# Null Geodesics, Raychaudhuri Equation, Trapped Surfaces, and Penrose Singularity Theorem 

Miguel Socolovsky ${ }^{\text {© }}$<br>Instituto de Ciencias Nucleares, Universidad Nacional Autónoma de México, Cd. Universitaria, Ciudad de México, México<br>Email: socolovs@nucleares.unam.mx

How to cite this paper: Socolovsky, M. (2022) Null Geodesics, Raychaudhuri Equation, Trapped Surfaces, and Penrose Singularity Theorem. Journal of High Energy
Physics, Gravitation and Cosmology, 8, 536-557.
https://doi.org/10.4236/jhepgc.2022.83039

Received: April 11, 2022
Accepted: June 26, 2022
Published: June 29, 2022

Copyright © 2022 by author(s) and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
http://creativecommons.org/licenses/by/4.0/


Open Access


#### Abstract

We review the concept of congruence of null geodesics, the Raychaudhuri equation for the expansion, its harmonic oscillator version and associated "quantum" propagator, the role of the equation in the derivation of the Pe nrose singularity theorem, the definition of trapped surfaces, and the derivation of the theorem itself.


## Keywords

Null Geodesics, Trapped Surfaces, Singularity Theorem

## 1. Introduction

The Penrose singularity theorem (P.S.T.) of 1965 [1] is one of the most important steps in the development of General Relativity theory since its inception by Einstein in 1915, just half a century before. Important for the conception of the theorem, though not cited by its author in its original formulation, was the Raychaudhuri equation (R.E.) [2] for the expansion function associated with the evolution of non-spacelike geodesic congruences (the null case being relevant for the P.S.T.) and its prediction of the existence of caustics or focal points: singularities of the congruences but not of the spacetime itself (points where the expansion diverges and the R.E. looses its validity).

The P.S.T. strongly involves global Lorentzian geometry and topology [3] [4] [5] [6] [7]. An extensive review of the theorem with all of its details can be found in [8]; with a more informative (i.e. less detailed) but equally conceptual version by the same author in [9]. The two crucial new concepts of the theorem are: the definition of spacetime singularity itself as geodesic incompleteness i.e. the impossibility of an affinely parametrized geodesic to extend beyond a finite value of
the affine parameter, and therefore the abrupt disappearance of matter or radiation into the "nothing"; and the introduction of the concept of trapped surface: a co-dimension 2 (2-dimensional surface in the case of a 4 -dimensional spacetime) spacelike submanifold (compact or not) on which the expansions of both ingoing and outgoing orthogonally emitted geodesic congruences have the same sign (negative for the case of gravitational collapse or positive for the case of matter or radiation creation (big bang)).

It is important to emphasize the well-known fact that all the "classical" singularity theorems, P.S.T., as well as the theorems in [10] and [11], are purely classical in the technical sense i.e. they exclude any reference to quantum mechanics, obviously due to the still missing complete theory of quantum gravity [12]. Some attempt to see a modification of the classical predictions with, however, opposite conclusions are, e.g., those in [13] and [14].

The purpose of the present review is to shortcircuit as much as possible the path from the formulation of the R.E. to the announcement and proof of the P.S.T. with the minimum necessary ingredients from causality theory and global Lorentzian geometry explicitly stated. Of great help in this path have been the already cited reference [8] and the textbook [15]. An extra ingredient is the resume of an attempt of the author [16] to define Feynman propagators associated with the "quantum" evolution of the expansion coefficient corresponding to null geodesic congruences in black holes; in particular in the simplest example: the Schwarzschild-Kruskal-Szekeres (S.K.S.) black hole.

Section 2 is devoted to the R.E.; Section 3 to concepts (definitions and propositions) of causality theory; and Section 4 to the proof of the P.S.T.

## 2. Null Geodesics Congruences. Raychaudhuri Equation

Let $(M, \mathbf{g})$ be a 4-dimensional time-oriented globally hyperbolic spacetime; that is, a connected, Hausdorff, pseudo-Riemannian manifold $M$ with metric $\mathbf{g}$, in local coordinates $x^{\mu}, \mu=0,1,2,3$, given by $\mathbf{g}=g_{\mu \nu} d x^{\mu} d x^{\nu}$. Let $\mathscr{U}^{2}$ be an open subset in $M$. A congruence in $\mathscr{U}$ is a family (set) of curves in $\mathscr{U}$ such that through each point $p \in \mathscr{C}$ passes one and only one curve of the family. As a consequence, curves of the congruence do not intersect each other; when this happens i.e. there exists a focal or conjugate point (see below), the congruence breaks down.

Consider a congruence of null geodesics $\gamma^{\prime} s$ in $\mathbb{Z}$, each affinely parametrized with parameter $\lambda$ and tangent velocity vector field $k^{\mu}=\frac{\partial x^{\mu}}{\partial \lambda}$. The geodesics in the congruence are labelled by a parameter $s$ which allows to define the deviation or separation vector field $\xi^{\nu}=\frac{\partial x^{\nu}}{\partial s}$. The vector fields $k^{\mu}$ and $\xi^{v}$ satisfy the following set of equations:

$$
\begin{gather*}
k^{2}=0  \tag{1}\\
(k \cdot D)\left(k^{\mu}\right)=0 \tag{2}
\end{gather*}
$$

$$
\begin{equation*}
k \cdot \xi=0 \tag{3}
\end{equation*}
$$

and

$$
\begin{equation*}
(k \cdot D)\left(\xi^{\mu}\right)=(\xi \cdot D)\left(k^{\mu}\right) \tag{4}
\end{equation*}
$$

where, for arbitrary vectors $V$ and $W, V \cdot W=V^{\mu} W_{\mu}=g_{\mu \nu} V^{\mu} W^{\nu}$, $(V \cdot D)\left(W^{\mu}\right)=V^{\nu} D_{\nu} W^{\mu} \equiv V^{\nu} W_{; \nu}^{\mu}$, and $D$ is the covariant derivative with respect to the Levi-Civita (L.C.) connection $\Gamma_{v \rho}^{\mu}=\frac{1}{2} g^{\mu \sigma}\left(\partial_{\nu} g_{\rho \sigma}+\partial_{\rho} g_{v \sigma}-\partial_{\sigma} g_{v \rho}\right)$ associated with the metric $\mathbf{g}$.
(1) is the equation that characterizes null geodesics: lightlike tangent vectors.
(2) is the affinely parametrized geodesic equation:

$$
\begin{align*}
(k \cdot D)\left(k^{\mu}\right) & =k^{v} D_{v}\left(k^{\mu}\right)=k^{v}\left(k_{, v}^{\mu}+\Gamma_{v \rho}^{\mu} k^{\rho}\right) \\
& =\frac{\partial x^{v}}{\partial \lambda} \frac{\partial}{\partial x^{v}}\left(\frac{\partial x^{\mu}}{\partial \lambda}\right)+\Gamma_{v \rho}^{\mu} \frac{\partial x^{v}}{\partial \lambda} \frac{\partial x^{\rho}}{\partial x^{\lambda}}  \tag{5}\\
& =\frac{\partial^{2} x^{\mu}}{\partial \lambda^{2}}+\Gamma_{v \rho}^{\mu} \frac{\partial x^{v}}{\partial \lambda} \frac{\partial x^{\rho}}{\partial \lambda} \\
& \equiv \ddot{x}^{\mu}+\Gamma_{v \rho}^{\mu} \dot{x}^{\prime} \dot{x}^{\rho}=0
\end{align*}
$$

with $\quad \dot{x}^{\alpha}=\left.\frac{\partial x^{\alpha}}{\partial \lambda}\right|_{s}$.
(3) expresses the fact that coordinates can be chosen such that the tangent and deviation vectors become orthogonal. In fact,

$$
\begin{aligned}
\frac{\partial}{\partial \lambda}\left(\xi^{\alpha} k_{\alpha}\right) & =\frac{\partial}{\partial x^{\mu}}\left(\xi^{\alpha} k_{\alpha}\right) \frac{\partial x^{\mu}}{\partial \lambda}=k^{\mu}\left(\xi^{\alpha} k_{\alpha}\right)_{; \mu}=k^{\mu} \xi_{; \mu}^{\alpha} k_{\alpha}+k^{\mu} \xi^{\alpha} k_{\alpha ; \mu} \\
& =\xi_{; \mu}^{\alpha} k^{\mu} k_{\alpha}+\xi_{\alpha} k \cdot D\left(k^{\alpha}\right)=\xi_{; \mu}^{\alpha} k^{\mu} k_{\alpha}=k_{; \mu}^{\alpha} \xi^{\mu} k_{\alpha} \\
& =\frac{1}{2} \xi^{\mu}\left(k^{\alpha} k_{\alpha}\right)_{; \mu}=\frac{1}{2} \xi^{\mu}\left(k^{2}\right)_{, \mu}=0
\end{aligned}
$$

where in the 6th equality we used (4) and in the 7th equality we used the Leibniz rule and the metric character of the L.C. connection. Then $\xi \cdot k$ is constant along each geodesic. It is clear that coordinates can be chosen such that the constant is zero.
(4), the deviation vector equation, which states that the parallel transport of the deviation vector $\xi$ in the direction of the tangent vector $k$ equals the parallel transport of $k$ in the direction of $\xi$, amounts to the statement that the deviation vector is Lie transported along the geodesic, and viceversa, the geodesic velocity is Lie transported along the deviation vector. In fact, from the definition of the Lie derivative and the symmetry of the connection,
$\left(\mathscr{L}_{k} \xi\right)^{\mu}=\xi_{; v}^{\mu} k^{\nu}-k_{j \nu}^{\mu} \xi^{\nu}=\xi_{, v}^{\mu} k^{\nu}-k_{, \nu}^{\mu} \xi^{v}=k \cdot D \xi^{\mu}-\xi \cdot D k^{\mu}$, and $\left(\mathscr{L}_{\xi} k\right)^{\mu}=k_{; \nu}^{\mu} \xi^{\nu}-\xi_{; v}^{\mu} k^{v}=k_{, \nu}^{\mu} \xi^{\nu}-\xi_{, v}^{\mu} k^{\nu}=\xi \cdot D k^{\mu}-k \cdot D \xi^{\mu}$, and using $k_{, s}^{\mu}=\xi_{, \lambda}^{\mu}$, $\xi_{, v}^{\mu} k^{\nu}=\partial_{v} \xi^{\mu} \partial_{\lambda} x^{\nu}=\partial_{\lambda} \xi^{\mu}=\partial_{s} k^{\mu}$, while $k_{, v}^{\mu} \xi^{\nu}=\partial_{v} k^{\mu} \partial_{s} x^{\nu}=\partial_{s} k^{\mu}$ i.e. the first terms of the r.h.s.'s of $\left(\mathscr{L}_{k} \xi\right)^{\mu}$ and $\left(\mathscr{L}_{\xi} k\right)^{\mu}$ are equal; the same holds for the second terms, so $\left(\mathscr{L}_{k} \xi\right)^{\mu}=\left(\mathscr{L}_{\xi} k\right)^{\mu}$; but $\mathscr{L}_{k} \xi=[k, \xi]=-[\xi, k]=-\mathscr{L}_{\xi} k$, so

$$
\begin{equation*}
\mathscr{L}_{k} \xi=\mathscr{L}_{\xi} k=0 \tag{6}
\end{equation*}
$$

and then $(k \cdot D)\left(\xi^{\mu}\right)=(\xi \cdot D)\left(k^{\mu}\right)$. (In terms of $\frac{\partial x^{\mu}}{\partial \lambda}$ and $\frac{\partial x^{\mu}}{\partial s}$, Equation (4) is the identity $X^{\alpha}=X^{\alpha}$ with $X^{\alpha}=\frac{\partial^{2} x^{\alpha}}{\partial t \partial s}+\Gamma_{v \rho}^{\alpha} \frac{\partial x^{\nu}}{\partial t} \frac{\partial x^{\rho}}{\partial s}$.)

## Geodesic deviation equation

(4) can be used to derive the geodesic deviation equation

$$
\begin{equation*}
\frac{D^{2} \xi^{\mu}}{d \lambda^{2}}+R_{\alpha \rho \sigma}^{\mu} k^{\alpha} \xi^{\rho} k^{\sigma}=0 \tag{7}
\end{equation*}
$$

where $\frac{D^{2}}{d \lambda^{2}}=\frac{D}{d \lambda}\left(\frac{D}{d \lambda}\right)=\left(\frac{D}{d \lambda}\right)^{2}$ with $\frac{D}{d \lambda}=k \cdot D=k^{\sigma} D_{\sigma}$ and so $\frac{D^{2} \xi^{\mu}}{d \lambda^{2}}=(k \cdot D)^{2} \xi^{\mu} ;$ and

$$
\begin{equation*}
R_{\alpha \rho \sigma}^{\mu}=\partial_{\sigma} \Gamma_{\rho \alpha}^{\mu}-\partial_{\rho} \Gamma_{\sigma \lambda}^{\mu}+\Gamma_{\sigma \varphi}^{\mu} \Gamma_{\rho \alpha}^{\varphi}-\Gamma_{\rho \varphi}^{\mu} \Gamma_{\sigma \alpha}^{\varphi} \tag{8}
\end{equation*}
$$

is the curvature tensor. (7) is a 2 nd . order linear ordinary differential Jacobi equation for the deviation vector $\xi^{\mu}$ (Jacobi field).

Proof of (7): Define the covariant gradient of the velocity field,

$$
\begin{equation*}
B_{v}^{\mu}:=k_{; v}^{\mu} \tag{9}
\end{equation*}
$$

It is easy to proof that it is orthogonal to $k^{\mu}$ :

$$
\begin{equation*}
k^{\mu} B_{\mu \nu}=k^{\mu} k_{\mu ; \nu}=\frac{1}{2}\left(k^{\mu} k_{\mu}\right)_{; \nu}=\frac{1}{2}\left(k^{2}\right)_{, \nu}=0 \tag{10}
\end{equation*}
$$

Then

$$
\begin{equation*}
\frac{D k^{\mu}}{d \lambda}=\partial_{\lambda} x^{\nu} D_{v} k^{\mu}=k^{\nu} k_{; v}^{\mu}=0 \tag{11}
\end{equation*}
$$

and

$$
\begin{equation*}
\frac{D \xi^{\mu}}{d \lambda}=\partial_{\lambda} x^{v} \xi_{; v}^{\mu}=k^{v} \xi_{; v}^{\mu}=\xi^{v} k_{; v}^{\mu}=\xi^{v} B_{v}^{\mu} \tag{12}
\end{equation*}
$$

so

$$
\begin{equation*}
k_{\mu} \frac{D \xi^{\mu}}{d \lambda}=k_{\mu} B_{v}^{\mu} \xi^{\nu}=0 \tag{13}
\end{equation*}
$$

i.e. $k_{\mu}$ is orthogonal to $\frac{D \xi^{\mu}}{d \lambda}$. (This is consistent with the constancy of $\xi \cdot k$ along each geodesic, which implies $\frac{\partial}{\partial \lambda}(\xi \cdot k)=0$, and so

$$
\begin{gathered}
\left.\frac{D}{d \lambda}(\xi \cdot k)=\frac{D \xi^{\mu}}{d \lambda} \cdot k_{\mu}+\xi^{\mu} \frac{D k_{\mu}}{d \lambda}=k^{\rho}(\xi \cdot k)_{; \rho}=k^{\rho}(\xi \cdot k)_{, \rho}=\frac{\partial}{\partial \lambda}(\xi \cdot k)=0 .\right) \\
\text { Then } \frac{D^{2}}{d \lambda^{2}} \xi^{\mu}=\frac{D}{d \lambda} \frac{D}{d \lambda} \xi^{\mu}=\frac{D}{d \lambda}\left(B_{v}^{\mu} \xi^{v}\right)=\left(\frac{D}{d \lambda} B_{v}^{\mu}\right) \xi^{v}+B_{v}^{\mu} \frac{D}{d \lambda} \xi^{v} \\
\\
=\left(\frac{D}{d \lambda} B_{\sigma}^{\mu}\right) \xi^{\sigma}+B_{v}^{\mu} B_{\sigma}^{v} \xi^{\sigma}=\left(\frac{D}{d \lambda} B_{\sigma}^{\mu}+B_{v}^{\mu} B_{\sigma}^{v}\right) \xi^{\sigma}
\end{gathered}
$$

which implies

$$
\begin{align*}
k_{\mu}\left(\frac{D}{d \lambda}\right)^{2}\left(\xi^{\mu}\right) & =\left(k^{\mu} \frac{D}{d \lambda} B_{v}^{\mu}+k_{\mu} B_{\rho}^{\mu} B_{v}^{\rho}\right) \xi^{v}=k_{\mu}\left(\frac{D}{d \lambda} B_{v}^{\mu}\right) \xi^{v}  \tag{14}\\
& =\left(\frac{D}{d \lambda}\left(k_{\mu} B_{v}^{\mu}\right)-\left(\frac{D}{d \lambda} k_{\mu}\right) B_{v}^{\mu}\right) \xi^{v}=0
\end{align*}
$$

i.e. $k_{\mu}$ is also orthogonal to $\frac{D^{2}}{d \lambda^{2}} \xi^{\mu}$.

Let us compute the r.h.s. in the expresion for $\frac{D^{2}}{d \lambda^{2}} \xi^{\mu}$ :

$$
\begin{aligned}
\frac{D}{d \lambda} B_{v}^{\mu} & =\frac{D}{d \lambda}\left(D_{v} k^{\mu}\right)=k^{\sigma}\left(D_{v} k^{\mu}\right)_{; \sigma}=k^{\sigma} D_{\sigma}\left(D_{v} k^{\mu}\right)=k^{\rho}\left(D_{\rho} D_{v}\left(k^{\mu}\right)\right), \\
B_{\rho}^{\mu} B_{v}^{\rho} & =k_{; \rho}^{\mu} k_{; v}^{\rho}=\left(k^{\rho} k_{; \rho}^{\mu}\right)_{; v}-k^{\rho} k_{; \rho ; v}^{\mu}=\left(k^{\rho} D_{\rho}\left(k^{\mu}\right)\right)_{; v}-k^{\rho} D_{v} D_{\rho}\left(k^{\mu}\right) \\
& =\left((k \cdot D)\left(k^{\mu}\right)\right)_{; v}-k^{\rho} D_{v} D_{\rho}\left(k^{\mu}\right)=-k^{\rho} D_{v} D_{\rho}\left(k^{\mu}\right)
\end{aligned}
$$

where we used the geodesic Equation (2). Then

$$
\begin{aligned}
\frac{D^{2}}{d \lambda^{2}} \xi^{\mu} & =\left(\frac{D}{d \lambda}\right)\left(\frac{D}{d \lambda} \xi^{\mu}\right)=\left(k^{\rho} D_{\rho} D_{v}\left(k^{\mu}\right)-k^{\rho} D_{v} D_{\rho}\left(k^{\mu}\right)\right) \xi^{v} \\
& =k^{\rho}\left(D_{\rho} D_{v}-D_{v} D_{\rho}\right)\left(k^{\mu}\right) \xi^{v}=\left[D_{\rho}, D_{v}\right]\left(k^{\mu}\right) k^{\rho} \xi^{v}
\end{aligned}
$$

which is (7) after using the tensor identity

$$
\begin{equation*}
\left[D_{\mu}, D_{v}\right]\left(W^{\alpha}\right)=R_{\sigma \mu \nu}^{\alpha} W^{\sigma} \tag{15}
\end{equation*}
$$

valid for an arbitrary vector field $W^{\alpha}$.
If

$$
\begin{equation*}
V^{\mu}=\frac{D}{d \lambda} \xi^{\mu} \tag{16}
\end{equation*}
$$

is the relative velocity between nearby geodesics, then

$$
\begin{equation*}
A^{\mu}=\frac{D}{d \lambda} V^{\mu} \tag{17}
\end{equation*}
$$

is the relative acceleration which, according to the geodesic deviation Equation (7) is given by

$$
\begin{equation*}
A^{\mu}=-R_{\alpha \rho \sigma}^{\mu} k^{\alpha} k^{\rho} \xi^{\sigma} \tag{18}
\end{equation*}
$$

$A^{\mu}$ is known as the tidal acceleration between geodesics. According to (18) it is proportional to the curvature.

Since $k^{2}=0, \xi^{\alpha}$ can have a component along $k^{\alpha}$. In fact: $\xi^{\alpha}:=\xi^{\perp \alpha}+$ const. $k^{\alpha}$ implies $0=k^{\mu}\left(\xi_{\mu}^{\perp}+\right.$ const. $\left.k_{\mu}\right)=k^{\mu} \xi_{\mu}^{\perp}$, and the orthogonality of $\xi_{\mu}^{\perp}$ with $k^{\mu}$ subsists. To isolate $\xi^{\perp \mu}$ from $\xi^{\mu}$ we construct the transverse metric, that is, the part of the metric $g_{\mu v}$ orthogonal to $k^{\mu}$. With this aim, we take a null vector field $n=\left(n^{\mu}\right)$ in a direction such that $n \cdot k=+1$. I.e.

$$
\begin{equation*}
n^{2}=g_{\mu \nu} n^{\mu} n^{\nu}=0, g_{\mu \nu} n^{\mu} k^{\nu}=+1 \tag{19}
\end{equation*}
$$

and define the (symmetric) tensor

$$
\begin{equation*}
h_{\mu \nu}:=g_{\mu \nu}-k_{\mu} n_{v}-k_{\nu} n_{\mu} \tag{20}
\end{equation*}
$$

The computation of its trace,
$\operatorname{tr}\left(h_{\mu \nu}\right)=g^{\mu \nu} h_{\mu \nu}=g^{\mu \nu} g_{\mu \nu}-g^{\mu \nu} k_{\mu} n_{v}-g^{\mu \nu} k_{\nu} n_{\mu}=4-2 k \cdot n=4-2=2$, shows that $h_{\mu \nu}$ is the metric of a 2-dimensional surface $\Sigma^{\perp}$ to which $n$ and $k$ are orthogonal:

$$
\begin{gather*}
k^{\mu} h_{\mu \nu}=k^{\mu} g_{\mu \nu}-k^{2} n_{v}-k_{v} k \cdot n=k_{v}-k_{v}=0  \tag{21}\\
n^{\mu} h_{\mu \nu}=n^{\mu} g_{\mu \nu}-(n \cdot k) n_{v}-k_{v} n^{2}=n_{v}-n_{v}=0 \tag{22}
\end{gather*}
$$

The mixed tensor

$$
\begin{equation*}
h_{v}^{\mu}=g^{\mu \alpha} h_{\alpha v}=g^{\mu \alpha}\left(g_{\alpha v}-k_{\alpha} n_{v}-k_{v} n_{\alpha}\right)=\delta_{v}^{\mu}-k^{\mu} n_{v}-k_{v} n^{\mu} \tag{23}
\end{equation*}
$$

acts as a projector operator on $\Sigma^{\perp}$ since

$$
\begin{equation*}
h_{v}^{\mu} h_{\rho}^{\nu}=h_{\rho}^{\mu}, h_{v}^{\mu} k^{\nu}=k^{\mu}-k^{\mu}=0, \text { and } h_{v}^{\mu} n^{v}=n^{\mu}-n^{\mu}=0 . \tag{24}
\end{equation*}
$$

In terms of $B_{v}^{\mu}$, Equation (4) is

$$
\begin{equation*}
\xi_{; v}^{\mu} k^{\nu}=B_{v}^{\mu} \xi^{v} \tag{25}
\end{equation*}
$$

So, $B_{v}^{\mu}$ measures the obstruction for the deviation vector $\xi^{\mu}$ to be parallel transported along the geodesic. Also, $B_{v}^{\mu}$ is orthogonal to the velocity vector $k^{\mu}$ i.e. to the geodesics: $k^{\beta} B_{\alpha \beta}=k^{\beta} k_{\alpha ; \beta}=k \cdot D\left(k_{\alpha}\right)=0$, while $k^{\alpha} B_{\alpha v}=-k^{\alpha} B_{\alpha v}$ since $k^{2}=0$, but not orthogonal to $n^{\mu}$ since $B_{\alpha \beta} n^{\beta}=k_{\alpha ; \beta} n^{\beta}=(n \cdot D)\left(k_{\alpha}\right)$ and $B_{\alpha \beta} n^{\alpha}=k_{\alpha ; \beta} n^{\alpha}=(n \cdot k)_{; \beta}-k_{\alpha} n_{; \beta}^{\alpha}=-k_{\alpha}^{\alpha} B_{\alpha \beta}=0, n^{\beta} n_{; \beta}^{\alpha}$ are $\neq 0$ in general. I.e.

$$
\begin{equation*}
k^{\nu} B_{\mu \nu}=0, k^{\mu} B_{\mu \nu}=0, n^{\alpha} B_{\alpha \beta} \neq 0, n^{\beta} B_{\alpha \beta} \neq 0 \tag{26}
\end{equation*}
$$

As a consequence of (26), $B_{v}^{\mu} \xi^{\nu}$ and by (25) $\xi_{; \nu}^{\mu} k^{\nu}$ have non-vanishing components along $n$ i.e. $B_{v}^{\mu} \xi_{v}$ is not contained in $\Sigma^{\perp}$. This non-transverse part of $\xi_{; v}^{\mu} k^{\nu}$ will be later eliminated projecting $B_{v}^{\mu} \xi^{\nu}$ with $h_{v}^{\mu}$ on this 2-surface.

The transverse part of $\xi^{\mu}$ is given by

$$
\begin{equation*}
\xi^{\perp \mu}=h_{v}^{\mu} \xi^{\nu}=\left(\delta_{v}^{\mu}-k^{\mu} n_{v}-k_{v} n^{\mu}\right) \xi^{\nu}=\xi^{\mu}-(\xi \cdot n) k^{\mu} \tag{27}
\end{equation*}
$$

with covariant derivative along $k^{v}$

$$
\begin{equation*}
(k \cdot D)\left(\xi^{\perp \mu}\right)=\left(\xi^{\perp \mu}\right)_{; v} k^{\nu}=\left(h_{\alpha}^{\mu} \xi^{\alpha}\right)_{; \nu} k^{v}=h_{\alpha}^{\mu} B_{v}^{\alpha} \xi^{v}-n_{\alpha ; \nu} \xi^{\alpha} k^{\nu} k^{\mu} \tag{28}
\end{equation*}
$$

where we have used (3), (26), and the covariant derivative of the transverse metric:

$$
\begin{equation*}
h_{\mu \alpha ; v}=-\left(k_{\mu ; \nu} n_{\alpha}+n_{\alpha ; v} k_{\mu}+k_{\alpha ; \nu} n_{\mu}+k_{\alpha} n_{\mu ; v}\right) \tag{29}
\end{equation*}
$$

So $(k \cdot D)\left(\xi^{\perp \mu}\right)$ also has a non-transverse part given by $-n_{\alpha ; \nu} \xi^{\alpha} k^{\nu} k^{\mu}$. Its projection on $\Sigma^{\perp}$ is

$$
\begin{equation*}
\left((k \cdot D)\left(\xi^{\perp \alpha}\right)\right)^{\perp}=\left(\xi_{; \nu}^{\perp \alpha} k^{\nu}\right)^{\perp}=h_{\mu}^{\alpha}\left(\xi^{\perp \mu}\right)_{; \nu} k^{\nu}:=\left(B^{\perp}\right)_{\sigma}^{\alpha} \xi^{\perp \sigma} \tag{30}
\end{equation*}
$$

where $\left(B^{\perp}\right)_{\sigma}^{\alpha}$ is the transverse part of $B_{v}^{\rho}$ :

$$
\begin{equation*}
\left(B^{\perp}\right)_{\sigma}^{\alpha}=h_{\rho}^{\alpha} h_{\sigma}^{v} B_{v}^{\rho} . \tag{31}
\end{equation*}
$$

(30) governs the purely tranverse behavior of the null geodesic congruence. From (23) and (26),

$$
\begin{equation*}
\left(B^{\perp}\right)_{\sigma}^{\alpha}\left(B^{\perp}\right)_{\alpha}^{\sigma}=B_{\sigma}^{\alpha} B_{\alpha}^{\sigma} \tag{32}
\end{equation*}
$$

In terms of $B, n$, and $k, B^{\perp}$ is given by

$$
\begin{equation*}
\left(B^{\perp}\right)_{\alpha \sigma}=g_{\alpha \beta}\left(B^{\perp}\right)_{\sigma}^{\beta}=B_{\alpha \sigma}-B_{\alpha \nu} n^{\nu} k_{\sigma}-k_{\alpha} n^{\rho} B_{\rho \sigma}+k_{\alpha} k_{\sigma} B_{\rho \nu} n^{\rho} n^{\nu} . \tag{33}
\end{equation*}
$$

$\left(B^{\perp}\right)_{\alpha \sigma}$ can be expanded in a part containing its trace $\Theta$ (1 component), its symmetric traceless part $\sigma_{\mu \nu}$ ( 9 components), and its antisymmetric part $\omega_{\mu \nu}$ (6 components):

$$
\begin{equation*}
\left(B^{\perp}\right)_{\mu \nu}:=\frac{1}{2} \Theta h_{\mu \nu}+\sigma_{\mu \nu}+\omega_{\mu \nu} \tag{34}
\end{equation*}
$$

with

$$
\begin{align*}
& \sigma_{\mu \nu}=\sigma_{v \mu}=\left(B^{\perp}\right)_{(\mu \nu)}-\frac{1}{2} \Theta h_{\mu \nu}, \operatorname{tr}(\sigma)=g^{\mu \nu} \sigma_{\mu \nu}=\sigma_{\mu}^{\mu}=0,  \tag{35}\\
& \text { and } \omega_{\mu \nu}=-\omega_{v \mu}=\left(B^{\perp}\right)_{[\mu \nu]}
\end{align*}
$$

respectively called the shear tensor, which measures the distortion in shape without change in volume, and the rotation tensor, which measures rotation without change in shape and volume. Also,

$$
\begin{align*}
& \operatorname{tr}\left(B^{\perp}\right)=g^{\mu \nu}\left(B^{\perp}\right)_{\mu \nu}=\frac{1}{2} \Theta g^{\mu \nu} h_{\mu \nu}= \Theta, \text { i.e. } \\
& \Theta=\left(B^{\perp}\right)_{\mu}^{\mu} \tag{36}
\end{align*}
$$

called the expansion scalar, which in what follows will be a fundamental quantity in the theory.

Using (32) and (26), one can easily show that $\operatorname{tr}\left(B^{\perp}\right)=\operatorname{tr}(B)=B_{\sigma}^{\sigma}=k_{; \sigma}^{\sigma} \quad$ i.e.

$$
\begin{equation*}
\Theta=k_{; \mu}^{\mu}=\frac{1}{\sqrt{-g}} \partial_{\mu}\left(\sqrt{-g} k^{\mu}\right), g=\operatorname{det}\left(g_{\mu \nu}\right) \tag{37}
\end{equation*}
$$

so that the expansion is nothing but the covariant divergence of the geodesic velocity at each of its points. We also see that $\Theta$ does not depend on the arbitrary choice of the null vector $n^{\mu}$. From the transversality of $B_{\mu \nu}$ and $h_{\mu \nu}$, we also have the transversality of $\left(B^{\perp}\right)_{\mu \nu}, \sigma_{\mu \nu}$, and $\omega_{\mu \nu}$ :

$$
\begin{align*}
& k^{\mu}\left(B^{\perp}\right)_{\mu \nu}=k^{\mu}\left(B^{\perp}\right)_{\nu \mu}=n^{\mu}\left(B^{\perp}\right)_{\mu \nu}=n^{\mu}\left(B^{\perp}\right)_{v \mu}=0,  \tag{38}\\
& k^{\mu} \sigma_{\mu \nu}=n^{\mu} \sigma_{\mu \nu}=0, k^{\mu} \omega_{\mu \nu}=n^{\mu} \omega_{\mu \nu}=0 .
\end{align*}
$$

The transverse character of $\sigma_{\mu \nu}$ and $\omega_{\mu \nu}$ implies that

$$
\begin{equation*}
\sigma_{\mu \nu} \sigma^{\mu \nu} \geq 0, \omega_{\mu \nu} \omega^{\mu \nu} \geq 0 \tag{39}
\end{equation*}
$$

## Raychaudhuri equation for the expansion scalar

From (15), contracting $\alpha$ with $\mu,\left(D_{\mu} D_{v}-D_{v} D_{\mu}\right) W^{\mu}=R_{\sigma v} W^{\sigma}$ where $R_{\sigma v}=R_{v \sigma}$ is the Ricci tensor associated with the curvature tensor $R_{\sigma \mu \nu}^{\alpha}$. Multiplying by $W^{\nu}$ one has $W^{\nu} D_{\mu} D_{v} W^{\mu}-W^{\nu} D_{v} D_{\mu} W^{\mu}=R_{\sigma v} W^{\sigma} W^{\nu}$; from Leibniz rule $D_{\mu}\left(W^{\nu} D_{v} W^{\mu}\right)=\left(D_{\mu} W^{\nu}\right)\left(D_{v} W^{\mu}\right)+W^{\nu} D_{\mu} D_{v} W^{\mu}$, and choosing $W^{\nu} \equiv k^{v}$ affinely parametrized according to (2), one obtains

$$
\begin{align*}
\frac{d}{d \lambda} \Theta & =-\left(D_{\mu} k^{v}\right)\left(D_{v} k^{\mu}\right)-R_{\sigma v} k^{\sigma} k^{\nu}=-B_{\mu}^{v} B_{v}^{\mu}-R_{\sigma v} k^{\sigma} k^{v}  \tag{40}\\
& =-\left(B^{\perp}\right)_{\mu}^{v}\left(B^{\perp}\right)_{v}^{\mu}-R_{\sigma v} k^{\sigma} k^{v}=-\frac{1}{2} \Theta^{2}-\sigma^{2}+\omega^{2}-R_{\mu v} k^{\mu} k^{v}
\end{align*}
$$

which is the Raychaudhuri equation for the expansion $\Theta$ : 1st. order non-linear differential equation (Riccati equation).

## Frobenius theorem

A hypersurface $S$ in $\mathscr{U} \subset M$ is given by an equation of the form

$$
\begin{equation*}
\phi\left(x^{\mu}\right)=\text { const. } \tag{41}
\end{equation*}
$$

where $\phi$ is a scalar.
A normal vector field to $S, n^{\mu}=g^{\mu v} n_{v}$ is given by the gradient

$$
\begin{equation*}
n_{v} \sim \partial_{v} \phi \tag{42}
\end{equation*}
$$

Theorem: A congruence of curves (timelike, spacelike, or null) in $\mathscr{Z} \subset M$ is hypersurface orthogonal if and only if

$$
\begin{equation*}
u_{[\mu ; \nu} u_{\gamma]}=\frac{2}{3!}\left(u_{[\mu ; \nu]} u_{\gamma}+u_{[v ; \gamma]} u_{\mu}+u_{[\gamma ; \mu]} u_{\nu}\right)=0 \tag{43}
\end{equation*}
$$

where $u_{\alpha}$ is the tangent vector to the curve at each of its points. (In our case, $u^{\mu}=k^{\mu}$.)

Proof: $\Rightarrow)$ The congruence is orthogonal to $S$ if $u_{\mu} \sim n_{\mu}$ i.e. $u_{\mu}=C \phi_{, \mu}$, where the scalar $C=C\left(x^{\alpha}\right)$ is constant on $S$. Then $u_{\mu ; \nu}=\left(C \phi_{, \mu}\right)_{; \nu}=C_{, \nu} \phi_{, \mu}+C\left(\phi_{, \mu}\right)_{; \nu}$. Since $\phi_{, \mu}$ is a 1-form, then $\left(\phi_{, \mu}\right)_{; \nu}=\phi_{, \mu, \nu}-\Gamma_{\mu \nu}^{\rho} \phi_{, \rho} \quad$ and $\quad\left(\phi_{, \nu}\right)_{; \mu}=\phi_{, v, \mu}-\Gamma_{\nu \mu}^{\rho} \phi_{, \rho}=\left(\phi_{, \mu}\right)_{; \nu} \quad$ from $\quad \Gamma_{\mu \nu}^{\rho}=\Gamma_{v \mu}^{\rho}$. Since $\phi$ is a scalar, $\phi_{, \alpha}=\phi_{; \alpha}$ and so $\phi_{; \mu ; \nu}=\phi_{; v ; \mu}$. Then,

$$
\begin{equation*}
u_{\mu ; \nu}=C_{, v} \phi_{, \mu}+C \phi_{; \mu v} \tag{44}
\end{equation*}
$$

with $\phi_{; \mu \nu} \equiv \phi_{; \mu ; \nu}=\phi_{; \nu \mu}$, is the equation obeyed by the tangent vectors $u^{\mu}$ to a congruence of curves which are orthogonal to a family of hypersurfaces $\mathcal{S}$ s. From (44) one obtains (43).
$\Leftarrow)$ Exercise.
Remarks. i) The fact that a congruence of curves is hypersurface orthogonal is determined only from the knowledge of the tangent vector field $u^{\mu}$. ii) Neither the normalization of $u^{\mu}$ nor the geodesic equation were used in the proof of the theorem. Then it is valid for arbitrary curves (geodesics or non-geodesics, timelike, null, or spacelike). iii) By definition, any 4-vector $t^{\mu}$ orthogonal to $k^{\mu}$ i.e. such that $k \cdot t=k^{\mu} t_{\mu}=0$ is tangent to the hypersurface $S$. Since $k \cdot k=k^{2}=0$, $k^{\mu}$ is also tangent to $S$ i.e. $k^{\mu}$ is both orthogonal and tangent to the hypersurface $S$. So, a hypersurface orthogonal null geodesic congruence is part of $S$.

The geodesics are called the generators of $S$. Also, from (21), $k^{\mu}$ is orthogonal to the 2-dimensional surface $\Sigma^{\perp} \subset S$.

Corollary: A hypersurface orthogonal congruence of null geodesics has no rotation, and viceversa. (The result is also valid for timelike geodesics.)

Proof: $\Rightarrow$ )

$$
\begin{align*}
\omega_{\mu \nu} & =\frac{1}{2}\left(B_{\mu \nu}^{\perp}-B_{v \mu}^{\perp}\right)=\frac{1}{2}\left(h_{\mu}^{\alpha} h_{v}^{\beta}-h_{v}^{\alpha} h_{\mu}^{\beta}\right) B_{\alpha \beta}=\frac{1}{2} h_{\mu}^{\alpha} h_{v}^{\beta}\left(B_{\alpha \beta}-B_{\beta \alpha}\right)  \tag{45}\\
& =h_{\mu}^{\alpha} h_{v}^{\beta} k_{[\alpha ; \beta]}=h_{\mu}^{\alpha} h_{v}^{\beta} k_{[\alpha, \beta]}=h_{\mu}^{\alpha} h_{v}^{\beta} \phi_{[\alpha, \beta]}=0 .
\end{align*}
$$

$\Leftarrow)$ Let $k_{[\alpha ; \beta]}=0$; one can then prove that $k=d \phi$ for some function $\phi$ (exercise).

So, the Raychaudhuri equation for the expansion scalar of a hypersurface orthogonal null geodesic congruence is

$$
\begin{equation*}
\frac{d}{d \lambda} \Theta=-\frac{1}{2} \Theta^{2}-\sigma^{2}-R_{\mu \nu} k^{\mu} k^{\nu} \tag{46}
\end{equation*}
$$

## Geodesic focussing

If the null convergence condition (N.C.C.):

$$
\begin{equation*}
R_{\mu \nu} k^{\mu} k^{\nu} \geq 0 \tag{47}
\end{equation*}
$$

holds, which by Einstein's equations $G_{\mu \nu}=R_{\mu \nu}-\frac{1}{2} g_{\mu \nu} R=$ const. $T_{\mu \nu}$ is related to the condition on the energy-momentum tensor $T_{\mu \nu}$ :

$$
\begin{equation*}
T_{\mu v} k^{\mu} k^{v} \geq 0 \tag{48}
\end{equation*}
$$

then

$$
\begin{equation*}
\frac{d}{d \lambda} \Theta=-\frac{1}{2} \Theta^{2}-\sigma^{2}-R_{\mu \nu} k^{\mu} k^{\nu} \leq 0 \tag{49}
\end{equation*}
$$

meaning that the expansion decreases during the congruence evolution: if at $\lambda=\lambda_{0}, \Theta=\Theta_{0}>0$ (initially divergent congruence), then the congruence will diverge less rapidly in the future $\left(\lambda>\lambda_{0}\right)$; if at $\lambda=\lambda_{0}, \Theta=\Theta_{0}<0$ (initially convergent congruence), then the congruence will converge more rapidly in the future $\left(\lambda>\lambda_{0}\right)$. So, for hypersurface orthogonal null geodesic congruences, if the N.C.C. holds, gravity is atractive, i.e. geodesics are focussed (focussing theorem).

For vacuum solutions, $R_{\mu \nu}=0$, and then

$$
\begin{equation*}
\frac{d}{d \lambda} \Theta=-\frac{1}{2} \Theta^{2}-\sigma^{2} \tag{50}
\end{equation*}
$$

So, from (39),

$$
\begin{equation*}
\frac{d}{d \lambda} \Theta \leq-\frac{1}{2} \Theta^{2} \tag{51}
\end{equation*}
$$

with equality if and only if the shear vanishes.
The integration of (51) is

$$
\begin{equation*}
\frac{1}{\Theta(\lambda)} \geq \frac{1}{\Theta_{0}}+\frac{\lambda}{2}, \lambda \geq 0 \tag{52}
\end{equation*}
$$

where $\Theta_{0}=\Theta(0)$ is the expansion at the hypersurface $S$ (we choosed $\lambda_{0}=0$ ). If $\Theta_{0}<0, \frac{1}{\Theta(\lambda)} \geq-\frac{1}{\left|\Theta_{0}\right|}+\frac{\lambda}{2}$ and so $\Theta(\lambda) \rightarrow-\infty$ and the congruence
breaks down since it converges to a point p: focal point or conjugate point to $S$ ([8] [17]; see also M. Basquens, Singularity theorems, Universitat Politécnica de Catalunya, 2016) when $\lambda \rightarrow \lambda^{\prime} \in\left(0, \frac{2}{\left|\Theta_{0}\right|}\right] .\left(\lambda^{\prime}<\frac{2}{\left|\Theta_{0}\right|}\right.$ if and only if $\sigma_{\mu \nu}$ and/or $R_{\mu v} k^{\mu} k^{\nu}$ do not vanish.) This focussing process is ilustrated in Figure 1.

It should be stressed that $p$ is not a singularity of the spacetime, but a singularity of the congruence, i.e. a point where the Raychaudhuri equation loses its validity.

## Interpretation of $\Theta$

If we call $\mu$ the parameter of the curves with tangents $n^{v}$ (the null vector field defined by (19)), each constant value of $\mu$ defines a hypersurface $S^{\prime}$ in $M$. Let $\delta S(\lambda, \mu)$ (fixed $\lambda$ and $\mu$ ) be an infinitesimal (2-dimensional) surface element, contained in the intersection $S \cap S^{\prime} \subset S$ of the hypersurfaces $\mu=$ const. and $\lambda=$ const.'. So, in particular, $\delta S \subset S$. If $\delta A$ is the area of $\delta S(\lambda, \mu)$, then $\Theta$ is the fractional rate of change of $\delta A$ along the geodesics, i.e.

$$
\begin{equation*}
\Theta=\frac{1}{\delta A} \frac{d}{d \lambda}(\delta A) \tag{53}
\end{equation*}
$$

Proof: Let $\vec{z}=\left(z^{A}\right), A=1,2$ be coordinates on $\delta S$. Together with $\lambda$ and $\mu$ one has a 4-dimensional coordinate system $x^{\alpha}=x^{\alpha}\left(\lambda, \mu, z^{1}, z^{2}\right)$ with $k^{\alpha}=\left(\frac{\partial x^{\alpha}}{\partial \lambda}\right)_{\mu, \bar{z}}$ and tangent vectors to $\delta S, \quad e_{A}, \quad A=1,2, \quad e_{A}^{\alpha}=\left(\frac{\partial x^{\alpha}}{\partial z^{A}}\right)_{\lambda, \mu}$. Then, the metric on $\delta S$ is

$$
\begin{equation*}
\sigma_{A B}=g_{\alpha \beta} e_{A}^{\alpha} e_{B}^{\beta} \tag{54}
\end{equation*}
$$

In fact, on $\delta S, d \lambda=d \mu=0$, i.e. $x^{\alpha}=x^{\alpha}\left(z^{A}\right)$ and so $d s^{2}=g_{\alpha \beta} d x^{\alpha} d x^{\beta}=g_{\alpha \beta} \frac{\partial x^{\alpha}}{\partial z^{A}} \frac{\partial x^{\beta}}{\partial z^{B}} d z^{A} d z^{B}=g_{\alpha \beta} e_{A}^{\alpha} e_{B}^{\beta} d z^{A} d z^{B}:=\sigma_{A B} d z^{A} d z^{B}$. The infinitesimal area element is $\delta A=\sqrt{\sigma} d^{2} z$ with $\sigma=\operatorname{det}\left(\sigma_{A B}\right)$. Let $\sigma^{A B}$ be the inverse metric of $\sigma_{A B}$, i.e. $\sigma^{A B} \sigma_{B C}=\delta_{C}^{A}$; then $\sigma^{A B} \sigma_{B A}=\delta_{A}^{A}=2$. It is easy to see that the $e_{A}$ 's vectors are Lie transported along $k^{\mu}$ :

$$
\begin{align*}
\left(\mathscr{L}_{k} e_{A}\right)^{\alpha} & =k \cdot \partial\left(e_{A}^{\alpha}\right)-e_{A} \cdot \partial\left(k^{\alpha}\right) \\
& =\left(\frac{\partial x^{\beta}}{\partial \lambda}\right)_{\mu, \bar{\imath}} \frac{\partial}{\partial x^{\beta}}\left(\frac{\partial x^{\alpha}}{\partial z^{A}}\right)_{\mu, \lambda}-\left(\frac{\partial x^{\beta}}{\partial z^{A}}\right)_{\mu, \lambda} \frac{\partial}{\partial x^{\beta}}\left(\frac{\partial x^{\alpha}}{\partial \lambda}\right)_{\mu, \bar{\imath}}  \tag{55}\\
& =\frac{\partial^{2} x^{\alpha}}{\partial \lambda \partial z^{A}}-\frac{\partial^{2} x^{\alpha}}{\partial z^{A} \partial \lambda}=0,
\end{align*}
$$



Figure 1. Focussing of hypersurface orthogonal null geodesics congruence.
with the necessary differentiability. Since with a symmetric connection $\left(\Gamma_{v \rho}^{\mu}=\Gamma_{\rho v}^{\mu}\right)$, in the Lie derivative $\partial$ can be replaced by $D$, one has

$$
\begin{equation*}
0=\left(\mathscr{L}_{k} e_{A}\right)^{\alpha}=k \cdot D\left(e_{A}^{\alpha}\right)-e_{A} \cdot D\left(k^{\alpha}\right)=k^{\sigma}\left(e_{A}^{\alpha}\right)_{; \sigma}-e_{A}^{\sigma} k_{; \sigma}^{\alpha} \tag{56}
\end{equation*}
$$

i.e. $k^{\sigma}\left(e_{A}^{\alpha}\right)_{; \sigma}=e_{A}^{\sigma} k_{; \sigma}^{\alpha}$. From (54) and (56) one can prove that

$$
\begin{equation*}
\frac{d}{d \lambda} \sigma_{A B}=\left(B_{\alpha \beta}+B_{\beta \alpha}\right) e_{A}^{\alpha} e_{B}^{\beta} \tag{57}
\end{equation*}
$$

Then

$$
\begin{align*}
\sigma^{A B} \frac{d}{d \lambda} \sigma_{A B} & =\left(B_{\alpha \beta}+B_{\beta \alpha}\right) \sigma^{A B} e_{A}^{\alpha} e_{B}^{\beta}=\left(B_{\alpha \beta}+B_{\beta \alpha}\right) h^{\alpha \beta}=2 B_{\alpha \beta} h^{\alpha \beta}  \tag{58}\\
& =2 B_{\alpha \beta}\left(g^{\alpha \beta}-k^{\alpha} n^{\beta}-k^{\beta} n^{\alpha}\right)=2 B_{\alpha}^{\alpha}=2 \Theta,
\end{align*}
$$

where we used

$$
\begin{equation*}
h^{\alpha \beta}=e_{A}^{\alpha} e_{B}^{\beta} \sigma^{A B} \tag{59}
\end{equation*}
$$

which can be proved contracting $h^{\alpha \beta}$ with $h_{\gamma \alpha}$, using (20), and the fact that $n_{\alpha} e_{A}^{\alpha}=k_{\alpha} e_{A}^{\alpha}=0$. This of course implies that $\delta S \cap \Sigma^{\perp} \neq \phi$.

Along the geodesics, $\vec{z}=\left(z^{A}, z^{B}\right)$ are constant, and so the change in $\delta A$ comes only from the change in $\sigma_{A B}$; so

$$
\begin{align*}
\frac{1}{\delta A} \frac{d}{d \lambda} \delta A & =\frac{1}{\sqrt{\sigma} d^{2} z} \frac{d}{d \lambda}\left(\sqrt{\sigma} d^{2} z\right)=\frac{1}{\sqrt{\sigma}} \frac{d}{d \lambda} \sqrt{\sigma}=\frac{1}{2 \sigma} \frac{d}{d \lambda} \sigma  \tag{60}\\
& =\frac{1}{2} \operatorname{det}\left(\sigma_{A B}\right) \frac{d}{d \lambda}\left(\operatorname{det}\left(\sigma_{A B}\right)\right)=\frac{1}{2} \sigma^{A B} \frac{d}{d \lambda} \sigma_{A B}
\end{align*}
$$

(for non-singular matrices $N, \operatorname{det}(N)=\exp (\operatorname{tr}(\ln N))$. Together with (58), we obtain (53).

Feynman propagators for null geodesic congruences
In terms of the function $F(\lambda)$ defined by ([18] [19])

$$
\begin{equation*}
\Theta(\lambda)=2 \frac{\dot{F}(\lambda)}{F(\lambda)} \tag{61}
\end{equation*}
$$

the Raychaudhuri Equation (40) becomes

$$
\begin{equation*}
\ddot{F}(\lambda)+(\Omega(\lambda))^{2} F(\lambda)=0 \tag{62}
\end{equation*}
$$

with

$$
\begin{equation*}
(\Omega(\lambda))^{2}=\frac{1}{2}\left(\sigma^{2}-\omega^{2}+R_{\mu \nu} k^{\mu} k^{\nu}\right) \tag{63}
\end{equation*}
$$

which is nothing but the equation of a classical 1-dimensional harmonic oscillator with $\lambda$ ("time")-dependent frequency $\Omega$. After Hill ([20]), (62) is known as a "Hill-type" equation. If at $\lambda=\bar{\lambda}$ the congruence has a focal point, i.e. $\Theta(\lambda) \rightarrow-\infty$ as $\lambda \rightarrow \bar{\lambda}$, then $\bar{\lambda}$ must be a zero of $F(\lambda)$ if $\dot{F}(\bar{\lambda})$ is finite.
(62) is the Euler-Lagrange equation of the $\lambda$-dependent Lagrangian

$$
\begin{equation*}
\mathscr{L}(F, \dot{F}, \lambda)=\frac{1}{2}\left(\dot{F}^{2}-\Omega^{2} F^{2}\right) . \tag{64}
\end{equation*}
$$

For a suitable domain of definition of $\lambda,(62)$ admits a solution $\bar{F}(\lambda)$ subject to the boundary conditions $F^{\prime}=\bar{F}\left(\lambda^{\prime}\right)$ and $F^{\prime \prime}=\bar{F}\left(\lambda^{\prime \prime}\right)$ with, e.g. $\lambda^{\prime}<\lambda^{\prime \prime}$. (Notice that if $[\lambda]=[L]$, then $[F]=[L]^{1 / 2}$ since $[$ action $]=\left[\int d \lambda \mathscr{L}\right]=[L][\mathscr{L}]=[L]^{0}$ and $[\mathscr{L}]=[L]^{-1}$.)

It is well known ([21] [22]) that a classical Lagrangian of the form

$$
\begin{equation*}
\mathscr{L}(x, \dot{x}, t)=\frac{1}{2}\left((\dot{x}(t))^{2}-b(t)(x(t))^{2}\right) \tag{65}
\end{equation*}
$$

has associated with it a Feynman propagator $K\left(x^{\prime \prime}, t^{\prime \prime} ; x^{\prime}, t^{\prime}\right)$ given by the path integral

$$
\begin{equation*}
\int_{x\left(t^{\prime}\right)=x^{\prime}}^{x\left(t t^{\prime \prime}\right)=x^{\prime \prime}} \mathscr{S}(t) \mathrm{e}^{i \int_{t^{\prime \prime}}{ }^{\prime \prime} d t(x, \dot{x}, t)} \tag{66}
\end{equation*}
$$

( $\hbar=1$ ) where, formally,

$$
\begin{equation*}
\int_{x\left(t^{\prime}\right)=x^{\prime}}^{x\left(t^{\prime \prime}\right)=x^{\prime \prime}} \mathscr{} x(t) \ldots=\Pi_{t \in\left(t^{\prime}, t^{\prime \prime}\right)} \int_{-\infty}^{+\infty} d x(t) \ldots \tag{67}
\end{equation*}
$$

The result is

$$
\begin{equation*}
K\left(x^{\prime \prime}, t^{\prime \prime} ; x^{\prime}, t^{\prime}\right)=\left(2 \pi i f\left(t^{\prime \prime}, t^{\prime}\right)\right)^{-1 / 2} \mathrm{e}^{i S[\bar{x}]} \tag{68}
\end{equation*}
$$

with $S[\bar{x}]=\int_{t^{\prime}}^{t^{\prime \prime}} \mathscr{L}(\bar{x}(t), \dot{\bar{x}}(t), t)$ and $\bar{x}(t)$ solution of

$$
\begin{equation*}
\ddot{x}(t)+b(t) x(t)=0 \tag{69}
\end{equation*}
$$

with $x\left(t^{\prime}\right)=x^{\prime}, x\left(t^{\prime \prime}\right)=x^{\prime \prime}$, and $f\left(t^{\prime \prime}, t^{\prime}\right)$ solution of

$$
\begin{equation*}
\frac{\partial^{2} f\left(t, t^{\prime}\right)}{\partial t^{2}}+b(t) f\left(t, t^{\prime}\right)=0 \tag{70}
\end{equation*}
$$

with $f\left(t^{\prime}, t^{\prime}\right)=0$ and $\left.\frac{\partial f\left(t, t^{\prime}\right)}{\partial t}\right|_{t=t^{\prime}}=0$.
Since (64) and (65) (and therefore (62) and (69)) have the same form, the path integral

$$
\begin{equation*}
K\left(F^{\prime \prime}, \lambda^{\prime \prime} ; F^{\prime}, \lambda^{\prime}\right)=\int_{F\left(\lambda^{\prime}\right)=F^{\prime}}^{F\left(\lambda^{\prime \prime}\right)=F^{\prime \prime}} \mathscr{} F(\lambda) \mathrm{e}^{i \int_{\lambda^{\prime \prime}}^{\lambda^{\prime \prime}} d \lambda(F, \dot{F}, \lambda)}=\left(2 \pi i f\left(\lambda^{\prime \prime}, \lambda^{\prime}\right)\right)^{-1 / 2} \mathrm{e}^{i S[\bar{F}]}(7 \tag{71}
\end{equation*}
$$

with

$$
\begin{equation*}
S[\bar{F}]=\int_{\lambda^{\prime}}^{\lambda^{\prime \prime}} \mathscr{L}(\bar{F}(\lambda), \dot{\bar{F}}(\lambda), \lambda) \tag{72}
\end{equation*}
$$

and $f\left(\lambda, \lambda^{\prime}\right)$ solution of (70) with $t$ 's replaced by $\lambda$ 's, can be formally considered the Feynman propagator describing the "quantum" flow of the geodesic congruence from $\lambda^{\prime}$ to $\lambda^{\prime \prime}$ or, equivalently, the quantity which includes all admisible "quantum" fluctuations of the expansion $\Theta$ from $\Theta\left(\lambda^{\prime}\right)$ (corres-
ponding to $\left(F^{\prime}, \lambda^{\prime}\right)$ ) to $\Theta\left(\lambda^{\prime \prime}\right)$ (corresponding to $\left(F^{\prime \prime}, \lambda^{\prime \prime}\right)$ ). The quotation marks in "quantum" is due to the still missing existence of a final theory of quantum gravity ([12]).
As a concrete example ([16]), we consider the evolution of the ingoing radial null geodesic congruences within the black hole (B.H.) region of the Schwarz-schild-Kruskal-Szekeres spacetime, from the future ( $\mathrm{H}^{+}$) and past ( $\mathrm{H}^{-}$) horizons to the future singularity at $r=0 \quad\left(\left.0\right|_{+}\right)$. In both cases the affine parameter is $\lambda=-r$ ([15]). In this case, since $\sigma^{2}=\omega^{2}=R_{\mu v} k^{\mu} k^{\nu}=0$, the Raychaudhuri equation reduces to $\frac{d \Theta}{d r}=\frac{1}{2} \Theta^{2}$, the focal points are at the singularity, and the propagators result

$$
\begin{equation*}
K^{x}\left(\left.H^{ \pm} \rightarrow 0\right|_{+}\right)=K^{x}\left(0,\left.0\right|_{+} ; F^{x}, 2 M\right)=\frac{\mathrm{e}^{-\frac{i\left(F^{x}\right)^{2}}{4 M}}}{\sqrt{8 \pi M}} \tag{73}
\end{equation*}
$$

where $M$ is the energy of the B.H., $x=a, b$ with $a \leftrightarrow H^{+}$and $b \leftrightarrow H^{-}$, i.e. $F^{a}=F(r=2 M)$ at $H^{+}, F^{b}=F(r=2 M)$ at $H^{-}$, and $F^{x}(r=0)=0$. Even if $\Theta^{x}(r)=-\frac{2}{r}$ and therefore $\Theta^{x}(0)=-\infty$, finite initial values for $F^{x}$ guarantee finite values for the propagators, which suggests that the introduction of a quantum description should smooth or even disappear the singularities of the classical theory ([13]) (See Figure 2).

## 3. Elements of Causality Theory

Let $(M, \mathbf{g})$ be a 4-dimensional pseudo-Riemannian (Lorentzian) manifold.

1. A vector field $X$ over $M$ is causal if $X_{p}$ is nonspacelike (temporal or null) $\forall p \in M$.
2. A curve $c:(a, b) \rightarrow M$ is causal if its tangent vector is nonspacelike at each of its points.
3. A Lorentzian manifold $(M, \mathbf{g})$ is temporally orientable if it admits a smooth global temporal vector field $T$, locally $T=T^{\mu} \partial_{\mu}, T^{2}>0$, called a temporal orientation. The triplet $(M, \mathbf{g}, T)$ is called a Lorentz oriented manifold. $-T$ defines the opposite temporal orientation. A non-zero causal vector field $X$ is future (past) directed if $\mathbf{g}(X, T)>0 \quad(\mathbf{g}(X, T)<0)$.


Figure 2. Future directed null ingoing geodesic propagators: (a): $\left.H^{+} \rightarrow 0\right|_{+}$, (b): $\left.\mathrm{H}^{-} \rightarrow 0\right|_{+}$.
4. It can be shown that any compact without boundary time oriented spacetime contains closed timelike curves ([6]). To avoid such causality violation situations, all spacetimes considered here are non-compact.
5. Two spacetimes $(M, \mathbf{g})$ and $\left(M^{\prime}, \mathbf{g}^{\prime}\right)$ are isometric if there exists a diffeomorphism $\varphi: M \rightarrow M^{\prime}$ such that $\mathbf{g}=\varphi^{*}\left(\mathbf{g}^{\prime}\right)$, where $\varphi^{*}$ denotes the pull-back operation.

## Extendibility

A spacetime $(M, \mathbf{g})$ is extendible if it is isometric to a proper subset of a spacetime $\left(M^{\prime}, \mathbf{g}^{\prime}\right) .\left(M^{\prime}, \mathbf{g}^{\prime}\right)$ is called an extension of $(M, \mathbf{g})$. For example, the Schwarzschild-Kruskal-Szekeres (SKS) spacetime is an extension (the maximal extension) of the Schwarzschild spacetime $S_{4}$ with the isometry generated by the inclusion $t: S_{4} \rightarrow$ SKS.
6. Let $\gamma:(a, b) \rightarrow M$ be a future directed causal curve in $M . p \in M$ is a future final point of $\gamma$ if for any open neighborhood $\mathscr{U}_{p}$ of $p$ there exists $t_{0} \in(a, b)$ such that $\gamma(t) \in \mathscr{U}_{p}, \forall t>t_{0} . \gamma$ is future inextendible if it has no future final point. Example: Let $\gamma:(-\infty, 0) \rightarrow \operatorname{Mink}^{4}, t \mapsto \gamma(t)=(t, \overrightarrow{0})$. It is clear that $(0, \overrightarrow{0})$ is a future final point of $\gamma$. So $\gamma$ is not future inextendible; that is, $\gamma$ is future extendible. Instead, let $\gamma^{\prime}:(-\infty, 0) \rightarrow \operatorname{Mink}^{4} \backslash(0, \overrightarrow{0})$, $\gamma^{\prime}(t)=(t, \overrightarrow{0}) . \gamma^{\prime}$ has no future final point since $(0, \overrightarrow{0})$ has been eliminated from the spacetime; so, $\gamma^{\prime}$ is future inextendible.

Changing "future" by "past" the previous definitions pass to past directed causal curves, and one has the analogous concepts of past final point and past inextendible causal curves.

A causal curve is inextendible if it is both future and past inextendible, i.e. it has neither a future nor a past final point. Geodesic completeness
7. A geodesic is complete if for it there exists an affine parameter $\lambda$ which extends from $-\infty$ to $+\infty$.

A spacetime is geodesically complete if all its causal (timelike or null) inextendible geodesics are complete.

A spacetime is singular if: i) as a spacetime it is inextendible, and ii) it is geodesically incomplete. So, an inextendible spacetime is singular if it has at least one causal inextendible geodesic which does not admit an affine parameter $\lambda$ extending from $-\infty$ to $+\infty$.

Examples. i) Mink ${ }^{4}$ is not singular. ii) SKS is singular since it has causal geodesics which die at $r=0$ at a finite value of their affine parameter.

Domains of dependence. Cauchy surfaces. Cauchy horizons
8. Let $(M, \mathbf{g})$ be a spacetime. $\Sigma \subset M$ is a partial Cauchy surface if $\Sigma$ is an hypersurface of $M$ such that no pair of points $p, q \in \Sigma, \quad p \neq q$, can be joined by a causal curve in $M$ (then in particular in $\Sigma$ ).

The future domain of dependence of $\Sigma, D^{+}(\Sigma)$, is the set of points $p \in M$ such that any past inextendible causal curve $\gamma$ which passes through $p$ intersects $\Sigma$. In particular $\Sigma \subset D^{+}(\Sigma)$.

The past domain of dependence of $\Sigma, D^{-}(\Sigma)$, is the set of points $q \in M$
such that any future inextendible causal curve $\gamma^{\prime}$ which passes through $q$ intersects $\Sigma$. In particular $\Sigma \subset D^{-}(\Sigma)$. So, $D^{+}(\Sigma) \cap D^{-}(\Sigma)=\Sigma$.

Since $\Sigma$ is acausal (see below), the number of intersections of $\gamma$ and $\gamma^{\prime}$ with $\Sigma$ is 1 i.e. $|\gamma \cap \Sigma|=\left|\gamma^{\prime} \cap \Sigma\right|=1$.

The domain of dependence of $\Sigma, D(\Sigma)$, is the union of $D^{+}(\Sigma)$ and $D^{-}(\Sigma)$ i.e.

$$
\begin{equation*}
D(\Sigma)=D^{+}(\Sigma) \cup D^{-}(\Sigma) \tag{74}
\end{equation*}
$$

Example: Let $\Sigma$ in Mink ${ }^{2}$ be given by the positive $x$-axis i.e. $\operatorname{Mink}^{2} \supset \Sigma=\{(0, x), x>0\} . \Sigma$ is closed in Mink ${ }^{2}$ since $\operatorname{Mink}^{2} \backslash \Sigma$ is open. Also, $\Sigma$ is achronal (see below 15.) and spacelike i.e. acausal: no two of its points can be joined by a causal curve. Therefore $\Sigma$ is a partial Cauchy surface $D^{+}(\Sigma)=\{(t, x) \mid 0 \leq t<x\}, D^{-}(\Sigma)=\{(t, x) \mid-x<t \leq 0\}, \quad \Sigma \subset D^{ \pm}(\Sigma)$ with $\Sigma=D^{+}(\Sigma) \cap D^{-}(\Sigma)$ and $D(\Sigma)=D^{+}(\Sigma) \cup D^{-}(\Sigma)$. Also, $\partial(D(\Sigma))=\{t= \pm x, x \geq 0\}$; clearly, $D(\Sigma) \neq$ Mink $^{2}$. If instead, $\Sigma=\Sigma^{\prime}=\{x$-axis $\}$, then $D\left(\Sigma^{\prime}\right)=$ Mink $^{2}$.

A partial Cauchy surface $\Sigma$ in $M$ is a Cauchy surface if $D(\Sigma)=M$. In this case the spacetime is called globally hyperbolic.

If $\Sigma \subset M$ is a partial Cauchy surface but not necessarily a Cauchy surface, the future (past) boundary of $D(\Sigma), \partial(D(\Sigma))^{+}\left(\partial(D(\Sigma))^{-}\right)$is called a future (past) Cauchy horizon of $\Sigma$. In the previous example in Mink ${ }^{2}$, $\partial(D(\Sigma))^{ \pm}=\{(t, x) \mid t= \pm x, x \geq 0\} ;$ one has $\partial(D(\Sigma))^{+} \cap \partial(D(\Sigma))^{-}=\{(0,0)\}$ and $\partial(D(\Sigma))^{+} \cup \partial(D(\Sigma))^{-}=\{t=x, x \geq 0\} \cup\{t=-x, x \geq 0\}$.

## Chronological and causal conditions

9. A time oriented spacetime satisfies the chronological condition if it has no closed future (past) directed timelike curves; (in particular geodesics). Such spacetime is said to be chronological.

A time oriented spacetime satisfies the causality condition if it has no closed future (past) directed causal curves; in particular geodesics. Such a spacetime is said to be causal.

So, a causal spacetime is a chronological spacetime, but not necessarily the other way around.
10. A spacetime $M$ is strongly causal if $\forall p \in M$ and $\forall \mathscr{U}_{p}$ (open neighborhood of $p$ ), $\exists \mathscr{y}_{p} \subset \mathscr{U}_{p}$ such that if $q, q^{\prime} \in \mathscr{Y}_{p}$ and $\gamma: q \rightarrow q^{\prime}$ is a causal curve, then $\gamma \subset \mathscr{q}_{p}$. Since $\%_{p}$ is arbitrarily small, taking $p=q=q$, the unique causal curve $\gamma: p \rightarrow p$ is the trivial one $\gamma \equiv p=$ const. So, in a strongly causal spacetime there are no nontrivial closed causal curves.

So, a strongly causal spacetime is a causal spacetime.
11. Proposition: A globally hyperbolic spacetime has no closed causal curves.

Proof. Let $\gamma$ with local coordinates $x^{\mu}(\lambda)$ be a closed causal curve in the spacetime. $\lambda$ should be periodic i.e. $x^{\mu}(\lambda+$ const. $)=x^{\mu}(\lambda)$ or, equivalently, $\lambda \in \mathbb{R}$. So, $\gamma$ would be an inextendible causal curve intersecting the Cauchy surface $\Sigma \infty$-many times, in contradiction with the definition 8 .

It can be shown that global hyperbolicity implies strong causality ([7]: p. 11)
and so we have the chain of implications:

$$
\begin{align*}
& \text { global hyperbolicity } \Rightarrow \text { strongly causal } \Rightarrow \nexists \text { closed causal curves }  \tag{75}\\
& \Rightarrow \text { causal condition. }
\end{align*}
$$

## Chronological and causal futures and pasts; horisms

Let $p \in M$.
The chronological future (past) of $p$ is given by the set

$$
\begin{equation*}
I^{ \pm}(p):=\{q \in M \mid \exists \gamma: p \rightarrow q, \text { future (past) directed timelike curve }\} . \tag{76}
\end{equation*}
$$

One writes $p \ll q$.
The causal future (past) of $p$ is given by the set

$$
\begin{equation*}
J^{ \pm}(p)=\{q \in M \mid \exists \gamma: p \rightarrow q \text {, future (past) directed causal curve }\} . \tag{77}
\end{equation*}
$$

One writes $p<q$.
Clearly,

$$
\begin{equation*}
I^{ \pm}(p) \subset J^{ \pm}(p), \text { and so } p \ll q \Rightarrow p<q \tag{78}
\end{equation*}
$$

but not viceversa.
The set of future (past) horisms of $p$ is given by

$$
\begin{equation*}
E^{ \pm}(p):=J^{ \pm}(p) \backslash I^{ \pm}(p) \tag{79}
\end{equation*}
$$

If $N$ is a subset of $M$, the corresponding quantities are defined by:

$$
\begin{equation*}
I^{ \pm}(N):=\bigcup_{p \in N} I^{ \pm}(p), J^{ \pm}(N):=\bigcup_{p \in N} J^{ \pm}(p), E^{ \pm}(N)=\bigcup_{p \in N} E^{ \pm}(p) . \tag{80}
\end{equation*}
$$

For any $N \subset M, N \subset J^{ \pm}(N)$ since the constant curve $p \rightarrow p$ for any $p \in N$ is null and so causal (a point).
12. It can be proved ([5]):
i) $I^{+}(p)$ is open (rough argument: a sufficient small deformation of a timelike curve remains timelike)
ii) $I^{+}(N)$ is open (union of open sets is open)
iii) $J^{+}(N) \subset \overline{I^{+}(N)}$
iv) $\overline{J^{+}(N)}=\overline{I^{+}(N)} \cdot\left(J^{+}(N)\right.$ is not necessarily closed; if $J^{+}(N)=\overline{I^{+}(N)}$ then $J^{+}(N)$ is closed i.e. $\left.J^{+}(N)=\overline{J^{+}(N)}\right)$
v) $\partial I^{+}(N)=\partial J^{+}(N)$
vi) $\operatorname{int} J^{+}(N)=I^{+}(N)$
13. Theorem: If a spacetime $M$ is strongly causal and $\forall p, q \in M$,
$J^{+}(p) \cap J^{-}(q)$ is compact, then $M$ is globally hyperbolic.
Proof. [8]: p. 734.
14. A spacetime is stably causal if it exists a global function $t: M \rightarrow \mathbb{R}$ such that $\operatorname{grad}(t) \in \operatorname{Vect}(M)$ is timelike (i.e. $\left.\operatorname{grad}(t)^{2}>0\right) .(\operatorname{grad}(t)$ is the vector field associated with the 1-form $d t$ by $g$, locally $d t=\Sigma_{\mu=0}^{3} \frac{\partial t}{\partial x^{\mu}} d x^{\mu}$.) $t$ is called a time function. A stably causal spacetime is time orientable. The preimage of $\lambda \in \mathbb{R}$ by $t$, $t^{-1}(\{\lambda\})$, is called a level set of $t$ and is denoted by $\Sigma_{\lambda}$. If the domain of dependence of each of the level sets is $M$ i.e. $D\left(\Sigma_{\lambda}\right)=M, \forall \lambda \in \mathbb{R}$,
then all level sets are homeomorphic to each other, each level set is a Cauchy surface, the spacetime is globally hyperbolic and, topologically, $M \cong \mathbb{R} \times \Sigma_{\lambda}$ for any $\lambda \in \mathbb{R}$. So,

$$
\begin{equation*}
\text { global hyperbolicity } \Rightarrow \text { stable causality, } \tag{81}
\end{equation*}
$$

but not the other way around.
Proposition: A stably causal spacetime satisfies the chronological condition.
Proof: $t$ increases in the direction of $\operatorname{grad}(t)$, so $t$ increases in the direction of any forward directed timelike curve. So, there is no closed forward directed timelike curve.

Theorem: A stably causal spacetime is strongly causal.
Proof: [7]: p. 11; [17]: p. 199.
One has the chain of implications:

$$
\begin{align*}
& \text { global hyperbolicity } \Rightarrow \text { stably causal } \Rightarrow \text { strongly causal }  \tag{82}\\
& \Rightarrow \text { causal condition } \Rightarrow \text { chronological condition. }
\end{align*}
$$

## Achronal and future sets, achronal boundaries, edges

15. An achronal set $N \subset M$ is a subset of $M$ such that no two of its points can be joined by a timelike curve. (Notice that if a set $N$ is acausal, then it is achronal, but not the other way around.)
i) Proposition: $N \subset M$ is achronal $\Leftrightarrow I^{ \pm}(N) \cap N=\phi$.

Proof. $\Rightarrow$ ) Suppose that $I^{ \pm}(N) \cap N \neq \phi ; \Rightarrow \exists p \in N$ and
$p \in I^{ \pm}(N)=\cup_{q \in N} I^{+}(q)$, which implies that $\exists q \in N$ such that $p \in I^{+}(q)$ i.e. $\exists$ a forward directed timelike curve $\gamma: q \rightarrow p$, which is a contradiction with the achronality of $N$.
$\Leftarrow)$ Let $I^{ \pm}(N) \cap N=\phi ; \Rightarrow$ : if $p \in N \Rightarrow p \notin I^{+}(N)$ i.e. $\nexists q \in N$ such that $\rho: q \rightarrow p$ is a forward directed timelike curve.
(In words, if a subset of $M$ is achronal, the subset and its chronological future and past are disjoint; and viceversa.)
ii) An achronal boundary is a subset of $M$ of the form $\partial I^{+}(N)$ (or $\partial I^{-}(N)$ ) for some subset $N \subset M$.
iii) Lemma: Let $N$ be a subset of $M$. If $p \in \partial I^{+}(N) \Rightarrow I^{+}(p) \subset I^{+}(N)$.

Proof. Let $q \in I^{+}(p) ; \Rightarrow p \in I^{-}(q)$ and so $I^{-}(q)$ is an open neighborhood of $p$. Since $p \in \partial I^{+}(N) \Rightarrow I^{-}(q) \cap I^{+}(N) \neq \phi$ and $\Rightarrow q \in I^{+}(N)$. $\Rightarrow I^{+}(p) \subset I^{+}(N)$.
iv) Proposition: An achronal boundary is achronal.

Proof. Let $p, q \in \partial I^{+}(N)$ with $q \in I^{+}(p)$. Because of the Lemma, $q \in I^{+}(N) . \Rightarrow I^{+}(N) \cap \partial I^{+}(N) \neq \phi$, what is a contradiction since $I^{+}(N)$ is open. Then $I^{+}(N) \cap \partial I^{+}(N)=\phi$ and $\Rightarrow \partial I^{+}(N)$ is achronal.
v) It can also be shown that an achronal boundary $\partial I^{+}(N)$ is a closed, continuous hypersurface of $M$.
vi) Let $N \subset M$ be an achronal subset of $M . \quad p \in \bar{N}$ is an edge point of $N$ if any $\mathbb{U}_{p}$ (open neighborhood of $p$ ) contains a timelike curve $\gamma$ from $I^{-}\left(p, \mathscr{U}_{p}\right)$ to $I^{+}\left(p, \mathscr{U}_{p}\right)$ that does not meet $N .\left(I^{ \pm}\left(p, \mathscr{U}_{p}\right)\right.$ is the chronological future (past) of $p$ within $\mathscr{V}_{p}$.) We denote

$$
\begin{equation*}
\operatorname{edge}(N)=\{p \in \bar{N} \mid p \text { is an edge point of } N\} . \tag{83}
\end{equation*}
$$

It is clear that $\bar{N} \backslash N \subset e \operatorname{edge}(N) \subset \bar{N}$. If edge $(N)=\phi$ one says that $N$ is edgeless. If $\bar{N} \backslash N=\{p \in \bar{N} \mid p \notin N\}=\phi, \Rightarrow \forall p \in \bar{N}, \quad p \in N$, which amounts to $\bar{N}=N$, i.e. $N$ is closed.
vii) Proposition: An achronal boundary is edgeless.

Proof. Let $N$ be a subset of $M ; \partial I^{+}(N)$ is an associated boundary (we could choose $\partial I^{-}(N)$ ). Let $p \in \partial I^{+}(N)$. Let $\gamma$ be any timelike curve from $I^{-}(p)$ to $I^{+}(p)$ i.e. $\gamma: q \rightarrow q^{\prime}$ with $q \in I^{-}(p)$ and $q^{\prime} \in I^{+}(p)$. $\Rightarrow \gamma \cap \partial I^{+}(N) \neq \phi$. Then $\partial I^{+}(N)$ is edgeless.
viii) Let $N$ be a subset of $M . N$ is a future set if $I^{+}(N) \subset N$ i.e. $N$ is enough big to contain its chronological future.
ix) Proposition: Let $N$ be a subset of $M$ with with $I^{+}(N) \neq \phi$. If $N$ is achronal $\Rightarrow N$ is not a future set.

Proof. $N$ achronal $\Rightarrow I^{+}(N) \cap N=\phi$ and $N$ future set $\Rightarrow I^{+}(N) \subset N$. $\Rightarrow I^{+}(N) \cap N=I^{+}(N)=\phi$, which is a contradiction.
x) The above proposition is equivalent to: If $N$ is a future set $\Rightarrow N$ can not be achronal (unless $I^{+}(N)=\phi$ ).
xi) Proposition: For any $N \subset M, I^{+}(N)$ is a future set, i.e. $I^{+}\left(I^{+}(N)\right) \subset I^{+}(N)$.
Proof. [8]: p. 731: $I^{+}\left(I^{+}(N)\right)=I^{+}(N)$.
The same occurs for causal sets: since $J^{+}\left(J^{+}(N)\right)=J^{+}(N)$ (same Ref.), $J^{+}(N)$ is a future set, i.e. $J^{+}\left(J^{+}(N)\right) \subset J^{+}(N)$.
xii) Proposition: $\forall N \subset M, \partial J^{+}(N)$ is achronal.

Proof: $\partial J^{+}(N)=\partial I^{+}(N)$, and $\partial I^{+}(N)$ is achronal.

## Trapped surfaces

16. A future (past) trapped surface $S$ in $M$ is a 2-dimensional spacelike submanifold of $M$ such that for both outgoing ( + ) and ingoing ( - ) forward directed null geodesic curves emitted orthogonally from $S$, the expansions $\Theta_{ \pm}$are both negative (positive) on $S$.

A closed trapped surface is a compact without boundary trapped surface.
17. Facts [8]:
i) If $\exists$ a forward directed causal curve from $p$ to $q$ but $\nexists$ a forward directed timelike curve from $p$ to $q, \Rightarrow$ every forward directed causal curve joining $p$ to $q$ must be a null geodesic segment (prop. 2.13, p. 729).
ii) Given a causal curve $\gamma$ from $p$ to $q$, (a): there is no neighborhood of $\gamma$ containing a timelike curve from $p$ to $q \Leftrightarrow(\mathrm{~b}): \gamma$ is a null geodesic segment from $p$ to $q$ without any point conjugate to $p$ between $p$ and $q$ (prop. 2.14, p. 729). In particular - (b) $\Rightarrow-(\mathrm{a})$ : If $\gamma$ is a null geodesic segment from $p$ to $q$ with a focal point conjugate to $p$ between $p$ and $q, \Rightarrow \exists \mathscr{V}$ open neighborhood of $\gamma$ which contains a timelike curve from $p$ to $q$.
iii) Given a causal curve $\gamma$ from a spacelike surface $S$ to $q$, (a): there is no neighborhood of $\gamma$ containing a timelike curve from $S$ to $q \Leftrightarrow$ (b): $\gamma$ is a null geodesic segment orthogonal to $S$ to $q$ without any point focal to $S$ between
$S$ and $q$ (prop. 2.14, p. 729). In particular, $-(\mathrm{b}) \Rightarrow-(\mathrm{a})$ : If $\gamma$ is a null geodesic segment emanating orthogonally from $S$ to $q$ with a focal point to $S$ between $S$ and $q, \Rightarrow \exists \mathscr{V}$ open neighborhood of $\gamma$ which contains a timelike curve from $S$ to $q$.
iv) If $N \subset M$ and $q \in E^{+}(N), \Rightarrow q$ lies on a future directed null geodesic segment from $N$.
18. Proposition: Let $S \subset M$ be a closed trapped surface and assume that the N.C.C. holds, i.e. $R_{\mu \nu} k^{\mu} k^{\nu} \geq 0, \forall$ null vectors $k^{\mu} . \Rightarrow(a): E^{+}(S)$ is compact, or (b): the spacetime is null geodesically incomplete to the future (past) ([8]: prop. 4.1, p. 780). (The disjunction connector "or" is exclusive.)

Proof. (We prove the Prop. "to the future"; the proof "to the past" is analogous.) Let us assume that the spacetime is null geodesically complete to the future, i.e. we assume - (b). Since $S$ is a future trapped surface, $\Rightarrow$ the expansions $\Theta_{ \pm}$ of both null geodesic congruences emanating orthogonally from $S$ are negative on $S$. Since $S$ is compact, the maximum and the minimum of $\Theta_{ \pm}$are atained on $S$. Let $\Theta_{M}$ be the maximum value of both $\Theta_{ \pm}$on $S\left(\Theta_{M}=-\left|\Theta_{M}\right|<0\right)$. Since the N.C.C. holds and the spacetime is null geodesically complete, then the Raychaudhuri equation implies that there is a focal point at a finite value of the affine parameter $\lambda_{M} \leq-\frac{2}{\Theta_{M}}=\frac{2}{\left|\Theta_{M}\right|}$. Let $K$ be the subset of $M$ containing all these null geodesics in both orthogonal congruences from $S$ up to $\lambda_{M}$ included. By construction, $K$ is compact and closed. Given that $E^{+}(S) \subset K$ (since if $q \in E^{+}(S) \Rightarrow q$ lies in a forward directed null geodesic segment from $S$ (by 17.iv)), to see that $E^{+}(S)$ is compact it is enough to show that $E^{+}(S)$ is closed (since any closed set in a compact set is compact). Let $\left\{p_{n}\right\}$ be a sequence of points in $E^{+}(S)$ converging to $p$, i.e. $p_{n} \rightarrow p$ as $n \rightarrow+\infty$. By construction, $E^{+}(S) \subset K \subset J^{+}(S)$. By the closedness of $K$, $p \in K \subset J^{+}(S)=E^{+}(S) \cup I^{+}(S)$. So, it is enough to prove that $p \notin I^{+}(S)$. If $p \in I^{+}(S) \Rightarrow \exists \mathscr{U}_{p} \subset I^{+}(S)$, what $\Rightarrow$ that some $p_{m} \in \mathscr{U}_{p}$ which is imposible since $p_{n} \in E^{+}(S)$ and $E^{+}(S) \cap I^{+}(S)=\phi . \Rightarrow p \in E^{+}(S)$ i.e. $E^{+}(S)$ is closed and $\Rightarrow$ compact.

Note: We have proved that $-(\mathrm{b}) \Rightarrow(\mathrm{a})$. This amounts to $-(\mathrm{a}) \Rightarrow$ (b), i.e.
If $E^{+}(S)$ is not compact $\Rightarrow$ the spacetime is null geodesically incomplete to the future.
19. Proposition: Let $p \in M$ and assume that the N.C.C. holds. If the expansion $\Theta$ of the forward directed null geodesic congruence emanating from $p$ becomes $<0$ along any geodesic of the congruence, $\Rightarrow(a): E^{+}(p)$ is compact, or (b): the spacetime is null geodesically incomplete ([8]: prop. 4.2, p. 780). (The disjunction connector "or" is exclusive.)

Proof. Assume that the spacetime is null geodesically complete, i.e. we assume -(b). If $\Theta<0$ at some point in each null geodesic from $p$ and the N.C.C. holds, by geodesic completeness we know that there will be a conjugate point to $p$ along each geodesic before or at the finite value $\frac{2}{|\Theta|}$ of the affine parameter. $\Rightarrow$, as
in the previous Proposition, there exists a compact set $K$ and, with analogous following steps, one arrives at the conclusion that $E^{+}(p)$ is compact.

Note: We proved that $-(\mathrm{b}) \Rightarrow$ (a). This is equivalent to $-(\mathrm{a}) \Rightarrow$ (b), i.e.
If $E^{+}(p)$ is not compact $\Rightarrow$ the spacetime is null geodesically incomplete to the future.
20. A non-empty achronal subset $N$ of $M$ is called future (past) trapped if $E^{+}(N)\left(E^{-}(N)\right)$ is compact.
A proper achronal boundary is the boundary of a future set.
21. Proposition: Let $N$ be a subset of $M$, and let $J^{+}(N)$ be closed, i.e. $\overline{J^{+}(N)}=J^{+}(N) . \Rightarrow$ : i) $E^{+}(N)=\partial I^{+}(N)$; ii) $E^{+}(N)$ is boundaryless; iii) $E^{+}(N)$ is achronal; iv) $E^{+}(N)$ is a proper achronal boundary.

Proof:
i) $E^{+}(N)=J^{+}(N) \backslash I^{+}(N)=J^{+}(N) \backslash \operatorname{int} J^{+}(N)=\partial J^{+}(N)$.
ii) $\partial E^{+}(N)=\partial\left(\partial J^{+}(N)\right)=\left(\partial^{2}\right) J^{+}(N)=\phi$.
iii) $\partial J^{+}(N)=\partial I^{+}(N)$, and $I^{+}(N)$ is achronal.
iv) $J^{+}(N)$ is a future set.

## 4. Penrose Singularity Theorem

Theorem: Let $M$ be a spacetime. Assume the N.C.C. holds. If there exists a non-compact Cauchy surface $\Sigma$ and a closed trapped surface $S$, then the spacetime is null geodesically incomplete.

Notes: i) Recall that $\operatorname{dim} \Sigma=3, \operatorname{dim} S=2$. ii) The three conditions stated in the theorem are respectively known as the curvature condition, the causality condition, and the initial/boundary condition. iii) Since by 14 ., $M \cong \mathbb{R} \times \Sigma, M$ connected $\Rightarrow \Sigma$ connected.

Proof: i) Suppose the spacetime is null geodesically complete. Since $S$ is a closed trapped surface (future or past) and the N.C.C. holds, then by Prop. 18, $E^{+}(S)$ is compact. By Prop. 15.iv and 15.xii, $\partial J^{+}(S)$ is achronal; on the other hand, by Prop. 21.i, $E^{+}(S)=J^{+}(S) \backslash I^{+}(S)=\partial J^{+}(S)$, and by Prop. 21.iii, $E^{+}(S)$ is achronal. Moreover, from Prop. 21.iv, $E^{+}(S)$ is a proper achronal boundary.
ii) Consider a timelike geodesic congruence in the spacetime. Since $\Sigma$ is a Cauchy surface i.e. the spacetime is globally hyperbolic, then every curve of the congruence intersects $\Sigma$ exactly once, and since $E^{+}(S)$ is achronal, then every curve of the congruence intersects $E^{+}(S)$ at most once. We take the timelike geodesic congruence such that each of its curves intersects $E^{+}(S)$ exactly once.
iii) Define the map $f: E^{+}(S) \rightarrow \Sigma$ which transports the points in $E^{+}(S)$ to $\Sigma$ along the curves of the timelike geodesic congruence. Since $f$ is continuous (because of the continuity of the congruence) and 1-1, then $E^{+}(S)$ is homeomorphic to its image $T$ on $\Sigma\left(f\left(E^{+}(S)\right) \equiv T \subset \Sigma\right)$, then $T$ is compact since $E^{+}(S)$ is compact. But $E^{+}(S)$ (or $T$ ), being a proper achronal boundary, is an embedded 3-dimensional submanifold (and therefore open and without boun-
dary) of $\Sigma$ ([8]: prop. 2.16, p. 732). Then $E^{+}(S)$ is an open subset of $\Sigma$. Since $\Sigma$ is Hausdorff and, by Prop. 18, $E^{+}(S)$ is compact, then $E^{+}(S)$ must also be closed ([23]: Thm. (3.6)). On the other hand, since $\Sigma$ is connected, then $E^{+}(S)=\phi$ or $E^{+}(S)=\Sigma$. But $E^{+}(S) \neq \phi$ because $M$ is null geodesically complete; then $E^{+}(S)=\Sigma$, what is a contradiction since $\Sigma$ is non-compact. So, $E^{+}(S)$ can not be compact. The contradiction comes from the assumption that the spacetime is null geodesically complete. So, $M$ is null geodesically incomplete.

## Acknowledgements

The author thanks for hospitality to the Instituto de Astronomía y Física del Espacio (IAFE) of the Universidad de Buenos Aires and CONICET, Argentina, where part of this work was done; and to Oscar Brauer at the University of Leeds, U.K., for the drawing of Figure 1 and Figure 2.

## Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

## References

[1] Penrose, R. (1965) Gravitational Collapse and Space-Time Singularities. Physical Review Letters, 14, 57-59. https://doi.org/10.1103/PhysRevLett.14.57
[2] Raychaudhuri, A. (1955) Relativistic Cosmology I. Physical Review, 98, 1123-1126. https://doi.org/10.1103/PhysRev.98.1123
[3] Hawking, S.W. and Ellis, G.F.R. (1973) The Large Scale Structure of Space-Time. Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9780511524646
[4] Joshi, P.S. (1993) Global Aspects in Gravitation and Cosmology. Clarendon Press, Oxford.
[5] Godinho, L. and Natário, J. (2014) An Introduction to Riemannian Geometry: With Applications to Mechanics and Relativity (Universitext). Springer, Berlin. https://doi.org/10.1007/978-3-319-08666-8
[6] Galloway, G.J. (2014) Notes on Lorentzian Causality. ESI-EMS-IAMP Summer School on Mathematical Relativity, Vienna.
[7] Witten, E. (2020) Light Rays, Singularities, and All That. Reviews of Modern Physics, 92, Article ID: 045004. https://doi.org/10.1103/RevModPhys.92.045004
[8] Senovilla, J.M.M. (1998) Singularity Theorems and Their Consequences. General Relativity and Gravitation, 30, 701-848. https://doi.org/10.1023/A:1018801101244
[9] Senovilla, J.M.M. (2015) The 1965 Penrose Singularity Theorem. Classical and Quantum Gravity, 32, Article ID: 124008. https://doi.org/10.1088/0264-9381/32/12/124008
[10] Hawking, S.W. (1965) Occurrence of Singularities in Open Universes. Physical Review Letters, 15, 689-690. https://doi.org/101103/PhysRevLett.15.689
[11] Hawking, S.W. and Penrose, R. (1970) The Singularities of Gravitational Collapse and Cosmology. Proceedings of the Royal Society of London. Series A, 314, 529-548.
https://doi.org/10.1098/rspa.1970.0021
[12] Álvarez, E. (2021) Windows on Quantum Gravity. Fortschritte der Physik, 69, Article ID: 2000080. https://doi.org/10.1002/prop. 202000080
[13] Modesto, L. (2004) Disappearance of the Black Hole Singularity in Loop Quantum Gravity. Physical Review D, 70, Article ID: 124009.
https://doi.org/10.1103/PhysRevD.70.124009
[14] Wall, A.C. (2013) The Generalized Second Law Implies a Quantum Singularity Theorem. Classical and Quantum Gravity, 30, Article ID: 165003. https://doi.org/10.1088/0264-9381/30/16/165003
[15] Poisson, E. (2004) A Relativist's Toolkit: The Mathematics of Black Hole Mechanics. Cambridge University Press, Cambridge.
[16] Socolovsky, M. (2021) Quantum Propagators for Geodesic Congruences. Theoretical Physics, 6, 9-17.
[17] Wald, R.M. (1984) General Relativity. The University of Chicago Press, Chicago. https://doi.org/10.7208/chicago/9780226870373.001.0001
[18] Kar, S. and Sengupta, S. (2007) The Raychaudhuri Equations: A Brief Review. Pramana, 69, 49-76. https://doi.org/10.1007/s12043-007-110-9
[19] Tipler, F.J. (1978) General Relativity and Conjugate Ordinary Differential Equations. Journal of Differential Equations, 30, 165-174. https://doi.org/10.1016/0022-0396(78)90012-8
[20] Teschl, G. (2012) Ordinary Differential Equations and Dynamical Systems. American Mathematical Society, Providence. https://doi.org/10.1090/gsm/140
[21] Feynman, R.P. and Hibbs, A.R. (1965) Quantum Mechanics and Path Integrals. McGraw-Hill, New York.
[22] Schulman, L.S. (1981) Techniques and Applications of Path Integration. Wiley, New York. https://doi.org/10.1063/1.2914703
[23] Amstrong, M.A. (1983) Basic Topology, Undergraduate Texts in Mathematics. Springer, Berlin. https://doi.org/10.1007/978-1-4757-1793-8.

