (16)$$

where $h_r = (h_r)_{ij} d\theta_i d\theta_j$ and all the components $(h_r)_{ij}$, A_i, are Lipschitz in (r, θ).

For any chosen sufficiently small radius $\bar{r}$ new Lipschitz coordinates $(r, \alpha_1, \ldots, \alpha_{n-1})$ can be found such that $\{\alpha_i\}$ provide a $C^{1,1}$ chart on $D_p^{-1}(\bar{r})$ and the metric takes the direct sum form

$$g = dr^2 + (h_r(\alpha))_{ij} \, d\alpha^i d\alpha^j,$$

where the components $(h_r(\alpha))_{ij}$ are defined almost everywhere and are bounded (this holds also for spherical normal coordinates) and for $r = \bar{r}$ they are defined everywhere and are Lipschitz in α.

1.6 Convexity on Lorentzian manifolds

In this section we study the problem of the existence of convex/concave functions and sets in Lorentzian manifolds. The Lorentzian case is more involved that the Riemannian but is essential for the understanding of Lorentzian manifolds under low differentiability conditions.

In the next theorem $\eta_{\alpha\beta}$ is the usual Minkowski metric in diagonal form ($\eta_{00} = -1$, $\eta_{ii} = +1$ for $i \geq 1$).

Theorem 10 *Let M be a $C^{2,1}$-manifold endowed with a $C^{1,1}$ Lorentzian metric g and corresponding Levi-Civita Lipschitz connection. Let $p \in M$ and let $\epsilon > 0$. Let $x^\mu : U \to \mathbb{R}^n$ be a chart in a neighborhood of p such that, $g_{\alpha\beta}(p) = \eta_{\alpha\beta}$, $\Gamma_{\alpha\beta}^\mu(p) = 0$. Let $C \subset U$ be a coordinate ball around p which we already know to be convex normal for sufficiently small radius. We also have that for sufficiently small radius, for every geodesic $x : [0, 1] \to C$, and for every $q \in C$*

$$\left| \left. \frac{d}{dt} D_q^2(x(t)) \right|_{t=1} - \left. \frac{d}{dt} D_q^2(x(t)) \right|_{t=0} - 2D^2(x(0), x(1)) \right| \leq \epsilon \|x(1) - x(0)\|^2, \quad (17)$$

where the Euclidean and affine structure of the coordinate chart is used just on the right-hand side.

Observe that in the next theorem there is no mention to the Euclidean or affine structures induced by a coordinate chart (we stress once again that D_p^2 can be negative).

Theorem 11 *Let M be a $C^{2,1}$-manifold endowed with a $C^{1,1}$ Lorentzian metric g and corresponding locally Lipschitz Levi-Civita connection. Let $p \in M$ and let $\epsilon > 0$. Let $\gamma : I \to M, t \mapsto \gamma(t)$, be a timelike geodesic such that $p = \gamma(0)$. The convex normal set $C \ni p$ of Theorem 4 can be chosen so small that once I is redefined to be the connected component of $\gamma^{-1}(C)$ containing 0, the following property holds.*

For every $q, r \in \gamma(I)$, $q \neq r$, there is a strictly convex normal set $O \ni r$, $\bar{O} \subset C$, such that all the points of $\bar{O}$ are either in the chronological future or in the chronological past of q; every geodesic $x : J \to O$, connecting two points on the same level set $(D_q^2)^{-1}(c)$, $c < 0$, of the function $D_q^2 : C \times C \to \mathbb{R}$, satisfies, once reparametrized with respect to g-arc length (x is necessarily spacelike by Theorem 6, thus $D(x(a), x(b)) \geq 0$), for every $a, b \in J$,

$$|\nabla_{\dot{x}(a)} D_q^2(x(a)) - \nabla_{\dot{x}(b)} D_q^2(x(b)) - 2D(x(a), x(b))| \leq \epsilon D(x(a), x(b)). \quad (18)$$

In particular, $D_q^2(x(t))$ is strongly convex with parameter $2 - \epsilon$

$$D_q^2(x((1-\alpha)a + \alpha b)) \leq (1-\alpha) D_q^2(x(a)) + \alpha D_q^2(x(b))$$
$$- \left(1 - \frac{\epsilon}{2}\right) \alpha(1-\alpha) D(x(a), x(b))^2,$$

and the sets of the form $(D_q^2)^{-1}((-\infty, c)) \cap O$ for $c < 0$ are strictly geodesically convex.

Remark 13 Mimicking the proof of [3, Prop. 3.1] it is easy to show that the Lorentzian distance on O from q, $D_q^L = \sqrt{-D_q^2}$ is semiconvex and hence almost everywhere first and twice differentiable. However, in order to prove the geodesic convexity of $(D_q^2)^{-1}((-\infty, c)) \cap O$ we need a result on the convexity of D_q^2 rather than on the convexity of D_q^L. Furthermore, we stress that the notion of semiconvexity, due to Rockafellar, is a rather weak notion. For instance the concave function $-x^2$ on $\mathbb{R}$ is semiconvex (thus the terminology *semiconvexity* can be misleading).

We recall that a subset $A \subset M$ is *causally convex* in $B \subset M$, with $A \subset B$, if every C^1 causal curve contained in B and joining two points of A is necessarily contained in A. An open subset B is *strongly causal* if every point $p \in B$ admits arbitrarily small open neighborhoods which are causally convex in B. An open subset S of M is called *causally simple* if it is strongly causal and $J_S^+ \subset S \times S$ is closed in the product topology [34,48]. A causally simple subset S is *globally hyperbolic* if for every $p, q \in S$, $J_S^+(p) \cap J_S^-(p)$ is compact.

The following claim was known under stronger differentiability assumptions. It can be found in a footnote of [48].

Theorem 12 *Convex normal subsets are causally simple.*

Proof Let T be the global timelike vector field which provides the time orientation. Let C be a convex normal subset, and let $f_1, f_2 : C \times C \to \mathbb{R}$ be the functions $f_1(p, q) := g(\exp_p^{-1} q, \exp_p^{-1} q)$, $f_2(p, q) := g(\exp_p^{-1} q, T(p))$ since $\exp^{-1}$ and g are continuous f_1 and f_2 are continuous, and hence $J_C^+ = f_1^{-1}((-\infty, 0]) \cap f_2^{-1}((-\infty, 0))$ is closed.

A spacetime C is strongly causal if and only if for every $p, q \in C$, $(p, q) \in J_C^+$ and $(q, p) \in \overline{J_C^+}$ imply $p = q$ (see [44]). Suppose that there are $p, q \in C$, $p \neq q$, such that $p \leq_C q$ and $q \leq_C p$ (we just proved $\overline{J_C^+} = J_C^+$). Let γ_1 be the future directed

causal geodesic connecting p to q and let γ_2 be the future directed causal geodesic connecting q to p. Then the images of γ_1 and γ_2 differ (otherwise there would be a geodesic which is both future and past directed), and hence there are two geodesics connecting p to q, a contradiction to the uniqueness of the connecting geodesic in convex normal sets (Theor. 4). □

Corollary 2 *Let M be a $C^{2,1}$-manifold endowed with a $C^{1,1}$ Lorentzian metric g and corresponding locally Lipschitz Levi-Civita connection. Let $p \in M$ then there is a local base $\{C_i\}$ for the topology at p such that for every i, C_i is strictly convex normal, globally hyperbolic, $\bar{C}_{i+1} \subset C_i$, and C_{i+1} is causally convex in C_i (and hence C_1).*

Part of the previous result was already known [48], what was open was the result on the convexity of C_i. Observe that if C_i is globally hyperbolic then any two causally related events are connected by a causal geodesic contained in C_i (by the Avez-Seifert theorem [34]). Thus what was really missing was the convexity with respect to spacelike geodesics. These globally hyperbolic convex normal sets behave pretty well, indeed this property is left invariant under finite intersections. As a consequence, by Lebesgue's covering lemma, for any compact set K on spacetime, for any metric $\rho : M \times M \to \mathbb{R}$ inducing the manifold topology, and finite covering of K by globally hyperbolic convex normal sets, there is an $\epsilon > 0$ such that for any pair of points p, q such that $\rho(p, q) < \epsilon$, points p and q are contained in one element of the covering and hence connected by some geodesic.

Remark 14 A spacetime is strongly causal if every point admits an arbitrarily small causally convex set (in M). In a strongly causal spacetime, we can find a causally convex (in M) open neighborhood Y of p contained in C_1, and for sufficiently large i, $C_i \subset Y$. As a consequence, the sets C_i for sufficiently large i are also causally convex in M. Thus for strongly causal spacetimes we can include in the previous Corollary the causal convexity of C_i with respect to M among the properties of these sets.

1.7 Two variations on the main theme

Our results admit a number of variations obtained replacing the base point p of the pointed exponential map with an embedded manifold. For instance, consider a pseudo-Riemannian manifold endowed with a $C^{1,1}$ metric and a Lipschitz metric compatible connection. Let $\phi : S \to M$, $k \geq 1$, be a $C^{1,1}$ k-dimensional embedding, such that the induced metric is pseudo-Riemannian. Its non-degeneracy implies that at each point $p \in \phi(S)$ the tangent space $T_p M$ is the direct sum of the tangent space to $\phi(S)$ and it normal space. Let $\nu(S)$ be the corresponding n-dimensional normal bundle with base $\phi(S)$.

We shall prove the following theorem analogous to Theorem 3:

Theorem 13

$(\exp_{\nu(S)})$ *The vector bundle $\nu(S)$ is Lipschitz. Moreover, the map $\exp_{\nu(S)}$ is strongly differentiable on the image of the zero section of $\pi|_{\nu(S)} : \nu(S) \to \phi(S)$, and establishes a Lipeomorphism between a neighborhood of the image of the zero section and a neighborhood of $\phi(S)$.*

By the Lipschitzness of $\nu(S)$ we can apply to $\exp_{\nu(S)} := \exp|_{\nu(S)}$ the results of Theorem 1. This theorem can also be used to construct tabular neighborhoods. The pointed exponential map can be regarded as a special case of this type of construction for $k = 0$.

We can also consider a function H^μ dependent on time provided it is Lipschitz. Let $\gamma_{(t_0, v)}(t)$ be any solution of

$$\frac{dx^\mu}{dt} = v^\mu, \tag{19}$$

$$\frac{dv^\mu}{dt} = H^\mu(t, x, v), \tag{20}$$

with initial condition $v \in TM$ at time t_0.

Theorem 14 *The solution $\gamma_{(t_0, v)}(t)$ is locally Lipschitz in (t, t_0, v) and $\gamma_{(t_0, v)}(\cdot)$ is $C^{2,1}_{loc}$. Let $(\hat{t}, \hat{p}) \in \mathbb{R} \times M$, then for every $\epsilon > 0$ there is an open neighborhood C of $\hat{p}$ such that for every $p_1, p_2 \in C$, $t_1, t_2 \in (t_0 - \epsilon, t_0 + \epsilon)$, $t_1 < t_2$, there is one and only one solution starting from p_1 at time t_1 and reaching p_2 at time t_2 entirely contained in C.*

The idea is to rewrite the system of first order ODE as follows

$$\frac{dt}{ds} = t',$$

$$\frac{dx^\mu}{ds} = x'^\mu,$$

$$\frac{dt'}{ds} = 0,$$

$$\frac{dx'^\mu}{ds} = (t')^2 H^\mu(t, x, x'/t') = H^\mu(t, x, x'),$$

where $x' = vt'$, so as to reduce the problem to the (s-)time independent case. The previous theorem is then a corollary of Theorem 4 whenever in the proof provided for that theorem in place of the Banach space $(\mathbb{R}^n, \|\ \|)$ we consider the Banach space $(\mathbb{R}^{n+1}, \max(\|\ \|, |\ |))$.

1.8 Some technical preliminary results

Let us first recall that in finite dimensions and for Lipschitz functions the notions of Gâteaux differential and Frechet differential coincide [19, p. 158].

Proposition 2 *Any Lipschitz function $f : O \to \mathbb{R}$, $O \subset \mathbb{R}^n$, which is Gâteaux differentiable at p is differentiable at p.*

Proof Let G be the Gâteaux differential at p. By contradiction, suppose that G is not the differential of f at p, then there is $\epsilon > 0$ and a sequence $v_n \in \mathbb{R}^n$, $v_n \neq 0$, $v_n \to 0$, such that

$$\|f(p + v_n) - f(p) - G(v_n)\| > \epsilon\|v_n\|. \tag{21}$$

Let $e_n = v_n/\|v_n\|$, by the compactness of S^{n-1} we can assume without loss of generality that $e_n \to e$, with $\|e\| = 1$. Let us decompose v_n in components parallel and perpendicular to e

$$v_n = a_n e + b_n$$

where a_n is a scalar and b_n is a vector. We have

$$
\begin{aligned}
\|f(p + v_n) - f(p) - G(v_n)\| &\le \|f(p + v_n) - f(p + a_n e)\| \\
&\quad + \|f(p + a_n e) - f(p) - G(a_n e)\| \\
&\quad + \|G(a_n e) - G(v_n)\| \\
&= K\|b_n\| + o_e(a_n) + |G|\,\|b_n\|,
\end{aligned}
$$

where K is the Lipschitz constant. The condition $e_n \to e$ reads $a_n/\|v_n\| \to 1$, $\|b_n\|/\|v_n\| \to 0$. For sufficiently large n the previous inequality contradicts Eq. (21). $\qquad\square$

In order to prove Theorem 1 we shall need the following result on differentiation under the integral sign which, we believe, is interesting in its own right.

Theorem 15 *Let $f : [a, b] \times \mathbb{R}^k \to \mathbb{R}$, $(x, y) \mapsto f(x, y)$, be a continuous function which is locally Lipschitz in y uniformly in x. Then for almost every y, the differential $d_2 f$ exists at (u, y) for almost every $u \in [a, b]$, it is summable in $[a, b]$, and for every $x \in [a, b]$:*

$$\left(d_2 \int_a^x f(u, y)\,du\right) = \int_a^x d_2 f(u, y)\,du.$$

Proof Let $\bar{y} \in \mathbb{R}^n$, there is a relatively compact neighborhood $O \ni \bar{y}$ which is a product of open intervals. The function $f(x, \cdot)|_O$ is K-Lipschitz in y where K does not depend on x. In particular, for any $x \in [a, b]$ the function $f(x, \cdot)$ is differentiable almost everywhere (Rademacher's theorem). Let $E(x) \subset O$ be the subset where $f(x, \cdot)|_O$ is differentiable. Fubini's theorem applied to the characteristic function of

$$A = \cup_{x \in [a, b]}[\{x\} \times E(x)] \subset [a, b] \times O,$$

states that for almost every $y \in O$, the differential $d_2 f$ exists at (x, y) for almost every $x \in [a, b]$. From now on let y be one of these special values.

As a consequence, for almost every $y \in O$, and for every vector $v \in \mathbb{R}^k$, the partial derivative $\partial_v f(x, y)$ exists for almost every $x \in [a, b]$. Because of the Lipschitz condition we have $|\partial_v f(x, y)| \le K\|v\|$. Let $\epsilon_n \to 0$ then

$$\partial_v f(x, y) = \lim_{n \to \infty} f_n^y(x),$$

where

$$f_n^y(x) = \frac{1}{\epsilon_n}[f(x, y + v\epsilon_n) - f(x, y)], \qquad |f_n^y(x)| \leq K\|v\|.$$

By the dominated convergence theorem $\partial_v f(x, y)$ is summable and

$$\partial_v \int_a^x f(u, y) du = \int_a^x \partial_v f(u, y) du.$$

This equation proves that the linear operator $G(v)$ on the right-hand side is the Gâteaux differential at y of $F(x, v) := \int_a^x f(u, v) du$. Moreover, this function is Lipschitz in v thus $F(x, \cdot)$ is differentiable and G coincides with its differential (Prop. 2). □

Alternatively, we could use a theorem [46] which improves Schwarz's theorem on the equality of mixed partial derivatives. The reader is referred to [46] for further details. A result similar to the next one was proved by Federer [24, Lemma 4.7].

Theorem 16 *Let $f \colon O \to \mathbb{R}$ be a Lipschitz function defined on an open subset of $\mathbb{R}^n$. Let $J \colon O \to \mathbb{R}^n$ be a continuous function. If the differential of f exists and coincides with J almost everywhere then f is C^1 and its differential is J.*

Proof For $n = 1$ this statement follows from inspection of the function

$$f(x) - f(a) - \int_a^x J(u) du,$$

$a \in O$. This function is Lipschitz, thus absolutely continuous, and hence has a derivative which vanishes almost everywhere, thus it is a constant. Taking $x = a$ we find that this constant is zero, thus $f(x) = f(a) + \int_a^x J(u) du$ which implies that f is C^1 with derivative J.

We can assume without loss of generality, $O = \mathbb{R}^n$. Let $v \in \mathbb{R}^n$, we want to show that at every $x \in O$, $\partial_v f = J(v)$. This fact would imply that f is Gâteaux differentiable at each point with Gâteaux differential J, thus the desired conclusion would follow from Prop. 2.

By Fubini's and Rademacher's theorem for almost every hyperplane perpendicular to v, the function f is almost everywhere differentiable at the points belonging to the hyperplane.

Let us introduce coordinates $\{y^0, \ldots, y^n\}$ such that one such hyperplane has equation $y^0 = 0$ and $v = \partial_{y^0}$. Let $g(y_1, \cdot, y_n) := f(0, y_1, \ldots, y_n)$ be the Lipschitz restriction of f to the hyperplane.

The function $J(t, y)$ is continuous and hence uniformly continuous over compact subsets. As a consequence, the function

$$F(t, y) = g(y) + \int_0^t J(t, y; e_0) dt$$

is continuous in (t, y) and C^1 in t.

Let $(t, y) \in \mathbb{R}^n$ and let $U \ni (t, y)$ be an open neighborhood. The function f is almost everywhere differentiable with differential J. By Fubini's theorem for almost every line parallel to the y^0-axis (the line is determined by its intersection with $y^0 = 0$ and "a.e." in this statement is meant in the Lebesgue $(n-1)$-dimensional measure of this hyperplane), f is almost everywhere differentiable with differential J. Thus there is some $(t', y') \in U$ passing through one such line. The function $f - F$ over such line is Lipschitz and differentiable almost everywhere with zero derivative, thus it is a constant. But $f - F = 0$ at $y^0 = 0$ thus that constant vanishes and hence $f(t', y') = F(t', y')$. As both functions are continuous and U is arbitrary, $f(t, y) = F(t, y)$ that is $f = F$. We conclude that $\partial_v f = \partial_v F = \partial_{y^0} F = J(e_0)$, which is what we wanted to prove. $\qquad\square$

2 Proofs I: a Picard-Lindelöf analysis

This section is devoted to the proofs of the results stated in the previous section and, in particular, to the proof of Theorem 4. In order to prove the strong differentiability of exp we shall make a Picard-Lindelöf analysis of the geodesic equation in the small. We shall also pass through the proof of existence and Lipschitz dependence on initial conditions. Though they are already known they will be useful to fix the notation and introduce the bounds used in the last step of the proof.

Let $x^\mu : U \to \mathbb{R}^n$ be a coordinate chart in a neighborhood U of $p \in M$. Without loss of generality let us assume that $x^\mu(p) = 0$ and let r be such that the closed ball $\bar{B}(p, r) = \{q : \|q - p\| \le r\}$ is contained in U, where $\| \ \|$ is the Euclidean coordinate norm. On the tangent bundle we introduce the local coordinate system $\{x^\mu, \dot{x}^\mu\}$. We shall regard the coordinate chart image $x^\mu(U)$ as an open subset of the normed space $(\mathbb{R}^n, \| \ \|)$.

Let us consider the system (1, 2) where H^μ is homogeneous of second degree in v and Lipschitz

$$\|H(x_2, v_2) - H(x_1, v_1)\| \le \alpha \|x_2 - x_1\| + \beta \|v_2 - v_1\|$$

in the domain $\bar{B}(p, r) \times \{v : \|v\| \le 1\}$.

Suppose that v_1, v_2 are not bounded as stated. Let V be any constant such that

$$V > \max(\|v_1\|, \|v_2\|). \tag{22}$$

The Lipschitz conditions is then better rewritten in the following form:

$$\begin{aligned}
\|H(x_2, v_2) - H(x_1, v_1)\| &= V^2 \left\| H(x_2, \tfrac{v_2}{V}) - H(x_1, \tfrac{v_1}{V}) \right\| \\
&\le V^2 \left\{ \alpha \|x_2 - x_1\| + \beta \| \tfrac{v_2}{V} - \tfrac{v_1}{V} \| \right\} \\
&\le \alpha V^2 \|x_2 - x_1\| + \beta V \|v_2 - v_1\|.
\end{aligned}$$

Let

$$M = \sup_{x \in \bar{B}(p,r)} \sup_{\|v\|=1} \|H(x,v)\|. \tag{23}$$

We rewrite (1, 2) in integral form

$$x^\mu(t, x_0, \dot{x}_0) = x_0^\mu + \int_0^t \dot{x}^\mu(t, x_0, \dot{x}_0) \, dt, \tag{24}$$

$$\dot{x}^\mu(t, x_0, \dot{x}_0) = \dot{x}_0^\mu + \int_0^t H^\mu\Big(x(t, x_0, \dot{x}_0), \dot{x}(t, x_0, \dot{x}_0)\Big) \, dt, \tag{25}$$

where $(x_0, \dot{x}_0)$ is the initial condition at time $t = 0$. We have included in the above expression the dependence of $(x, \dot{x})$ on the initial conditions. Unless needed we shall remove it from the expressions below.

Let $(x_0, \dot{x}_0)$ belong to the domain

$$\max\{\|x_0\|, \|\dot{x}_0\|\} < \delta, \tag{26}$$

where δ is a positive constant such that (the expression makes sense for $M = 0$)

$$\delta < \frac{1}{M}\left(1 - e^{-Mr/2}\right) \leq \frac{r}{2}, \tag{27}$$

and is sufficiently small that

$$\frac{\delta}{1 - \delta M} \leq 1, \qquad \frac{\beta\delta}{2(1 - \delta M)}\left(1 + \sqrt{1 + 4\alpha/\beta^2}\right) \leq 1. \tag{28}$$

Starting from $k = -1$ with

$$x_{-1}(t) = x_0, \quad \dot{x}_{-1}(t) = 0,$$

we define inductively the next two functions defined over $[0, 1]$ which, in the induction hypothesis are well defined and C^1 and are such that $(x_k, \dot{x}_k)(t) \in \pi^{-1}(\bar{B}(p,r))$ for every $t \in [0, 1]$

$$x_{k+1}^\mu(t) = x_0^\mu + \int_0^t \dot{x}_k^\mu(t) \, dt, \tag{29}$$

$$\dot{x}_{k+1}^\mu(t) = \dot{x}_0^\mu + \int_0^t H^\mu\Big(x_k(t), \dot{x}_k(t)\Big) \, dt. \tag{30}$$

In particular observe that they imply for $k = 0$

$$x_0(t) = x_0, \quad \dot{x}_0(t) = \dot{x}_0,$$

while for $k \geq 1$ they will generically depend on t. From Eqs. (29, 30) we obtain (we shall repeatedly use the inequality $\|\int v^\mu(t)dt\| \leq \int \|v^\mu(t)\|dt$ which is well known

once interpreted as the fact that the length of a C^1 curve on $\mathbb{R}^n$ is greater than the distance between its endpoints)

$$\|x_{k+1}(t)\| \leq \|x_0\| + \int_0^t \|\dot{x}_k(t)\| \, dt,$$

$$\|\dot{x}_{k+1}(t)\| \leq \|\dot{x}_0\| + \int_0^t M \|\dot{x}_k(t)\|^2 \, dt.$$

The second inequality implies inductively the following bound

$$\|\dot{x}_{k+1}(t)\| \leq \frac{\|\dot{x}_0\|}{1 - \|\dot{x}_0\| M t},$$

which is clearly satisfied for $k = -1$ and which, replaced into the first inequality gives

$$\|x_{k+1}(t)\| \leq \|x_0\| - \frac{1}{M} \ln(1 - \|\dot{x}_0\| M t) \leq \delta - \frac{1}{M} \ln(1 - \delta M).$$

By (27) we have for every $k \geq 0$,

$$\|x_k(t)\| < r,$$

$$\|\dot{x}_k(t)\| < \frac{\delta}{1 - \delta M}$$

thus these functions define indeed points belonging to $\pi^{-1}(\bar{B}(p, r))$. We observe that for any instants $t_1, t_2 \in [0, 1]$ and any pair $v_1 := \dot{x}_i(t_1)$, $v_2 := \dot{x}_j(t_2)$

$$V := \frac{\delta}{1 - \delta M} \tag{31}$$

satisfies the condition of Eq. (22). From now on V will be given by this equation.

2.1 Existence of geodesics

We have:

$$\|x_{k+1}(t) - x_k(t)\| \leq \int_0^t \|\dot{x}_k(t) - \dot{x}_{k-1}(t)\| \, dt,$$

$$\|\dot{x}_{k+1}(t) - \dot{x}_k(t)\| \leq \int_0^t \|H(x_k(t), \dot{x}_k(t)) - H(x_{k-1}(t), \dot{x}_{k-1}(t))\| \, dt,$$

$$\leq \int_0^t \{A\|x_k(t) - x_{k-1}(t)\| + B\|\dot{x}_k(t) - \dot{x}_{k-1}(t)\|\} \, dt,$$

where $A = [\frac{\delta}{1-\delta M}]^2 \alpha$ and $B = \frac{\beta \delta}{1-\delta M}$. There is a positive constant D such that

$$\frac{A}{D} + B = D, \tag{32}$$

namely

$$D = \frac{1}{2}\left(B + \sqrt{B^2 + 4A}\right) = \frac{\beta\delta}{2(1 - \delta M)}\left(1 + \sqrt{1 + 4\alpha/\beta^2}\right). \tag{33}$$

By Eq. (28)

$$D \leq 1.$$

By induction we have the bounds

$$\|x_k(t) - x_{k-1}(t)\| \leq \frac{1}{D}\frac{(Dt)^k}{k!}$$
$$\|\dot{x}_k(t) - \dot{x}_{k-1}(t)\| \leq \frac{(Dt)^k}{k!}$$

They imply that the series

$$x_k^\mu(t) = x_0^\mu + \sum_{k=0}^{n}\left[x_{k+1}^\mu(t) - x_k^\mu(t)\right],$$
$$\dot{x}_k^\mu(t) = \dot{x}_0^\mu + \sum_{k=0}^{n}\left[\dot{x}_{k+1}^\mu(t) - \dot{x}_k^\mu(t)\right],$$

define a succession of continuous functions which converge uniformly to (continuous) functions $x^\mu(t)$ and $\dot{x}^\mu(t)$ over [0, 1]. By uniform convergence we can pass to the limit in Eqs. (29, 30). Indeed, observe that $H^\mu(x, \dot{x})$ is continuous over the compact set $\bar{B}(p, r) \times \{v : \|v\| \leq V\}$ and hence uniformly continuous. Thus these limits indeed solve the system (24, 25). In particular, Eq. (24) proves that indeed $\dot{x}^\mu(t) = \frac{d}{dt}x^\mu(t)$.

For the proof of the uniqueness of geodesics the reader can consult [31, Theor. 1.1, Chap. 2].

2.2 Lipschitz dependence on initial conditions (Theorem 1)

The Lipschitz dependence on the initial condition can be proved using the Gronwall's inequality. It is rather easy to obtain it directly using the Picard–Lindelöf approximation.

Let us consider two solutions $x^\mu(t)$ and $y^\mu(t)$ of the geodesic equation with initial conditions $(x_0^\mu, \dot{x}_0^\mu)$, $(y_0^\mu, \dot{y}_0^\mu)$, which belong to the domain given by Eq. (26). Starting from these initial conditions we define inductively functions $x_k^\mu(t)$, $y_k^\mu(t)$ as above, and subtract the corresponding Eqs. (29, 30).

$$\|x_{k+1}(t) - y_{k+1}(t)\| \leq \|x_0 - y_0\| + \int_0^t \|\dot{x}_k(t) - \dot{y}_k(t)\| \, dt,$$

$$\|\dot{x}_{k+1}(t) - \dot{y}_{k+1}(t)\| \leq \|\dot{x}_0 - \dot{y}_0\| + \int_0^t \{A\|x_k(t) - y_k(t)\| + B\|\dot{x}_k(t) - \dot{y}_k(t)\|\} \, dt.$$

We now regard the chart-trivialized tangent bundle of coordinates $\{x^\mu, \dot{x}^\mu\}$ as the direct sum of vector spaces $\mathbb{R}^n \oplus \mathbb{R}^n$ endowed with the norm

$$\|(x, \dot{x})\| := \max\{\|x\|, \|\dot{x}\|\}.$$

By induction we obtain

$$\|x_k(t) - y_k(t)\| \leq \max\{D\|x_0 - y_0\|, \|\dot{x}_0 - \dot{y}_0\|\} \frac{e^{Dt}}{D},$$

$$\|\dot{x}_k(t) - \dot{y}_k(t)\| \leq \max\{D\|x_0 - y_0\|, \|\dot{x}_0 - \dot{y}_0\|\} e^{Dt}.$$

Clearly, the induction hypothesis is satisfied for $k = 0$. Taking the limit $k \to \infty$

$$\|x(t) - y(t)\| \leq \max\{D\|x_0 - y_0\|, \|\dot{x}_0 - \dot{y}_0\|\} \frac{e^{Dt}}{D},$$

$$\|\dot{x}(t) - \dot{y}(t)\| \leq \max\{D\|x_0 - y_0\|, \|\dot{x}_0 - \dot{y}_0\|\} e^{Dt}.$$

Recalling that $D \leq 1$ these inequalities imply

$$\|(x, \dot{x})(t) - (y, \dot{y})(t)\| \leq \|(x_0, \dot{x}_0) - (y_0, \dot{y}_0)\| \frac{e^{Dt}}{D},$$

which proves the Lipschitz dependence on the initial conditions (the exponential map is obtained for $t = 1$).

The joint dependence on t and $(x_0, \dot{x}_0)$ is also locally Lipschitz, indeed since the time dependence of $(x, \dot{x})(t)$ is $C^{1,1}$ it is locally Lipschitz, thus for $t, t' \in [a, b]$

$$\|(x, \dot{x})(t') - (y, \dot{y})(t)\| \leq \|(x, \dot{x})(t') - (x, \dot{x})(t)\| + \|(x, \dot{x})(t) - (y, \dot{y})(t)\|,$$

from which we infer the local Lipschitzness of the dependence on $(t, x_0, \dot{x}_0)$.

As a consequence, the geodesic flow over TM is locally Lipschitz because every geodesic segment can be covered with a finite number of coordinate patches. By continuity the domain W where it is defined is open and satisfies the conditions of Theorem 1. Analogously, Ω is open.

Let us rewrite the system (24, 25) reintroducing the dependence on the initial conditions

$$(x, \dot{x})(t, x_0, \dot{x}_0) = (x_0, \dot{x}_0) + \int_0^t (f_1, f_2)(t, x_0, \dot{x}_0) dt, \tag{34}$$

where

$$f_1^\mu(t, x_0, \dot{x}_0) = \dot{x}^\mu(t, x_0, \dot{x}_0),$$
$$f_2^\mu(t, x_0, \dot{x}_0) = H^\mu\big(x(t, x_0, \dot{x}_0), \dot{x}(t, x_0, \dot{x}_0)\big).$$

Since the dependence $(x, \dot{x})(t, x_0, \dot{x}_0)$ is Lipschitz we can apply Theorem 15 with the replacements $x \to t$, $y \to (x_0, \dot{x}_0)$.

We conclude that for almost every $(x_0, \dot{x}_0)$ the differential $\mathrm{d}_{(x_0,\dot{x}_0)}(f_1, f_2)(u)$ exists for almost every u, it is summable and for every $t \in [0, 1]$

$$\mathrm{d}_{(x_0,\dot{x}_0)}(x, \dot{x}) = \mathrm{d}_{(x_0,\dot{x}_0)}(x_0, \dot{x}_0) + \int_0^t \mathrm{d}_{(x_0,\dot{x}_0)}(f_1, f_2)(u, x_0, \dot{x}_0)\mathrm{d}u. \qquad (35)$$

In particular, for almost every $(x_0, \dot{x}_0)$ satisfying Eq. (27) the differential $\mathrm{d}_{(x_0,\dot{x}_0)}(x, \dot{x})$ exists for every $t \in [0, 1]$ and is given by this equation. Let us call U this special subset of the initial conditions.

Since the functions f_1 and f_2 are locally Lipschitz the differential on the right-hand side is bounded and so, for every $(x_0, \dot{x}_0) \in U$, the quantity $\mathrm{d}_{(x_0,\dot{x}_0)}(x, \dot{x})$ has a Lipschitz dependence on t where the Lipschitz constant does not vary in a small relatively compact neighborhood of $(x_0, \dot{x}_0)$. In other words, $\mathrm{d}_{(x_0,\dot{x}_0)}(x, \dot{x})$ is a function dependent on $(t, x_0, \dot{x}_0)$ which is Lipschitz in t uniformly in those $(x_0, \dot{x}_0)$ belonging to U.

Differentiating Eq. (35) with $(x_0, \dot{x}_0) \in U$ we get that for almost every t

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\mathrm{d}_{(x_0,\dot{x}_0)}(x, \dot{x}) = \mathrm{d}_{(x_0,\dot{x}_0)}(f_1, f_2) = \mathrm{d}_{(x_0,\dot{x}_0)}\frac{\mathrm{d}}{\mathrm{d}t}\,(x, \dot{x}).$$

Though we performed just a local analysis, the conclusion does not change in the setting of Theorem 1 where U is replaced by $\tilde{\Omega}$ since, as observed above, every geodesic segment can be covered with a finite number of coordinate patches. The fact that U and $\tilde{\Omega}$ are star-shaped is a consequence of Eq. (9).

In order to prove the last statement of Theorem 1 it suffices to observe that we can restart the last argument beginning with Eq. (34) by introducing the Lipschitz dependence $(x_0, \dot{x}_0)(z)$ where z are local coordinates on N. Since $(x, \dot{x})(t, z)$ is locally Lipschitz the whole argument stills works where it is understood that $\partial_z \varphi$ exists almost everywhere in the Lebesgue m-dimensional measure of N (This measure can be equivalently defined either as done here using a chart of N or in a more intrinsic way regarding N as a subset of TM, see [22, Sect. 3.3.3]). Theorem 1 is proved.

Remark 15 We pause for a moment to outline how to generalize this proof for $C^{1,1}$ sprays over $C^{3,1}$ manifold, the further generalization to the $C^{k,1}$, $k \geq 0$, spray case being then analogous.

The idea is to introduce variables $x^\mu_{k,\beta}$, $x^\mu_{k,\dot{\beta}}$, $\dot{x}^\mu_{k,\beta}$, $\dot{x}^\mu_{k,\dot{\beta}}$ and add to the system (29, 30) the equations obtained (formally) differentiating the right-hand side (29, 30) with respect to the initial conditions

$$x^\mu_{k+1,\beta}(t) = \delta^\mu_\beta + \int_0^t \dot{x}^\mu_{k,\beta}(t)\,\mathrm{d}t,$$

$$x^\mu_{k+1,\dot{\beta}}(t) = \int_0^t \dot{x}^\mu_{k,\dot{\beta}}(t)\,\mathrm{d}t,$$

$$\dot{x}^\mu_{k+1,\beta}(t) = \int_0^t (\partial_{x^\alpha} H^\mu) x^\alpha_{k,\beta} + (\partial_{\dot{x}^\alpha} H^\mu) \dot{x}^\alpha_{k,\beta} \, dt,$$

$$\dot{x}^\mu_{k+1,\dot{\beta}}(t) = \delta^\mu_\beta + \int_0^t (\partial_{x^\alpha} H^\mu) x^\alpha_{k,\dot{\beta}} + (\partial_{\dot{x}^\alpha} H^\mu) \dot{x}^\alpha_{k,\dot{\beta}} \, dt.$$

The proof of the convergence for $k \to \infty$ proceeds as above. From here it follows that $x^\mu_{,\beta}$, $x^\mu_{,\dot{\beta}}$, $\dot{x}^\mu_{,\beta}$, $\dot{x}^\mu_{,\dot{\beta}}$ are Lipschitz as they solve a first order Lipschitz ODE. Since the solution is unique $x^\mu_{,\beta}$ coincides with $\partial_{x^\beta_0} x^\mu$ and so on.

2.3 Strong differentiability of exp (Theorems 3 and 13)

Let us consider two solutions $x^\mu(t)$ and $y^\mu(t)$ of the geodesic equation with initial conditions $(x^\mu_0, \dot{x}^\mu_0)$, $(y^\mu_0, \dot{y}^\mu_0)$, which belong to the domain given by Eq. (26). Starting from these initial conditions we define inductively functions $x^\mu_k(t)$, $y^\mu_k(t)$ as above, and subtract the corresponding Eqs. (29, 30), rearranging them as follows

$$x_{k+1}(t) - y_{k+1}(t) - (x_0 - y_0) - (\dot{x}_0 - \dot{y}_0)t = \int_0^t [\dot{x}_k(t) - \dot{y}_k(t) - (\dot{x}_0 - \dot{y}_0)] \, dt,$$

$$\dot{x}_{k+1}(t) - \dot{y}_{k+1}(t) - (\dot{x}_0 - \dot{y}_0) = \int_0^t \left[H(x_k(t), \dot{x}_k(t)) - H(y_k(t), \dot{y}_k(t)) \right] dt.$$

Thus

$$\|x_{k+1}(t) - y_{k+1}(t) - (x_0 - y_0) - (\dot{x}_0 - \dot{y}_0)t\| \le \int_0^t \|\dot{x}_k(t) - \dot{y}_k(t) - (\dot{x}_0 - \dot{y}_0)\| \, dt,$$

and

$$\|\dot{x}_{k+1}(t) - \dot{y}_{k+1}(t) - (\dot{x}_0 - \dot{y}_0)\| \le \int_0^t \|H(x_k(t), \dot{x}_k(t)) - H(y_k(t), \dot{y}_k(t))\| \, dt$$

$$\le \int_0^t \{A\|x_k - y_k\| + B\|\dot{x}_k - \dot{y}_k\|\} \, dt$$

$$\le \int_0^t \{A[\|x_k - y_k - (x_0 - y_0) - (\dot{x}_0 - \dot{y}_0)t\| + \|x_0 - y_0\| + \|\dot{x}_0 - \dot{y}_0\|t]$$
$$+ B[\|\dot{x}_k - \dot{y}_k - (\dot{x}_0 - \dot{y}_0)\| + \|\dot{x}_0 - \dot{y}_0\|]\} \, dt.$$

By induction it follows that

$$\|x_k(t) - y_k(t) - (x_0 - y_0) - (\dot{x}_0 - \dot{y}_0)t\| \le \max\{D\|x_0 - y_0\|, \|\dot{x}_0 - \dot{y}_0\|\}$$
$$\times \left(\frac{e^{Dt} - 1}{D} - t \right),$$

$$\|\dot{x}_k(t) - \dot{y}_k(t) - (\dot{x}_0 - \dot{y}_0)\| \le \max\{D\|x_0 - y_0\|, \|\dot{x}_0 - \dot{y}_0\|\} \left(e^{Dt} - 1 \right).$$

The induction hypothesis for the former equation and $k = 1$ is satisfied because of Eq. (29). The induction hypothesis for the latter equation and $k = 1$ is satisfied because using Eq. (32)

$$\|\dot{x}_1(t) - \dot{y}_1(t) - (\dot{x}_0 - \dot{y}_0)\| \leq \int_0^t A\|x_0 - y_0\| + B\|\dot{x}_0 - \dot{y}_0\| dt$$
$$\leq \max\{D\|x_0 - y_0\|, \|\dot{x}_0 - \dot{y}_0\|\} Dt.$$

Taking the limit $k \to \infty$ we obtain

$$\|x(t) - y(t) - (x_0 - y_0) - (\dot{x}_0 - \dot{y}_0)t\| \leq \max\{D\|x_0 - y_0\|, \|\dot{x}_0 - \dot{y}_0\|\}$$
$$\times \left(\frac{e^{Dt} - 1}{D} - t\right), \qquad (36)$$

$$\|\dot{x}(t) - \dot{y}(t) - (\dot{x}_0 - \dot{y}_0)\| \leq \max\{D\|x_0 - y_0\|, \|\dot{x}_0 - \dot{y}_0\|\} \left(e^{Dt} - 1\right), \qquad (37)$$

Let us disregard for the moment the last inequality. We have the trivial inequality

$$\|x_0 - y_0 - (x_0 - y_0)\| \leq \max\{D\|x_0 - y_0\|, \|\dot{x}_0 - \dot{y}_0\|\}(e^{Dt} - 1 - Dt). \qquad (38)$$

On $X = \mathbb{R}^{2n}$ let us consider a function $f : X \to X$ defined as follows

$$f(x_0, \dot{x}_0) = (x_0, x(1)).$$

Clearly, f is the coordinate expression of the exponential map. Let $L : X \to X$ be the linear map given by the matrix

$$L = \begin{pmatrix} I & 0 \\ I & I \end{pmatrix} \qquad (39)$$

with I the $n \times n$ identity matrix. Recalling that $D \leq 1$ the inequalities (38) and (36) can be rewritten for $t = 1$

$$\|f(x_0, \dot{x}_0) - f(y_0, \dot{y}_0) - L((x_0, \dot{x}_0) - (y_0, \dot{y}_0))\|$$
$$\leq \|(x_0, \dot{x}_0) - (y_0, \dot{y}_0)\| \left(\frac{e^D - 1}{D} - 1\right)$$

for every $(x_0, \dot{x}_0)$ and $(y_0, \dot{y}_0)$ such that $\|(x_0, \dot{x}_0)\| < \delta$ and $\|(y_0, \dot{y}_0)\| < \delta$.

But Eq. (33) shows that D as a function of δ satisfies $\lim_{\delta\to0} D(\delta) = 0$, thus f is strongly differentiable at $(0, 0)$ with strong differential L. Since any point $p \in M$ corresponds to zero coordinates for some chart compatible with the atlas, we have that exp is differentiable at any point in the image of $p \mapsto 0_p$.

Let us observe that if a strongly differentiable function $f(x, y)$ is strongly differentiable then keeping x fixed we obtain a function $f(x, \cdot)$ which is strongly differentiable. This fact follows easily from the definition of strong differentiability.

Furthermore, the composition of strongly differentiable functions is strongly differentiable. As a consequence, the pointed exponential map $\exp_p := \pi_2(f(0, \cdot))$ is strongly differentiable at the origin. From Eq. (39) we read that the Jacobian is given by the $n \times n$ identity matrix. Since it is invertible $\exp_p$ establishes a local Lipeomorphism.

2.3.1 Normal vector bundle case (Theorem 13)

Let us prove Theorem 13. Let $p \in \phi(S)$, let $\{s^k; k = 1, \ldots, l\}$ be coordinates on S at $\phi^{-1}(p)$ and let $x_0^\mu(s) := \phi^\mu(s)$. The tangent vectors $(\partial_k x_0^\mu)\partial_\mu, k = 1, \ldots, l$, provide a (Lipschitz) base for the tangent space at any point in a neighborhood of p. Applying the Gram-Schmidt procedure to

$$\left(\partial_1 x_0^\mu\right) \partial_\mu, \ldots, \left(\partial_l x_0^\mu\right) \partial_\mu, \partial_1, \ldots, \partial_n,$$

discarding the last found l null vectors, and keeping the last $n - l$ non-trivial vectors, we are left with a Lipschitz base of the normal space. Call this base $e_k^\mu \partial_\mu, k = 1, \ldots, n - l$. By construction it is Lipschitz. Thanks to this base we can introduce a chart of coordinates $(s, y) \in \mathbb{R}^l \times \mathbb{R}^{n-l}$ over $\nu(S)$, so as to represent each $v \in \nu(S)$, $v = (x_0, \dot{x}_0)$, as follows

$$x_0^\mu(s, y) = x_0^\mu(s^i), \tag{40}$$
$$\dot{x}_0^\mu(s, y) = y^j e_j^\mu(s^i). \tag{41}$$

This system of equations gives the map between $\nu(S)$ and TM expressed through the respective coordinate charts. A naive calculation would suggest that the Jacobian of this transformation is given by the $(n + n) \times (l + (n - l))$ matrix

$$J = \begin{pmatrix} \partial_k x_0^\mu & 0 \\ y^j \partial_i e_j^\mu & e_j^\mu \end{pmatrix}.$$

However, since e_j^μ is just Lipschitz the $n \times l$ block matrix $y^j \partial_i e_j^\mu$ is not well-defined. Nevertheless, this expression suggests that J should be as given for $y = 0$, namely on $\phi(S)$, for in this case the ill defined block matrix vanishes.

Let us prove that for $y = 0$, J is the strong differential of the map $(s, y) \to (x_0, \dot{x}_0)$. Let $\epsilon > 0$. We need only to show that for y_1, y_2 sufficiently close to zero and for s_1, s_2 sufficiently close to s

$$\|x_0(y_2, s_2) - x_0(y_1, s_1) - (\partial_i x_0)(s_2^i - s_1^i)\| \leq \epsilon(\|s_2 - s_1\| + \|y_2 - y_1\|)$$

$$\|\dot{x}_0(y_2, s_2) - \dot{x}_0(y_1, s_1) - e_j(s)(y_2^j - y_1^j)\| \leq \epsilon(\|s_2 - s_1\| + \|y_2 - y_1\|)$$

The former inequality is a consequence of the fact that $x_0(s)$ is C^1 hence strongly differentiable. The latter inequality can be rewritten using Eq. (41)

$$\|y_2^j e_j(s_2) - y_1^j e_j(s_1) - e_j(s)(y_2^j - y_1^j)\| \leq \epsilon(\|s_2 - s_1\| + \|y_2 - y_1\|)$$

which is a consequence of the Lipschitzness of e_j^μ.

The map

$$(s, y) \rightarrow (x_0, \dot{x}_0) \rightarrow \exp(x_0, \dot{x}_0) \rightarrow \pi_2(\exp(x_0, \dot{x}_0))$$

being the composition of (1) a strongly differentiable map at $y = 0$, (2) a strongly differentiable map at $\dot{x}_0 = 0$, (3) the strongly differentiable projection map $(x, y) \rightarrow y$, and being such that $\dot{x}_0(s, 0) = 0$, is strongly differentiable at $y = 0$, with strong differential given by the $n \times n$ matrix $J_{\pi_2} L J = (\partial_k x_0^\mu, e_j^\mu)(s)$ which is invertible because its columns are linearly independent vectors. Thus the hypothesis of Leach's inverse function theorem are satisfied.

2.4 Convex neighborhoods (Theorem 4)

Let us prove that for sufficiently small δ, $\bar{B}(p, \delta)$ is reversible strictly convex normal.

By the strong differentiability of exp there is an open neighborhood $O \ni (p, p)$, $O \subset M \times M$, which is Lipeomorphic to an open neighborhood U of $0_p \in TM$. Let $\delta > 0$ be sufficiently small that $\bar{B}(p, \delta)^2 \subset O$ and δ satisfies Eqs. (27, 28). Let $q \in B(p, \delta)$, and let

$$\begin{aligned} N_q &= (\exp^{-1} B(p, \delta)^2) \cap \pi^{-1}(q) = [\exp^{-1}(\{q\} \times B(p, \delta))] \cap \pi^{-1}(q) \\ &= \exp_q^{-1} B(p, \delta). \end{aligned}$$

The first defining equality shows that this set is open in the topology of $\pi^{-1}(q) = T_q M$. Furthermore, $\exp_q |_{N_q} : N_q \rightarrow B(p, \delta)$ is injective because $\{q\} \times N_q \subset U$ and so is $\exp |_U$. It is surjective, for if $r \in B(p, \delta)$ then $(q, r) \in B(p, \delta)^2 \subset O$, thus there is some $v \in T_q M$, $v \in U$, such that $\exp_q v = r$. Finally, $\exp_q |_{N_q}$ is Lipschitz because so is $\exp |_U$, and $\exp_q^{-1} |_{B(p,\delta)}$ is Lipschitz because so is $\exp^{-1} |_{B(p,\delta)^2}$. Thus we have proved that for each $q \in B(p, \delta)$, there is an open set N_q such that $\exp_q : N_q \rightarrow B(p, \delta)$ is a Lipeomorphism. We stress that we have not yet shown that N_q is star-shaped.

Analogously, for sufficiently small δ, for each $q \in B(p, \delta)$, there is an open set $\tilde{N}_q (= \tilde{\exp}_q^{-1} B(p, \delta))$ such that $\tilde{\exp}_q : \tilde{N}_q \rightarrow B(p, \delta)$ is a Lipeomorphism.

Our choice of δ allows us to prove the strict convexity of $\bar{B}(p, \delta)$ for both the spray and the reverse spray, and hence that each N_q and $\tilde{N}_q$ are star-shaped. The key observation is that the continuous function defined on $\pi^{-1}(\bar{B}(p, \delta))$ ("·" is the

Euclidean scalar product in $\mathbb{R}^n$ induced by the chart)

$$z^{\pm}(x, v) := \|v\|^2 + x \cdot H(x, \pm v)$$

is positive if restricted to the unit tangent bundle $\bar{B}(p, \delta) \times S^{n-1}$. Indeed,

$$z^{\pm}(x, e) := 1 + x \cdot H(x, \pm e) \geq 1 - \delta M > 0$$

where the last inequality is a consequence of Eq. (27). Let us consider a geodesic segment contained in $\bar{B}(p, \delta)$. Any geodesic $x(t)$ is $C^{2,1}$ thus $\|x\|^2(t)$ is $C^{2,1}$ and

$$
\begin{aligned}
\frac{d^2\|x\|^2}{dt^2} &= 2 \left(\|\dot{x}\|^2 + x \cdot \frac{d^2 x}{dt^2} \right) = 2 \left(\|\dot{x}\|^2 + x \cdot H(x, \dot{x}) \right) \\
&= 2z^+ \left(x, \frac{\dot{x}}{\|\dot{x}\|} \right) \|\dot{x}\|^2 > 0,
\end{aligned}
$$

where we used the fact that by definition a geodesic is regular, i.e. $\dot{x} \neq 0$.

Analogously, if we consider a geodesic for the reverse spray

$$
\begin{aligned}
\frac{d^2\|x\|^2}{dt^2} &= 2 \left(\|\dot{x}\|^2 + x \cdot \frac{d^2 x}{dt^2} \right) = 2 \left(\|\dot{x}\|^2 + x \cdot H(x, -\dot{x}) \right) \\
&= 2z^- \left(x, \frac{\dot{x}}{\|\dot{x}\|} \right) \|\dot{x}\|^2 > 0.
\end{aligned}
$$

As a consequence $\|x\|^2(t)$ takes its maximum value at the boundary of its interval of definition, that is, in correspondence of the endpoints of the geodesic segment which, therefore, must be contained in $\bar{B}(p, \delta)$. Furthermore, if its endpoints are at the boundary of the ball then its interior points stay in $B(p, \delta)$ because the inequality is strict. Thus $C := \bar{B}(p, \delta)$ is reversible strictly convex normal. The fact that exp establishes a Lipeomorphism between an open subset of TC and $B(p, \delta)^2$ is immediate from the inclusion $B(p, \delta)^2 \subset O$. Analogously, the same result holds for $\tilde{\exp}$. The proof of Theorem 4 is complete.

2.5 Role of the coordinate affine structure and position vector

We shall need the following result on the behavior of the position vector on a convex neighborhood.

Theorem 17 *Under the assumptions of Theorem 4, for every $\epsilon > 0$ we have for any sufficiently small δ that $C = B(p, \delta)$ not only satisfies the conclusions of Theorem 4 but also that for every $q_1, q_2, q_1', q_2', q \in C$ we have, interpreting the minus sign as that given by the affine structure induced by the coordinate chart*

$$
\begin{aligned}
&\|[P(q_1', q_2') - (q_2' - q_1')] - [P(q_1, q_2) - (q_2 - q_1)]\| \leq \epsilon \max\{\|q_1' - q_1\|, \|q_2' - q_2\|\}, \\
&\|P(q, q_2) - P(q, q_1) - (q_2 - q_1)\| \leq \epsilon \|q_2 - q_1\|, \tag{42}
\end{aligned}
$$

$$\|P(q_1, q_2) - (q_2 - q_1)\| \leq \epsilon\|q_2 - q_1\|, \tag{43}$$
$$\|P(q_1, q_2)\| \leq \epsilon. \tag{44}$$

In the connection case, if the coordinate chart is chosen so as to make $\Gamma^{\mu}_{\alpha\beta}(p)$ vanish, then δ can be also chosen such that

$$\|P(q_1, q_2) + P(q_2, q_1)\| \leq \epsilon\|q_2 - q_1\|^2. \tag{45}$$

This section will be devoted to the proof of this result.

Our coordinate system $\{x^{\mu}\}$ in a neighborhood U of p induces an affine structure which allows us to compare tangent vectors at different points of U. Let us consider the geodesic $x(t)$ such that its initial condition $(x_0, \dot{x}_0)$ satisfy Eq. (26). We ask if, keeping x_0 fixed, the map $\dot{x}_0 \mapsto \dot{x}(1)$ is injective (observe that $\dot{x}(1) = P(x_0, x(1)) = P(x_0, \exp_{x_0}\dot{x}_0)$.

Lemma 3 *Let $\|(x_0, \dot{x}_0)\| \leq \delta$. For fixed base point x_0, the map $\dot{x}_0 \mapsto \dot{x}(1)$ is strongly differentiable at the origin, the strong differential being the $n \times n$ identity matrix I. The map $(x_0, \dot{x}_0) \mapsto (x_0, \dot{x}(1))$ is also strongly differentiable wherever $\dot{x}_0 = 0$, the strong differential being the matrix*

$$\begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}.$$

Thus for sufficiently small δ both maps are injective (and bi-Lipschitz) with inverse strongly differentiable wherever $\dot{x}(1) = 0$.

Proof Let us consider the exponential map of the vectors $\dot{x}_0$, $\dot{y}_0$ with base point x_0 (thus $y_0 = x_0$). By Eq. (37) setting $t = 1$

$$\|\dot{x}(1) - \dot{y}(1) - (\dot{x}_0 - \dot{y}_0)\| \leq \|\dot{x}_0 - \dot{y}_0\|(e^D - 1).$$

Since $D(\delta) \to 0$ for $\delta \to 0$, the map $\dot{x}_0 \mapsto \dot{x}(1)$ is strongly differentiable at the origin, with strong differential the identity matrix.

As for the map $(x_0, \dot{x}_0) \mapsto (x_0, \dot{x}(1))$ it suffices to include in the previous analysis the trivial inequality

$$\|x_0 - y_0 - (x_0 - y_0)\| \leq \|x_0 - y_0\|(e^D - 1),$$

and recall that on $\mathbb{R}^n \oplus \mathbb{R}^n$ we use the norm $\max\{\| \|, \| \|\}$. The last claim follows from Leach's inverse function theorem. □

Proposition 3 *Let $p \in M$, and let $\{x^{\mu}\}$ be a chart in a neighborhood U of p. Let $P(r, q)$ be the position vector of q with respect to r. There is an open convex neighborhood $C \ni p$, such that for every $q_2, q_1 \in C$, the map $(q_1, q_2) \mapsto P(q_1, q_2) - (q_2 - q_1)$, interpreted with the affine structure induced by $\{x^{\mu}\}$, is strongly differentiable on the diagonal $q_1 = q_2$, with zero strong differential.*

Proof Let C be a convex neighborhood of p such that exp establishes a Lipeomorphism between an open subset of TC and $C \times C$. Let $x^\mu : C \mapsto \mathbb{R}^n$ be $(C^{2,1})$ coordinates on C, such that $x^\mu(p) = 0$.

We proved that exp is strongly differentiable on the zero section of TC with differential L on a suitable trivialization [Eq. (39)], thus by Leach's inverse function theorem, $\exp^{-1}$ is strongly differentiable on the diagonal of $C \times C$ with differential L^{-1}. Stated in another way, the map in coordinates given by $(q_1, q_2) \mapsto (q_1, \exp_{q_1}^{-1} q_2)$ is strongly differentiable on the diagonal with differential L^{-1}.

Lemma 3 proves that the coordinate map $(q_1, v) \mapsto P(q_1, \exp_{q_1} v)$ is strongly differentiable at the origin with strong differential the identity , thus the coordinate map $(q_1, q_2) \mapsto (q_1, P(q_1, q_2))$ is strongly differentiable on the diagonal with strong differential

$$L^{-1} = \begin{pmatrix} I & 0 \\ -I & I \end{pmatrix}.$$

This is the same strong differential of the map $(q_1, q_2) \mapsto (q_1, q_2 - q_1)$ where the difference makes sense using the affine structure induced by the coordinate chart. Thus the map $(q_1, q_2) \mapsto (q_1, P(q_1, q_2) - (q_2 - q_1))$ has vanishing strong differential on the diagonal of $C \times C$ which implies that the map $(q_1, q_2) \mapsto P(q_1, q_2) - (q_2 - q_1)$ has vanishing strong differential on the diagonal of $C \times C$.   □

In particular the map $(q_1, q_2) \mapsto P(q_1, q_2) - (q_2 - q_1)$ is strongly differentiable at (p, p) thus for every $\epsilon > 0$ the constant $\delta > 0$ and hence the convex neighborhood $C = B(p, \delta)$ can be chosen such that for every $q_1, q_2, q_1', q_2' \in C$

$$\|[P(q_1', q_2') - (q_2' - q_1')] - [P(q_1, q_2) - (q_2 - q_1)]\| \leq \epsilon \max\{\|q_1' - q_1\|, \|q_2' - q_2\|\}.$$

Thus for every $q_1, q_2, q \in C$ (set $q_1' \to q, q_2' \to q_2, q_2 \to q_1, q_1 \to q$)

$$\|P(q, q_2) - P(q, q_1) - (q_2 - q_1)\| \leq \epsilon \|q_2 - q_1\|,$$

and (set $q = q_1$)

$$\|P(q_1, q_2) - (q_2 - q_1)\| \leq \epsilon \|q_2 - q_1\|. \tag{46}$$

Thus $\|P(q_1, q_2)\| \leq (1 + \epsilon)\|q_2 - q_1\|$ and the diameter of C can be chosen sufficiently small that $\|P(q_1, q_2)\| \leq \epsilon$.

Suppose now that the spray is a connection (hence reversible) and assume to have chosen the coordinate system in such a way that for every every normalized vector e, $H^\mu(p, e) = 0$. This is always possible through an invertible quadratic coordinate change, so that the new coordinate system is still $C^{2,1}$. By continuity C can be chosen sufficiently small that on C, $\|H\| := \sup_{x \in C} \sup_{\|e\|=1} \|H(x, e)\| \leq \epsilon$.

Let $x : [0, 1] \to C$, $x(0) = q_1', x(1) = q_1$, be a geodesic. Let $q_2, q_2' = x(t)$, $t \in [0, 1]$ then the above 4-points inequality gives

$$\|P(x(0), x(t)) - P(x(1), x(t)) - (x(1) - x(0))\| \leq \epsilon \|x(1) - x(0)\|.$$

We observe that $P(x(0), x(t)) = t\dot{x}(t)$ and $P(x(1), x(t)) = -(1 - t)\dot{x}(t)$ thus for every $t \in [0, 1]$

$$\|\dot{x}(t) - (x(1) - x(0))\| \leq \epsilon \|x(1) - x(0)\|.$$

We have $P(x(0), x(1)) = \dot{x}(1)$ and $P(x(1), x(0)) = -\dot{x}(0)$ thus setting $r(t) = \dot{x}(t) - (x(1) - x(0))$,

$$\|P(x(0), x(1)) + P(x(1), x(0))\| = \|\dot{x}(1) - \dot{x}(0)\| = \|\int_0^1 H(x(s), \dot{x}(s)) ds\|$$

$$\leq \int_0^1 \|H(x(s), \hat{\dot{x}}(s))\| \, \|\dot{x}(s)\|^2 ds \leq \epsilon \int_0^1 \|\dot{x}(s)\|^2 ds$$

$$\leq \epsilon [\int_0^1 \|r(s)\|^2 ds + \int_0^1 \|x(1) - x(0)\|^2 ds + 2[x(1) - x(0)] \cdot \int_0^1 r(s) ds]$$

$$\leq \epsilon(1 + \epsilon^2)\|x(1) - x(0)\|^2,$$

thus a redefinition of ϵ gives the the last inequality of Theorem 17.

2.6 Local lipeomorphisms (Theorems 3 and 13)

Let $\varphi(v, t) = \gamma_v'(t)$ be the geodesic flow on TM. We have, $\exp v = (\pi(v), \pi(\gamma_v(1)))$. If $\varphi(v, 1)$ is well defined then, by the continuity of the geodesic flow, so is $\varphi(w, 1)$ for w near v. Thus Ω is open and, analogously, so is Ω_p.

The map $\exp$ is locally Lipschitz wherever it is defined because π is locally Lipschitz and if $v \in \Omega$, $\varphi(v, 1)$ is Lipschitz for w near v by the local Lipschitzness of the geodesic flow (that is by the dependence on initial conditions of solutions to the geodesic equation, see Sect. 2.2). Analogously, $\exp_p$ is locally Lipschitz on Ω_p.

We have shown that for each p there is a convex normal relatively compact open neighborhood C_p, such that $\exp$ provides a Lipeomorphism between the star-shaped relatively compact open set $\exp^{-1} C_p^2$ and C_p^2.

Let $\{\exp^{-1} C_{p_i}^2\}$ be a locally finite covering of the image of the zero section Z of TM, and let $N = \cup_{p_i} \exp^{-1} C_{p_i}^2$. Observe that if $w \in N$ then $w \in \exp^{-1} C_{p_i}^2$ for some i and so it cannot be $\exp w = \pi(w)$ unless $w \in Z$ by the injectivity of $\exp$ on C_p. By construction for every compact set $K \subset M$, we have that $\pi^{-1}(K) \cap \bar{N}$ is compact.

We want to show that there is an open subset $E \subset N$, containing Z such that $\exp|_E$ is injective and hence a Lipeomorphism on its image. Let K_i, $K_i \subset \text{Int}K_{i+1} \subset M$, be a sequence of compact sets such that $\cup_i K_i = M$, and let $\bar{N}_j \subset N$, $\bar{N}_{j+1} \subset N_j$, be open neighborhoods of Z such that for every compact set K, $\pi^{-1}(K) \cap \bar{N}_j$ is compact and $\cap_j \bar{N}_j = Z$. Clearly $\exp$ is injective on Z. By induction, suppose that there is an increasing map σ defined on $\{1, 2, \ldots, k\}$ such that the map $\exp$ is injective on the set $(E_0 := Z)$

$$E_k := Z \cup \bigcup_{i=1}^{k} \pi^{-1}(\text{Int}K_i) \cap N_{\sigma(i)}.$$

and hence E_j for $j \leq k$. Then we can define $\sigma(k+1)$ such that the same property holds with k replaced by $k+1$. For if not we could find sequence $v_j, w_j \in E_k \cup (\pi^{-1}(\mathrm{Int}K_{k+1}) \cap N_j)$, $v_j \neq w_j$, such that for every j, $\exp v_j = \exp w_j$. On the first component the last equality reads $\pi(v_j) = \pi(w_j)$. However, by injectivity of exp on E_k, v_j, w_j do not both belong to E_k thus passing to a subsequence we can assume without loss of generality that $v_j \in \pi^{-1}(\mathrm{Int}K_{k+1}) \cap N_j$. Observe that v_j belongs to the compact set $\pi^{-1}(K_{k+1}) \cap \bar{N}$, and passing to the limit we obtain that up to subsequences v_j converges to

$$v \in \cap_j [\pi^{-1}(K_{k+1}) \cap \bar{N}_j] \subset \pi^{-1}(K_{k+1}) \cap Z.$$

In particular, $\pi(w_j) = \pi(v_j)$ converges to some point of K_{k+1}, thus for sufficiently large j, w_j is contained in a compact set $\pi^{-1}(K_{k+2}) \cap \bar{N}_1 \subset N$ and so converges up to subsequences to some vector $w \in N$. Using the continuity of π we obtain $\pi(v) = \pi(w)$. By the continuity of exp, $\exp v = \exp w$, and using $v \in Z$, $\pi(v) = \exp w$. Thus $\pi(w) = \exp w$, and since $w \in N$ we have by the observation above that $w \in Z$, thus $w = v = 0_{\pi(v)}$. Let C be a convex normal neighborhood of $r := \pi(v)$, $(v = 0_r)$ then $\exp^{-1} C^2$ is a neighborhood of $v \in TC$ and the vectors v_j and w_j for sufficiently large j enter it which contradicts the injectivity of exp on $\exp^{-1} C^2$. Finally, $E = \cup_j E_j$ gives the searched open set.

The proof in the normal bundle case is analogous. We have to start from a sequence $p_i \in \phi(S)$ such that C_{p_i} is a locally finite covering of $\phi(S)$, define $N = \cup_{p_i} \exp^{-1} C_{p_i} \subset \nu(S)$ and proceed as in [51, Prop. 26, Chap. 7] to prove the injectivity of $\exp_{\nu(S)}$ on an open subset $E \subset N$.

3 Proofs II: Pseudo-Finsler sprays and connections

In the next section we prove Gauss' Lemma for sprays which come from a pseudo-Finsler metric L.

3.1 Gauss' Lemma (Theorem 5)

Let us prove Eq. (13). Since $\exp_p^{-1}$ is Lipschitz, D_p^2 is Lipschitz. By Theorem 1 the function $2g_{P(p,q)}(P(p,q), \cdot)$ where $P(p,q) := \gamma'_{\exp_p^{-1} q}(1)$, is Lipschitz in q. By Theorem 16 we need only to show that the differential of D_p^2 exists and coincides with the previous expression almost everywhere on N.

However, we know that exp is differentiable almost everywhere over the star-shaped open set $\exp^{-1} N \subset T_p M$, and hence, by Fubini's theorem, that it is almost everywhere differentiable on almost every radial line passing through the origin (the expression a.e. here refers to the $(n-1)$-dimensional Lebesgue measure of a Euclidean sphere contained in $\exp^{-1} N$). It suffices to take q on the exponential map of one of these geodesics (Lipeomorphisms preserve zero measure sets [22, Sect. 2.4]).

Let $w \in T_q M$ and let $\sigma : [-a, a] \to N$ be a $(C^{2,1})$ geodesic segment such that $\sigma'(0) = w$. Let $x^{(s)} : [0, 1] \to N$, $t \mapsto x^{(s)}(t)$, be the unique geodesic such

that $x^{(s)}(0) = p, x^{(s)}(1) = \sigma(s)$. Let $v(s) = \exp_p^{-1}\sigma(s)$, since $\exp_p^{-1}$ is Lipschitz $v : [-a, a] \to T_p M$ is Lipschitz. In particular $S = \{u \in T_p M : u = tv(s), t \in [0, 1], s \in [-a, a]\}$ is a Lipschitz submanifold of $T_p M$, thus by Theorem 1 $x(t, s) := x^{(s)}(t) = \exp_p(tv(s))$ is Lipschitz in (t, s). Furthermore, still by Theorem 1

(*): for almost every s, for every $t \in [0, 1]$, $x(t, \cdot)$ is differentiable at s and the derivative $\partial_2 x(t, s)$ is locally Lipschitz in t, locally uniformly with respect to those s where it is defined. Finally, for any such s we have that for almost every $t \in [0, 1]$, the mixed partial derivatives $\partial_1\partial_2 x(t, s), \partial_2\partial_1 x(t, s)$ are locally bounded and coincide.

The quantity $L(x^{(s)}(t), x_t^{(s)}(t))$ is independent of t hence coincident with $L(p, v(s))$. The function $v(s)$ being Lipschitz is differentiable almost everywhere. Let s be such that $v'(s)$ exists and (*) holds

$$\frac{1}{2}\partial_{\sigma'(s)}D_p^2 = \frac{\mathrm{d}}{\mathrm{d}s}L(p, v(s)) = \frac{\partial}{\partial s}L(x(t, s), x_t(t, s)) = x_s^\mu\frac{\partial L}{\partial x^\mu} + \frac{\partial x_t^\mu}{\partial s}\frac{\partial L}{\partial x_t^\mu}$$
$$= \frac{\partial}{\partial t}\left(x_s^\mu\frac{\partial L}{\partial x_t^\mu}\right) - x_s^\mu\left(\frac{\partial}{\partial t}\frac{\partial L}{\partial x_t^\mu} - \frac{\partial L}{\partial x^\mu}\right) + \left(\frac{\partial x_t^\mu}{\partial s} - \frac{\partial x_s^\mu}{\partial t}\right)\frac{\partial L}{\partial x_t^\mu}$$

where t can be chosen arbitrarily. By (*) for almost every t the last term vanishes. Moreover, the second term of the right-hand side vanishes because $x^{(s)}(t)$ is a geodesic, hence it solves the Euler-Lagrange equation for the Lagrangian L. Integrating in t over the interval $[0, 1]$ we obtain, taking into account that the left-hand side does not depend on t and using Eq. (7)

$$\frac{1}{2}\partial_{\sigma'(s)}D_p^2 = g_{x_t(1,s)}(x_t(1, s), \sigma'(s)) - g_{x_t(0,s)}(x_t(0, s), x_s(0, s))$$
$$= g_{\gamma_{v(s)}'(1)}(\gamma_{v(s)}'(1), \sigma'(s)), \tag{47}$$

where we used the fact that since $x(0, s) = p$, it is $x_s(0, s) = 0$. Evaluated at $s = 0$ the previous expression proves Eq. (13) and hence the first claim of the theorem.

Let $\alpha: [0, a) \to N, t \mapsto \alpha(t)$, be an integral curve of P. Let us differentiate D_p^2 along it

$$\frac{\mathrm{d}D_p^2(\alpha(t))}{\mathrm{d}t} = 2g_{P(\alpha(t))}(P(\alpha(t)), \dot\alpha(t)) = 2g_{P(\alpha(t))}(P(\alpha(t)), P(\alpha(t))) = 2D_p^2(\alpha(t)).$$

Then $D_p^2(\alpha(t)) = D_p^2(\alpha(0))\exp(2t)$, and since $\exp^{-1} N$ is star-shaped the $-t$-time flow maps $(D_p^2)^{-1}(s)$ to $(D_p^2)^{-1}(se^{-2t})$ for $t > 0$. Since P is Lipschitz, by the mentioned result on the dependence of solutions to first order ODE on the initial conditions, the flow map is Lipschitz, and since it is injective and can be inverted it is actually a Lipeomorphism between $(D_p^2)^{-1}(s)$ and its image on $(D_p^2)^{-1}(se^{-2t})$.

Finally, suppose that $\exp_p$ is differentiable at $v \in T_p M\backslash 0$. We observe that $\gamma_v'(1) = (d\exp_p)_v v$. Let $w \in T_v(T_p M) \sim T_p M$ and define $v(s) := v + sw$ and $\sigma(s) := \exp_p(v + sw)$, so that by the differentiability assumption $\sigma'(s) = (d\exp_p)_v w$. Due to

the choice of curve $\sigma(s)$ we have $D_p^2(\sigma(s)) = 2L(p, \exp_p^{-1}(\sigma(s))) = 2L(p, v+sw)$ thus from Eq. (7)

$$\frac{1}{2}\frac{\mathrm{d}D_p^2(\sigma(s))}{\mathrm{d}s}|_{s=0} = g_v(v, w)$$

while from Eq. (47)

$$\frac{1}{2}\frac{\mathrm{d}D_p^2(\sigma(s))}{\mathrm{d}s}|_{s=0} = g_{(d \exp_p)_v v}((d \exp_p)_v v, (d \exp_p)_v w).$$

3.2 Local properties of geodesics in pseudo-Finsler geometry (Theorem 6)

Let $\sigma: [0, 1] \to N$ be an AC-curve starting from p and ending at $q \in N$.

Let us consider the Finsler case. By continuity there is a last value $\hat{s} \geq 0$, such that $D_p(\sigma(\hat{s})) = 0$ (possibly $\hat{s} = 0$ or $\hat{s} = 1$).

The length of the geodesic connecting p to q is $D_p(q)$ and is positive if and only if $q \neq p$. The statement of the theorem is trivial for $q = p$ thus let us assume $q \neq p$ (so that $\hat{s} \neq 1$). Since D_p^2 is $C^{1,1}$, D_p is $C^{1,1}$ in the region $N\backslash\{p\}$. Thus $D_p(\sigma(s))$ being the composition of a locally Lipschitz and an absolutely continuous function is absolutely continuous. We have for $s \geq \hat{s}$

$$D_p(\sigma(s)) = \int_{\hat{s}}^s \frac{\mathrm{d}D_p(\sigma(s))}{\mathrm{d}s}\,\mathrm{d}s = \int_{\hat{s}}^s \frac{1}{D_p(\sigma(s))}\,g_{P(p,\sigma(s))}(P(p, \sigma(s)), \sigma'(s))\,\mathrm{d}s$$

$$= \int_{\hat{s}}^s g_{\hat{P}(p,\sigma(s))}(\hat{P}(p, \sigma(s)), \sigma'(s))\,\mathrm{d}s$$

$$\leq \int_{\hat{s}}^s \sqrt{g_{\sigma'(s)}(\sigma'(s), \sigma'(s))}\,\mathrm{d}s \leq l[\sigma],$$

where $\hat{P} := P/\sqrt{g_P(P, P)}$. In the last step we used the analog of the Cauchy-Schwarz inequality for Finsler geometry [5, Theor. 1.2.2]. For $s = 1$ the above inequality proves that the length of σ is no smaller than that of the geodesic connecting its endpoints. If they are equal then for almost every $s \in [0, 1]$, we have the equality $g_{\hat{P}(p,\sigma(s))}(\hat{P}(p, \sigma(s)), \sigma'(s)) = \sqrt{g_{\sigma'(s)}(\sigma'(s), \sigma'(s))}$, thus, by the equality case in the Finslerian Cauchy-Schwarz inequality we have that for almost every $s \in [0, 1]$, $\sigma' \propto P(p, \sigma(s))$. If we introduce spherical normal coordinates $(r, \theta_1, ..., \theta_{n-1})$, as this coordinate chart is Lipschitz related to those of M, σ is still absolutely continuous in this chart. Thus since $\sigma' \propto \partial_r$ almost everywhere, the angular coordinates cannot change over σ, for otherwise since $\theta_i(\sigma(s))$ is the integral of its own derivative one would gets that σ' is not radial in a set of non-vanishing measure, a contradiction. Thus the image of σ coincides with the image of an integral curve of P and hence coincides with the image of the geodesic $\eta(r)$ connecting $p = \sigma(0)$ with $q = \sigma(1)$. Since the coordinates of the spherical normal chart are Lipschitz functions, the composition $r(s)$ is absolutely continuous. By definition r is an affine parameter over the geodesic which has the same image of σ. The map is necessarily increasing, for if $r(s_2) \leq r(s_1)$

for $s_1 < s_2$, then we would have $r' < 0$ (by definition of AC-curve $r' \neq 0$ almost everywhere) in a subset of measure different from zero on $[s_1, s_2]$, and it would be easy to obtain a shorter curve cutting a piece of domain from σ, a contradiction to the length minimization assumption.

Let us consider now the Lorentzian–Finsler case with σ causal and future directed. We recall that the analog to the reverse Cauchy–Schwarz inequality for Finsler space-times reads [47]:

Let v_1, v_2 be causal and future directed then

$$- g_{v_1}(v_1, v_2) \geq \sqrt{-g_{v_1}(v_1, v_1)} \sqrt{-g_{v_2}(v_2, v_2)}, \qquad (48)$$

with equality if and only if v_1 and v_2 are proportional.

Suppose that for some $\tilde{s}$, $D_p^2(\tilde{s}) < 0$. The Lorentzian–Finsler length of the geodesic connecting p to q is: $D_p^L(q) := (-D_p^2(q))^{1/2}$. Since D_p^2 is $C^{1,1}$, D_p^L is $C^{1,1}$ in the region $D_p^2 < 0$. Thus $D^L(\sigma(s))$ being the composition of a locally Lipschitz and an absolutely continuous function is absolutely continuous. We know that $D_p^2(\tilde{s}) < 0$ and by continuity the same inequality holds in an interval $[\tilde{s}, s]$ provided s is sufficiently close to $\tilde{s}$. We have

$$
\begin{aligned}
D_p^L(\sigma(s)) - D_p^L(\sigma(\tilde{s})) &= \int_{\tilde{s}}^{s} \frac{\mathrm{d}D_p^L(\sigma(s))}{\mathrm{d}s}\, \mathrm{d}s \\
&= - \int_{\tilde{s}}^{s} \frac{1}{D_p^L(\sigma(t))}\, g_{P(p,\sigma(s))}(P(p,\sigma(s)), \sigma'(s))\, \mathrm{d}s \\
&= - \int_{\tilde{s}}^{s} g_{\hat{P}(p,\sigma(s))}(\hat{P}(p,\sigma(s)), \sigma'(s))\, \mathrm{d}s \geq \int_{\hat{s}}^{s} \sqrt{-g_{\sigma'(s)}(\sigma'(s), \sigma'(s))}\, \mathrm{d}s \\
&\geq l[\sigma], \qquad (49)
\end{aligned}
$$

where $\hat{P} := P/\sqrt{-g_P(P, P)}$. In the last inequality we used the above Finslerian reverse Cauchy–Schwarz inequality.

The equality so obtained proves that once σ enters a region with $D_p^2 < 0$ (the chronological future of p) it remains in that region.

Now let $\eta \colon [-\epsilon, 0] \to N$, $\eta(0) = p$, be a small future directed timelike geodesic contained in a reversible convex normal neighborhood of p, $C \subset N$. For sufficiently small s, $\sigma(s) \in C$, and the curve obtained concatenating η with σ which connects $\eta(-\epsilon)$ to $\sigma(s)$ starts with a timelike geodesic, hence it enters the chronological future of $\eta(-\epsilon)$, and hence, by the above argument there is a future directed timelike geodesic $\nu^{(\epsilon)}$ connecting $\eta(-\epsilon)$ with $\sigma(s)$. Letting $\epsilon \to 0$, and using the continuity of the exponential map $\tilde{\exp}$ for the reverse spray at $\sigma(s)$ we infer the existence of a geodesic connecting p to $\sigma(s)$, which by the continuity $g_v(v, v)$ at $T_{\sigma(s)}M$ must be future directed causal. As s is arbitrary we have shown that in a maximal closed interval $[0, b] \subset [0, 1]$, $b > 0$, we have $D_p^2(\sigma(s)) \leq 0$.

Let us prove that if for $a \in (0, b]$, $D_p^2(\sigma(a)) = 0$ then $\sigma|_{[0,a]}$ is a lightlike geodesic up to parametrizations and hence that $D_p^2 = 0$ over $[0, a]$.

Observe that D_p^2 is Lipschitz thus $D_p^2(\sigma(s))$ is absolutely continuous

$$D_p^2(\sigma(a)) = \int_0^a \frac{\mathrm{d} D_p^2(\sigma(s))}{\mathrm{d} s} \, \mathrm{d} s = 2 \int_0^a g_{P(p,\sigma(s))}(P(p,\sigma(s)), \sigma'(s)) \, \mathrm{d} s.$$

Since on the region $D_p^2 \leq 0$, we have $g_{P(p,\sigma(s))}(P(p,\sigma(s)), \sigma'(s)) \leq 0$ for almost every s (by the Finslerian reverse Cauchy-Schwarz inequality since σ' is future directed causal almost everywhere), thus we can have $D_p^2(\sigma(a)) = 0$ only if $\sigma' \propto P$ for almost every s in $[0, a]$. Introducing a Euclidean scalar product on $T_p M$, associated spherical normal coordinates over N, and arguing as above for the Finsler case we obtain that $\sigma|_{[0,a]}$ is an integral curve of P, hence a lightlike geodesic issued from p.

From now on let a be the maximum value of s for which $D_p^2(\sigma(s)) = 0$.

It remains only to prove that $b = 1$. Suppose not then $a = b$ otherwise $D_p^2(b) < 0$ which would imply the same inequality also in $(b, 1]$, a contradiction to $b < 1$. Set $p' = \sigma(b)$ and take a reversible convex normal neighborhood $C' \ni p'$, $C' \subset N$. Arguing as above proves that for any sufficiently small δ, p' is connected to $\sigma(b + \delta)$ by a future directed causal geodesic $\eta : [0, 1] \to C'$. This geodesic cannot be the prolongation of the lightlike geodesic $\sigma_{[0,b]}$ for we would get $D_p^2(\sigma(b + \alpha\delta)) \leq 0, \alpha \in [0, 1]$, a contradiction to the maximality of b. Thus the scalar product $g_{P(p,\eta(t))}(P(p,\eta(t)), \eta'(t))$ is negative for $t = 0$ and hence in a neighborhood of $t = 0$. Now observe that D_p^2 is Lipschitz thus $D_p^2(\eta(t))$ is absolutely continuous, and for sufficiently small t

$$\begin{aligned} D_p^2(\eta(t)) &= D_p^2(\sigma(b)) + \int_0^t \frac{\mathrm{d} D_p^2(\eta(t))}{\mathrm{d} t} \, \mathrm{d} t \\ &= 2 \int_0^t g_{P(p,\eta(t))}(P(p,\eta(t)), \eta'(t)) \, \mathrm{d} t < 0. \end{aligned}$$

As the concatenation of $\sigma_{|[0,b]}$ with η is a causal AC-curve and on it D_p^2 becomes negative at some point, and it remains so, we have at the endpoint $D_p^2(\sigma(b + \delta)) = D_p^2(\eta(1)) < 0$. As δ is arbitrarily we get a contradiction to the maximality of b. The contradiction proves that $b = 1$.

If σ is a lightlike geodesic up to parametrization, then clearly its Lorentzian-Finsler length vanishes and the inequality $D_p^L(\sigma(1)) \geq l(\sigma)$ is satisfied. Suppose that σ is not a lightlike geodesic up to parametrizations then $a < 1$, and its Lorentzian-Finsler length is given just by the contribution of $\sigma_{[a,1]}$. Let $\tilde{s} \in [a, 1]$ so that $D_p^2(\sigma(\tilde{s})) < 0$. By (49)

$$D_p^L(\sigma(1)) \geq l(\sigma_{[\tilde{s},1]})$$

and taking the limit $\tilde{s} \to a$ we obtain $D_p^L(\sigma(1)) \geq l(\sigma)$. This proves that σ has a Lorentzian-Finsler length no larger than that of the geodesic connecting its endpoints.

Now, suppose by contradiction that they have the same Lorentzian-Finsler length and that σ is not a causal geodesic up to parametrizations. Then necessarily $a < 1$,

for otherwise it would be a lightlike geodesic. But then from (49), for $\tilde{s} > a$,

$$D_p^L(\sigma(1)) \geq D_p^L(\sigma(\tilde{s}) + l(\sigma_{[\tilde{s},1]}) \geq l(\sigma_{[0,\tilde{s}]}) + l(\sigma_{[\tilde{s},1]}) = l(\sigma).$$

Thus the equality implies that the first inequality is actually an equality which implies that $g_{P(p,\sigma(s))}(P(p,\sigma(s)), \sigma'(s)) = 0$ for almost every $s \in [\tilde{s}, 1]$, and hence, by the arbitrariness of $\tilde{s}$, $\sigma' \propto P$ for almost every $s \in [a, 1]$. Introducing again spherical normal coordinates and arguing as above proves that the image of $\sigma|_{[a,1]}$ is an integral curve of P (and hence the prolongation of $\sigma|_{[0,a]}$ if $a \neq 0$) thus it is the image of a geodesic.

Finally, suppose that the image of σ coincides with that of a causal geodesic η. Since the coordinates of the spherical normal chart are Lipschitz functions, the composition $r(s)$ is absolutely continuous. By definition r is an affine parameter over the geodesic η. The map $r(s)$ is necessarily increasing, for if $r(s_2) \leq r(s_1)$ for $s_1 < s_2$, then we would have $r' \leq 0$ and hence in a subset of measure different from zero on $[s_1, s_2]$, which would imply that $\frac{d}{ds}\sigma = (\frac{d}{dr}\eta)r'$ is not future directed causal in a set of measure different from zero, a contradiction to the definition of future directed causal AC-curve.

3.3 Strong convexity of squared Riemannian distance (Theorem 8)

Let C be a convex neighborhood of p as in Theorems 4 and 17 where $\epsilon \in (0, 1)$ and where the coordinate chart is chosen so that $\Gamma_{\alpha\beta}^{\mu}(p) = 0$.

Let us consider the Riemannian case. From Eq. (42) using the the Cauchy-Schwarz inequality

$$|P(q, q_2) \cdot (q_2 - q_1) - P(q, q_1) \cdot (q_2 - q_1) - (q_2 - q_1)^2| \leq \epsilon \|q_2 - q_1\|^2. \quad (50)$$

Let g_q be the matrix of g at q. Since g is C^1 it is strongly differentiable and by the choice of coordinate system its strong derivative vanishes at p. Thus C can be chosen sufficiently small that for every $q_1, q_2 \in C$

$$\|g_{q_2} - g_{q_1}\| \leq \epsilon \|q_2 - q_1\|$$

Moreover, C can be chosen sufficiently small that once expressed in the coordinate chart $\|g - I\| \leq \epsilon$ on C.

Thus if D_q^2 is the squared distance function from q, using also Eqs. (42), (43) and (44) we are able to prove Eq. (14)

$$\frac{1}{2}\left|\left[dD_q^2(q_2) - dD_q^2(q_1)\right](q_2 - q_1) - 2(q_2 - q_1)^2\right|$$
$$= |g_{q_2}(P(q, q_2), (q_2 - q_1)) - g_{q_1}(P(q, q_1), (q_2 - q_1)) - (q_2 - q_1)^2|$$
$$= |g_{q_2}(P(q, q_2) - P(q, q_1), (q_2 - q_1)) - (g_{q_1} - g_{q_2})(P(q, q_1),$$
$$\times (q_2 - q_1)) - (q_2 - q_1)^2|$$
$$\leq |[P(q, q_2) - P(q, q_1)] \cdot (q_2 - q_1) - (q_2 - q_1)^2| + \epsilon \|P(q, q_2) - P(q, q_1)\|$$

$$\|q_2 - q_1\| + \epsilon \|P(q, q_2)\| \, \|q_2 - q_1\|^2$$
$$\leq 2\epsilon(1 + \epsilon)\|q_2 - q_1\|^2,$$

which proves that D_q^2 is strongly convex with respect to the affine structure induced by the coordinate chart.

Let us give a geodesic version. This time we shall need to use Eq. (45). Also observe that from Eq. (43) we have

$$\|P(q_1, q_2) - (q_2 - q_1)\| \leq \epsilon \|q_2 - q_1\| \leq \epsilon \|P(q_1, q_2) - (q_2 - q_1)\| + \epsilon \|P(q_1, q_2)\|,$$

and hence

$$\|P(q_1, q_2) - (q_2 - q_1)\| \leq \frac{\epsilon}{1 - \epsilon} \|P(q_1, q_2)\|,$$

and

$$\|q_2 - q_1\| \leq \frac{1}{1 - \epsilon} \|P(q_1, q_2)\|. \tag{51}$$

Similarly, let g be a metric such that at any point of C, $\|g - I\| \leq \epsilon$, and let v be any vector. Then

$$|(g - I)(v, v)| \leq \epsilon v \cdot v = \epsilon|(I - g)(v, v) + g(v, v)| \leq \epsilon|(g - I)(v, v)| + \epsilon g(v, v),$$

from which we obtain

$$|(g - I)(v, v)| \leq \frac{\epsilon}{1 - \epsilon} \, |g(v, v)|,$$

and

$$|v \cdot v| \leq |(I - g)(v, v)| + |g(v, v)| \leq \frac{1}{1 - \epsilon} \, |g(v, v)|.$$

Let $x : [0, 1] \to C$ be a geodesic and let $q_1 := x(0)$, $q_2 := x(1)$. We are ready to prove Eq. (15).

$$\frac{1}{2}|\frac{\mathrm{d}}{\mathrm{d}t}D(q, x(t))^2|_{t=1} - \frac{\mathrm{d}}{\mathrm{d}t}D(q, x(t))^2|_{t=0} - 2D(q_1, q_2)^2|$$
$$= |g_{q_2}(P(q, q_2), P(q_1, q_2)) + g_{q_1}(P(q, q_1), P(q_2, q_1))$$
$$\quad - g_{q_1}(P(q_2, q_1), P(q_2, q_1))|$$
$$\leq |g_{q_1}(P(q, q_2), P(q_1, q_2)) + g_{q_1}(P(q, q_1), P(q_2, q_1))$$
$$\quad - g_{q_1}(P(q_2, q_1), P(q_2, q_1))| + \epsilon^2(1 + \epsilon)\|q_2 - q_1\|^2$$
$$\leq |g_{q_1}(P(q, q_2), -P(q_2, q_1)) + g_{q_1}(P(q, q_1), P(q_2, q_1))$$
$$\quad - g_{q_1}(P(q_2, q_1), P(q_2, q_1))| + 2(1 + \epsilon)\epsilon^2\|q_2 - q_1\|^2$$
$$\leq |g_{q_1}(P(q, q_1) - P(q, q_2), P(q_2, q_1)) - g_{q_1}(q_1 - q_2, P(q_2, q_1))|$$

$$+ (2(1+\epsilon)\epsilon^2 + \epsilon(1+\epsilon)^2)\|q_2 - q_1\|^2$$
$$\leq |g_{q_1}(P(q, q_1) - P(q, q_2) - (q_1 - q_2), P(q_2, q_1))|$$
$$+ (2(1+\epsilon)\epsilon^2 + \epsilon(1+\epsilon)^2)\|q_2 - q_1\|^2$$
$$\leq 2\epsilon(1+\epsilon)(1+2\epsilon)\|q_2 - q_1\|^2 \leq 2\epsilon \frac{(1+\epsilon)(1+2\epsilon)}{(1-\epsilon)^2}\|P(q_1, q_2)\|^2$$
$$\leq 2\epsilon \frac{(1+\epsilon)(1+2\epsilon)}{(1-\epsilon)^3} g_{q_2}(P(q_1, q_2), P(q_1, q_2)) \leq 2\epsilon \frac{(1+\epsilon)(1+2\epsilon)}{(1-\epsilon)^3} D(q_1, q_2)^2.$$

A reparametrization of x with arc-length and a redefinition of ϵ gives Eq. (15).

The statement of Theorem 8 concerning the strong convexity of D_q^2 is immediate from the triangle inequality and from the equivalences recalled in Sect. 1.5.

3.4 Splitting of the metric at a given radius (Theorem 9)

On the coordinate ball let us introduce radial coordinates $(\rho, \theta_1, \ldots, \theta_{n-1})$. Each ball $(D_p)^{-1}([0, r])$ is convex with respect to the affine structure induced by the coordinate chart $(x_1, \ldots, x_n)$, thus the radial lines issued from p intersect the boundary of the ball only once. Let $r = D_p$, there is therefore a function $\rho(r, \theta)$ establishing the dependence of the radial coordinate on the angular ones. We known that $r(q(\rho, \theta))$ is $C^{1,1}$, and Eq. (43) and $\|g - I\| \leq \epsilon$ imply that $g(P(p, q), (q - p)) \neq 0$, namely $\partial r/\partial \rho \neq 0$. By the usual implicit function theorem $\rho(r, \theta)$ is $C^{1,1}$, thus the components of g in coordinates $(r, \theta_1, \ldots, \theta_{n-1})$ are locally Lipschitz. In particular, for any given $r > 0$ the map which sends S^{n-1} (i.e. θ) to $D_p^{-1}(r)$ is $C^{1,1}$ thus differentiable.

Taking $r = cnst.$ in $g_{ij}(r, \theta)$ shows that the metric induced on each hypersurface $D_p^{-1}(r)$ is locally Lipschitz.

The function r is $C^{1,1}$ and by Theorem 5 ∇r is the normalized geodesic field orthogonal to the level sets of constant r. Thus $g^{-1}(\mathrm{d}r, \mathrm{d}r) = g(\nabla r, \nabla r) = 1$ and Eq. (16)

$$g^{-1} = (\partial_r + A_i(r, \theta)\partial_i)^2 + (h_r^{-1})_{ij}\partial_i \otimes \partial_j$$

holds for some Lipschitz components A_i, $(h_r^{-1})_{ij}$. Its inverse is

$$g = \mathrm{d}r^2 + (h_r)_{ij}(\mathrm{d}\theta_i - A_i \mathrm{d}r)(\mathrm{d}\theta_j - A_j \mathrm{d}r),$$

thus h_r is the metric induced on the level set $D_p^{-1}(r)$.

Let us fix $\bar{r}$ so that $D_p^{-1}(\bar{r}) \subset C$. We know that $\{\theta_i\}$ provides a $C^{1,1}$ chart over $D_p^{-1}(\bar{r})$. Let us rename these coordinates $\{\alpha_i\}$ and let us extend them in a neighborhood of $D_p^{-1}(\bar{r})$ solving the differential equation $\partial\theta_i/\partial r = A_i(r, \theta)$, $\theta_i(\bar{r}, \alpha) = \alpha_i$. Since A_i is Lipschitz the solution $\theta(r, \alpha)$ is Lipschitz in α and $C^{1,1}$ in r [31, Ex. 1.2, Chap. 2] [13, Prop. 1.10.1] [38, Cor. 1.6]. Thus g can be brought to a direct sum form almost everywhere where the Jacobian $\partial\theta/\partial\alpha$ exists. This Jacobian exists at $r = \bar{r}$ and equals

the identity matrix, thus

$$h'_{ij}(\bar{r}, \alpha) = h_{ks}(\bar{r}, \theta(\bar{r}, \alpha)) J_i^k J_j^s = h_{ij}(\bar{r}, \theta(\bar{r}, \alpha)) = h_{ij}(\bar{r}, \alpha).$$

This equality proves that the components $h'_{ij}(\bar{r}, \cdot)$ exist and are Lipschitz in α.

Of course, since the geodesic flow is Lipschitz we could deduce immediately that normal spherical coordinates are Lipschitz and hence that g can be brought to a direct sum form almost everywhere, but we wanted to construct a coordinate system for which the direct sum form was valid everywhere at a given radius.

3.5 Some local results on the strong concavity of the squared Lorentzian distance (Theorems 10, 11, Corollary 2)

Proof (Proof of Theorem 10) Let C be a convex neighborhood of p as in Theorems 4 and 17 where $\epsilon \in (0, 1/3)$ and the coordinate system is such that $g_{\alpha\beta}(p) = \eta_{\alpha\beta}$, $\Gamma^\mu_{\alpha\beta}(p) = 0$. The metric defined by

$$r(v_1, v_2) = g(v_1, v_2) + 2(\partial_0 \cdot v_1)(\partial_0 \cdot v_2) \tag{52}$$

is positive definite at p and hence C can be chosen sufficiently small that it is positive definite everywhere in C. Observe that at p the metric r once expressed in components coincides with the identity matrix. Thus C can be chosen sufficiently small that

$$\|r - I\| \le \epsilon.$$

From Eqs. (42, 43)

$$\|P(q, q_1) - P(q, q_2) - P(q_2, q_1)\| \le 2\epsilon \|q_2 - q_1\|. \tag{53}$$

Let $x : [0, 1] \to C$ be a geodesic, let $q_1 := x(0)$ and $q_2 := x(1)$, and let

$$V = P(q, q_1) - P(q, q_2) - P(q_2, q_1).$$

We have

$$\begin{aligned}
|g_{q_1}(V, P(q_2, q_1))| &= |r(V, P(q_2, q_1)) - 2(\partial_0 \cdot V)(\partial_0 \cdot P(q_2, q_1))| \\
&\le \|V\| \|P(q_2, q_1)\| [\|r\| + 2]] \le \|V\| \|P(q_2, q_1)\| (3 + \epsilon) \\
&\le 2\epsilon(1 + \epsilon)(3 + \epsilon) \|q_2 - q_1\|^2
\end{aligned}$$

Since g_q is $C^{1,1}$ in q it is strongly differentiable at p with zero strong differential, thus C can be chosen sufficiently small that for every $q_1, q_2 \in C$,

$$\|g_{q_1}\| \le 1 + \epsilon, \qquad \|g_{q_2} - g_{q_1}\| \le \epsilon \|q_2 - q_1\|. \tag{54}$$

Thus

$$
\frac{1}{2}\left|\frac{\mathrm{d}}{\mathrm{d}t}D_q^2(x(t))\right|_{t=1} - \frac{\mathrm{d}}{\mathrm{d}t}D_q^2(x(t))|_{t=0} - 2D_{q_1}^2(q_2)|
$$

$$
= |g_{q_2}(P(q,q_2), P(q_1,q_2)) + g_{q_1}(P(q,q_1), P(q_2,q_1))
$$
$$
- g_{q_1}(P(q_2,q_1), P(q_2,q_1))|
$$
$$
\leq |g_{q_1}(P(q,q_2), P(q_1,q_2)) + g_{q_1}(P(q,q_1), P(q_2,q_1))
$$
$$
- g_{q_1}(P(q_2,q_1), P(q_2,q_1))| + \epsilon^2(1+\epsilon)\|q_2 - q_1\|^2
$$
$$
\leq |g_{q_1}(P(q,q_2), -P(q_2,q_1)) + g_{q_1}(P(q,q_1), P(q_2,q_1))
$$
$$
- g_{q_1}(P(q_2,q_1), P(q_2,q_1))| + 2(1+\epsilon)\epsilon^2\|q_2 - q_1\|^2
$$
$$
\leq |g_{q_1}(V, P(q_2,q_1))| + 2(1+\epsilon)\epsilon^2\|q_2 - q_1\|^2
$$
$$
\leq 2\epsilon(1+\epsilon)(3+2\epsilon)\|q_2 - q_1\|^2
$$

A redefinition of ϵ proves Eq. (17).   □

Proof (Proof of Theorem 11) It is sufficient to prove Eq. (18) for $\epsilon < 1/9$. Let us parametrize γ with respect to g-arc length (i.e. proper time), $g(\dot\gamma, \dot\gamma) = -1$. As a first step let us introduce, through a quadratic locally invertible coordinate transformation, a coordinate system such that $\dot\gamma(0) = \partial_0$, $g_{\alpha\beta}(p) = \eta_{\alpha\beta}$ and $\Gamma_{\alpha\beta}^\mu(p) = 0$, so that the hypothesis of Theorem 10 apply.

Let us consider the spacelike subspace of $T_{\gamma(t)}M$, given by

$$
S(t) := \operatorname{Ker} g(\dot\gamma(t), \cdot).
$$

Let $L(t)$ be the subspace of $T_{\gamma(t)}M$ spanned by $\{\partial_i, i \geq 1\}$. For $t = 0$, $S(0) = L(0)$, thus by continuity there is a neighborhood of 0, such that for t belonging to this neighborhood $S(t)$ makes with $L(t)$ an (Euclidean) angle smaller than $1/16$ rad. Thus C can be taken sufficiently small that this property holds for every $t \in \gamma^{-1}(C)$.

Let $x : [0,1] \to C$ be a geodesic such that $x(0)$ and $x(1)$ belong to the same level set of $(D_q^2)^{-1}(c)$, $c < 0$, for some $q \in C$. By Theorem 6 x cannot be future directed causal otherwise $D_q^2(x(1)) > D_q^2(x(0))$, and it cannot be past directed causal otherwise $D_q^2(x(1)) < D_q^2(x(0))$, thus x is spacelike. Let $a, b \in [0,1]$, $a < b$, then $y(t) = x((b-a)t + a)$ is such that $y(0) = x(a)$ and $y(1) = x(b)$. Let us use Eq. (51) in Eq. (17) for the geodesic y

$$
|\frac{\mathrm{d}}{\mathrm{d}t}D_q^2(y(t))|_{t=1} - \frac{\mathrm{d}}{\mathrm{d}t}D_q^2(y(t))|_{t=0} - 2D^2(x(a), x(b))| \leq \frac{\epsilon}{(1-\epsilon)^2}\|P(x(a), x(b))\|^2.
\tag{55}
$$

Now we have to impose some constraint to $x(a)$ and $x(b)$ so as to obtain an inequality of the form

$$
\|P(x(a), x(b))\|^2 \leq \sigma(\epsilon)g_{x(b)}(P(x(a), x(b)), P(x(a), x(b)))
$$

where $\epsilon\sigma(\epsilon) \to 0$ for $\epsilon \to 0$.

We recall that given two Lorentzian metrics g_1, g_2 on a differentiable manifold, $g_1 < g_2$ means that at each point the timelike cone of g_2 contains the causal cone of g_1. Let me consider the metric

$$\eta^+ := -\frac{1 + 2\sqrt{\epsilon}}{1 - 2\sqrt{\epsilon}}\, (dx^0)^2 + (d\mathbf{x})^2$$

which satisfies $g < \eta^+$ at p. By continuity we can choose C so small that it holds anywhere in C.

Suppose that $v \in TC$ is a spacelike vector for η^+, that is $\eta^+(v, v) \geq 0$, then a little algebra shows that this condition can be rewritten

$$v \cdot v \leq \frac{1}{\sqrt{\epsilon}}(-(1 + \sqrt{\epsilon})(v^0)^2 + (1 - \sqrt{\epsilon})(\mathbf{v})^2).$$

Now observe that at p for every $v \in T_p C$, $\|v\| = 1$,

$$-(1 + \sqrt{\epsilon})(v^0)^2 + (1 - \sqrt{\epsilon})(\mathbf{v})^2 < g_p(v, v) = -(v^0)^2 + (\mathbf{v})^2,$$

thus by continuity the same holds at any point in a neighborhood of p, and we can choose C sufficiently small that for every $v \in TC$, and $q \in C$

$$(-(1 + \sqrt{\epsilon})(v^0)^2 + (1 - \sqrt{\epsilon})(\mathbf{v})^2) \leq g_q(v, v).$$

As a consequence, for every $q \in C$ and for every $v \in TC$ such that $\eta^+(v, v) \geq 0$

$$v \cdot v \leq \frac{1}{\sqrt{\epsilon}}\, g_q(v, v). \tag{56}$$

Let $q = \gamma(t_q), r = \gamma(t_r), t_q, t_r \in I, t_q \neq t_r$.

In Sect. 2.4 we proved that $O := \bar{B}(r, \delta)$ is strictly convex normal for any sufficiently small δ, and in Sect. 2.1 through Eq. (31) we proved that the Euclidean velocities $\dot{x}(t), t \in [0, 1]$ of any geodesic $x : [0, 1] \to \bar{O}$ are bounded by a constant $V(\delta)$ which goes to zero for $\delta \to 0$. Since $\Gamma^\mu_{\alpha\beta}$ is bounded in a neighborhood of r, the geodesic equation implies that there is $M > 0$ such that $\frac{dv^\mu}{dt} \leq M\|v\|^2$, which is the same as saying that the (Euclidean) radius of curvature is greater than $1/M$ and hence $\dot{x}$ can vary in an subinterval of $[0, 1]$ of an angle of at most $V(\delta)M$. Thus if we take δ sufficiently small we can make the variation of angle on the tangent to any geodesic $x : [0, 1] \to \bar{O}$ to be bounded by $\frac{1}{8}$ rad.

We have already shown in Theorem 5 that D_q^2 is $C^{1,1}$ on C with a differential at r given by $2g(P(q, r), \cdot)$. Since the differential is continuous, D_q^2 is strongly differentiable thus for every β we can write for sufficiently small $\delta > 0$, for $q_1, q_2 \in \bar{O}$

$$|D_q^2(q_2) - D_q^2(q_1) - 2g(P(q, r), q_2 - q_1)| \leq \beta\sqrt{-g(P(q, r), P(q, r))}\, \|q_2 - q_1\|.$$

If q_1 and q_2 belong to the same level set, recalling that

$$\dot{\gamma}(t_r) = P(q, r)/\sqrt{-g(P(q, r), P(q, r))}$$

we obtain

$$|g_r(\dot{\gamma}(t_r), q_2 - q_1)| \leq \beta \|q_2 - q_1\|.$$

As $\beta \to 0$, the direction of $q_2 - q_1$ is constrained to approach $S(t_r)$. As the space of directions is a sphere and hence compact there is a value of β and a corresponding O such that whenever $q_1, q_2 \in \bar{O}$ belong to the same level set of D_q, $q_2 - q_1$ makes with $S(t_r)$ an (Euclidean) angle smaller than $1/16$ rad.

Since $\|P(q_1, q_2) - (q_2 - q_1)\| \leq \epsilon \|q_2 - q_1\|$ and $\epsilon < 1/9$, the vector $P(q_1, q_2)$ makes with $q_2 - q_1$ an Euclidean angle smaller than $2 \arcsin(\epsilon/2) < 1/8$ rad (because $\epsilon < 1/9$). Thus $P(q_1, q_2)$ makes with $S(t_r)$ an (Euclidean) angle smaller than $\frac{3}{16}$ rad, and with $L(t_r)$ an (Euclidean) angle smaller than $\frac{2}{8}$ rad. If $x(a)$ and $x(b)$ are any two points in the geodesic x joining q_1 to q_2 then $P(x(a), x(b))$ makes an angle with $P(q_1, q_2)$ of at most $\frac{1}{8}$ rad, thus it makes an angle with $L(t_r)$ smaller than $\frac{3}{8}$ rad. Moreover,

$$\arctan\left(\left[\frac{1 - 2\sqrt{\epsilon}}{1 + 2\sqrt{\epsilon}}\right]^{1/2}\right) > \arctan\left(\left[\frac{1 - 2/3}{1 + 2/3}\right]^{1/2}\right) > \frac{3}{8} \text{ rad}$$

thus $P(x(a), x(b))$ is η^+-spacelike (we knew already that it was g-spacelike). Since $P(x(a), x(b))$ is η^+-spacelike we have using Eqs. (55) and (56)

$$|\nabla_{(b-a)\dot{x}(b)} D_q^2 - \nabla_{(b-a)\dot{x}(a)} D_q^2 - 2D^2(x(a), x(b))|$$
$$\leq \frac{\sqrt{\epsilon}}{(1 - \epsilon)^2} g_{x(b)}(P(x(a), x(b)), P(x(a), x(b))) = \frac{\sqrt{\epsilon}}{(1 - \epsilon)^2} D^2(x(a), x(b)),$$

which a redefinition of ϵ, and a redefinition of parametrization such that $(b - a) \to D(x(a), x(b))$ (namely the g-arc length parametrization) brings to the form of Eq. (18).

The statement of concerning the strong convexity of $D_q^2 \circ x$ is immediate from the equivalences recalled in Sect. 1.5.

Let us prove the the strict convexity of $(D_q^2)^{-1}((-\infty, c)) \cap O$. Suppose that $c < 0$ is such that $G := (D_q^2)^{-1}((-\infty, c)) \cap O \neq \emptyset$, otherwise there is nothing to prove. Let $x : [a, b] \to C$ be a geodesic such that $x(a), x(b) \in \bar{G}$, then $x(a), x(b) \in \bar{O}$, and since O is strictly geodesically convex x is contained in O but for the endpoints. Since $D_q^2 : C \times C \to \mathbb{R}$ is continuous, $x(a), x(b) \leq c$ and we have to show that for every $t \in (a, b)$, $D_q^2(x(t)) < c$.

Suppose not, then there is some $t_{max} \in (a, b)$ such that $D_q^2(x(t_{max}))$ is the maximum of $D_q^2(x(\cdot))$ over $[a, b]$ and $D_q^2(x(t_{max})) \geq c$. In particular,

$$\frac{d(D_p^2 \circ x)}{dt}|_{t=t_{max}} = \nabla_{\dot{x}(t_{max})} D_q^2(x(t_{max})) = 0.$$

However, by Eq. (18) no two values of t can attain this maximum, thus in any neighborhood E of t_{max} we can find $t_1, t_2 \in E \setminus \{t_{max}\}$, $t_1 < t_{max} < t_2$ such that $D_q^2(x(t_1)) = D_q^2(x(t_2)) < D_q^2(x(t_{max}))$, thus we can apply once again the equation following Eq. (18) using these two values as endpoints of the spacelike geodesic. But that equation implies that $D_q^2(x(t_{max})) < D_q^2(x(t_2))$, a contradiction. $\qquad \square$

Lemma 4 *Let $p \in M$ and let $\gamma : I \to M$, $t \mapsto \gamma(t)$, be a timelike geodesic such that $p = \gamma(0)$. The convex normal set $C \ni p$ can be taken sufficiently small that once I is redefined to be the connected component of $\gamma^{-1}(C)$ containing 0, the following property holds. We can find $q_1 = \gamma(t_1)$, and $q_2 = \gamma(t_2)$ with $t_1 < 0 < t_2$, and a strictly convex normal set $O \ni p$, such that introduced the constants $c_1 := D_{q_1}^2(p)$ and $c_2 := D_{q_2}^2(p)$, we have that, for any $c_1' > c_1$ sufficiently close to c_1 and for any $c_2' > c_2$ sufficiently close to c_2,*

$$S(c_1', c_2') = (D_{q_1}^2)^{-1}((-\infty, c_1')) \cap (D_{q_2}^2)^{-1}((-\infty, c_2')) \cap O$$

is strictly convex normal and globally hyperbolic.

Proof By Theorem 11 we can find a strictly convex relatively compact set $C \ni p$ with the property of that theorem. In particular, let $q_1 = \gamma(t_1)$, $q_1 \in C$, $t_1 < 0$, and let $q_2 = \gamma(t_2)$, $q_2 \in C$, $t_2 > 0$. There is a strictly convex normal set $O_1 \ni p$, $\bar{O}_1 \subset I_C^+(q_1)$, such that $D_{q_1}^2 : C \times C \to \mathbb{R}$ is strongly convex over the geodesics segments in O_1 connecting two points in its level surfaces. Similarly there is a strictly convex set O_2, $\bar{O}_2 \subset I_C^-(q_2)$, with an analogous property with respect to q_2. Let $O = O_1 \cap O_2$.

Let us introduce the closed sets

$$A(c_1', c_2') = (D_{q_1}^2)^{-1}((-\infty, c_1']) \cap (D_{q_2}^2)^{-1}((-\infty, c_2']) \cap \bar{O}.$$

The set $A(c_1, c_2)$ contains just p for otherwise there would be a different timelike curve made of two geodesic pieces of total Lorentzian length equal with that of $\gamma|_{[t_1,t_2]}$, a contradiction to Theorem 6. Since the intersection of a family of non-empty compact sets is non-empty, for sufficiently large i the set $A(c_1 + 1/i, c_2 + 1/i)$ must be disjoint from ∂O.

Thus for $c_1' \leq c_1 + 1/i$ and $c_2' \leq c_2 + 1/i$, $S(c_1', c_2')$ is the component of the open set $(D_{q_1}^2)^{-1}((-\infty, c_1')) \cap (D_{q_2}^2)^{-1}((-\infty, c_2'))$ contained in O and its closure is contained in O. As $D_{q_1}^2$ is decreasing over future directed causal curves and $D_{q_1}^2$ is increasing over future directed causal curves, no causal curve in C can leave and reenter $S(c_1', c_2')$, in particular since C is causally simple and relatively compact, $S(c_1', c_2')$ is globally hyperbolic.

The set $S(c_1', c_2')$ is strictly convex normal because it is the intersection of the strictly convex normal set $(D_{q_1}^2)^{-1}((-\infty, c_1')) \cap O_1$ and the strictly convex normal set $(D_{q_2}^2)^{-1}((-\infty, c_2')) \cap O_2$. □

Proof (Proof of Corollary 2) Just take $C_i = S(c_1 + 1/(k+i), c_2 + 1/(k+i))$ for sufficiently large k and use the results of Lemma 4 (see also its proof). □

References

1. Akbar-Zadeh, H.: Sur les espaces de Finsler a courbures sectionnelles constantes. Acad. Roy. Belg. Bull. Cl. Sci. **74**, 281–322 (1988)
2. Ambrose, W., Palais, R.S., Singer, I.M.: Sprays. An. Acad. Brasil. Ciênc. **32**(2), 163–178 (1960)
3. Andersson, L., Galloway, G.J., Howard, R.: The cosmological time function. Class. Quantum Grav. **15**, 309–322 (1998)
4. Antonelli, P.L., Ingarden, R.S., Matsumoto, M.: The theory of sprays and Finsler spaces with applications in physics and biology. Springer, Dordrecht (1993)
5. Bao, D., Chern, S.S., Shen, Z.: An introduction to Riemann-Finsler geometry. Springer, New York (2000)
6. Beem, J.K.: Indefinite Finsler spaces and timelike spaces. Can. J. Math. **22**, 1035–1039 (1970)
7. Beem, J.K.: Characterizing Finsler spaces which are pseudo-Riemannian of constant curvature. Pacific J. Math. **64**, 67–77 (1976)
8. Beem, J.K., Ehrlich, P.E., Easley, K.L.: Global Lorentzian geometry. Marcel Dekker Inc., New York (1996)
9. Berestovskij, V.N., Nikolaev, I.G., Reshetnyak, Y.G.: Geometry IV: non-regular Riemannian geometry. Springer, Berlin (1993)
10. Buttazzo, G., Giaquinta, M.: One-dimensional variational problems. Oxford University Press, Oxford (1998)
11. Candela, A.M., Flores, J.L., Sánchez, M.: Global hyperbolicity and Palais-Smale condition for action functionals in stationary spacetimes. Adv. Math. **218**, 515–536 (2008)
12. do Carmo, M.: Riemannian geometry. Birkhäuser, Boston (1992)
13. Cartan, H.: Differential calculus. Hermann, Paris (1971)
14. Chen, B.L., LeFloch, P.G.: Injectivity radius of Lorentzian manifolds. Commun. Math. Phys. **278**, 679–713 (2008)
15. Chruściel, P.T.: Elements of causality theory (2011). ArXiv:1110.6706v1
16. Chruściel, P.T., Grant, J.D.E.: On Lorentzian causality with continuous metrics. Class. Quantum Grav. **29**, 145001 (2012)
17. Clarke, F.H.: On the inverse function theorem. Pacific J. Math. **64**, 97–102 (1976)
18. DeTurck, D.M., Kazdan, J.L.: Some regularity theorems in Riemannian geometry. Ann. Sci. École Norm. Sup. **4**(14), 249–260 (1981)
19. Dieudonné, J.: Treatise on Analysis. Foundations of modern analysis. Academic Press, New York (1969)
20. Dolecki, S., Greco, G.H.: Amazing oblivion of Peano's legacy (2012)
21. Esser, M., Shisha, O.: A modified differentiation. Am. Math. Mon. **71**, 904–906 (1964)
22. Evans, L.C.: Partial differential equations. American Mathematical Society, Providence (1998)
23. Fan, L., Liu, S., Gao, S.: Generalized monotonicity and convexity of non-differentiable functions. J. Math. Anal. Appl. **279**, 276–289 (2003)
24. Federer, H.: Curvature measures. Trans. Am. Math. Soc. **93**, 418–491 (1959)

25. Foertsch, T.: Ball versus distance convexity of metric spaces. Beiträge Algebra Geom. **45**, 481–500 (2004)
26. Gallot, S., Hulin, D., Lafontaine, J.: Riemannian geometry. Springer, Berlin (1987)
27. Garg, K.M.: Theory of differentiation. Wiley, New York (1998)
28. Harris, S.G.: A triangle comparison theorem for Lorentz manifolds. Indiana Univ. Math. J. **31**, 289–308 (1982)
29. Hartman, P.: On the local uniqueness of geodesics. Am. J. Math. **72**, 723–730 (1950)
30. Hartman, P.: On geodesic coordinates. Am. J. Math. **73**, 949–954 (1951)
31. Hartman, P.: Ordinary differential equations. Wiley, New York (1964)
32. Hartman, P.: Remarks on geodesics. Proc. Am. Math. Soc. **89**, 467–472 (1983)
33. Hartman, P., Wintner, A.: On the problems of geodesics in the small. Am. J. Math. **73**, 132–148 (1951)
34. Hawking, S.W., Ellis, G.F.R.: The large scale structure of space-time. Cambridge University Press, Cambridge (1973)
35. Krantz, S.G., Parks, H.R.: The implicit function theorem: history, theory, and applications. Birkhäuser, Boston (2000)
36. Kunzinger, M., Steinbauer, R., Stojković, M.: The exponential map of a $C^{1,1}$-metric. Differential Geom. Appl. 34, 14–24 (2014). ArXiv:1306.4776v1
37. Kunzinger, M., Steinbauer, R., Stojković, M., Vickers, J.A.: A regularisation approach to causality theory for $C^{1,1}$-Lorentzian metrics. Gen. Relativ. Gravit. **46**, 1738 (2014). ArXiv:1310.4404v2
38. Lang, S.: Differential and Riemannian manifolds. Springer, New York (1995)
39. Lang, S.: Fundamentals of differential geometry. Springer, New York (1999)
40. Leach, E.B.: A note on inverse function theorems. Proc. Am. Math. Soc. **12**, 694–697 (1961)
41. Leach, E.B.: On a related function theorem. Proc. Am. Math. Soc. **14**, 687–689 (1963)
42. Matsumoto, M.: Foundations of Finsler geometry and special Finsler spaces. Kaseisha Press, Tokio (1986)
43. Milnor, J.: Morse theory. Princeton University Press, Princeton (1969)
44. Minguzzi, E.: The causal ladder and the strength of K-causality. I. Quantum Grav. **25**, 015009 (2008)
45. Minguzzi, E.: Limit curve theorems in Lorentzian geometry. J. Math. Phys. **49**, 092501 (2008)
46. Minguzzi, E.: The equality of mixed partial derivatives under weak differentiability conditions. Real Anal. Exchange (2014) (To appear). E-print Archive ArXiv:1309.5841
47. Minguzzi, E.: Light cones in Finsler spacetime. Comm. Math. Phys. (2014). ArXiv:1403.7060. doi:10.1007/s00220-014-2215-6
48. Minguzzi, E., Sánchez, M.: The causal hierarchy of spacetimes, ESI Lect. Math. Phys., vol. Baum, H., Alekseevsky, D. (eds.) Recent developments in pseudo-Riemannian geometry, pp. 299–358. Eur. Math. Soc. Publ. House, Zurich (2008). ArXiv:gr-qc/0609119
49. Nijenhuis, A.: Strong derivatives and inverse mappings. Am. Math. Mon. **81**, 969–980 (1974)
50. Nomizu, K.: Remarks on sectional curvature of an indefinite metric. Proc. Am. Math. Soc. **89**, 473–476 (1983)
51. O'Neill, B.: Semi-Riemannian geometry. Academic Press, San Diego (1983)
52. Peano, G.: Sur la définition de la dérivée. Mathesis **2**, 12–14 (1892)
53. Penrose, R.: Techniques of differential topology in relativity. Cbms-Nsf regional conference series in applied mathematics. SIAM, Philadelphia (1972)
54. Perelman, G.: Elements of Morse theory on Aleksandrov spaces (russian). Algebra i Analiz 5(1), 232–241 (1993). Engl. transl. St. Peterburg Math. J. 5:1, 205–213 (1994)
55. Perelman, G., Petrunin, A.: Extremal subsets in Aleksandrov spaces and the generalized Lieberman theorem (in russian). Algebra i Analiz 5(1), 242–256 (1993). Engl. transl. St. Peterburg Math. J. 5:1, 215–227 (1994)
56. Perlick, V.: Fermat principle in Finsler spacetimes. Gen. Relativ. Gravit. **38**, 365–380 (2006)
57. Rudin, W.: Real and complex analysis. McGraw-Hill, London (1970)
58. Sakai, T.: Riemannian geometry. American Mathematical Society, Providence (1996)
59. Senovilla, J.M.M.: Singularity theorems and their consequences. Gen. Relativ. Gravit. **30**, 701–848 (1998)
60. Villani, C.: Optimal transport. Springer, Berlin (2009)
61. Wang, Z.: Notes on differential geometry (2012)

62. Whitehead, J.H.C.: Convex regions in the geometry of paths. Quart. J. Math. Oxford Ser. **3**, 33–42 (1932)
63. Whitehead, J.H.C.: Convex regions in the geometry of paths—addendum. Quart. J. Math. Oxford Ser. **4**, 226–227 (1933)