EPFL– Fall 2025
SOLUTIONS: Series 9

Differential Geometry IV: General relativity

G. Moschidis
19 Nov. 2025

9.1 In this exercise, we will establish *Birkhoff's theorem* for spherically symmetric solutions to the vacuum Einstein equations in $3+1$-dimensions.

(a) Let $(\mathcal{M}^{3+1}, g)$ be a Lorentzian manifold such that $\mathcal{M} = \mathcal{Q}^{1+1} \times \mathbb{S}^2$ and, in any local coordinate chart (x^0, x^1) on $\mathcal{Q}$ and using the standard (θ, ϕ) coordinates on $\mathbb{S}^2$, g takes the form

$$g = \tilde{g}_{AB} dx^A dx^B + r^2 \big(d\theta^2 + \sin^2\theta d\phi^2 \big)$$

with $A, B \in \{0, 1\}$ and:

- $\tilde{g}_{AB}$ and r depend only on x^0, x^1,
- $r > 0$.

Deduce that $(\mathcal{M}, g)$ is spherically symetric, i.e. $SO(3)$ acts isometrically on $(\mathcal{M}, g)$ with spherical orbits. Show also that, around any point $p \in \mathcal{Q}$, there exists a local coordinate system (u, v, θ, ϕ) around $\{p\} \times \mathbb{S}^2$ such that

$$g = -\Omega^2(u,v) du dv + r^2(u,v) \big(d\theta^2 + \sin^2\theta d\phi^2 \big).$$

(such a coordinate system is called *double null*). *Hint: Use Exercise 2.3.*

Remark. It can be shown that any spherically symmetric spacetime can be expressed locally in the above form.

(b) Assume that $(\mathcal{M}, g)$ above satisfies the vacuum Einstein equations $Ric_{\alpha\beta} = 0$. In double null coordinates, it can be easily calculated that this system of equations takes the following form in terms of the metric components Ω and r:

$$\begin{aligned}
\partial_u \partial_v (r^2) &= -\frac{1}{2}\Omega^2, \\
\partial_u \partial_v \log(\Omega^2) &= \frac{\Omega^2}{2r^2}\big(1 + 4\Omega^{-2}\partial_u r \partial_v r\big), \\
\partial_u(\Omega^{-2}\partial_u r) &= 0, \\
\partial_v(\Omega^{-2}\partial_v r) &= 0.
\end{aligned}$$

(note that this is an overdetermined system; this is why, at the end of the day, Birkhoff's theorem holds). Show that the quantity $m : \mathcal{Q} \to \mathbb{R}$ defined by

$$m \doteq \frac{r}{2}\big(1 - g^{\alpha\beta}\partial_\alpha r \partial_\beta r\big) = \frac{r}{2}\big(1 + 4\Omega^{-2}\partial_u r \partial_v r\big)$$

(which is known as the *Hawking mass* of the sphere $\{p\} \times \mathbb{S}^2$) is locally constant on $\mathcal{Q}$.

(c) Let g_M be the Schwarzschild metric for $M \in \mathbb{R}$. Show that, in this case, $m = M$.

(d) Let $p \in \mathcal{Q}$ and assume, without loss of generality, that $(u(p), v(p)) = 0$. Show that there exists an open neighborhood $\mathcal{U}$ of $\{p\} \times \mathbb{S}^2$ in $\mathcal{M}$ and an open neighborhood $\mathcal{U}_{Sch}$ of a point q in the maximally extended Schwarzschild spacetime with $M = m(p)$ (chosen so that $r(q) = r(p)$) which are isometric. *Hint: Choose coordinates u, v on $\mathcal{U}_{Sch}$ so that the functions $\partial_u r(u, 0)$ and $\partial_v r(0, v)$ are the same in both spacetime domains. Deduce that the functions $r(u, v)$ and $\Omega(u, v)$ are the same for both spacetime domains, using the system of equations.*

Solution. (a) Since, in the (x^1, x^2, θ, ϕ) coordinate system, the metric g takes the form

$$g = \tilde{g}_{AB} dx^A dx^B + r^2 g_{\mathbb{S}^2}(\theta, \phi),$$

where $g_{\mathbb{S}^2} = d\theta^2 + \sin^2\theta d\phi^2$ and $\tilde{g}_{AB}, r$ are independent of θ, ϕ, any map $F : \mathbb{S}^2 \to \mathbb{S}^2$, $(\theta, \phi) \to (\theta', \phi')$, which is an isometry of the spherical metric $g_{\mathbb{S}^2}$ extends to a map $(x^1, x^2, \theta, \phi) \to (x^1, x^2, \theta', \phi')$ which is an isometry of $(\mathcal{M}, g)$. Therefore, the group of isometries $SO(3)$ of $(\mathbb{S}^2, g_{\mathbb{S}^2})$ acts isometrically on $(\mathcal{M}, g)$ with orbits of the form $\{p\} \times \mathbb{S}^2$ for each $p \in \mathcal{Q}$.

Note that the components $\tilde{g}_{AB}$ in the expression above define a Lorentzian metric $\tilde{g}$ on $\mathcal{Q}$ (it is easy to check that it is symmetric and has Lorentzian signature, since the $3+1$-metric g has Lorentzian signature; the transformation formulas for g under changes of coordinates $(x^1, x^2, \theta, \phi) \to (y^1, y^2, \theta, \phi)$ imply that $\tilde{g}$ indeed transforms as a $(0, 2)$-tensor). Using Exercise 2.3, for any $p \in \mathcal{Q}$, there exists a coordinate system (u, v) in a neighbohood $\mathcal{U} \subset \mathcal{Q}$ of p in which the metric $\tilde{g}$ takes the form

$$\tilde{g} = -\Omega^2(u, v) du dv$$

for some $\Omega \in C^\infty(\mathcal{U})$. Therefore, in the (u, v, θ, ϕ) coordinate system on $\mathcal{U} \times \mathbb{S}^2 \subset \mathcal{M}$, the metric g takes the form

$$g = -\Omega^2(u, v) du dv + r^2(u, v)\big(d\theta^2 + \sin^2\theta d\phi^2\big).$$

(b) We can readily compute using the expression for m:

$$\begin{aligned}\partial_u m &= \partial_u\Big(\frac{r}{2}\big(1 + 4\Omega^{-2}\partial_u r \partial_v r\big)\Big)\\ &= \frac{\partial_u r}{2}\big(1 + 4\Omega^{-2}\partial_u r \partial_v r\big) + 2r\partial_u(\Omega^{-2}\partial_u r)\partial_v r + 2r\Omega^{-2}\partial_u r \partial_u \partial_v r.\end{aligned}$$

Using the Einstein vacuum equations for the (Ω, r) pair (as listed in the exercise), among them, in particular, the relations

$$\partial_u(\Omega^{-2}\partial_u r) = 0 \quad \text{and} \quad \partial_u\partial_v r = -\frac{1}{4r}\Omega^2 - \frac{1}{r}\cdot\partial_u r \partial_v r,$$

we obtain that $\partial_u m = 0$. Similarly, $\partial_v m = 0$. Therefore, m is locally constant on $\mathcal{Q}$ (i.e. it is constant in every connected component of $\mathcal{Q}$).

(c) Using the more geometric relation

$$m = \frac{r}{2}\big(1 - g^{\alpha\beta}\partial_\alpha r \partial_\beta r\big),$$

we can compute for the Schwarzschild metric

$$g_M = -\Big(1 - \frac{2M}{r}\Big)dt^2 + \Big(1 - \frac{2M}{r}\Big)^{-1} dr^2 + r^2\big(d\theta^2 + \sin^2\theta d\phi^2\big),$$

in the (t, r, θ, ϕ) coordinate system covering region I of the maximal extension:

$$m = \frac{r}{2}\big(1 - g^{\alpha\beta}\partial_\alpha r \partial_\beta r\big) = \frac{r}{2}\big(1 - g^{rr}\partial_r r \partial_r r\big) = \frac{r}{2}\big(1 - (1 - \frac{2M}{r})\big) = M.$$

Since m is locally constant and the maximally extended Schwarzschild spacetime is connected, we deduce that $m = M$ on the whole of the Schwarzschild spacetime.

(d) For this part, we will use the subscript S (i.e. Ω_S and r_S) to denote quantities associated to the Schwarzschild metric. Let us consider the maximally extended Schwarzschild metric in the Kruskal coordinates (u, v, θ, ϕ) (the precise choice of these coordinates is not important; any double null coordinate system covering the whole Schwarzschild manifold would suffice):

$$g_M = -\Omega_S^2 du dv + r_S^2(u,v)\big(d\theta^2 + \sin^2\theta d\phi^2\big),$$

where

$$\Omega_S^2 = \frac{32M^3}{r_S}\exp\Big(-\frac{r_S}{2M}\Big)$$

and $r_S(u,v)$ is defined by the implicit relation:

$$\Big(\frac{r_S(u,v)}{\frac{2M}{r}-1}\Big)\exp\Big(\frac{r_S}{2M}\Big) = -uv$$

(recall that, in this case, the range of the coordinates (u,v) is $\mathbb{R}^2 \cap \{uv < 1\}$).

Our aim is to show that, after applying a coordinate transformation of the form $(u,v,\theta,\phi) \to (u' = u'(u), v' = v'(v), \theta, \phi)$ in a neighborhood $\mathcal{U} \times \mathbb{S}^2$ of the point $p \in \mathcal{M} = \mathcal{Q} \times \mathbb{S}^2$, we can achieve

$$\Omega(u,v) = \Omega_S(u,v) \quad \text{and} \quad r(u,v) = r_S(u,v) \quad \text{for all } (u,v) \in \mathcal{U} \tag{1}$$

(this will imply, in particular, that $(\mathcal{M}, g)$ is locally isometric, around p, to the open domain in Schwarzschild spacetime parametrized by $(u,v,\theta,\phi) \in \mathcal{U} \times \mathbb{S}^2$). To this end, let us fix the Schwarzschild mass parameter M so that

$$M = m(p)$$

(recall that $m_S = M$ on the Schwarzschild spacetime). Note that since the metric coefficient Ω^2 can be expressed in terms of the Hawking mass m and the sphere radius r by the relation

$$\Omega^2\big(1 - \frac{2m}{r}\big) = -4\partial_u r \partial_v r, \tag{2}$$

showing (1) reduces, in this case, to simply establishing that

$$r(u,v) = r_S(u,v) \quad \text{for } (u,v) \in \mathcal{U}.$$

Let $q = (u_0, v_0, \theta, \phi)$ be a point on the maximally extended Schwarzschild spacetime satisfying

$$r_S(q) = r(p).$$

We will distinguish two cases:

1. In the case when $r_S(q) = r(p) \neq 2M$, we will make no further assumptions on the point q.

2. In the case when $r_S(q) = r(p) = 2M$, the expression (2) (together with our assumption that $m(p) = M = m_S(q)$) implies that

$$\partial_u r(p) \partial_v r(p) = 0 = \partial_u r_S(q) \partial_v r_S(q).$$

Therefore, in this case, we choose the point q on the subset $\{r = 2M\} = \{u = 0\} \cup \{v = 0\}$ of Schwarzschild spacetime so that $\partial_u r_S(q) = 0$ if and only if $\partial_u r(p) = 0$ and $\partial_v r_S(q) = 0$ if and only if $\partial_v r(p) = 0$.

By applying a coordinate transformation of the form $(u, v) \to (u + u_0, v + v_0)$ on $(\mathcal{M}, g)$, we can assume without loss of generality that

$$\big(u(p), v(p)\big) = \big(u(q), v(q)\big).$$

Our choice of the point q above then implies that the functions $\frac{\partial_u r(u,v)}{\partial_u r_S(u,v)}$ and $\frac{\partial_v r(u,v)}{\partial_v r_S(u,v)}$ are well-defined and continuous in (u, v) a neighborhood of (u_0, v_0).

Let $\mathcal{U} = (u_0 - \delta, u_0 + \delta) \times (v_0 - \delta, v_0 + \delta)$ for some $\delta > 0$ small enough. We will show that there exists a coordinate transformation of the form $(u, v) \to (u', v') = (F(u), G(v))$ on $\mathcal{U}$ with $F(u_0) = u_0$ and $G(v_0) = v_0$ such that, in the new coordinate system on $\mathcal{M}$, we have

$$r(u_0, v) = r_S(u_0, v) \quad \text{and} \quad r(u, v_0) = r_S(u, v_0) \quad \text{for } v \in (v_0 - \delta, v_0 + \delta),\ u \in (u_0 - \delta, u_0 + \delta), \text{ respectively.}$$

Since $r(u_0, v_0) = r_S(u_0, v_0)$, it suffices to have

$$\partial_v r(u_0, v) = \partial_v r_S(u_0, v) \quad \text{and} \quad \partial_u r(u, v_0) = \partial_u r_S(u, v_0) \quad \text{for } v \in (v_0 - \delta, v_0 + \delta),\ u \in (u_0 - \delta, u_0 + \delta), \text{ respectively.}$$

Thus, the coordinate transformation functions $F(u)$ and $G(v)$ are uniquely determined by the following conditions (with respect to the old coordinates):

$$\begin{cases} \frac{dF}{du}(u) = \frac{\partial_u r(u,v_0)}{\partial_u r_S(u,v_0)}, \quad \frac{dG}{dv}(v) = \frac{\partial_v r(u_0,v)}{\partial_v r_S(u_0,v)}, \\ F(u_0) = u_0, \quad G(v_0) = v_0 \end{cases}$$

With respect to the new coordinate system on $\mathcal{M}$, we have the property that both the functions $r(u, v)$ and $r_S(u, v)$ satisfy the non-linear wave equation

$$\partial_u \partial_v r = \frac{2M}{r(r - 2M)} \partial_u r \partial_v r \quad \text{on} \quad (u_0 - \delta, u_0 + \delta) \times (v_0 - \delta, v_0 + \delta)$$

(this is simply equation $\partial_u \partial_v (r^2) = \frac{1}{2}\Omega^2$ for the metrics g and g_M, where we have used again the relation (2) to express Ω^2 in terms of (m, r) together with the condition that $m = M$ in a neighborhood of $p \in \mathcal{M}$) and satisfy $r = r_S$ along $\{u = u_0\}$ and $\{v = v_0\}$. Therefore, the uniqueness property for non-linear wave equations implies that

$$r = r_S \quad \text{on} \quad (u_0 - \delta, u_0 + \delta) \times (v_0 - \delta, v_0 + \delta),$$

as desired.

9.2 Let $(\mathcal{M}, g)$ be a Lorentzian manifold and $S \subset \mathcal{M}$ be a submanifold. For any vector field W along S which is orthogonal to S, we will define the associated second fundamental form $\chi^{(W)} : \Gamma(S) \times \Gamma(S) \to \mathbb{R}$ by the relation

$$\chi^{(W)}(X, Y) \doteq g(\nabla_X W, Y),$$

where ∇ denotes the connection of g and we think of X, Y as being extended to vector fields in $\mathcal{M}$.

(a) Show that $\chi^{(W)}$ is well defined independently of the choice of extensions of X, Y. Show also that it is a symmetric $(0, 2)$-tensor field.

(*b) Assume that S is *spacelike*; we will also denote the induced (Riemannian) metric on S by h. Let W be a non-vanishing vector field on $\mathcal{M}$ which is orthogonal to S and let $\Phi_t^{(W)}$ be the flow map of W. For the one parameter family of surfaces $S_t = \Phi_t^{(W)}(S)$, with induced metrics h_t, show that, in any coordinate chart (x^1, x^2) on $S_t$which is transported along the flow of W:

$$\frac{d}{dt}\sqrt{\det(h_t)}\Big|_{t=0} = \mathrm{tr}_h \chi^{(W)} \cdot \sqrt{\det(h)},$$

where $\mathrm{tr}_h \chi^{(W)} \doteq h^{AB} \chi^{(W)}_{AB}$ For this reason, $\mathrm{tr}_h \chi^{(W)}$ is usually called the *expansion* in the direction of W, since it measures the rate of change of the volume form of S. *(Hint: You might want to use Jacobi's formula from linear algebra:* $\frac{d}{dt}\log(\det M) = \mathrm{tr}(M^{-1}\frac{d}{dt}M)$ *for a square-matrix valued function* $M(t)$*.)*

(c) We will now restrict to the case when M is $3 + 1$ dimensional and time oriented and that S is a 2-dimensional surface.. in that case, at each point $p \in S$, the normal bundle $TS^\perp$ is spanned by two **future directed null** vector fields along S, which we will denote with L and $\underline{L}$. We will also denote the induced (Riemannian) metric on S by h. We will say that such a surface S is **trapped** if it is compact and, at every point on S, both null expansions are negative, i.e.

$$\mathrm{tr}_h \chi^{(L)}, \mathrm{tr}_h \chi^{(\underline{L})} < 0.$$

Show that, on the maximally extended Schwarzschild spacetime, the spheres of symmetry are trapped if and only if they correspond to points in the region II of the Penrose diagram (i.e. the black hole region).

Remark. We will later see in class that, as a consequence of Penrose's incompleteness theorem, if an asymptotically flat spacetime contains a trapped surface S, then this is necessarily inside a black hole, i.e. $J^+[S]$ does not reach future null infinity $\mathscr{I}^+$. Since the condition defining a trapped surface is anopen condition, a trapped surface remains trapped even under small changes of the metric; thus, small perturbations of Schwarzschild spacetime still contain a black hole.

Solution.

(a) Let X, Y be tangent vector fields to S, extended arbitrarily to vector fields on $\mathcal{M}$. Let us also extend W arbitrarily to a vector field on $\mathcal{M}$. In view of the fact that $W|_S$ is orthogonal to S and X is tangential to S, we have

$$g(W,Y)|_S = 0 \Rightarrow X\big(g(W,Y)\big)\big|_S = 0 \Rightarrow g\big(\nabla_X W, Y\big)\big|_S + g\big(W, \nabla_X Y\big)\big|_S 0 =,$$

from which we deduce that:

$$\chi^{(W)}(X,Y) = -g\big(W, \nabla_X Y\big)|_S. \tag{3}$$

The fact that $\chi^{(W)}(X,Y)$ is independent of the choice of extension of X, Y, W now follows from the fact that $\nabla_X Y|_S$ is independent of the choice of extensions from X, Y (which can be readily verified in any local coordinate system $(x^1, \dots, x^{dim(\mathcal{M})})$ on $\mathcal{U} \subset \mathcal{M}$ in which $S \cap U$ is the set $\{x^1 = \dots x^{dim(\mathcal{M})-dim(S)} = 0\}$).

To prove that $\chi^{(W)}$ is a symmetric $(0,2)$-tensor field, we will show that

- For all smooth functions $f_1, f_2 : \mathcal{M} \to \mathbb{R}$ and smooth vector fields X, Y_1 and Y_2, we have
$$\chi^{(W)}(X, f_1 Y_1 + f_2 Y_2) = f_1 \chi^{(W)}(X, Y_1) + f_2 \chi^{(W)}(X, Y_2). \tag{4}$$
- For all smooth vector fields X and Y, we have
$$\chi^{(W)}(X,Y) = \chi^{(W)}(Y,X). \tag{5}$$

Relation (4), follows immediately from the fact that g is a tensor: For all smooth functions f_1, f_2 and smooth tangent vectors X, Y_1, Y_2, we have

$$\begin{aligned}\chi^{(W)}(X, f_1 Y_1 + f_2 Y_2) &= g(\nabla_X W, f_1 Y_1 + f_2 Y_2)\\ &= f_1 g(\nabla_X W, Y_1) + f_2 g(\nabla_X W, Y_2)\\ &= f_1 \chi^{(W)}(X, Y_1) + f_2 \chi^{(W)}(X, f Y_2).\end{aligned}$$

Relation (5) follows immediately from (3), the fact that the Levi-Civita connection is symmetric and the fact that, if $X, Y \in \Gamma(\mathcal{M})$ are tangent to S, then $[X,Y]$ is also tangent to S (which can be verified in local coordinates as above):

$$\chi^{(W)}(X,Y) - \chi^{(W)}(Y,X) = -g\big(W, \nabla_X Y\big) + g\big(W, \nabla_Y X\big) = -g\big(W, \nabla_X Y - \nabla_Y X\big) = -g(W, [X,Y]) = 0.$$

(b) Let $S \subset \mathcal{M}$ be a spacelike hypersurface and $W \in \Gamma(\mathcal{M})$ be as in the statement of the exercise; recall that W doesn't vanish anywhere. Let also $n = dim(S)$, $m = dim(\mathcal{M})$. Note that, for any $p \in S$, since $W|_p \perp T_p S$ and $T_p S$ is spacelike, we have that $W|_p$ is transversal to $T_p S$.[1]

For any $p \in S$, let $(x^1, \dots, x^n)$ be a local coordinate system in a neighborhood $\mathcal{U} \subset S$ of p. Since $W|_p \neq 0$ and is transversal to $T_p S$, the flow map $\Phi^{(W)} : (-\delta, \delta) \times U \to \mathcal{M}$, $(t, x^1, \dots, x^n) \to \Phi^{(W)}_t(x^1, \dots, x^n)$ is an embedding on a small neighborhood of $(0, p)$ (to see this, use the implicit function theorem: note that the differential $d\Phi^{(W)}|_{(0,p)}$ satisfies $d\Phi^{(W)}|_{(0,p)}(X) = X$ for any $X \in T_p S$

---

[1] That wouldn't necessarily be the case if $T_p S$ was null, for instance.

and $d\Phi^{(W)}|_{(0,p)}(\partial_t) = W|_p$, hence $d\Phi^{(W)}|_{(0,p)}$ is injective since $W|_p$ is transversal to T_pS). Without loss of generality, we will assume that $\delta > 0$ has been fixed small enough and $\mathcal{U}$ is a small enough neighborhood of p in S so that $\Phi^{(W)} : (-\delta,\delta)\times U \to \mathcal{M}$ is an embedding.

Let us denote with $\mathcal{N} = \Phi^{(W)}\Big((-\delta,\delta)\times U\Big) \subset \mathcal{M}$. Note that, through the map $\Phi^{(W)}$, $(t, x^1, \dots, x^n)$ defines a coordinate system on $\mathcal{N}$, such that $\partial_t = W$ and $(x^1,\dots,x^n)$ is a coordinate system on the slices $U_t = \Phi^{(W)}_t(U)$ for each $t \in (-\delta,\delta)$. Thus, if h_t denotes the induced metric on U_t (so that $h = h_0$), we can compute:

$$\begin{aligned}\partial_t(h_t)_{ij} &= \partial_t\big(h_t(\partial_i,\partial_j)\big) = h_t\big(\nabla_{\partial_t}\partial_i,\partial_j\big) + h_t\big(\nabla_{\partial_t}\partial_j,\partial_i\big)\\ &\overset{[\partial_t,\partial_i]=0}{=} h_t\big(\nabla_{\partial_i}\partial_t,\partial_j\big) + h_t\big(\nabla_{\partial_j}\partial_i,\partial_i\big).\end{aligned}$$

Evaluating the above at $t = 0$ and using the fact that $\partial_t = W$ and $\chi^{(W)}_{ij} = h(\nabla_{\partial_i}W,\partial_j)$, we get

$$\partial_t(h_t)_{ij}|_{t=0} = \chi^{(W)}_{ij} + \chi^{(W)}_{ji} = 2\chi^{(W)}_{ij}. \tag{6}$$

Using Jacobi's formula $\frac{d}{dt}\log(\det M) = \mathrm{tr}(M^{-1}\frac{d}{dt}M)$, we then get:

$$\partial_t\sqrt{\det(h_t)}\Big|_{t=0} = \frac{1}{2}\frac{\partial_t\big(\det(h_t)\big)\big|_{t=0}}{\sqrt{\det(h)}} = \frac{1}{2}\frac{\det(h)\cdot\big(h_t^{ij}\cdot\partial_t(h_t)_{ij}\big)|_{t=0}}{\sqrt{\det(h)}} = \frac{1}{2}\sqrt{\det(h)}\cdot 2h^{ij}\chi^{(W)}_{ij} = \sqrt{\det(h)}\mathrm{tr}_h\chi^{(W)}.$$

Let us denote by S_t the image $\Phi^{(W)}_t(S)$. We will fix a coordinate system $(x^1,\dots,x^n)$ on $\Phi^{(W)}_t(U)$ so that the functions x^i are transported along the flow of W.

(c) Recall that the maximally extended Schwarzschild spacetime $(\mathcal{M}^{(M)}_{Sch}, g_M)$ is covered by the Kruskal–Szekeres double null coordinate system (U, V, θ, φ), in which $\mathcal{M}^{(M)}_{Sch} \simeq \mathcal{U}^{1+1}\times\mathbb{S}^2$ with $\mathcal{U} = \big\{UV < 1\big\}$ and

$$g_M = -\frac{32M^3}{r}e^{-\frac{r}{2M}}dUdV + r^2\big(d\theta^2 + \sin^2\theta d\varphi^2\big),$$

where $r = r(U,V)$ is defined implicitly by the relation

$$U\cdot V = \Big(1 - \frac{r}{2M}\Big)e^{\frac{r}{2M}} \tag{7}$$

(any local choice of a spherically symetric double null coordinate system would work equally well for the computation below). We fix a time orientation so that the null vectors ∂_U, ∂_V are future directed. The black hole interior (the region II in the Penrose diagram) corresponds to the region $\big\{U > 0,\ V > 0\big\}$.

The spheres of symmetry correspond to the 2-surfaces $\{U, V = \mathit{const}\}$. Note that the induced metric on those surfaces is

$$h = r^2\big(d\theta^2 + \sin^2\theta d\varphi^2\big) = r^2 g_{\mathbb{S}^2},$$

while the future directed null normal vector fields $L, \underline{L}$ can be simply chosen to be $L = \partial_V$, $\underline{L} = \partial_U$. Note that, in this case, the flow map of L is simply a translation in the V coordinate (and similarly for $\underline{L}$); therefore, from the formula (6) (with the role of t now played by V), we have (with $(x^1, x^2) = (\theta, \varphi)$)

$$\chi_{ij}^{(L)} = \frac{1}{2}\partial_V h_{ij} = \frac{1}{2}\partial_V\Big(r^2 (g_{\mathbb{S}^2})_{ij}\Big) = r\partial_V r (g_{\mathbb{S}^2})_{ij}$$

(since the components of $(g_{\mathbb{S}^2})_{ij}$ only depend on (x^1, x^2)) and, similarly,

$$\chi_{ij}^{(\underline{L})} = r\partial_U r (g_{\mathbb{S}^2})_{ij}.$$

Therefore, the sphere of symmetry is trapped if and only if $\partial_U r < 0$ and $\partial_V r < 0$. Differentiating the implicit relation (7) with respect to U and V, it is easy to verify that this only holds when $U > 0$ and $V > 0$.

9.3 Let

$$T_{\mu\nu}[\phi] = \partial_\mu\phi\partial_\nu\phi - \frac{1}{2}g_{\mu\nu}g^{\alpha\beta}\partial_\alpha\phi\partial_{\beta\phi}$$

be the energy momentum tensor associated to the scalar wave equation $\Box_g\phi = 0$ on $(\mathcal{M}, g)$ (recall that $\Box_g \doteq g^{\alpha\beta}\nabla_\alpha\nabla_\beta\phi = \frac{1}{\sqrt{-\mathrm{det}g}}\partial_\alpha\big(\sqrt{-\mathrm{det}g}g^{\alpha\beta}\partial_\beta\big)$). For the first two questions, we will not assume that $\phi : \mathcal{M} \to \mathbb{R}$ solves any particular equation.

(a) Show that, for any $\phi \in C^\infty(\mathcal{M})$, any $p \in \mathcal{M}$ and any two future oriented causal vectors $V, W \in T_p\mathcal{M}$:

$$T_{\mu\nu}[\phi]V^\mu W^\nu \geqslant 0$$

(*Hint: Choose a suitable double null frame in* $T_p\mathcal{M}$). If V, W are moreover timelike, show that

$$T_{\mu\nu}[\phi]V^\mu W^\nu \geqslant c\sum_{i=0}^{n}|\partial_i\phi|^2,$$

with the constant $c > 0$ depending on V, W, g and the choice of local coordinates (but is independent of ϕ).

(b) Assume, now, that ϕ solves $\Box_g\phi = 0$. Show that

$$(\mathrm{div}T[\phi])_\nu \doteq g^{\alpha\beta}\nabla_\alpha T_{\beta\nu}[\phi] = 0.$$

(c) Show that, if, in addition, V is a Killing vector field of $(\mathcal{M}, g)$, then the 1-form $J^V_\nu[\phi] \doteq T_{\mu\nu}[\phi]V^\mu$ is divergence free, i.e.

$$\mathrm{div}J^V[\phi] \doteq g^{\alpha\beta}\nabla_\alpha J^V_\beta[\phi] = 0.$$

Solution.

(a) Note that, for any $V, W \in \Gamma(\mathcal{M})$:

$$T[\phi](V, W) = T_{\mu\nu}[\phi]V^\mu W^\nu$$

$$
\begin{aligned}
&= (V^\mu \partial_\mu \phi)(W^\nu \partial_\nu \phi) - \frac{1}{2}(g_{\mu\nu} V^\mu W^\nu)(g^{\alpha\beta}\partial_\alpha \phi \partial_\beta \phi) \\
&= V(\phi) W(\phi) - \frac{1}{2} g(V,W) g^{\alpha\beta}(\partial_\alpha \phi)(\partial_\beta \phi).
\end{aligned}
$$

Assume that both V and W are causal and future directed (so that, in particular, $g(V,W) \leqslant 0$). The tangent vectors V and W span the space $\Pi := \mathrm{span}(\{V,W\}) \subset T\mathcal{M}$. If the dimension of Π is one, namely if $W = \lambda V$ for some $\lambda > 0$ (the sign condition follows from the assumption that both vector fields are future directed), the claim follows easily: If V is null, then $T(V,W) = \lambda\big(V(\phi)\big)^2 \geqslant 0$; if V is timelike then, with respect to an orthonormal frame $\{e_0, e_1, \dots, e_n\}$ such that $V = \lambda' e_0$, we can compute from the above expression:

$$
T(V,W) = \lambda T(V,V) = \lambda(\lambda')^2 T(e_0,e_0) = \frac{1}{2}\lambda(\lambda')^2 \Big(\big(e_0(\phi)\big)^2 + \dots + \big(e_n(\phi)\big)^2 \Big).
$$

Thus, from now on, we will assume without loss of generality that V, W are not collinear, so that Π is a 2 dimensional plane. Note that Π is a *timelike* plane, since it contains at least one timelike vector.

We choose two future directed null vectors L and $\underline{L}$ spanning Π, normalized so that

$$
g(L,L) = 0, \quad g(\underline{L},\underline{L}) = 0, \quad g(L,\underline{L}) = -2 \tag{8}
$$

and express V and W with respect to this basis,

$$
V = \alpha_1 L + \alpha_2 \underline{L}, \quad W = \beta_1 L + \beta_2 \underline{L}.
$$

Note that, since $V, W, L, \underline{L}$ are all future directed (and hence all their pairwise inner products are non-negative), we must have

$$
\alpha_1, \alpha_2, \beta_1, \beta_2 \geqslant 0. \tag{9}
$$

Moreover, V, W are both strictly timelike if and only if

$$
\alpha_1, \alpha_2, \beta_1, \beta_2 > 0 \tag{10}
$$

(since if one of these coefficients (say α_1) is 0, then the corresponding vector (in this case, V) is equal to a null vector (in this case, $\underline{L}$).

Since Π is a 2-dimensional subspace, we can split the tangent space as

$$
T_p\mathcal{M} = \Pi \oplus \Pi^\perp,
$$

where

$$
\begin{aligned}
&\Pi := \mathrm{span}(\{V,W\}), \\
&\Pi^\perp := \{X \in T_p\mathcal{M} : g(X,Y) = 0,\ \forall Y \in \Pi\}
\end{aligned}
$$

Note that $\Pi^\perp$ is *spacelike* and of dimension $n-1$; let $\{e_1, \dots, e_{n-1}\}$ be an orthonormal basis of $\Pi^\perp$. Thus,

$$
g(L,e_i) = 0, \quad g(\underline{L},e_i) = 0, \quad g(e_i,e_j) = \delta_{ij} \tag{11}
$$

for all $i = 1, 2, \dots, n-1$. We can easily compute

$$\begin{aligned} T[\phi](L,L) &= L(\phi)L(\phi) - \frac{1}{2}g(L,L)g^{\alpha\beta}(\partial_\alpha\phi)(\partial_\beta\phi) = (L(\phi))^2, \\ T[\phi](\underline{L},\underline{L}) &= \underline{L}(\phi)\underline{L}(\phi) - \frac{1}{2}g(\underline{L},\underline{L})g^{\alpha\beta}(\partial_\alpha\phi)(\partial_\beta\phi) = (\underline{L}(\phi))^2 \end{aligned}$$

Moreover, there exist real ceofficients b, c and $\{d_i : i = 1, \dots, n-1\}$ so that

$$\mathrm{grad}\phi = bL + c\underline{L} + \sum_{i=1}^{n-1} d_i e_i$$

(where $\mathrm{grad}\phi^\alpha = \partial^\alpha\phi = g^{\alpha\beta}\partial_\beta\phi$). In fact, due to (8) and (11), these ceofficients are given by

$$\begin{aligned} b &= -\frac{1}{2}g(\mathrm{grad}\phi, \underline{L}) = -\frac{1}{2}d\phi(\underline{L}) = -\frac{1}{2}\underline{L}(\phi), \\ c &= -\frac{1}{2}g(\mathrm{grad}\phi, L) = -\frac{1}{2}d\phi(L) = -\frac{1}{2}L(\phi), \\ d_i &= g(\mathrm{grad}\phi, e_i) = d\phi(e_i) = e_i(\phi). \end{aligned}$$

Now, these yield that

$$\begin{aligned} T[\phi](L,\underline{L}) &= L(\phi)\underline{L}(\phi) - \frac{1}{2}g(L,\underline{L})g^{\alpha\beta}(\partial_\alpha\phi)(\partial_\beta\phi) \\ &= L(\phi)\underline{L}(\phi) + g^{\alpha\beta}(\partial_\alpha\phi)(\partial_\beta\phi) \\ &= L(\phi)\underline{L}(\phi) + g_{\alpha\beta}(\partial^\alpha\phi)(\partial^\beta\phi) \\ &= L(\phi)\underline{L}(\phi) + g\left(bL + c\underline{L} + \sum_{i=1}^{n-1} d_i e_i\right)\left(bL + c\underline{L} + \sum_{j=1}^{n-1} d_j e_j\right) \\ &= L(\phi)\underline{L}(\phi) + 2bcg(L,\underline{L}) + \sum_{i,j=1}^{n-1} d_i d_j g(e_i, e_j) \\ &= L(\phi)\underline{L}(\phi) - 4bc + \sum_{i=1}^{n-1} d_i^2 \\ &= L(\phi)\underline{L}(\phi) - 4\left(-\frac{1}{2}\underline{L}(\phi)\right)\left(-\frac{1}{2}L(\phi)\right) + \sum_{i=1}^{n-1} d_i^2 \\ &= L(\phi)\underline{L}(\phi) - \underline{L}(\phi)L(\phi) + \sum_{i=1}^{n-1} d_i^2 \\ &= \sum_{i=1}^{n-1} d_i^2 = \sum_{i=1}^{n-1} (e_i(\phi))^2. \end{aligned}$$

Since $T[\phi]$ is a tensor (multilinear map) and symmetric, we infer

$$T[\phi](V,W) = T[\phi](\alpha_1 L + \alpha_2\underline{L}, \beta_1 L + \beta_2\underline{L})$$

$$\begin{aligned}
&= \alpha_1\beta_1 T[\phi](L,L) + \alpha_1\beta_2 T[\phi](L,\underline{L}) + \alpha_2\beta_1 T[\phi](\underline{L},L) + \alpha_2\beta_2 T[\phi](\underline{L},\underline{L})\\
&= \alpha_1\beta_1 T[\phi](L,L) + (\alpha_1\beta_2 + \alpha_2\beta_1) T[\phi](L,\underline{L}) + \alpha_2\beta_2 T[\phi](\underline{L},\underline{L})\\
&= \alpha_1\beta_1 (L(\phi))^2 + (\alpha_1\beta_2 + \alpha_2\beta_1)\sum_{i=1}^{n-1}(e_i(\phi))^2 + \alpha_2\beta_2(\underline{L}(\phi))^2
\end{aligned}$$

Thus, if V and W are causal, the assumption (9) yields that $T[\phi](V,W) \geqslant 0$, whereas, if V and W are timelike, the same assumption (10) yields that

$$T[\phi](V,W) \geqslant C_{\alpha_1\alpha_2\beta_1\beta_2}\left((L(\phi))^2 + (\underline{L}(\phi))^2 + \sum_{i=1}^{n-1}(e_i(\phi))^2\right),$$

for some strictly positive constant $C_{\alpha_1\alpha_2\beta_1\beta_2}$ (independent of ϕ but dependent on the constants $\alpha_1, \alpha_2, \beta_1, \beta_2$ which depend on V and W), that completes the proof.

(b) First, since the Levi-Civita connection is compatible with the metric, we get

$$\begin{aligned}
&\nabla_\alpha \nabla^\beta \phi = \nabla_\alpha (g^{\beta\gamma}\nabla_\gamma \phi) = (\nabla_\alpha g^{\beta\gamma})\nabla_\gamma\phi + g^{\beta\gamma}\nabla_\alpha\nabla_\gamma\phi = g^{\beta\gamma}\nabla_\alpha\nabla_\gamma\phi,\\
&\nabla^\beta\nabla_\alpha\phi = g^{\beta\gamma}\nabla_\gamma\nabla_\alpha\phi = g^{\beta\gamma}\nabla_\alpha\nabla_\gamma\phi
\end{aligned}$$

and hence it follows

$$\nabla_\alpha\nabla^\beta\phi = \nabla^\beta\nabla_\alpha\phi. \tag{12}$$

Now, using (12), we compute

$$\begin{aligned}
\nabla_\mu(|d\phi|^2_{g^{-1}}) &= \nabla_\mu(\partial^\sigma\phi\partial_\sigma\phi) = \nabla_\mu(\nabla^\sigma\phi\nabla_\sigma\phi) = (\nabla_\mu\nabla^\sigma\phi)\nabla_\sigma\phi + \nabla^\sigma\phi(\nabla_\mu\nabla_\sigma\phi)\\
&= (\nabla_\mu\nabla^\sigma\phi)\nabla_\sigma\phi + \nabla^\sigma\phi(\nabla_\mu(g_{\sigma\lambda}\nabla^\lambda\phi))\\
&= (\nabla_\mu\nabla^\sigma\phi)\nabla_\sigma\phi + \nabla^\sigma\phi(\nabla_\mu g_{\sigma\lambda})\nabla^\lambda\phi + (g_{\sigma\lambda}\nabla^\sigma\phi)(\nabla_\mu\nabla^\lambda\phi)\\
&= (\nabla_\mu\nabla^\sigma\phi)\nabla_\sigma\phi + (g_{\sigma\lambda}\nabla^\sigma\phi)(\nabla_\mu\nabla^\lambda\phi)\\
&= (\nabla_\mu\nabla^\sigma\phi)\nabla_\sigma\phi + \nabla_\lambda\phi(\nabla_\mu\nabla^\lambda\phi) = 2\nabla_\rho\phi(\nabla_\mu\nabla^\rho\phi).
\end{aligned} \tag{13}$$

Using (13), we get that

$$\begin{aligned}
\nabla^\nu T_{\mu\nu}[\phi] &= \nabla^\nu\left(\partial_\mu\phi\partial_\nu\phi - \frac{1}{2}g_{\mu\nu}|d\phi|^2_{g^{-1}}\right)\\
&= (\nabla^\nu\partial_\mu\phi)\partial_\nu\phi + \partial_\mu\phi(\nabla^\nu\partial_\nu\phi) - \frac{1}{2}(\nabla^\nu g_{\mu\nu})|d\phi|^2_{g^{-1}} - \frac{1}{2}g_{\mu\nu}\nabla^\nu(|d\phi|^2_{g^{-1}})\\
&= (\nabla^\nu\partial_\mu\phi)\partial_\nu\phi + \partial_\mu\phi(\nabla^\nu\partial_\nu\phi) - \frac{1}{2}g_{\mu\nu}\nabla^\nu(|d\phi|^2_{g^{-1}})\\
&= (\nabla^\nu\nabla_\mu\phi)\nabla_\nu\phi + \nabla_\mu\phi(\nabla^\nu\nabla_\nu\phi) - \frac{1}{2}g_{\mu\nu}\nabla^\nu(|d\phi|^2_{g^{-1}})\\
&= (\nabla_\mu\nabla^\nu\phi)\nabla_\nu\phi + \nabla_\mu\phi(\nabla^\nu\nabla_\nu\phi) - \frac{1}{2}g_{\mu\nu}\nabla^\nu(|d\phi|^2_{g^{-1}})
\end{aligned}$$

$$
\begin{aligned}
&= (\nabla_\mu \nabla^\nu \phi)\nabla_\nu \phi + \nabla_\mu \phi(\nabla^\nu \nabla_\nu \phi) - \frac{1}{2}\nabla_\mu(|d\phi|^2_{g^{-1}})\\
&= (\nabla_\mu \nabla^\nu \phi)\nabla_\nu \phi + \nabla_\mu \phi(\nabla^\nu \nabla_\nu \phi) - \nabla_\rho \phi(\nabla_\mu \nabla^\rho \phi)\\
&= \nabla_\mu \phi(\nabla^\nu \nabla_\nu \phi) = (\nabla_\mu \phi)(\Box_g \phi) = 0,
\end{aligned}
$$

that completes the proof.

(c) Let V be a Killing vector field. Then, by definition, we have $\pi[V] = 0$. We define the 1-form

$$J^V_\mu[\phi] := T_{\mu\nu}[\phi]V^\mu$$

We need to compute

$$\nabla^\mu J^V_\mu[\phi] = \nabla^\mu(T_{\mu\nu}[\phi]V^\nu) = (\nabla^\mu T_{\mu\nu}[\phi])V^\nu + T_{\mu\nu}[\phi](\nabla^\mu V^\nu).$$

Observe that, since T is symmetric (meaning $T_{\mu\nu}[\phi] = T_{\nu\mu}[\phi]$), we have

$$
\begin{aligned}
T_{\mu\nu}[\phi]\nabla^\mu V^\nu &= \frac{1}{2}\left(T_{\mu\nu}[\phi]\nabla^\mu V^\nu + T_{\mu\nu}[\phi]\nabla^\mu V^\nu\right) = \frac{1}{2}\left(T_{\mu\nu}[\phi]\nabla^\mu V^\nu + T_{\nu\mu}[\phi]\nabla^\nu V^\mu\right)\\
&= \frac{1}{2}\left(T_{\mu\nu}[\phi]\nabla^\mu V^\nu + T_{\mu\nu}[\phi]\nabla^\nu V^\mu\right) = \frac{1}{2}T_{\mu\nu}[\phi]\left(\nabla^\mu V^\nu + \nabla^\nu V^\mu\right)\\
&= \frac{1}{2}T_{\mu\nu}[\phi]\pi^{\mu\nu}[V] = 0.
\end{aligned}
$$

Consequently, we infer

$$
\begin{aligned}
\nabla^\mu J^V_\mu[\phi] &= V^\nu(\nabla^\mu T_{\mu\nu}[\phi]) + T_{\mu\nu}[\phi]\nabla^\mu V^\nu = V^\nu(\nabla^\mu T_{\nu\mu}[\phi]) + T_{\mu\nu}[\phi]\nabla^\mu V^\nu\\
&= V^\nu \partial_\nu \phi(\Box_g \phi) + T_{\mu\nu}[\phi]\nabla^\mu V^\nu = V(\phi)\Box_g \phi + \frac{1}{2}T_{\mu\nu}[\phi]\pi^{\mu\nu}[V] = 0,
\end{aligned}
$$

that completes the proof.

9.4 Let $f : [0,T] \to [0,+\infty)$ satisfy

$$f(t) \leqslant A(t) + \int_0^t M(s)f(s)\,ds$$

for some non-negative functions A, M on $[0,T]$. Show that

$$f(t) \leqslant A(t) + \int_0^t e^{\int_s^t M(x)\,dx} M(s)A(s)\,ds.$$

In particular, if $A(t) = A$ is constant, show that

$$f(t) \leqslant e^{\int_0^t M(s)\,ds} A.$$

This is known as *Gronwall's inequality*; this inequality will play a crucial role in establishing energy-type estimates for hyperpolic PDEs. (*Hint: You might want to first consider the differential inequality satisfied by $F'(t)$ for $F(t)$ being the right hand side of the inequality we start with.*)

Solution.

Let us set

$$F(t) \doteq \int_0^t M(s)f(s)ds.$$

By differentiating F and using thefact that the given inequality can be reexpressed as

$$f(t) \leqslant A(t) + F(t), \tag{14}$$

we have that:

$$\begin{aligned}
F'(t) = M(t)f(t) \leqslant M(t)A(t) + M(t)F(t) \Longrightarrow \\
e^{-\int_0^t M(s)ds}F'(t) \leqslant e^{-\int_0^t M(s)ds}M(t)A(t) + e^{-\int_0^t M(s)ds}M(t)F(t) \Longrightarrow \\
e^{-\int_0^t M(s)ds}F'(t) - e^{-\int_0^t M(s)ds}M(t)F(t) \leqslant e^{-\int_0^t M(s)ds}M(t)A(t) \Longrightarrow \\
\left(e^{-\int_0^t M(s)ds}F(t)\right)' \leqslant e^{-\int_0^t M(s)ds}A'(t) \Longrightarrow \\
e^{-\int_0^s M(\tau)d\tau}F(s)\big|_{s=0}^{s=t} \leqslant \int_0^t e^{-\int_0^s M(\tau)d\tau}M(s)A(s)ds.
\end{aligned}$$

Noting that $F(0) = 0$, we therefore get:

$$e^{-\int_0^t M(\tau)d\tau}F(t) \leqslant\leqslant \int_0^t e^{-\int_0^s M(\tau)d\tau}M(s)A(s)ds \Longrightarrow$$

$$F(t) \leqslant e^{\int_0^t M(\tau)d\tau}\int_0^t e^{-\int_0^s M(\tau)d\tau}M(s)A(s)ds = \int_0^t e^{\int_s^t M(\tau)d\tau}M(s)A(s)ds$$

Going back to (14), we therefore get

$$f(t) \leqslant A(t) + F(t) = A(t) + \int_0^t e^{\int_s^t M(\tau)d\tau}M(s)A(s)ds,$$

that completes the proof of the first part.

For the second part, assume that $A(t) = A$, for all $0 \leqslant t \leqslant T$. Then, from the previous computation, we have that

$$\begin{aligned}
f(t) &\leqslant A + A\int_0^t e^{\int_s^t M(\tau)d\tau}M(s)ds \\
&\leqslant A - A\int_0^t \frac{d}{ds}\left(e^{\int_s^t M(\tau)d\tau}\right)ds \\
&= A - A\left(e^{\int_s^t M(\tau)d\tau}\Big|_{s=0}^{s=t}\right) \\
&= A - A\left(1 - e^{\int_0^t M(\tau)d\tau}\right) \\
&= A - A + Ae^{\int_0^t M(\tau)d\tau} \\
&= Ae^{\int_0^t M(\tau)d\tau}
\end{aligned}$$

, that completes the proof of the second part.