Class. Quantum Grav. **15** (1998) 309–322. Printed in the UK PII: S0264-9381(98)86379-X

The cosmological time function

Lars Andersson†‖, Gregory J Galloway‡¶ and Ralph Howard§+

† Department of Mathematics, Royal Institute of Technology, S-100 44 Stockholm, Sweden
‡ Department of Mathematics, University of Miami, Coral Gables, FL 33124, USA
§ Department of Mathematics, University of South Carolina, Columbia, SC 29208, USA

Received 29 July 1997

Abstract. Let (M, g) be a time-oriented Lorentzian manifold and d the Lorentzian distance on M. The function $\tau(q) := \sup_{p<q} d(p, q)$ is the cosmological time function of M, where as usual $p < q$ means that p is in the causal past of q. This function is called regular iff $\tau(q) < \infty$ for all q and also $\tau \to 0$ along every past inextendible causal curve. If the cosmological time function τ of a spacetime (M, g) is regular it has several pleasant consequences: (i) it forces (M, g) to be globally hyperbolic; (ii) every point of (M, g) can be connected to the initial singularity by a rest curve (i.e. a timelike geodesic ray that maximizes the distance to the singularity); (iii) the function τ is a time function in the usual sense; in particular, (iv) τ is continuous, in fact, locally Lipschitz and the second derivatives of τ exist almost everywhere.

PACS numbers: 0420C, 0420G, 9880H

1. Introduction

Time functions play an important role in general relativity. They arise naturally in the global causal theory of spacetime and they permit a decomposition of spacetime into space and time which is useful, for example, in the study of the solution of the Einstein equation. The choice of a time function, however, can be rather arbitrary and a given time function may have little physical significance. Very few situations have been identified which lead to a canonically defined time function. In this paper we introduce and study what may be viewed in the cosmological setting as a canonical time function.

Let (M, g) be a spacetime (i.e. a time-oriented Lorentzian manifold) and let $d: M \times M \to [0, \infty]$ be the Lorentzian distance function. Define the cosmological time function $\tau: M \to (0, \infty]$ by

$$\tau(q) := \sup_{p<q} d(p, q). \tag{1.1}$$

If c is a causal curve in M denote by $L(c)$ the Lorentzian length of c and for $q \in M$, let $\mathcal{C}^-(q)$ be the set of all past-directed causal curves c in M that start at q. Then we have the

‖ Visiting scientist: Max-Planck-Institut für Gravitationsphysik (Albert-Einstein-Institut) Schlaatzweg 1, D-14473 Potsdam, Germany. E-mail address: larsa@math.kth.se
¶ E-mail address: galloway@math.miami.edu
+ E-mail address: howard@math.sc.edu

alternative definition

$$\tau(q) := \sup\{L(c) : c \in \mathcal{C}^-(q)\}.$$

The number $\tau(q)$ can be thought of as the length of time that point q has been in existence.

In general the function τ need not be at all nice; for example, in the case of flat Minkowski space $\tau \equiv \infty$. We will give examples below where (M, g) is globally hyperbolic, $\tau(q) < \infty$ for all q but τ is discontinuous.

Definition 1.1. The cosmological time function τ of (M, g) is regular if and only if

(i) (M, g) has finite existence times, i.e. $\tau(q) < \infty$ for all $q \in M$.

(ii) $\tau \to 0$ along every past inextendible causal curve.

The first of these conditions is an assertion that the spacetime has an initial singularity in the strong sense that for each point of the spacetime any particle that passes through q has been in existence for at most a time $\tau(q)$. The second condition is a weak completeness assumption. It asserts that if we believe that the condition $\tau = 0$ defines the initial singularity and that worldlines of particles are inextendible, then every particle came into existence at the initial singularity.

Our main result is that if the cosmological time function is regular then the spacetime is quite well behaved.

Theorem 1.2. Suppose (M, g) is a spacetime such that the function $\tau : M \to (0, \infty)$ defined by (1.1) is regular. Then the following properties hold.

(i) (M, g) is globally hyperbolic.

(ii) τ is a time function in the usual sense, i.e. τ is continuous and is strictly increasing along future-directed causal curves.

(iii) For each $q \in M$ there is a future-directed timelike ray $\gamma_q : (0, \tau(q)] \to M$ that realizes the distance from the 'initial singularity' to q, that is, γ_q is a future-directed timelike unit speed geodesic, which is maximal on each segment, such that

$$\gamma_q(\tau(q)) = q, \qquad \tau(\gamma_q(t)) = t, \qquad \text{for} \quad t \in (0, \tau(q)]. \tag{1.2}$$

(iv) The tangent vectors $\{\gamma_q'(\tau(q)) : q \in M\}$ are locally bounded away from the light cones. More precisely, if $K \subseteq M$ is compact then $\{\gamma_q'(\tau(q)) : q \in K\}$ is a bounded subset of the tangent bundle $T(M)$.

(v) τ has the following additional regularity property: it is locally Lipschitz and its first and second derivatives exist almost everywhere.

Conditions similar to property (iv) have played an important role in the analysis of the regularity of Lorentzian Busemann functions and their level sets (cf [1, 8]). Here property (iv) will be used to establish property (v).

For regularity properties of the level sets $\{\tau = a\}$ see section 3 (as well as the corollary at the end of section 2). The various conclusions of the theorem will be proven as separate propositions in the following sections.

1.1. Terminology and notation

We use the standard terminology and notation from Lorentzian geometry, following for example [4, 9], in particular, if (M, g) is a spacetime then $p \ll q$ (respectively, $p < q$) means there is a future-directed timelike (respectively causal) curve from p to q. If $S \subset M$ then $I^+(S)$ is the chronological future of S and $J^+(S)$ is the causal future of S. Likewise

$I^-(S)$ and $J^-(S)$ are the chronological past and causal past of S. If $p < q$, then the Lorentzian distance $d(p, q)$ is the supremum of the lengths of all the future-directed causal curves from p to q and if $p \not< q$ then $d(p, q) = 0$. A fact that will be used repeatedly is that if $x < p < q$ then the reverse triangle inequality

$$d(x, q) \geqslant d(x, p) + d(p, q)$$

holds.

2. Proofs of the basic properties of the cosmological time function

2.1. Continuity of the cosmological time function

Proposition 2.1. If the cosmological time function τ of (M, g) is regular then it is continuous and satisfies the reverse Lipschitz inequality

$$p < q \qquad \text{implies} \quad \tau(p) + d(p, q) \leqslant \tau(q). \tag{2.1}$$

Proof. To prove the reverse Lipschitz inequality assume $p < q$ and let $x < p$. Then $x < q$ and so by the reverse triangle inequality (i.e. $d(x, p) + d(p, q) \leqslant d(x, q)$),

$$\tau(p) + d(p, q) = \sup_{x<p}(d(x, p) + d(p, q)) \leqslant \sup_{x<p} d(x, q) \leqslant \sup_{x<q} d(x, q) = \tau(q).$$

We now establish the continuity of τ. For any $p \in M$ the function $q \mapsto d(p, q)$ is lower semicontinuous on M. (That is $\liminf_{x\to q} d(p, x) \geqslant d(p, q)$.) For example, cf [9], p 215. Then $\tau(q) = \sup_{p<q} d(p, q)$ is a supremum of lower semicontinuous functions and therefore also lower semicontinuous. Thus to prove continuity of τ it is enough to show it is upper semicontinuous, that is $\limsup_{x\to q} \tau(x) \leqslant \tau(q)$.

Assume, toward a contradiction, that τ is not upper semicontinuous at $q \in M$. Then there is $\varepsilon > 0$ and a sequence $x_\ell \to q$ such that for each ℓ

$$\tau(x_\ell) \geqslant \tau(q) + \varepsilon.$$

For each ℓ we can choose p_ℓ with

$$d(p_\ell, x_\ell) \geqslant \tau(x_\ell) - \frac{1}{\ell}.$$

Moreover, by the regularity of τ, we can choose the sequence $\{p_\ell\}$ so that $\tau(p_\ell) \to 0$ as $\ell \to 0$. (To see this make any choice of $\hat{p}_\ell$ with $d(\hat{p}_\ell, x_\ell) \geqslant \tau(x_\ell) - 1/\ell$. Then choose a past-directed inextendible curve σ starting at $\hat{p}_\ell$. By the definition of regular there is a point p_ℓ on σ with $\tau(p_\ell) < 1/\ell$. Then $d(p_\ell, x_\ell) \geqslant d(\hat{p}_\ell, x_\ell) \geqslant \tau(x_\ell) - 1/\ell$ and $\lim_{\ell\to\infty} \tau(p_\ell) = 0$.) The condition $\tau(p_\ell) \to 0$ and the lower semicontinuity of τ implies that $\{p_\ell\}$ diverges to infinity, that is it has no convergent subsequences.

We now put a complete Riemannian metric h on M and assume that all causal curves (except possibly those arising as limit curves) are parametrized with respect to arc length in the metric h. Since $d(p_\ell, x_\ell) < \infty$ there is a past-directed causal curve $c_\ell: [0, a_\ell] \to M$ (parametrized with respect to arc length in h) from x_ℓ to p_ℓ such that

$$L(c_\ell) \geqslant d(p_\ell, x_\ell) - \frac{1}{\ell} \geqslant \tau(x_\ell) - \frac{2}{\ell} \geqslant \tau(q) + \varepsilon - \frac{2}{\ell} \tag{2.2}$$

where $L(\cdot)$ is the Lorentzian arc length functional. Since $\{p_\ell\}$ diverges, $a_\ell \to \infty$. Hence, by passing to a subsequence if necessary, we have that $\{c_\ell\}$ converges uniformly on compact sets to a past inextendible timelike or null ray (maximal half geodesic) $c: [0, \infty) \to M$

(cf [7], sections 2 and 3). Moreover, by the upper semicontinuity of the Lorentzian arclength functional (strong causality is not required, again cf [7]), for each $b \geqslant 0$

$$L(c|_{[0,b]}) \geqslant \limsup_{\ell\to\infty} L(c_\ell|_{[0,b]}). \tag{2.3}$$

Claim 1. The curve $c\colon [0,\infty) \to M$ is null.

If not then c is a timelike ray. Choose $t > 0$ and $\delta > 0$ so that

$$L(c|_{[0,t]}) + \delta \leqslant \tfrac{1}{2}\varepsilon.$$

By (2.3) there is an N such that for all $\ell \geqslant N$,

$$L(c_\ell|_{[0,t]}) \leqslant L(c|_{[0,t]}) + \delta \leqslant \tfrac{1}{2}\varepsilon.$$

Hence, by (2.2) and the above,

$$L(c_\ell|_{[t,a_\ell]}) = L(c_\ell) - L(c_\ell|_{[0,t]}) \geqslant \tau(q) + \frac{\varepsilon}{2} - \frac{2}{\ell}.$$

Thus when ℓ is sufficiently large,

$$L(c_\ell|_{[t,a_\ell]}) > \tau(q).$$

On the other hand, since c is timelike, we have that $c_\ell(t) \in I^-(q)$ for all ℓ sufficiently large. It follows that $\tau(q) \geqslant L(c_\ell|_{[t,a_\ell]}) > \tau(q)$. This contradiction establishes the claim.

Claim 2. $\tau(p) > \tau(q)$ for all $p \in I^+(c)$.

Choose y on c such that $y \in I^-(p)$.

We have $y = c(b)$ for some $b \geqslant 0$. Since c is null (2.3) implies,

$$L(c_\ell|_{[0,b]}) \to 0.$$

Let $y_\ell := c_\ell(b)$. Then,

$$\begin{aligned}\tau(y_\ell) &\geqslant L(c_\ell|_{[b,a_\ell]}) = L(c_\ell) - L(c_\ell|_{[0,b]})\\ &\geqslant \tau(q) + \varepsilon - \frac{2}{\ell} - L(c_\ell|_{[0,b]}) > \tau(q),\end{aligned}$$

for all sufficiently large ℓ. Moreover, by taking ℓ even larger if necessary, we have $y_\ell \in I^-(p)$, and hence $\tau(p) > \tau(y_\ell) > \tau(q)$, as claimed.

A past inextendible timelike curve $\tilde{c}$ asymptotic to c and contained in $I^+(c)$ may be constructed as follows.

(i) First choose $y_\ell = c(b_\ell)$, $b_\ell \to \infty$ so that $\lim_{\ell\to\infty} y_\ell$ does not exist.

(ii) Then choose $\{z_\ell\} \subset M$ so that: (a) $z_{\ell+1} \in I^-(z_\ell)$; (b) $z_\ell \in I^+(y_\ell)$; (c) $\lim_{\ell\to\infty} z_\ell$ does not exist.

Let $\tilde{c}$ be a past-directed timelike curve which threads through $z_1 \gg z_2 \gg z_3 \gg \cdots$.

As $\lim_{\ell\to\infty} z_\ell$ does not exist, the curve $\tilde{c}$ is past inextendible. Since by construction $\tilde{c} \subset I^+(c)$, claim 2 implies $\tau > \tau(q)$ along $\tilde{c}$, which contradicts the regularity condition. This completes the proof. □

2.2. Global hyperbolicity of (M, g)

Proposition 2.2. Let (M, g) be a spacetime so that the cosmological time function τ is regular. Then (M, g) is globally hyperbolic.

Proof. We have shown in proposition 2.1 that τ is continuous. Therefore, if $S_t := \{q \in M : \tau(q) = t\}$ then by elementary topological and causal considerations, S_t is closed, achronal and edgeless. (That S_t is achronal follows from the reverse Lipschitz inequality. That S_t is closed and edgeless follows from the continuity of τ.)

Recall that the future domain of dependence $D^+(S_t)$ of S_t is the set of all points $q \in M$ such that every past inextendible causal curve from q intersects S_t. The past domain of dependence $D^-(S_t)$ is defined time-dually. The domain of dependence of S_t is $D(S_t) = D^+(S_t) \cup D^-(S_t)$. From the definition of regularity and the continuity of τ we see that $D^+(S_t) = \{q : \tau(q) \geqslant t\}$. It follows that each point of M is contained in $\mathrm{int} D^+(S_t)$ for some t. Since strong causality holds at each point of $\mathrm{int} D^+(S_t)$ (cf [10], proposition 5.22, p 48), (M, g) is strongly causal.

Now let $p, q \in M$ with $p < q$. Then choose $t > 0$ with $t < \tau(p)$. Then $J^-(q) \cap J^+(p)$ is a subset of the open set $\{x : \tau(x) > t\} \subset D(S_t)$ and thus $J^-(q) \cap J^+(p)$ is contained in the interior of $D(S_t)$. This implies (cf [10], proposition 5.23, p 48) $J^-(q) \cap J^+(p)$ is compact. As (M, g) is strongly causal and p and q were arbitrary points of M with $p < q$, this verifies the definition of globally hyperbolic. □

2.3. Existence of maximizing rays to the initial singularity

Proposition 2.3. Let (M, g) be a spacetime with regular cosmological time function τ. Then for each $q \in M$ there is a future-directed timelike ray $\gamma_q : (0, \tau(q)] \to M$ that realizes the distance from the 'initial singularity' to q; that is, γ_q is a future-directed timelike unit speed geodesic that realizes the distance between any two of its points (for $0 < s < t \leqslant \tau(q)$, $d(\gamma_q(s), \gamma_q(t)) = t - s$) and satisfies,

$$\gamma_q(\tau(q)) = q, \qquad \tau(\gamma_q(t)) = t, \qquad \text{for} \quad t \in (0, \tau(q)]. \tag{2.4}$$

Proof. For the purpose of the proof we will parametrize curves with respect to a complete Riemannian metric h on M as in the proof of proposition 2.1. Fix $q \in M$. As in the proof of proposition 2.1, one can construct a sequence $\{y_\ell\} \subset I^-(q)$ that diverges to infinity and such that

$$d(y_\ell, q) \geqslant \tau(q) - \frac{1}{\ell} \qquad \text{and} \qquad \tau(y_\ell) < \frac{1}{\ell}.$$

By proposition 2.2, (M, g) is globally hyperbolic so there is a past-directed maximal geodesic segment $\gamma_\ell : [0, a_\ell] \to M$ from $q = \gamma_\ell(0)$ to $y_\ell = \gamma_\ell(a_\ell)$. Since $\{y_\ell\}$ diverges to infinity and the curves are parametrized with respect to h arc length we have $a_\ell \to \infty$. Hence, by passing to a subsequence if necessary, the sequence $\{\gamma_\ell\}$ converges to a past inextendible timelike or null ray $\gamma : [0, \infty) \to M$ based at $q = \gamma(0)$. Hence for all $b \in (0, \infty)$,

$$L(\gamma|_{[0,b]}) = d(\gamma(b), q). \tag{2.5}$$

Claim. γ is timelike and for each $b \in (0, \infty)$,

$$d(\gamma(b), q) = \tau(q) - \tau(\gamma(b)). \tag{2.6}$$

Hence by suitably reparametrizing γ we obtain a timelike ray γ_q that satisfies (2.4).

To see that the claim holds first note by the reverse Lipschitz inequality,

$$d(\gamma(b), q) \leqslant \tau(q) - \tau(\gamma(b)). \tag{2.7}$$

By the maximality of the segments γ_ℓ,

$$d(\gamma_\ell(b), q) = d(y_\ell, q) - d(y_\ell, \gamma_\ell(b)) \geqslant \left(\tau(q) - \frac{1}{\ell}\right) - \tau(\gamma_\ell(b)).$$

Letting $\ell \to \infty$ we obtain $d(\gamma(b), q) \geqslant \tau(q) - \tau(\gamma(b))$ which, together with (2.7), establishes (2.6). Moreover, since $\tau(\gamma(b)) \to 0$ as $b \to \infty$, by taking b large enough in (2.6) we see that $d(\gamma(b), q) > 0$ and thus γ must be timelike. This completes the proof of the claim and the proposition. □

Proposition 2.4. Assume the cosmological time function τ of M is regular and that $K \subset M$ is compact. For each $q \in K$ let $\gamma_q: (0, \tau(q)] \to M$ be a maximizing ray from the initial singularity to q in the sense that (2.4) holds. Then $\{\gamma_q'(\tau(q)): q \in K\} \subset T(M)$ is bounded in $T(M)$ (or, which is the same thing, $\{\gamma_q'(\tau(q)): q \in K\}$ has compact closure in $T(M)$).

Proof. The proof is similar to the last proposition and again we parametrize curves with respect to a complete Riemannian metric on M. If $\{\gamma_q'(\tau(q)): q \in K\}$ is not bounded then there exist past inextendible timelike rays $\gamma_\ell: [0, \infty) \to M$, parametrized with respect to h arc length, which satisfy

$$d(\gamma_\ell(b), \gamma_\ell(0)) = \tau(\gamma_\ell(0)) - \tau(\gamma_\ell(b)) \tag{2.8}$$

for all $b \in (0, \infty)$, such that $\gamma_\ell(0) \to q \in K$ and the h-unit vectors $\gamma_\ell'(0)$ converge to an h-unit vector X which is null in the Lorentzian metric. Let $\gamma: [0, \infty) \to M$ be the past inextendible null geodesic parametrized with respect to h arc length, satisfying $\gamma(0) = q$ and $\gamma'(0) = X$. Then γ is necessarily a null ray (otherwise the maximality of the γ_ℓ's would be violated). By (2.8) we have

$$\begin{aligned} d(\gamma(b), \gamma(0)) &= \lim_{\ell\to\infty} d(\gamma_\ell(b), \gamma_\ell(0)) \\ &= \lim_{\ell\to\infty} (\tau(\gamma_\ell(0)) - \tau(\gamma_\ell(b))) = \tau(\gamma(0)) - \tau(\gamma(b)) > 0 \end{aligned}$$

for sufficiently large b. But this contradicts that γ is a null ray. □

2.4. τ is strictly monotonic on causal curves

Proposition 2.5. If the cosmological time function τ is regular then it is a time function in the usual sense, that is, it is continuous and strictly increasing along future-directed causal curves.

Proof. We have already shown that τ is continuous. Let $\sigma: (a, b) \to M$ be a future-directed causal curve and $t_1, t_2 \in (a, b)$ with $t_1 < t_2$. Set $p := \sigma(t_1)$ and $q := \sigma(t_2)$. If $d(p, q) > 0$ then $\tau(q) \geqslant \tau(p) + d(p, q) > \tau(p)$ by the reverse Lipschitz inequality for τ. Thus assume $d(p, q) = 0$. Then there is a null geodesic ray η from p to q. Let γ_p be the timelike ray to p guaranteed by proposition 2.3. Choose a point x on γ_p to the past of p. Then by a 'cutting the corner' argument near p strict inequality holds in the reverse triangle inequality. This strict inequality and $d(p, q) = 0$ imply

$$d(x, q) > d(x, p) + d(p, q) = d(x, p).$$

Hence,

$$\tau(q) - \tau(p) \geqslant d(x, q) > d(x, p) = \tau(p) - \tau(x)$$

which implies $\tau(p) > \tau(q)$, as desired. □

Recall that for a closed subset $S \subset M$, the future Cauchy horizon $H^+(S)$ is by definition a future boundary of the domain of dependence $D^+(S)$,

$$H^+(S) = \overline{D^+(S)} - I^-(D^+(S)).$$

$H^-(S)$ is defined analogously. If $S \subset M$ is edgeless and acausal, then S is called a partial Cauchy surface and if in addition $H^+(S) = \emptyset$ then S is called a future Cauchy surface (see [9], ch 6 for details). We can now state the following corollary.

Corollary 2.6. If the cosmological time function τ is regular then the level sets $S_a := \{q : \tau(q) = a\}$ (if nonempty) are future Cauchy surfaces.

Proof. As observed in proposition 2.2, S_a is edgeless. The acausality of S_a follows immediately from proposition 2.5. Suppose $H^+(S_a) \neq \emptyset$. Let η be a past inextendible null geodesic generator of $H^+(S_a)$ with future end point $q \in H^+(S_a)$ (cf [9], proposition 6.5.3, p 203). Since $q \in I^+(S_a)$, $\tau(q) > a$. But then, since $\tau \to 0$ along η and τ is continuous, there is a point p on η such that $\tau(p) = a$, i.e. η meets S_a, which cannot happen. □

Simple examples show that the level sets S_a need not be Cauchy, i.e. $H^-(S_a)$ need not be empty.

3. Other regularity properties of τ and its level sets

A continuous function u defined on an open subset U of $\mathbb{R}^n$ is semiconvex if and only if for each point $x \in U$ there is a smooth function f defined near x so that $u + f$ is convex in a neighbourhood of x. Using lemma 3.2 below it is not hard to check that the class of semiconvex functions is closed under diffeomorphisms between open subsets of $\mathbb{R}^n$ and therefore the definition of semiconvex extends to smooth manifolds (cf [3]). By a well known theorem of Aleksandrov a convex function has first and second derivatives almost everywhere and thus a semiconvex function has the same property. (For a beautiful recent proof see [5] theorem A.2, p 56).

Proposition 3.1. If the cosmological time function τ is regular on (M, g) then it is semiconvex and thus its first and second derivatives exist at almost all points of M.

If f is a smooth function on an open subset of $\mathbb{R}^n$ then denote by $D^2 f$ the matrix of second partial derivatives of f. Let I be the $n \times n$ identity matrix. For a constant c let $D^2 f(x) \leqslant cI$ mean that $cI - D^2 f(x)$ is positive semidefinite. Also recall that if u is continuous then a smooth function φ is a lower support function for u at x_0 iff both u and φ are defined in a neighbourhood of x_0, $u(x_0) = \varphi(x_0)$ and $\varphi \leqslant u$ near x_0. The proof of the proposition is based on the following lemma.

Lemma 3.2. Let $U \subset \mathbb{R}^n$ be convex and let $u : U \to \mathbb{R}$ be continuous. Assume for some constant c and all $q \in U$ that u has a lower support function ϕ_q at q so that $D^2\varphi_q(x_0) \geqslant cI$. Then $u - c\|x\|^2/2$ is convex in U and therefore u is semiconvex.

Proof. While in some circles this is a well known folk theorem, the only explicit reference we know is [1], section 2. □

Proof of proposition 3.1. For any point $q \in M$ let $\gamma_q\colon (0, \tau(q)] \to M$ be a geodesic ray realizing the distance from the initial singularity to q as in proposition 2.3. Define a function ϕ_q on $I^+(\gamma_q(\tau(q)/2))$ by

$$\phi_q(x) := \tau(q)/2 + d(\gamma_q(\tau(q)/2), x).$$

By proposition 2.3, γ_q realizes the distance between any two of its points and thus $d(\gamma_q(t), q) = \tau(q) - t$ for $t \in (0, \tau(q)]$. Hence

$$\phi_q(q) = \tau(q)/2 + d(\gamma_q(\tau(q)/2), q) = \tau(q).$$

By the reverse Lipschitz inequality for τ, if $x \in I^+(\gamma_q(\tau(q)/2))$

$$\tau(x) - \tau(\gamma_q(\tau(q)/2)) \geqslant d(\gamma_q(\tau(q))/2, x),$$

which implies $\tau(x) \geqslant \phi_q(q)$ and thus ϕ_q is a lower support function for τ at q.

Also as γ_q maximizes the distance between its points the segment $\gamma_q|_{[\tau(q)/2,\tau(q)]}$ will be free of cut points. Thus the map $x \mapsto d(\gamma_q(\tau(q)/2), x)$ is smooth in a neighbourhood of q. This implies ϕ_q is smooth near q. By standard comparison theorems (see e.g. [2, 6]) it is possible to give upper and lower bounds for the Hessian (defined in terms of the metric connection of (M, g)) of $x \mapsto d(\gamma_q(\tau(q)/2), x)$ just in terms of upper and lower bounds of the timelike sectional curvatures of two planes containing $\gamma_q'(t)$ for $t \in [\tau(q)/2, \tau(q)]$ and the length $\tau(q)/2$ of $\gamma_q|_{[\tau(q)/2,\tau(q)]}$. The same Hessian bound will hold for ϕ_p.

Now let $K \subset M$ be compact. Then by proposition 2.4 the vectors $\gamma_q'(\tau(q))$ for $q \in K$ are all contained in some compact set $\widehat{K}$ of the tangent bundle of M. Therefore there is a compact set $K_1 \subset M$ that will contain all the segments $\gamma_q|_{[\tau(q)/2,\tau(q)]}$ with $q \in K$ and a compact set $\widehat{K}_1 \subset T(M)$ that will contain all the tangent vectors to these segments. Therefore there are uniform upper and lower bounds for both the sectional curvatures of two planes containing a tangent vector to all of the segments $\gamma_q|_{[\tau(q)/2,\tau(q)]}$ and also the lengths $\tau(q)/2$ of these segments. It follows that there are uniform two-sided bounds on the Hessians for the support functions ϕ_q for $q \in K$. Therefore given any point q_0 and a compact coordinate neighbourhood K of q_0, by writing out the two-sided Hessian bounds in terms of the coordinates we find that the lower support functions ϕ_q, $q \in K$, to τ will also satisfy two-sided bounds on the Hessian $D^2\phi_q(q)$ with respect to the coordinates. Therefore lemma 3.2 implies τ is semiconvex near q. As q was any point of M this completes the proof. □

We now consider further the regularity of the level sets $S_a := \{q\colon \tau(q) = a\}$ of the cosmological time function. To do this it is convenient to work in some special coordinate systems. Let q be any point of M and let N_0 be a smooth spacelike hypersurface passing through q. Let $\left(x^1, \ldots, x^{n-1}\right)$ be local coordinates on N_0 centred at q and let x^n be the signed Lorentzian distance (with x^n positive to the future of N_0 and negative to the past). Then near q, $\left(x^1, \ldots, x^n\right)$ is a local coordinate system so that the form of the metric in this coordinate system is

$$g = \sum_{A,B=1}^{n} g_{AB}\, \mathrm{d}x^A\, \mathrm{d}x^B = \sum_{i,j=1}^{n-1} g_{ij}\, \mathrm{d}x^i\, \mathrm{d}x^j - \left(\mathrm{d}x^n\right)^2.$$

Call such a coordinate system an adapted coordinate system centred at q. Then for any spacelike hypersurface N of M through q we have that locally N can be parametrized as the graph of a function f; that is,

$$F_f\left(x^1, \ldots, x^{n-1}\right) = \left(x^1, \ldots, x^{n-1}, f\left(x^1, \ldots, x^{n-1}\right)\right). \tag{3.1}$$

Proposition 3.3. Let the cosmological time function τ of (M, g) be regular and for $a \in (0, \infty)$ let $S_a = \{x : \tau(x) = a\}$ be a nonempty level set of τ. Then for any $q \in S_a$ and every adapted coordinate system $x^1, \dots, x^n$ centred at q there is a local parametrization of S_a of the form (3.1) for a unique function f defined on a neighbourhood of the origin in $\mathbb{R}^{n-1}$. This function f is semiconcave (that is $-f$ is semiconvex) and therefore it is locally Lipschitz and its first and second derivatives exist almost everywhere.

Proof. The existence and uniqueness of the function f is elementary, and follows from the fact that S_a is an acausal hypersurface. We are now going to construct upper support functions for S_a at each of its points. For any $p \in S_a$ let $\gamma_p : (0, \tau(p)] \to M$ be a ray that realizes the distance to the initial singularity of M in the sense of proposition 2.3. Then define

$$\Sigma_p := \{x \in I^+(\gamma_p(a/2)) : d(\gamma_p(a/2), x) = a/2\}.$$

That is, Σ_p is the future Lorentzian distance sphere of radius $a/2$ about the point $\gamma_p(a/2)$. Using the fact that γ_p realizes the distance between any two of its points and that $\gamma_p(\tau(a)) = p$ we see $d(\gamma_p(a/2), p) = d(\gamma_p(a/2), \gamma_p(a)) = a/2$ so that p is in Σ_p. Also using the reverse Lipschitz inequality for τ, if $x \in \Sigma_p$ then

$$\tau(x) \geqslant \tau(\gamma_p(a/2)) + d(\gamma_p(a/2), x) = \tfrac{1}{2}a + \tfrac{1}{2}a = a.$$

Thus every point of Σ_p is in the causal future of S_a. As γ_p is maximizing, the segment $\gamma_p|_{[a/2,a]}$ will be free of conjugate points and therefore the Σ_p is a smooth hypersurface in a neighbourhood of p. Now let $K \subset S_a$ be a compact set. Then by proposition 2.3 the set $\{\gamma_p'(a) : p \in K\}$ has compact closure in $T(M)$. Therefore an argument like that used in the proof of proposition 3.1 (based on elementary comparison theory) implies that if $h_p^{\Sigma_p}$ is the second fundamental form of Σ_p at the point p then $h_p^{\Sigma_p}$ satisfies a uniform two-sided bound for $p \in K$ (or, which is the same thing, the absolute values of the principal curvatures of Σ_p at the point p are uniformly bounded for $p \in K$).

For $p \in S_a$ sufficiently close to q we can parametrize Σ_p by a function F_{f_p} with F_{f_p} defined as in (3.1). As the hypersurfaces Σ_p are in the causal future of S_a the functions f_p satisfy $f_p \geqslant f$ near p and thus they are upper support functions for f near p. The bound on the second fundamental forms of the Σ_p's can be translated into a bound on the Hessians $D^2 f_p$ (for the details of this calculation see [1]). Therefore lemma 3.2 implies $-f$ is semiconvex. This completes the proof. □

4. Examples

4.1. A globally hyperbolic spacetime with τ finite valued but discontinuous

Let $\varphi : \mathbb{R} \to [0, \infty)$ be a smooth function with support in the interval $\left[\frac{1}{4}, \frac{1}{2}\right]$ and with $\int_{-\infty}^{\infty} \varphi(t)\,\mathrm{d}t = \int_{1/4}^{1/2} \varphi(t)\,\mathrm{d}t = 2$. Define a function Φ on the upper half-plane $M := \{(x, y) : y > 0\}$ by

$$\Phi(x, y) = \begin{cases} 1 + \dfrac{1}{x}\varphi\left(\dfrac{y}{x}\right), & x > 0 \\ 1, & x \leqslant 0. \end{cases}$$

Let g be the Lorentzian metric on M given by

$$g := \mathrm{d}x^2 - \Phi(x, y)^2\,\mathrm{d}y^2.$$

Then this metric is smooth on M and using $\int_{1/4}^{1/2} \varphi(t)\,\mathrm{d}t = 2$ it is not hard to check that for any $a > 0$ the length of the timelike curve $c_a\colon [a/4, a/2] \to M$ given by $c_a(t) := (a, t)$ has Lorentzian length $L(c_a) = \int_{a/4}^{a/2} \Phi(a, t)\,\mathrm{d}t = 2 + a/4$ (see figure 1). Let $g_0 = \mathrm{d}x^2 - \mathrm{d}y^2$ be the standard flat Lorentzian metric on M and let W be the open wedge $W := \{(x, y)\colon x > 0, y > 0, x/4 < y < x/2\}$. Then $g = g_0$ outside of W. If $F := I^+(W) \setminus W = \{(x, y)\colon y > 0, -y < x \leqslant y/2\}$ then, using that the segments c_a all have Lorentzian length greater than 2, we see that

$$\tau(x, y) > 2 \qquad \text{for all} \quad (x, y) \in F.$$

But for $(x, y) \notin I^+(W)$ the existence time of (x, y) is the distance of (x, y) from the x-axis in the usual metric g_0, that is

$$\tau(x, y) = y \qquad \text{for all} \quad (x, y) \in M \setminus I^+(W).$$

This implies that τ is discontinuous at each point of the segment $\{(x, y)\colon -2 < x < 0, y = -x\}$. But the spacetime (M, g) is globally hyperbolic and has finite existence times.

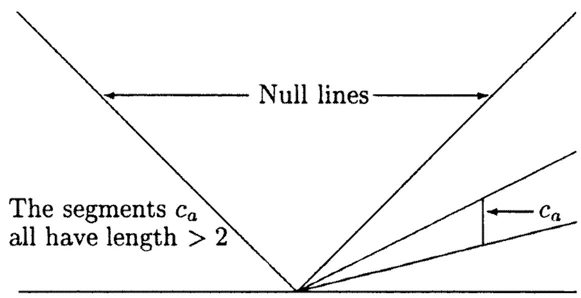

Figure 1. A globally hyperbolic spacetime with τ finite valued but discontinuous.

4.2. Non-strongly causal spacetimes with τ finite valued

Consider the well known example of a spacetime which is causal but not strongly causal (cf [9], p 193, figure 38). In this example, which is a cylinder with slits, it is easily verified that τ is finite valued.

If we are willing to drop the requirement that the metric of (M, g) is smooth, but only of class C^1 then there is an example of a Lorentzian metric on a cylinder that has τ finite valued, but which has a closed causal curve (which turns out to be a null geodesic). This example, which we now describe, is used in the next subsection to construct a spacetime with a nonregular τ such that $\tau \to 0$ along all past inextendible timelike geodesics.

Let the circle S^1 (which we think of as $\mathbb{R}$ modulo 2π) have coordinate x and for any $\alpha > \frac{1}{2}$ define a metric on the space $M := S^1 \times \mathbb{R}$ by

$$g := \mathrm{d}x\,\mathrm{d}t + |t|^{2\alpha}\mathrm{d}x^2 = \mathrm{d}x\big(\mathrm{d}t + |t|^{2\alpha}\,\mathrm{d}x\big).$$

At each point the null directions are defined by $\mathrm{d}x = 0$ and $\mathrm{d}t + |t|^{2\alpha}\,\mathrm{d}x = 0$. If the direction of $\partial/\partial t$ is used as the direction of increasing time then the only closed causal curve is the curve $\{t = 0\}$.

A past inextendible causal curve will either diverge along the cylinder to $t = -\infty$ or be asymptotic to the the null geodesic $\{t = 0\}$; see figure 2. We now show that any past inextendible causal curve asymptotic to $\{t = 0\}$ starting at (x_0, t_0) has length bounded just

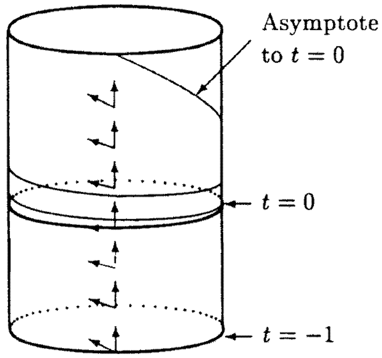

Figure 2. A non-causal spacetime with τ finite valued.

in terms of (x_0, t_0). In doing this it is convenient to work on the universal cover of the cylinder, that is $\mathbb{R}^2$. And in doing the preliminary part of the calculation it is no harder to work with a slightly more general class of metrics. Let $f(x)$ be any smooth positive function defined on the real line (in our example $f(x) \equiv 1$) and let $\varphi(t)$ be a C^1 function so that $t = 0$ is the only zero of φ (in our example $\varphi(t) = |t|^\alpha$) again defined on the real line. As $t = 0$ is the only zero of φ it does not change sign on $(0, \infty)$ and we assume that $\varphi(t) > 0$ on $(0, \infty)$. Define a Lorentzian metric on $\mathbb{R}^2$ by

$$g_0 = \mathrm{d}x\, \mathrm{d}t + \varphi(t)^2 f(x)^2\, \mathrm{d}x^2 = \mathrm{d}x\big(\mathrm{d}t + \varphi(t)^2 f(x)^2\, \mathrm{d}x\big)$$

and use the time orientation so that $\partial/\partial t$ points to the future. At each point the null directions are defined by $\mathrm{d}x = 0$ and $\mathrm{d}t + \varphi(t)^2 f(x)^2\, \mathrm{d}x = 0$. From this it follows that $\{t = 0\}$ is a null geodesic and that every past inextendible causal curve c either is divergent with $t \to -\infty$ along c or c remains in the closed upper half-plane defined by $t \geqslant 0$ and c is asymptotic to the null geodesic $\{t = 0\}$ in such a way that x is increasing monotonically along c.

Now let c be a past inextendible causal curve starting at the point (x_0, t_0) and so that c is asymptotic to the null geodesic $\{t = 0\}$. Then c has a parametrization of the form $c(t) = (x(t), t)$ defined on $(0, t_0]$. As this curve is causal we have (using the notation $\dot{x} = \mathrm{d}x/\mathrm{d}t$),

$$\begin{aligned} 0 &\geqslant g_0(c'(t), c'(t)) = \dot{x} + \varphi(t)^2 f(x)^2 \dot{x}^2 \\ &= \left(\dot{x}\varphi(t) f(x) + \frac{1}{2\varphi(t) f(x)}\right)^2 - \frac{1}{4\varphi(t)^2 f(x)^2} \\ &\geqslant \frac{-1}{4\varphi(t)^2 f(x)^2}, \end{aligned}$$

and thus

$$\left|\dot{x} + \varphi(t)^2 f(x)^2 \dot{x}^2\right| \leqslant \frac{1}{4\varphi(t)^2 f(x)^2}. \tag{4.1}$$

As c is asymptotic to $\{t = 0\}$ it follows that $t \geqslant 0$ and thus also $\varphi(t) \geqslant 0$ along c. Thus the Lorentzian length of c satisfies

$$L(c) = \int_0^{t_0} \sqrt{\dot{x} + \varphi(t)^2 f(x)^2 \dot{x}^2}\, \mathrm{d}t \leqslant \int_0^{t_0} \frac{\mathrm{d}t}{2\varphi(t) f(x(t))}$$

where the inequality follows from using the bound in (4.1). Now letting $\varphi(t) = |t|^\alpha$ with $\frac{1}{2} < \alpha < 1$ and $f(x) \equiv 1$ then this leads to the bound $L(c) \leqslant t_0^{1-\alpha}/(2(1-\alpha))$ as required.

Now if we let $M := \{(x,t) \in S^1 \times \mathbb{R}: t > -1\}$ then the bound on the length of curves asymptotic to $\{t = 0\}$ just given implies that if $\frac{1}{2} < \alpha < 1$ and M has the metric $g = \mathrm{d}x\,\mathrm{d}t + |t|^{2\alpha}\,\mathrm{d}x^2$ then (M, g) has τ finite valued, but τ does not go to zero along the inextendible causal curves asymptotic to $\{t = 0\}$. It is worth noting that in this example τ is continuous.

We know of no example where τ is finite, there are closed causal curves, and the metric is smooth.

4.3. A non-regular τ going to zero along all past inextendible causal geodesics

The definition of τ being regular requires that τ go to zero along all past inextendible causal curves. It is natural to ask whether this can be weakened to only requiring that τ go to zero along all past inextendible causal geodesics. Here we give an example to show that this is not the case. Like the example just given the metric in this example is of class C^1 but not C^2.

First let $(M_2, g_2) = \left(S^1 \times (-1, \infty), \mathrm{d}x\,\mathrm{d}t + |t|^{2\alpha}\,\mathrm{d}x^2\right)$ be the two-dimensional example just given (so that $\frac{1}{2} < \alpha < 1$) and set

$$f(y) = \mathrm{e}^{y^2} - 1.$$

Note that $f(0) = 0$ and $f(y) > 0$ for $y \neq 0$. Let $M := \{(x, y, t) \in S^1 \times \mathbb{R} \times \mathbb{R}: t > -1\}$ with the metric

$$g := \mathrm{d}y^2 + \mathrm{e}^{2y}\left(\mathrm{d}x\,\mathrm{d}t + \left(|t|^{2\alpha} + f(y)\right)\mathrm{d}x^2\right).$$

Then the two-dimensional submanifold defined by $y = 0$ is isometric to (M_2, g_2). Moreover this submanifold is totally umbilic in (M, g) and so no curve in (M_2, g_2) can be a geodesic in (M, g). Let η be the null geodesic defined by $\{t = 0, y = 0\}$. The following is easy to verify.

Lemma 4.1. Let c be a past inextendible causal curve in (M, g). Then one of the following holds:

(i) $t \to -1$ along c and c runs off of the 'bottom' of M (that is the part of the boundary defined by $t = -1$).

(ii) $t \to 0$ along c and c is asymptotic to the closed null curve η. Neither η nor any curve asymptotic to it are geodesics. □

Harder to show is:

Lemma 4.2. Let c be a past inextendible curve starting at the point (x_0, y_0, t_0) which is asymptotic to the null curve η. Then there is a finite upper bound on the length of c only depending on t_0.

Proof. Analogous to what was done in the previous example, there is a parametrization of c of the form $c(t) = (x(t), y(t), t)$ with $t \in (0, t_0]$. As c is causal $g(c'(t), c'(t)) \leqslant 0$ which implies

$$0 \geqslant g(c'(t), c'(t)) = \dot{y}^2 + \mathrm{e}^{2y}\big(\dot{x} + \big(|t|^{2\alpha} + f(y)\big)\dot{x}^2\big)$$
$$= \dot{y}^2 + \mathrm{e}^{2y}\left(\frac{1}{2\sqrt{|t|^{2\alpha} + f(y)}} + \sqrt{|t|^{2\alpha} + f(y)}\,\dot{x}\right)^2 - \frac{\mathrm{e}^{2y}}{4\big(|t|^{2\alpha} + f(y)\big)}$$
$$\geqslant -\frac{\mathrm{e}^{2y}}{4\big(|t|^{2\alpha} + f(y)\big)}$$

and thus

$$\sqrt{|g(c'(t), c'(t))|} \leqslant \frac{\mathrm{e}^y}{2\sqrt{|t|^{2\alpha} + f(y)}}.$$

If $y \leqslant 3$ then,

$$\frac{\mathrm{e}^y}{2\sqrt{|t|^{2\alpha} + f(y)}} \leqslant \frac{\mathrm{e}^3}{2\sqrt{|t|^{2\alpha}}} = \frac{\mathrm{e}^3}{2|t|^\alpha} \leqslant \frac{12}{|t|^\alpha}.$$

If $y \geqslant 3$ then $\mathrm{e}^y \leqslant \sqrt{\mathrm{e}^{y^2} - 1} = \sqrt{f(y)}$ and so

$$\frac{\mathrm{e}^y}{2\sqrt{|t|^{2\alpha} + f(y)}} \leqslant \frac{\mathrm{e}^y}{2\sqrt{f(y)}} \leqslant \frac{1}{2}.$$

Putting these together we have,

$$\sqrt{|g(c'(t), c'(t))|} \leqslant \max\left(\frac{12}{|t|^\alpha}, \frac{1}{2}\right),$$

which implies

$$L(c) = \int_0^{t_0} \sqrt{|g(c'(t), c'(t))|}\,\mathrm{d}t \leqslant \int_0^{t_0} \max\left(\frac{12}{|t|^\alpha}, \frac{1}{2}\right)\mathrm{d}t,$$

which is finite as $\frac{1}{2} < \alpha < 1$. This gives the required bound and completes the proof of the lemma. □

Therefore in the spacetime (M, g) all past inextendible curves c either have $t \to -1$ along c (in which case $\tau \to 0$ along c) or c is asymptotic to the null curve η (in which case τ does not go to zero). As no geodesics are asymptotic to η this gives an example of a spacetime where $\tau \to 0$ along all past inextendible causal geodesics, but which is not regular. It would be interesting to know whether there is a smooth example where this happens.

Acknowledgments

Joe Fu pointed out the relevance of [3] to the results of section 3. Some comments of Jeeva Anandan were also useful in writing this paper. LAL was supported in part by NFR, contract no F-FU 4873-307. RH supported in part by by DEPSCoR grants N00014-94-1-1163 and DAAH-04-96-1-0326.

Note added in proof.

Previously Wald and Yip introduced the cosmological time function (or rather its time dual, which they referred to as the 'maximum lifetime function') in order to study the existence of synchronous coordinates in a neighbourhood of a spacelike singularity, see 1981 *J. Math. Phys.* **22** 2659–65. We are grateful to R Wald for bringing this article to our attention.

References

[1] Andersson L, Galloway G J and Howard R 1996 A strong maximum principle for weak solutions of quasi-linear elliptic equations with applications to Lorentzian geometry *Commun. Pure Appl. Math.* to be published

[2] Andersson L and Howard R 1998 Comparison and rigidity theorems in semi-Riemannian geometry *Commun. Anal. Geom.* to appear

[3] Bangert V 1979 Analytische Eigenschaften konvexer Funktionen auf Riemannschen Mannigfaltigkeiten *J. Reine Angew. Math.* **307/308** 309–24

[4] Beem J K, Ehrlich P E and Easley K L 1996 *Global Lorentzian Geometry* 2nd edn (*Pure and Applied Mathematics 202*) (New York: Dekker)

[5] Crandall M G, Ishii H and Lions P-L 1992 Users guide to viscosity solutions of second order partial differential equations *Bull. Am. Math. Soc.* **27** 1–67

[6] Eschenburg J-H 1987 Comparison theorems and hypersurfaces *Manuscripta Math.* **59** 295–323

[7] Eschenburg J-H and Galloway G J 1992 Lines in spacetimes *Commun. Math. Phys.* **148** 209–16

[8] Galloway G J and Horta A 1996 Regularity of Lorentzian Busemann functions *Trans. Am. Math. Soc.* **348** 2063–83

[9] Hawking S and Ellis G F W 1973 *The Large Scale Structure of Space-Time* (Cambridge: Cambridge University Press)

[10] Penrose R 1972 *Techniques of Differential Topology in Relativity (CBMS Regional Conference Series in Applied Mathematics 7)* (Philadelphia, PA: SIAM)