LECTURE NOTES IN GENERAL RELATIVITY: Spring 2009

Sergiu Klainerman

Department of Mathematics, Princeton University, Princeton NJ 08544

E-mail address: seri@math.princeton.edu

CHAPTER 1

Special Relativity

1. Minkowski Space

The n+1 dimensional Minkowski space, which we denote by $\mathbb{R}^{n+1}$, consists of the manifold $\mathbb{R}^{n+1}$ together with a Lorentz metric $\mathbf{m}$ and a distinguished system of coordinates x^α, $\alpha = 0, 1, \dots n$, called inertial, relative to which the metric has the diagonal form $\mathbf{m}_{\alpha\beta} = \mathrm{diag}(-1, 1, \dots, 1)$. We write, splitting the spacetime coordinates x^α into the time component $x^0 = t$ and space components $x = x^i, \dots x^n$,

$$ds^2 = \mathbf{m}'_{\alpha\beta} dx^\alpha dx^\beta = -dt^2 + (dx^1)^2 + (dx^2)^2 + \dots + (dx^n)^2. \tag{1}$$

We use standard geometric conventions of lowering and raising indices relative to $\mathbf{m}$, and its inverse $\mathbf{m}^{-1} = \mathbf{m}$, as well as the usual summation convention over repeated indices. The coordinate vectorfields $\frac{\partial}{\partial x^\alpha}$ are denoted by ∂_α. The dual 1 forms are dx^α. An arbitrary vectorfield can be expressed as a linear combination of the coordinate vectorfields $X = X^\alpha \partial_\alpha$ with smooth functions $X^\alpha = X^\alpha(x^0, \dots, x^n)$. An arbitrary 1-form is a linear combination of the coordinate 1-forms $A = A_\mu dx^\mu$.

Under a change of coordinates $x^\alpha = x^\alpha(x'^{\,\mu})$ we obtain,

$$ds^2 = \mathbf{m}_{\mu\nu} dx'^{\,\mu} dx'^{\,\nu}, \qquad \mathbf{m}'_{\mu\nu} = \mathbf{m}_{\alpha\beta} \frac{\partial x^\alpha}{\partial x'^{\,\mu}} \frac{\partial x^\alpha}{\partial x'^{\,\mu}}$$

Two inertial systems of coordinates are connected to each other by translations $x^\alpha = x'^{\,\alpha} + x^\alpha_{(0)}$, Lorentz transformations,

$$x'^{\,\alpha} = \Lambda^\alpha_\beta x^\beta, \qquad \mathbf{m}_{\alpha\beta} = \mathbf{m}_{\mu\nu} \Lambda^\mu_\alpha \Lambda^\nu_\beta \tag{2}$$

and combinations of the two $x'^{\,\alpha} = \Lambda^\alpha_\beta x^\beta + x^\alpha_{(0)}$.

Exercise. Show that the Lorentz transformations $B(v) = B_{(0i)}(v) : \mathbb{R}^{1+n} \longrightarrow \mathbb{R}^{1+n}$, with $-1 < v < 1$, (called boosts) which rotate the axes $t = x^0$ and $x = x^i$, $i = 1, 2, \dots n$ and keep all other fixed have the form,

$$t' = \frac{t - vx}{(1 - v^2)^{1/2}}, \quad x' = \frac{x - vt}{(1 - v^2)^{1/2}} \tag{3}$$

Relative to null coordinates $u = t - x$, $v = t + x$ and $u' = t' - x'$, $v' = t' + x'$ we have,

$$u' = \lambda^{-1} u, \quad v' = \lambda v, \qquad \lambda = \frac{(1 - v)^{1/2}}{(1 + v)^{1/2}} \tag{4}$$

Show that $B(v)$, $|v| < 1$ forms a one parameter group of diffeomorphisms of the Minkowski space and find the *relativistic law of addition of velocities.*

$\mathbf{D}$ the flat covariant derivative of $\mathbb{R}^{n+1}$. Recall that covariant differentiation associates to any two smooth vectorfields X, Y another vectorfield $\mathbf{D}_X Y$ which verifies the following rules.

(1) Given three vectorfields X, Y, Z and scalar functions a, b we have

$$\begin{aligned} \mathbf{D}_{aX+bY} Z &= a\mathbf{D}_X Z + b\mathbf{D}_Y Z \\ \mathbf{D}_X(fY) &= X(f)Y + f\mathbf{D}_X Y \end{aligned}$$

(2) For any vectorfields X, Y, Z,

$$X\mathbf{m}(Y, Z) = \mathbf{m}(\mathbf{D}_X Y, Z) + \mathbf{m}(X, \mathbf{D}_Y Z)$$

Given an arbitrary 1-form $A = A_\alpha dx^\alpha$ we have $\mathbf{D}_\alpha \omega_\beta = \partial_\alpha \omega_\beta$.

A vector X is said to be timelike, null or spacelike according to whether $\mathbf{m}(X, X)$ is < 0, $= 0$ or > 0. Accordingly a smooth curve $x^\alpha(s)$ is said to be timelike, null or spacelike if its tangent vector $\frac{dx^\alpha}{ds}$ is timelike, null or spacelike at every one of its points. A causal curve can be either timelike and null at any of its points. The canonical time orientation of $\mathbb{R}^{n+1}$ is given by the vectorfield $T_0 = \partial_0$. A timelike vector X is said to be *future oriented* if $\mathbf{m}(X, T_0) < 0$ and *past oriented* if $\mathbf{m}(X, T_0) > 0$.

Similarly a hypersurface $u(x^0, \ldots x^n) = 0$ is said to be spacelike, null or timelike if its normal $N^\alpha = -\mathbf{m}^{\alpha\beta}\partial_\beta u$ is, respectively, timelike, null or spacelike. The metric induced by $\mathbf{m}$ on a spacelike hypersurface is necessarily positive definite, that is Riemannian. A function $\mathbf{t}(x^0, x^1, \ldots, x^n)$ is said to be a time function if its level hypersurfaces $\mathbf{t} = t$ are spacelike. On a null hypersurface the induced metric is degenerate relative to the normal direction, i.e. $\mathbf{m}(N, N) = 0$. In particular function $\mathbf{u} = \mathbf{u}(x^0, \ldots x^n)$ whose level surfaces $\mathbf{u} = u$ are null must verify the Eikonal equation

$$\mathbf{m}^{\alpha\beta}\partial_\alpha \mathbf{u} \partial_\beta \mathbf{u} = 0 \tag{5}$$

Equation (5) can also be written in the form $\mathbf{D}_N N = 0$. We call N a *null geodesic generator* of the level hypersurfaces of $\mathbf{u}$.

Definition Smooth solutions $u = u(x^0, \ldots x^n)$ of the eikonal equation (5) are called optical functions. Their level hypersurfaces are null.

Exercise. Show that the functions $t \pm r$, with $t = x^0$ and $r = \sqrt{(x^1)^2 + \cdots (x^n)^2}$ are optical functionsin $\mathbb{R}^{n+1} \setminus \{0\}$.

1.1. Physical Interpretation. Given a timelike curve $C : x^\alpha = x^\alpha(s)$ we define *proper time* along the curve,

$$\tau = \int (-\mathbf{m}_{\alpha\beta} \frac{dx^\alpha}{ds} \frac{dx^\beta}{ds})^{1/2} ds \tag{6}$$

According to special relativity τ is the time which would elapse on a clock carried along the curve. The tangent vector $U = u^\alpha \partial_\alpha$ to a timelike curve parametrized by its proper time τ is called the *four velocity* of the curve. Clearly,

$$u_\alpha u^\alpha = \mathbf{m}(U,U) = -1$$

Particles subject to no external forces travel along geodesics, i.e. its four velocity verifies

$$\mathbf{D}_U U = 0. \tag{7}$$

In coordinates $u^\alpha \partial_\alpha u^\beta = 0$. Clearly, all geodesics in Minkowski space are straight lines. All material particles have a *rest mass* m which appears as a parameter in the equations of motion in the presence of forces. The energy momentum 4-vector of a particle of mass m is defined to be,

$$P = mU \tag{8}$$

or in coordinates $p^\alpha = mu^\alpha$. By definition the energy E of the particle is defined to be p^0. In particular, in the rest fram of the particle we have,

$$E = m \tag{9}$$

or $E = mc^2$ if the original Minkowski metric is in fact $\mathbf{m}_{\alpha\beta} = \mathrm{diag}(-c^2, 1, \dots, 1)$.

In the rest frame of a particle its four-velocity has components $U^\mu = (1,0,0,0)$. In a frame ,with respect to which the particle is moving with velocity v along the x^1 axis, we find (by performing a Lorentz transformation to a frame relative to which the particle is at rest)

$$p^\mu(\gamma m, v\gamma m, 0, 0), \qquad \gamma = 1/\sqrt{1-v^2}.$$

For small v this gives, $p^0 = m + \frac{1}{2}mv^2$ (rest energy plus newtonian kinetic energy) and $p^1 = mv$ (the newtonian momentum of the particle).

More generally, given an observer, present at the site of the particle whose 4- velocity is $V = v^\alpha \partial_a$, we can decompose,

$$P = EV + P^\perp, \qquad E = -\mathbf{m}(P,V) = -mu_\alpha v^\alpha \tag{10}$$

where $P^\perp$ denotes the component of P perpendicular to V and E denotes the energy of the particle as measured by the observer. For a particle at rest with respect to the observer, i.e. $U = V$, we have again Einstein's famous formula $E = m$.

We define (as four-force) the four- vector,

$$f^\mu = \frac{d}{d\tau}p^\mu(\tau) = m\frac{d^2}{d\tau^2}x^\mu(\tau)$$

Gravity, which is the simplest example of force in newtonian mechanics, is manifested in relativity by the curvature of the spacetime itself. Another force which is important in electromagnetism is the Lorentz force. If $F_{\mu\nu}$ is a given electromagnetic force one defines the four-Lorentz force acting on a particle with four velocity U and charge q,

$$f^\mu = qU^\lambda F_\lambda{}^\mu \tag{11}$$

The equations of motion for the particle are,

$$m\frac{d^2}{d\tau^2}x^\mu(\tau) = f^\mu(\tau). \tag{12}$$

The causal future $J^+(S)$ of a set S consists of all points in $\mathbb{R}^{n+1}$ which can be connected to S by a future directed causal curve. The causal past $\mathcal{J}^-(S)$ is defined in the same way. Thus, for a point $p = (t, x)$, $\mathcal{J}^+(p) = \{(t \geq t_0, x)/|x - x_0| \leq t - t_0\}$.

1.2. Symmetries. Let x^μ be an inertial coordinate system of Minkowski space $\mathbb{R}^{n+1}$. The following are all the isometries and conformal isometries (see definition (3.3) in the Appendix) of $\mathbb{R}^{n+1}$.

1. Translations: for any given vector $a = (a^0, a^1, ..., a^n) \in \mathbb{R}^{n+1}$,

$$x^\mu \to x^\mu + a^\mu$$

2. Lorentz rotations: Given any $\Lambda = \Lambda^\rho_\sigma \in \mathbf{O}(1, n)$,

$$x^\mu \to \Lambda^\mu_\nu\, x^\nu$$

3. Scalings: Given any real number $\lambda \neq 0$,

$$x^\mu \to \lambda x^\mu$$

4. Inversion: Consider the transformation $x^\mu \to I(x^\mu)$, where

$$I(x^\mu) = \frac{x^\mu}{(x, x)}$$

defined for all points $x \in \mathbb{R}^{n+1}$ such that $(x, x) \neq 0$.

The first two sets of transformations are isometries of $\mathbb{R}^{n+1}$, the group generated by them is called the Poincarè group. The last two type of transformations are conformal isometries. the group generated by all the above transformations is the conformal group. In fact the Liouville theorem, whose infinitesimal version will be proved later on, states that it is the group of all the conformal isometries of $\mathbb{R}^{n+1}$.

Remark. The transformations mentioned above generate 1 parameter groups of transformations. Thus, corresponding to time translations, we associate the additive transformation group $U_{(\alpha)}(s)$ which fixes all coordinates x^β, $\beta \neq \alpha$ and takes x^α to $x^\alpha + s$. Corresponding to a Lorentz transformation which rotates the $t = x^0, x = x^1$ axes we associate the $1-$ parameter additive group $B_{(0i)}(v)$ given by (3). The scaling transformation generate a multiplicative. Finally, using the inversion I we can generate the $1-$ parameter additive groups $C_\alpha = I^{-1} \circ U_\alpha \circ I$.

Any additive $1-$ parameter group of transformations $U(s)$ generates a vectorfield X according to the formula,

$$Xf = \frac{d}{ds}(f \circ U(s))|_{s=0}.$$

Clearly the generators of the groups of translations and lorentz transformations are Killing while those generated by scalings and inverted translations are conformal Killing (see definition 3.4).

We list below the Killing and conformal Killing vector fields which generate the above transformations.

i. The generators of translations in the x^μ directions, $\mu = 0, 1, ..., n$:

$$\mathbf{T}_\mu = \frac{\partial}{\partial x^\mu}$$

ii. The generators of the Lorentz rotations in the (μ, ν) plane:

$$\mathbf{L}_{\mu\nu} = x_\mu \partial_\nu - x_\nu \partial_\mu$$

iii. The generators of the scaling transformations:

$$\mathbf{S} = x^\mu \partial_\mu$$

iv. The generators of the inverted translations [1]:

$$\mathbf{K}_\mu = 2x_\mu x^\rho \frac{\partial}{\partial x^\rho} - (x^\rho x_\rho)\frac{\partial}{\partial x^\mu}$$

Of particular importance is the vectorfield $\mathbf{K}_0 = (t^2+r^2)\partial_t + 2tx^i\partial_i$, which is causal. Here $r^2 = |x|^2 = (x^1)^2 + \cdots (x^n)^2$.

Remark. Observe that the vectorfields $\mathbf{T}, \mathbf{L}$ are all Killing while $\mathbf{S}, \mathbf{K}$ are conformal Killing. Recall that a vectorfield X is a Killing vectorfield for a metric $\mathbf{g}$ if $\mathcal{L}_X \mathbf{g} = 0$ or , equivalently if its deformation tensor ${}^{(X)}\pi_{\alpha\beta} := \mathbf{D}_a X_\beta + \mathbf{D}_\beta X_\alpha$ vanishes identically. The vectorfield X is called conformal Killing if ${}^{(X)}\pi$ is proportional to the metric $\mathbf{g}$.

We also list below the commutator relations between these vector fields,

$$\begin{aligned}
&[\mathbf{L}_{\alpha\beta}, \mathbf{L}_{\gamma\delta}] = \mathbf{m}_{\alpha\gamma}\mathbf{L}_{\beta\delta} - \mathbf{m}_{\beta\gamma}\mathbf{L}_{\alpha\delta} + \mathbf{m}_{\beta\delta}\mathbf{L}_{\alpha\gamma} - \mathbf{m}_{\alpha\delta}\mathbf{L}_{\beta\gamma}\\
&[\mathbf{L}_{\alpha\beta}, \mathbf{T}_\gamma] = \mathbf{m}_{\alpha\gamma}\mathbf{T}_\beta - \mathbf{m}_{\beta\gamma}\mathbf{T}_\alpha\\
&[\mathbf{T}_\alpha, \mathbf{T}_\beta] = 0\\
&[\mathbf{T}_\alpha, \mathbf{S}] = \mathbf{T}_\alpha\\
&[\mathbf{T}_\alpha, \mathbf{K}_\beta] = 2(\mathbf{m}_{\alpha\beta}\mathbf{S} + \mathbf{L}_{\alpha\beta})\\
&[\mathbf{L}_{\alpha\beta}, \mathbf{S}] = [\mathbf{K}_\alpha, \mathbf{K}_\beta] = 0\\
&[\mathbf{L}_{\alpha\beta}, \mathbf{K}_\gamma] = \mathbf{m}_{\alpha\gamma}\mathbf{K}_\beta - \mathbf{m}_{\beta\gamma}\mathbf{K}_\alpha
\end{aligned} \tag{13}$$

Denoting $\mathcal{P}(1,n)$ the Lie algebra generated by the vector fields $\mathbf{T}_\alpha, \mathbf{L}_{\beta\gamma}$ and $\underline{\mathcal{K}}(1,n)$ the Lie algebra generated by all the vector fields $\mathbf{T}_\alpha, \mathbf{L}_{\beta\gamma}, \mathbf{S}, \mathbf{K}_\delta$ we state the following version of the Liouville theorem,

THEOREM 1.3. *The following statements hold true.*

1) $\mathcal{P}(1,n)$ is the Lie algebra of all Killing vector fields in $\mathbb{R}^{n+1}$.

2) If $n > 1$, $\underline{\mathcal{K}}(1,n)$ is the Lie algebra of all conformal Killing vector fields in $\mathbb{R}^{n+1}$.

[1] Observe that the vector fields $\mathbf{K}_\mu$ can be obtained applying I_* to the vector fields $\mathbf{T}_\mu$.

3) If $n=1$, the set of all conformal Killing vector fields in $\mathbb{R}^{1+1}$ is given by the following expression

$$f(x^0+x^1)(\partial_0+\partial_1)+g(x^0-x^1)(\partial_0-\partial_1)$$

where f, g are arbitrary smooth functions of one variable.

Proof: The proof for part 1 of the theorem follows immediately, as a particular case, from Proposition (3.7). From (111) as $\mathbf{R}=0$ and X is Killing we have

$$D_\mu D_\nu X_\lambda = 0 \ .$$

Therefore, there exist constants $a_{\mu\nu}, b_\mu$ such that $X^\mu = a_{\mu\nu}x^\nu + b_\mu$. Since X is Killing $D_\mu X_\nu = -D_\nu X_\mu$ which implies $a_{\mu\nu} = -a_{\nu\mu}$. Consequently X can be written as a linear combination, with real coefficients, of the vector fields $T_\alpha, L_{\beta\gamma}$.

Let now X be a conformal Killing vector field. There exists a function Ω such that

$$^{(X)}\pi_{\rho\sigma} = \Omega \mathbf{m}_{\rho\sigma} \tag{14}$$

Using formulas (111) and (112), in the appendix, it follows that

$$D_\mu D_\nu X_\lambda = \frac{1}{2}\left(\Omega_{,\mu}\mathbf{m}_{\nu\lambda} + \Omega_{,\nu}\mathbf{m}_{\mu\lambda} - \Omega_{,\lambda}\mathbf{m}_{\nu\mu}\right) \tag{15}$$

Taking the trace with respect to μ, ν, on both sides of (15) we infer that

$$\begin{aligned} \Box X_\lambda &= -\frac{n-1}{2}\Omega_\lambda \\ D^\mu X_\mu &= \frac{n+1}{2}\Omega \end{aligned} \tag{16}$$

and applying D^λ to the first equation, $\Box$ to the second one and subtracting we obtain

$$\Box\Omega = 0 \tag{17}$$

Applying D_μ to the first equation of (16) and using (17) we obtain

$$\begin{aligned} (n-1)D_\mu D_\lambda \Omega &= \frac{n-1}{2}(D_\mu D_\lambda\Omega + D_\lambda D_\mu \Omega) = -\Box(D_\mu X_\lambda + D_\lambda X_\mu) \\ &= -(\Box\Omega)\mathbf{m}_{\mu\lambda} = 0 \end{aligned} \tag{18}$$

Hence for $n \neq 1$, $D_\mu D_\lambda \Omega = 0$. This implies that Ω must be a linear function of x^μ. We can therefore find a linear combination, with constant coefficients, $cS + d^\alpha K_\alpha$ such that the deformation tensor of $X - (cS + d^\alpha K_\alpha)$ must be zero. This is the case because $^{(S)}\pi = 2\mathbf{m}$ and $^{(K_\mu)}\pi = 4x_\mu \mathbf{m}$. Therefore $X - (cS + d^\alpha K_\alpha)$ is Killing which, in view of the first part of the theorem, proves the result.

Part 3 can be easily derived by solving (14). Indeed posing $X = a\partial_0 + b\partial_1$, we obtain $2D_0X_0 = -\Omega$, $2D_1X_1 = \Omega$ and $D_0X_1 + D_1X_0 = 0$. Hence a, b verify the system

$$\frac{\partial a}{\partial x^0} = \frac{\partial b}{\partial x^1} \ , \ \frac{\partial b}{\partial x^0} = \frac{\partial a}{\partial x^1} \ .$$

Hence the one form $adx^0 + bdx^1$ is exact, $adx^0 + bdx^1 = d\phi$, and $\frac{\partial^2 a}{\partial x^{0\,2}} = \frac{\partial^2 b}{\partial x^{1\,2}}$, that is $\Box\phi = 0$. In conclusion

$$X = \frac{1}{2}\left(\frac{\partial\phi}{\partial x^0} + \frac{\partial\phi}{\partial x^1}\right)(\partial_0 + \partial_1) + \frac{1}{2}\left(\frac{\partial\phi}{\partial x^0} - \frac{\partial\phi}{\partial x^1}\right)(\partial_0 - \partial_1)$$

which proves the result.

1.4. Null frames. A pair of null vectorfields $L, \underline{L}$ form a *null pair* if $\mathbf{m}(L, \underline{L}) = -2$. A null pair $e_n = L, e_{n+1} = \underline{L}$ together with vectorfields $e_1, \ldots e_{n-1}$ such that $\mathbf{m}(L, e_a) = \mathbf{m}(\underline{L}, e_a) = 0$ and $\mathbf{m}(e_a, e_b) = \delta_{ab}$, for all $a, b = 1, \ldots, n-1$, is called a *null frame*. The null pair,

$$L = \partial_t + \partial_r, \quad \underline{L} = \partial_t - \partial_r, \tag{19}$$

with $r = |x|$ and $\partial_r = x^i/r\partial_i$, is called *canonical*. Simmilarly a null frame $e_1, \ldots e_{n+1}$ with $e_n = L, e_{n+1} = \underline{L}$ is called a canonical null frame. In that case $e_1, \ldots, e_{n-1}$ form, at any point, an orthonormal basis for the the sphere $S_{t,r}$, of constant t and r, passing through that point. Observe also that L is the null geodesic generator associated to $\mathbf{u} = t - r$ while $\underline{L}$ the null geodesic of $\underline{u} = t + r$.

Remark. Expresse relative to the canonical null pair,

$$\mathbf{T}_0 = 2^{-1}(L + \underline{L}), \quad \mathbf{S} = 2^{-1}(\underline{u}\, L + \mathbf{u}\,\underline{L}), \quad \mathbf{K}_0 = 2^{-1}(\underline{u}^2\, L + \mathbf{u}^2\,\underline{L}). \tag{20}$$

Both $\mathbf{T}_0 = \partial_t$ and $\mathbf{K}_0 = (t^2 + |x|^2)\partial_t + 2tx^i\partial_i$ are causal. This makes them important in deriving energy estimates. Observe that $\mathbf{S}$ is causal only in $\mathcal{J}^+(0) \cup \mathcal{J}^-(0)$.

1.5. Conformal Compactifcation. In polar coordinates $x^1 = r\cos\theta^1, x^2 = rsin\theta^1\cos\theta^2, \ldots, x^n = rsin\theta^1\sin\theta^2\cdots\sin\theta^{n-1}$, the Minkowski metric takes the form,

$$-dt^2 + dr^2 + r^2 d\omega_{n-1}^2, \qquad r > 0$$

where,

$$d\omega_{n-1}^2 = (d\theta^1)^2 + \sin^2\theta^1(d\theta^2)^2 + \cdots + \sin^2\theta^1\cdots\sin^2\theta^{n-2}(d\theta^{n-1})^2$$

is the metric of the standard $n-1$ dimensional sphere $\mathbb{S}^{n-1}$. We introduced the advanced and retarded coordinates $u = t - r$, $\underline{u} = t + r$ and rewrite $\mathbf{m}$ in the form,

$$-dud\underline{u} + \frac{1}{4}(\underline{u} - u)^2 d\omega_{n-1}^2, \qquad -\infty < u < \underline{u} < \infty$$

We now make the change of variables,

$$u = \tan U, \qquad \underline{u} = \tan \underline{U}, \qquad -\frac{\pi}{2} < \underline{U} < U < \frac{\pi}{2}$$

and rewrite the Minkowski metric in the form,

$$\frac{1}{cos^2 U\cos^2\underline{U}}\big(-dUd\underline{U} + \frac{1}{4}\sin^2(\underline{U} - U)d\omega_{n-1}^2\big)$$

or, introducing the new metric $\widetilde{m}$,

$$-4dUd\underline{U} + \sin^2(\underline{U} - U)d\omega_{n-1}^2 \tag{21}$$

we derive,

$$\widetilde{\mathbf{m}} = \Omega^2 \mathbf{m},$$

with,

$$\Omega = \cos U \cos \underline{U} = \frac{1}{(1+u^2)^{1/2}(1+\underline{u}^2)^{1/2}} \tag{22}$$

Finally introducing the new variables,

$$T = U + \underline{U}, \qquad R = \underline{U} - U$$

the new metric $\widetilde{m}$ takes the form,

$$-dT^2 + dR^2 + \sin^2 R d\omega_{n-1}^2$$

Observe that $dR^2 + \sin^2 R d\omega_{n-1}^2$ is precisely the metric $d\omega_n^2$ of the standard sphere $\mathbb{S}^n$. Thus the metric $\widetilde{m}$ is precisely the standard Lorentz metric,

$$-dT^2 + dw_n^2 \tag{23}$$

of the cylinder $\mathbb{E}^{n+1} = \mathbb{R} \times \mathbb{S}^n$. The space-time thus obtained is called the Einstein cylinder. Consider the map $P : \mathbb{R}^{1+n} \to \mathbb{E}^{1+n}$ defined by,

$$(t, r, \omega) \to (T, R, \omega), \quad \omega \in \mathbb{S}^{n-1} \tag{24}$$

where,

$$\begin{aligned} T &= tan^{-1}(t+r) + \tan^{-1}(t-r) \\ R &= tan^{-1}(t+r) - \tan^{-1}(t-r) \end{aligned}$$

or,

$$U = \tan^{-1} u, \quad \underline{U} = \tan^{-1} \underline{u}$$

with $\tan^{-1} : \mathbb{R} \to (-\pi/2, \pi/2)$. We have established the following:

Proposition 1.6. *The map P establishes a conformal isometry between the minkowski space $\mathbb{R}^{n+1}$ with metric $\mathbf{m}$ and the Einstein cylinder $\mathbb{E}^{n+1}$ with metric[2] $\widetilde{m}$,*

$$P^*(\widetilde{m}) = \Omega^2 m. \tag{25}$$

with conformal factor Ω given by formula (22). The image $P(\mathbb{R}^{n+1})$ is the bounded region of $\mathbb{E}^{n+1}$ characterized by the conditions,

$$-\pi < T \pm R < \pi, \quad 0 \le R < \pi$$

Definition. The boundary of $P(\mathbb{R}^{n+1})$ in $\mathbb{E}^{n+1}$ is given by,

$$\partial P(\mathbb{R}^{n+1}) = \mathcal{S}^+ \cup \mathcal{S}^- \cup i^0 \cup i^+ \cup i^-$$

Here

$$\begin{aligned} \mathcal{S}^+ &= \{T + R = \pi, 0 < R < \pi\} = \{\underline{U} = \frac{\pi}{2}, -\frac{\pi}{2} < U < \frac{\pi}{2}\} \\ \mathcal{S}^- &= \{T - R = -\pi, 0 < R < \pi\} = \{U = -\frac{\pi}{2}, -\frac{\pi}{2} < \underline{U} < \frac{\pi}{2}\} \end{aligned}$$

are called the future and past null infinities of Minkowski space. The point

$$i^0 = \{T = 0, R = \pi\} = \{U = -\frac{\pi}{2}, \underline{U} = \frac{\pi}{2}\}$$

[2] Recall that, given $\Phi : \mathbf{M} \to \mathbf{M}'$ with $\mathbf{T}$ a covariant 2-tensor on $\mathbf{M}'$ one defines the pull back tensor $\Phi^* T$ on $\mathbf{M}$ by $\Phi^* T(X, Y) = T(\Phi_* X, \Phi_* Y)$.

is called the space-like infinity while the points,

$$
\begin{aligned}
i^+ &= = \{T = \pi, R = 0\} = \{U = \underline{U} = \frac{\pi}{2}\} \\
i^- &= \{T = -\pi, R = 0\} = \{U = \underline{U} = -\frac{\pi}{2}\}
\end{aligned}
$$

are called the time-like future and past infinities of the image $P(\mathbb{R}^{n+1}$ of Minkowski space in the Einstein cylinder. Note that all time-like geodesics of Minkowski space begin at i^- and end at i^+, all space-like geodesics begin and end at i^0 and all null geodesics start on $\mathcal{S}^-$ and end on $\mathcal{S}^+$.

Remark 1. Observe that a conformal isometry maps null hypersurfaces int null hypersurfaces. Thus, since $u = t - r$ and $\underline{u} = t = r$ are optical functions for $\mathbb{R}^{n+1}$ there is no surprise that $U = \frac{T-R}{2}$, $\underline{U} = \frac{T+R}{2}$ are null for $\mathbb{E}^{n+1}$. In particular we see that the future and past null infinity boundaries $\mathcal{S}^\pm$ are indeed null.

Remark 2. Observe that $\Omega > 0$ and vanishes at the boundary $\partial P(\mathbb{R}^{n+1}$ in $\mathbb{E}^{n+1}$. Also $(\partial_U \Omega, \partial_V \Omega) \neq 0$ on $\mathcal{S}^+ \cup \mathcal{S}^-$. In other words the differential $d\Omega$ of Ω, in $\mathbb{E}^{1+n}$ is non-vanishing along the null boundaries of $P(\mathbb{R}^{n+1})$. On the other hand $d\Omega$ vanishes at i^0. One can show only that the hessian $\tilde{\mathbf{D}}^2\Omega$ is non-degenerate, where $\tilde{\mathbf{D}}$ denotes the covariant derivative operator on $\mathbb{E}^{n+1}$. In fact,

$$
\tilde{\mathbf{D}}_\alpha \tilde{\mathbf{D}}_\beta \Omega = 2\widetilde{\mathbf{m}}_{\alpha\beta}
$$

Using these facts on can prove that the null boundary $\partial P(\mathbb{R}^{n+1}$ is of class C^2 at i^0 and C^∞, in fact real analytic, everywhere else.

Remark 3. Observe that the vectorfield ∂_T in $\mathbb{E}^{n+1}$ is a Killing vectorfield for the metric $\widetilde{m}$. It is in fact the image through P of the vectorfield $\mathbf{K}_0$, i.e. $P_*(\mathbf{K}_0) = \partial_T$.

Exercise. Verify all statements made in Remarks 2, 3.

2. Classical Field Theory

2.1. Basic Notions. In this section we will discuss some basic examples of nonlinear wave equations which arise variationally from a relativistic Lagrangian. The fundamental objects of a relativistic field theory are

- **Space-time** $(\mathbf{M}, \mathbf{g})$ which consists of an $n+1$ dimensional manifold $\mathbf{M}$ and a Lorentz metric $\mathbf{g}$; i.e . a nondegenerate quadratic form with signature $(-1, 1, \ldots, 1)$ defined on the tangent space at each point of $\mathbf{M}$. We denote the coordinates of a point in $\mathbf{M}$ by x^α, $\alpha = 0, 1, \ldots, n$.

  Throughout most of this chapter the space-time will in fact be the simplest possible example - namely, the Minkowski space-time in which the manifold is $\mathbb{R}^{n+1}$ and the metric is given by

$$
ds^2 = \mathbf{m}_{\alpha\beta} dx^\alpha dx^\beta = -dt^2 + \left(dx^1\right)^2 + \cdots + (dx^n)^2 \tag{26}
$$

with $t = x^0, m_{\alpha\beta} = \mathrm{diag}(-1, 1, \ldots, 1)$. Recall that any system of coordinates for which the metric has the form (26) is called inertial. Any two inertial coordinate systems are related by Lorentz transformations.

- **Collection of fields** $\psi = \psi^{(1)}, \psi^{(2)}, \ldots, \psi^{(p)}$ which can be scalars, tensors, or some other geometric objects[3] such as spinors, defined on $\mathbf{M}$.
- **Lagrangian density** L which is a scalar function on $\mathbf{M}$ depending only on the tensorfields ψ and the metric[4] $\mathbf{g}$.

We then define the corresponding *action* $\mathcal{S}$ to be the integral,

$$\mathcal{S} = \mathcal{S}[\psi, \mathbf{g} : \mathcal{U}] = \int_{\mathcal{U}} L[\psi] dv_{\mathbf{g}}$$

where $\mathcal{U}$ is any relatively compact set of $\mathbf{M}$. Here $dv_{\mathbf{g}}$ denotes the volume element generated by the metric $\mathbf{g}$. More precisely, relative to a local system of coordinates x^α, we have

$$dv_{\mathbf{g}} = \sqrt{-\mathbf{g}} dx^0 dx^1 \cdots dx^n = \sqrt{-\mathbf{g}} dx$$

with g the determinant of the matrix $(\mathbf{g}_{\alpha\beta})$.

By a compact variation of a field ψ we mean a smooth one-parameter family of fields $\psi_{(s)}$ defined for $s \in (-\epsilon, \epsilon)$ such that,

(1) At $s = 0$, $\quad \psi_{(0)} = \psi$.
(2) At all points $p \in \mathbf{M} \setminus \mathcal{U}$ we have $\psi_{(s)} = \psi$.

Given such a variation we denote $\delta\psi := \dot{\psi} := \frac{d\psi_{(s)}}{ds}\Big|_{s=0}$. Thus, for small s,

$$\psi_{(s)} = \psi + s\dot{\psi} + O(s^2)$$

A field ψ is said to be *stationary* with respect to $\mathcal{S}$ if, for any compact variation $(\psi_{(s)}, \mathcal{U})$ of ψ, we have

$$\frac{d}{ds}\mathbf{S}(s)\Big|_{s=0} = 0$$

where,

$$\mathbf{S}(s) = \mathbf{S}[\psi_{(s)}, \mathbf{g}; \mathcal{U}]$$

We write this in short hand notation as

$$\frac{\delta \mathbf{S}}{\delta \psi} = 0$$

Action Principle, also called the Variational Principle, states that an acceptable solution of a physical system must be stationary with respect to a given Lagrangian density called the Lagrangian of the system. The action principle allows us to derive partial differential equations for the fields ψ called the Euler-Lagrange equations. Here are some simple examples:

1. Scalar Field Equations.

---

[3]For simplicity we restrict ourselves to covariant tensors.
[4]as well as its inverse $\mathbf{g}^{-1}$

One starts with the Lagrangian density $L[\phi] = -\frac{1}{2}\mathbf{g}^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - V(\phi)$ where ϕ is a complex scalar function defined on $(\mathbf{M}, \mathbf{g})$ and $V(\phi)$ a given real function of ϕ.

Given a compact variation $(\phi_{(s)}, \mathcal{U})$ of ϕ, we set $\mathcal{S}(s) = \mathcal{S}[\phi_{(s)}, \mathbf{g}; \mathcal{U}]$. Integration by parts gives,

$$\begin{aligned}\frac{d}{ds}\mathcal{S}(s)\Big|_{s=0} &= \int_{\mathcal{U}} [-\mathbf{g}^{\mu\nu}\partial_\mu\dot{\phi}\partial_\nu\phi - V'(\phi)\dot{\phi}]\sqrt{-\mathbf{g}}dx \\ &= \int_{\mathcal{U}} \dot{\phi}[\Box_{\mathbf{g}}\phi - V'(\phi)]dv_{\mathbf{g}}]\end{aligned}$$

where $\Box_{\mathbf{g}}$ is the D'Alembertian,

$$\Box_{\mathbf{g}}\phi = \frac{1}{\sqrt{-\mathbf{g}}}\partial_\mu\Big(\mathbf{g}^{\mu\nu}\sqrt{-\mathbf{g}}\,\partial_\nu\phi\Big).$$

In view of the action principle and the arbitrariness of $\dot{\phi}$ we infer that ϕ must satisfy the following Euler-Lagrange equation

$$\Box_{\mathbf{g}}\phi - V'(\phi) = 0, \tag{27}$$

Equation (27) is called the *scalar wave equation* with potential $V(\phi)$.

2. Wave Maps.

The wave map equations will be defined in the context of a space-time $(\mathbf{M}, \mathbf{g})$, a Riemannian manifold N with metric h, and a mapping

$$\phi : \mathbf{M} \longrightarrow N.$$

We recall that if X is a vectorfield on $\mathbf{M}$ then $\phi_* X$ is the vectorfield on N defined by $\phi_* X(f) = X(f \circ \phi)$. If ω is a 1-form on N its pull-back $\phi^*\omega$ is the 1-form on $\mathbf{M}$ defined by $\phi^*\omega(X) = \omega(\phi_* X)$, where X is an arbitrary vectorfield on $\mathbf{M}$. Similarly the pull-back of the metric h is the symmetric 2-covariant tensor on $\mathbf{M}$ defined by the formula $(\phi^* h)(X, Y) = h(\phi_* X, \phi_* Y)$. In local coordinates x^α on $\mathbf{M}$ and y^a on N, if ϕ^a denotes the components of ϕ relative to y^a, we have,

$$(\phi^* h)_{\alpha\beta}(p) = \frac{\partial\phi^a}{\partial x^\alpha}\frac{\partial\phi^b}{\partial x^\beta}h_{ab}(\phi(p)) = \langle\frac{\partial\phi}{\partial x^\alpha}, \frac{\partial\phi}{\partial x^\beta}\rangle$$

where $< \cdot, \cdot >$ denotes the Riemannian scalar product on N.

Consider the following Lagrangian density involving the map ϕ,

$$L = -\frac{1}{2}\mathrm{Tr}_{\mathbf{g}}(\phi^* h)$$

where $\mathrm{Tr}_{\mathbf{g}}(\phi^* h)$ denotes the trace relative to $\mathbf{g}$ of $\phi^* h$. In local coordinates,

$$L[\phi] = -\frac{1}{2}\mathbf{g}^{\mu\nu}h_{ab}(\phi)\frac{\partial\phi^a}{\partial x^\mu}\frac{\partial\phi^b}{\partial x^\nu}.$$

By definition *wave maps* are the stationary points of the corresponding action. Thus by a a straightforward calculation,

$$\begin{aligned} 0 &= \frac{d}{ds}\mathcal{S}(s)\Big|_{s=0} = I_1 + I_2 && (28)\\ I_1 &= -\frac{1}{2}\int_{\mathcal{U}} \mathbf{g}^{\mu\nu}\frac{\partial h_{ab}(\phi)}{\partial\phi^c}\dot{\phi}^c\,\partial_\mu\phi^a\partial_\nu\phi^b\sqrt{-\mathbf{g}}dx\\ I_2 &= -\int_{\mathcal{U}} \mathbf{g}^{\mu\nu}h_{ab}(\phi)\partial_\mu\dot{\phi}^a\partial_\nu\phi^b\sqrt{-\mathbf{g}}dx \end{aligned}$$

After integrating by parts, relabelling and using the symmetry in b, c, we can rewrite I_2 in the form,

$$\begin{aligned} I_2 &= \int_{\mathcal{U}}\dot{\phi}^a\left(h_{ab}(\phi)\Box_{\mathbf{g}}\phi^b + \mathbf{g}^{\mu\nu}\frac{\partial h_{ab}}{\partial\phi^c}\partial_\mu\phi^c\partial_\nu\phi^b\right)dv_{\mathbf{g}} && (29)\\ &= \int_{\mathcal{U}}\dot{\phi}^a\left(h_{ab}(\phi)\Box_{\mathbf{g}}\phi^b + \frac{1}{2}\mathbf{g}^{\mu\nu}\left(\frac{\partial h_{ab}}{\partial\phi^c} + \frac{\partial h_{ac}}{\partial\phi^b}\right)\partial_\mu\phi^b\partial_\nu\phi^c\right)dv_{\mathbf{g}} \end{aligned}$$

Also, relabelling indices

$$I_1 = -\frac{1}{2}\int_{\mathcal{U}}\mathbf{g}^{\mu\nu}\frac{\partial h_{bc}}{\partial\phi^a}\dot{\phi}^a\,\partial_\mu\phi^b\partial_\nu\phi^c\,dv_{\mathbf{g}}.$$

Therefore,

$$\begin{aligned} 0 &= I_1 + I_2\\ &= \int_{\mathcal{U}}\dot{\phi}^a\left(h_{ab}\Box_{\mathbf{g}}\phi^b + \partial_\mu\phi^b\partial_\nu\phi^c\mathbf{g}^{\mu\nu}\frac{1}{2}\left(\frac{\partial h_{ab}}{\partial\phi^c} + \frac{\partial h_{ac}}{\partial\phi^b} - \frac{\partial h_{bc}}{\partial\phi^a}\right)\right)dv_{\mathbf{g}}\\ &= \int_{\mathcal{U}}\dot{\phi}^a\left(h_{ad}\Box_{\mathbf{g}}\phi^d + \partial_\mu\phi^b\partial_\nu\phi^c\mathbf{g}^{\mu\nu}\frac{1}{2}h^{ds}h_{ad}\cdot\left(\frac{\partial h_{sb}}{\partial\phi^c} + \frac{\partial h_{sc}}{\partial\phi^b} - \frac{\partial h_{bc}}{\partial\phi^s}\right)\right)dv_{\mathbf{g}}\\ &= \int_{\mathcal{U}}\dot{\phi}^a h_{ad}\left(\Box_{\mathbf{g}}\phi^d + \partial_\mu\phi^b\partial_\nu\phi^c\mathbf{g}^{\mu\nu}\Gamma^d_{bc}\right)dv_{\mathbf{g}} \end{aligned}$$

where $\Gamma^d_{bc} = \frac{1}{2}h^{ds}\left(\frac{\partial h_{sb}}{\partial\phi^c} + \frac{\partial h_{sc}}{\partial\phi^b} - \frac{\partial h_{bc}}{\partial\phi^s}\right)$ are the Christoffel symbols corresponding to the Riemannian metric h. The arbitrariness of $\dot{\phi}$ yields the following equation for wave maps,

$$\Box_{\mathbf{g}}\phi^a + \Gamma^a_{bc}\,\mathbf{g}^{\mu\nu}\partial_\mu\phi^b\partial_\nu\phi^c = 0 \qquad (30)$$

Example: Let N be a two dimensional Riemannian manifold endowed with a metric h of the form,

$$ds^2 = dr^2 + f(r)^2d\theta^2$$

Let ϕ be a wave map from $\mathbf{M}$ to N with components ϕ^1, ϕ^2, relative to the r, θ coordinates. Then, $\Gamma^1_{11} = \Gamma^2_{11} = \Gamma^1_{12} = \Gamma^2_{22} = 0$ and $\Gamma^1_{22} = -f'(r)f(r)$, $\Gamma^2_{12} = \frac{f'(r)}{f(r)}$. Therefore,

$$\begin{aligned} \Box_{\mathbf{g}}\phi^1 &= f'(r)f(r)\mathbf{g}^{\mu\nu}\partial_\mu\phi^2\partial_\nu\phi^2\\ \Box_{\mathbf{g}}\phi^2 &= -\frac{f'(r)}{f(r)}\mathbf{g}^{\mu\nu}\partial_\mu\phi^1\partial_\nu\phi^2 \end{aligned}$$

3. Maxwell equations.

An *electromagnetic field* F is an exact two form on a four dimensional manifold $\mathbf{M}$. That is, F is an antisymmetric tensor of rank two such that

$$F = dA \tag{31}$$

where A is a one-form on $\mathbf{M}$ called a *gauge potential* or connection 1-form. Note that A is not uniquely defined - indeed if χ is an arbitrary scalar function then the transformation

$$A \longrightarrow \tilde{A} = A + d\chi \tag{32}$$

yields another gauge potential $\tilde{A}$ for F. This degree of arbitrariness is called gauge freedom, and the transformations (32) are called gauge transformations.

The Lagrangian density for electromagnetic fields is

$$L[F] \quad = \quad -\frac{1}{4} F_{\mu\nu} F^{\mu\nu}.$$

Any compact variation $(F_{(s)}, \mathcal{U})$ of F can be written in terms of a compact variation $(A_{(s)}, \mathcal{U})$ of a gauge potential A, so that $F_{(s)} = dA_{(s)}$. Write

$$\dot{F} = \frac{d}{ds} F_{(s)} \Big|_{s=0}, \qquad \dot{A} = \frac{d}{ds} A_{(s)} \Big|_{s=0}$$

so that relative to a coordinate system x^α we have $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ and therefore $\dot{F}_{\mu\nu} = \partial_\mu \dot{A}_\nu - \partial_\nu \dot{A}_\mu$. The action principle gives

$$\begin{aligned} 0 &= \frac{d}{ds} \mathcal{S}(s) \Big|_{s=0} = -\frac{1}{2} \int_{\mathbf{M}} \dot{F}_{\mu\nu} F^{\mu\nu} dv_{\mathbf{g}} \\ &= -\frac{1}{2} \int_{\mathcal{U}} (\partial_\mu \dot{A}_\nu - \partial_\nu \dot{A}_\mu) F^{\mu\nu} dv_{\mathbf{g}} \\ &= -\int_{\mathcal{U}} \partial_\mu \dot{A}_\nu F^{\mu\nu} dv_{\mathbf{g}} = \int_{\mathcal{U}} \dot{A}_\nu \left(\frac{1}{\sqrt{-\mathbf{g}}} \partial_\nu \left(\sqrt{-\mathbf{g}} F^{\mu\nu} \right) \right) dv_{\mathbf{g}} \end{aligned}$$

Note that the second factor in the integrand is just $\mathbf{D}_\mu F^{\mu\nu}$ where $\mathbf{D}$ is the covariant derivative on $\mathbf{M}$ corresponding to $\mathbf{g}$. Hence the Euler-Lagrange equations take the form

$$\mathbf{D}_\nu F^{\mu\nu} = 0. \tag{33}$$

Together, (31) and (33) constitute the Maxwell equations.

Exercise. Given a vector field X^α on $\mathbf{M}$, show

$$\mathbf{D}_\alpha X^\alpha \quad = \quad \frac{1}{\sqrt{-\mathbf{g}}} \partial_\alpha \left(\sqrt{-\mathbf{g}} X^\alpha \right)$$

We can write the Maxwell equations in a more symmetric form by using the Hodge dual of F,

$${}^\star F_{\mu\nu} = \frac{1}{2} \in_{\mu\nu\alpha\beta} F^{\alpha\beta}$$

and by noticing that (33) is equivalent to $d\,{}^\star F = 0$. The Maxwell equations then take the form

$$dF = 0, \qquad d\,{}^\star F = 0 \tag{34}$$

or, equivalently,

$$\mathbf{D}_\nu F^{\mu\nu} = 0, \qquad \mathbf{D}_\nu \, {}^\star F^{\mu\nu} = 0 \tag{35}$$

Note that since Lorentz transformations commute with both the Hodge dual and exterior differentiation, the Lorentz invariance of the Maxwell equations is explicit in (34). Note also the very interesting duality symmetry of the equation,

$$F \to {}^\star F, \qquad {}^\star F \to -F \tag{36}$$

Definition. Given X an arbitrary vector field, we can define the contractions

$$\begin{aligned} E_\alpha &= (i_X F)_\alpha &= X^\mu F_{\alpha\mu} \\ H_\alpha &= (i_X \, {}^\star F)_\alpha &= X^\mu \, {}^\star F_{\alpha\mu} \end{aligned}$$

called, respectively, the electric and magnetic components[5] of F. Note that both these one-forms are perpendicular to X.

We specialize to the case when $\mathbf{M}$ is the Minkowski space $\mathbb{R}^{1+3}$ and $X = \frac{d}{dx^0} = \frac{d}{dt}$. As remarked, E, H are perpendicular to $\frac{d}{dt}$, so $E_0 = H_0 = 0$. The spatial components are by definition

$$E_i = F_{0i}, \qquad H_i = {}^\star F_{0i} = \frac{1}{2} \in_{0ijk} F^{jk} = \frac{1}{2} \in_{ijk} F^{jk}$$

We now use (34) to derive equations for E and H from above, which imply

$$\mathbf{D}_\nu \, {}^\star F^{\mu\nu} = 0 \tag{37}$$

and (33), respectively. Setting $\mu = 0$ in both equations of (35) we derive,

$$\partial^i E_i = 0, \qquad \partial^i H_i = 0$$

Setting $\mu = i$ and observing that $F_{ij} = \in_{ijk} H^k, \quad {}^\star F_{ij} = - \in_{ijk} E^k$ we write

$$\begin{aligned} 0 &= -\partial^0 E_i + \partial^j F_{ij} = \partial_0 E_i + \in_{ijk} \partial^j H^k = \partial_t E_i + (\nabla \times H)_i \\ 0 &= \partial_t H_i - \in_{ijk} \partial_j E_k = \partial_t H_i - (\nabla \times E)_i \end{aligned}$$

Therefore,

$$\partial_t E + \nabla \times H = 0, \quad \partial_t H - \nabla \times E = 0 \tag{38}$$

$$divE = 0, \quad \text{div } H = 0 \tag{39}$$

These are the classical Maxwell in vacuum (i.e. in the absence of sources). Alongside (38) and (39) we can assign data at time $t = 0$,

$$E_i(0, x) = E_i^{(0)}, \qquad H_i(0, x) = H_i^{(0)}$$

Exercise. Show that the equations (39) are preserved by the time evolution of the system (38. In other words if $E^{(0)}, H^{(0)}$ satisfy (39) then they are satisfied by E, H for all times $t \in \mathbb{R}$.

[5] In physics books the electric field $E_i = -F_{0i}$ and the magnetic field is dented by B. Thus the Maxwell equations look somewhat different from ours.

The Maxwell equations with sources have the form,

$$\begin{aligned}\partial_t E + \nabla \times H &= j_i \\ \nabla \cdot E &= \rho \\ \partial_t H - \nabla \times E &= 0 \\ \nabla \cdot H &= 0\end{aligned} \tag{40}$$

or, in space-time notation, with $J^0 = \rho$ and $J^i = j^i$,

$$\mathbf{D}^\nu F_{\mu\nu} = J^\mu \tag{41}$$

Observe that the 4-current J verifies the continuity equation,

$$\mathbf{D}_\mu J^\mu = 0. \tag{42}$$

Note that the symmetry (36) is broken by the presence of charges.

5. The Einstein Field Equations:

We now consider the action,

$$\mathcal{S} = \int_{\mathcal{U}} L dv_{\mathbf{g}}.$$

Here U is a relatively compact domain of $(\mathbf{M}, \mathbf{g})$ and L, the Lagrangian, is assumed to be a scalar function on $\mathbf{M}$ whose dependence on the metric should involve no more than two derivatives[6]. It is also assumed to depend on the matterfields $\psi = \psi^{(1)}, \psi^{(2)}, \dots \psi^{(p)}$ present in our space-time.

In fact we write,

$$\mathcal{S} = \mathcal{S}_G + \mathcal{S}_M$$

with,

$$\begin{aligned}\mathcal{S}_G &= \int_{\mathcal{U}} L_G dv_{\mathbf{g}} \\ \mathcal{S}_M &= \int_{\mathcal{U}} L_M dv_{\mathbf{g}}\end{aligned}$$

denoting, respectively, the actions for the gravitational field and matter. The matter Lagrangian L_M depends only on the matterfields ψ, assumed to be covariant tensorfields, and the inverse of the space-time metric $\mathbf{g}^{\alpha\beta}$ which appears in the contraction of the tensorfields ψ in order to produce the scalar L_M. It may also depend on additional positive definite metrics which are not to be varied [7].

Now the only possible candidate for the gravitational Lagrangian L_G, which should be a scalar invariant of the metric with the property that the corresponding Euler-Lagrange equations involve at most two derivatives of the metric, is given[8] by the

[6] In fact we only require that the corrsponding Euler-Lagrange equations should involve no more than two derivatives of the metric.

[7] This is the case of the metric h in the case of wave maps or the Killing scalar product in the case of the Yang-Mills equations.

[8] up to an additive constant

scalar curvature $\mathbf{R}$. Therefore we set,

$$L_G = \mathbf{R}.$$

Consider now a compact variation $(\mathbf{g}_{(s)}, \mathcal{U})$ of the metric $\mathbf{g}$. Let $\dot{\mathbf{g}}_{\mu\nu} = \frac{d}{ds}\mathbf{g}_{\mu\nu}|_{s=0}$. Thus for small s, $\mathbf{g}_{\mu\nu}(s) = \mathbf{g}_{\mu\nu} + s\dot{\mathbf{g}}_{\mu\nu} + O(s^2)$. Also, $\mathbf{g}^{\mu\nu}(s) = \mathbf{g}^{\mu\nu} - s\dot{\mathbf{g}}^{\mu\nu} + O(s^2)$ where $\dot{\mathbf{g}}^{\mu\nu} = \mathbf{g}^{\alpha\mu}\mathbf{g}^{\beta\nu}\dot{\mathbf{g}}_{\alpha\beta}$. Then,

$$\frac{d}{ds}\mathbf{S}_G(s)\Big|_{s=0} = \int_{\mathcal{U}} \dot{\mathbf{R}} dv_{\mathbf{g}} + \int_{\mathcal{U}} \mathbf{R} \dot{dv_{\mathbf{g}}}$$

Now,

$$\dot{dv_{\mathbf{g}}} = \frac{1}{2}\mathbf{g}^{\mu\nu}\dot{\mathbf{g}}_{\mu\nu} dv_{\mathbf{g}}$$

Indeed, relative to a coordinate system, $dv_{\mathbf{g}} = \sqrt{-\mathbf{g}}dx^0 dx^1 \dots dx^n$ Thus, the above equality follows from,

$$\dot{\mathbf{g}} = \mathbf{g}\mathbf{g}^{\alpha\beta}\dot{\mathbf{g}}_{\alpha\beta},$$

with $\mathbf{g}$ the determinant of $\mathbf{g}_{\alpha\beta}$. On the other hand, writing $\mathbf{R} = \mathbf{g}^{\mu\nu}\mathbf{R}_{\mu\nu}$ and using the formula $\frac{d}{ds}\mathbf{g}^{\mu\nu}_{(s)}\Big|_{s=0} = -\dot{\mathbf{g}}^{\mu\nu}$, we calculate, $\dot{\mathbf{R}} = -\dot{\mathbf{g}}^{\mu\nu}\mathbf{R}_{\mu\nu} + \mathbf{g}^{\mu\nu}\dot{\mathbf{R}}_{\mu\nu}$. Therefore,

$$\frac{d}{ds}\mathbf{S}_G(s)\Big|_{s=0} = -\int_{\mathcal{U}} (\mathbf{R}^{\mu\nu} - \frac{1}{2}\mathbf{g}^{\mu\nu}\mathbf{R})\dot{\mathbf{g}}_{\mu\nu} dv_{\mathbf{g}} + \int_{\mathcal{U}} \mathbf{g}^{\mu\nu}\dot{\mathbf{R}}_{\mu\nu} dv_{\mathbf{g}} \tag{43}$$

To calculate $\dot{\mathbf{R}}_{\mu\nu}$ we make use of the following Lemma,

LEMMA 2.2. *Let $\mathbf{g}_{\mu\nu}(s)$ be a family of space-time metrics with $\mathbf{g}(0) = \mathbf{g}$ and $\frac{d}{ds}\mathbf{g}(0) = \dot{\mathbf{g}}$. Set also, $\frac{d}{ds}\mathbf{R}_{\alpha\beta}(s)|_{s=0} = \dot{\mathbf{R}}_{\alpha\beta}$. Then,*

$$\dot{\mathbf{R}}_{\mu\nu} = \mathbf{D}_\alpha \dot{\Gamma}^\alpha_{\mu\nu} - \mathbf{D}_\mu \dot{\Gamma}^\alpha_{\alpha\nu}$$

where $\dot{\Gamma}$ is the tensor,

$$\dot{\Gamma}^\alpha_{\beta\gamma} = \frac{1}{2}\mathbf{g}^{\alpha\lambda}(\mathbf{D}_\beta \dot{\mathbf{g}}_{\gamma\lambda} + \mathbf{D}_\gamma \dot{\mathbf{g}}_{\beta\lambda} - \mathbf{D}_\lambda \dot{\mathbf{g}}_{\beta\gamma})$$

Proof : Since both sides of the identity are tensors it suffices to prove the formula at a point p relative to a particular system of coordinates for which the Christoffel symbols Γ vanish at p. Relative to such a coordinate system the Ricci tensor has the form $\mathbf{R}_{\mu\nu} = \mathbf{D}_\alpha \Gamma^\alpha_{\mu\nu} - \mathbf{D}_\mu \Gamma^\alpha_{\alpha\nu}$. ■

Returning to (43) we find that since $\mathbf{g}^{\mu\nu}\dot{\mathbf{R}}_{\mu\nu}$ can be written as a space-time divergence of a tensor compactly supported in U the corresponding integral vanishes identically. We therefore infer that,

$$\frac{d}{ds}\mathbf{S}_G(s)\Big|_{s=0} = -\int_U \mathbf{E}^{\mu\nu}\dot{\mathbf{g}}_{\mu\nu} dv_{\mathbf{g}} \tag{44}$$

where $\mathbf{E}^{\mu\nu} = \mathbf{R}^{\mu\nu} - \frac{1}{2}\mathbf{g}^{\mu\nu}\mathbf{R}$. We now consider the variation of the action integral $\mathbf{S}_M$ with respect to the metric. As remarked before L_M depends on the metric $\mathbf{g}$ through its inverse $\mathbf{g}^{\mu\nu}$. Therefore if we denote $\mathbf{S}_M(s) = \mathbf{S}_M[\psi, \mathbf{g}_{(s)}; \mathcal{U}]$ we have, writing $d\dot{v}_{\mathbf{g}} = \frac{1}{2}\mathbf{g}_{\mu\nu}\dot{\mathbf{g}}^{\mu\nu} dv_{\mathbf{g}}$,

$$\begin{aligned}\frac{d}{ds}\mathbf{S}_M(s)\Big|_{s=0} &= -\int_{\mathcal{U}} \frac{\partial L_M}{\partial \mathbf{g}^{\mu\nu}}\dot{\mathbf{g}}^{\mu\nu} dv_{\mathbf{g}} + \int_{\mathcal{U}} L_M d\dot{v}_{\mathbf{g}} \\ &= -\int_{\mathcal{U}} (\frac{\partial L_M}{\partial \mathbf{g}^{\mu\nu}} - \frac{1}{2}\mathbf{g}_{\mu\nu} L_M)\dot{\mathbf{g}}^{\mu\nu} dv_{\mathbf{g}}\end{aligned}$$

Definition. The symmetric tensor,

$$\mathbf{T}_{\mu\nu} = -\left(\frac{\partial L_M}{\partial \mathbf{g}^{\mu\nu}} - \frac{1}{2}\mathbf{g}_{\mu\nu} L_M\right)$$

is called the *energy-momentum tensor* of the action $\mathbf{S}_M$.

With this definition we write,

$$\frac{d}{ds}\mathbf{S}_M(s)\Big|_{s=0} = \int_{\mathcal{U}} \mathbf{T}^{\mu\nu}\dot{\mathbf{g}}_{\mu\nu} dv_{\mathbf{g}} \tag{45}$$

Finally, combining 44 with 45, we derive for the total action $\mathbf{S}$,

$$\frac{d}{ds}\mathbf{S}(s)\Big|_{s=0} = -\int_{\mathcal{U}} (\mathbf{E}^{\mu\nu} - \mathbf{T}^{\mu\nu})\dot{\mathbf{g}}_{\mu\nu} dv_{\mathbf{g}}$$

Since $\dot{\mathbf{g}}_{\mu\nu}$ is an arbitrary symmetric 2-tensor compactly supported in U we derive the *Einstein field equation*,

$$\mathbf{E}^{\mu\nu} = \mathbf{T}^{\mu\nu}$$

Recall that the Einstein tensor $\mathbf{E}$ satisfies the twice contracted Bianchi identity,

$$\mathbf{D}^{\nu}\mathbf{E}_{\mu\nu} = 0$$

This implies that the energy-momentum tensor $\mathbf{T}$ is also divergenceless,

$$\mathbf{D}_{\nu}\mathbf{T}^{\mu\nu} = 0 \tag{46}$$

which is the concise, space-time expression for the law of conservation of energy-momentum of the matter-fields.

Let us now replace $\mathbf{S}_G$ with the new action,

$$\mathbf{S}_{G,\Lambda} = \int (R - 2\Lambda) dv_{\mathbf{g}}$$

The resulting field equations are,

$$\mathbf{E}_{\mu\nu} + \Lambda\mathbf{g}_{\mu\nu} = \mathbf{T}_{\mu\nu} \tag{47}$$

where Λ is the cosmological constant. Observe that (47) can be written in the form,

$$\mathbf{E}_{\mu\nu} = \mathbf{T}_{\mu\nu} - \Lambda\mathbf{g}_{\mu\nu}$$

and interpret $-\Lambda\mathbf{g}_{\mu\nu}$ as the energy density of the vacuum, a source of energy and momentum present even in the absence of matterfields.

Another possible gravitational lagrangean, which depends on a scalar function ϕ

$$\mathbf{L}_G = f(\phi)\mathbf{R} + \frac{1}{2}\mathbf{g}^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - V(\phi)$$

where $f(\phi), V(\phi)$ are specified functions which define the theory. This leads to scalar-tensor theories of gravity.

3. The energy-momentum tensor

The conservation law (46) is a fundamental property of a matterfield. We now turn to a more direct derivation.

We consider an arbitrary Lagrangian field theory with stationary solution ψ. Let Φ_s be the one-parameter group of local diffeomorphisms generated by a given vectorfield X. We shall use the flow Φ to vary the fields ψ according to

$$\begin{aligned}\mathbf{g}_s &= (\Phi_s)_*\mathbf{g}\\ \psi_s &= (\Phi_s)_*\psi.\end{aligned}$$

From the invariance of the action integral under diffeomorphisms,

$$\mathbf{S}(s) = \mathbf{S}[\psi_s, \mathbf{g}_s; \mathbf{M}] = \mathbf{S}_M[\psi, \mathbf{g}; \mathbf{M}].$$

So that

$$0 = \frac{d}{ds}\mathbf{S}(s)\Big|_{s=0} = \int_{\mathbf{M}} \frac{\delta \mathbf{S}}{\delta \psi} dv_{\mathbf{g}} + \int_{\mathbf{M}} \mathbf{T}^{\mu\nu}\dot{\mathbf{g}}_{\mu\nu} dv_{\mathbf{g}} \tag{48}$$

The first term is clearly zero, ψ being a stationary solution. In the second term, which represents variations with respect to the metric, we have

$$\dot{\mathbf{g}}_{\mu\nu} = \frac{d}{ds}(\mathbf{g}_s)_{\mu\nu}\Big|_{s=0} = \mathcal{L}_X\mathbf{g}_{\mu\nu} = \mathbf{D}_\mu X_\nu + \mathbf{D}_\nu X_\mu$$

Therefore

$$0 = \int_{\mathbf{M}} \mathbf{T}^{\mu\nu}\mathcal{L}_X\mathbf{g}_{\mu\nu} dv_{\mathbf{g}} = 2\int_{\mathbf{M}} \mathbf{T}^{\mu\nu}\mathbf{D}_\nu X_\mu dv_{\mathbf{g}} = -2\int_{\mathbf{M}} \mathbf{D}_\nu\mathbf{T}^{\mu\nu}X_\mu dv_{\mathbf{g}}$$

As X was arbitrary, we conclude

$$\mathbf{D}_\nu\mathbf{T}^{\mu\nu} = 0. \tag{49}$$

This is again the law of conservation of energy-momentum.

We list below the energy-momentum tensors of the field theories discussed before. We leave it to the reader to carry out the calculations using the definition.

(1) The energy-momentum for the scalar field equation is,

$$\mathbf{T}_{\alpha\beta} = \frac{1}{2}\left(\phi_{,\alpha}\phi_{,\beta} - \frac{1}{2}\mathbf{g}_{\alpha\beta}(\mathbf{g}^{\mu\nu}\phi_{,\mu}\phi_{,\nu} + 2V(\phi))\right)$$

(2) The energy-momentum for wave maps is given by,

$$\mathbf{T}_{\alpha\beta} = \frac{1}{2}\left(< \phi_{,\alpha}, \phi_{,\beta} > - \frac{1}{2}\mathbf{g}_{\alpha\beta}(\mathbf{g}^{\mu\nu} < \phi_{,\mu}, \phi_{,\nu} >)\right)$$

where $< \, , \, >$ denotes the Riemannian inner product on the target manifold.

(3) The energy-momentum tensor for the Maxwell equations is,

$$\mathbf{T}_{\alpha\beta} = F_{\alpha}^{\cdot\mu} F_{\beta\mu} - \frac{1}{4}\mathbf{g}_{\alpha\beta}(F_{\mu\nu}F^{\mu\nu})$$

An acceptable notion of the energy-momentum tensor $\mathbf{T}$ must satisfy the following properties in addition of the conservation law (49),

(1) $\mathbf{T}$ is symmetric
(2) $\mathbf{T}$ satisfies the *positive energy condition* that is, $\mathbf{T}(X, Y) \geq 0$, for any future directed time-like vectors X, Y.

The symmetry property is automatic in our construction. The following proposition asserts that the energy-momentum tensors of the field theories described above satisfy the positive energy condition.

Proposition 3.1. *The energy-momentum tensor of the scalar wave equation satisfies the positive energy condition if V is positive. The energy- momentum tensors for the wave maps, Maxwell equations and Yang-Mills satisfy the positive energy condition.*

Proof : To prove the positivity conditions consider two vectors X, Y, at some point $p \in \mathbf{M}$, which are both causal future oriented. The plane spanned by X, Y intersects the null cone at p along two null directions[9]. Let $L, \underline{L}$ be the two future directed null vectors corresponding to the two complementary null directions and normalized by the condition

$$< L, \underline{L} >= -2$$

i.e. they form a null pair. Since the vectorfields X, Y are linear combinations with positive coefficients of $L, \underline{L}$, the proposition will follow from showing that $\mathbf{T}(L, L) \geq 0$, $\mathbf{T}(\underline{L}, \underline{L}) \geq 0$ and $\mathbf{T}(L, \underline{L}) \geq 0$. To show this we consider a frame at p formed by the vectorfields $E_{(n+1)} = L$, $E_{(n)} = \underline{L}$ and $E_{(1)}, \ldots, E_{(n-1)}$ with the properties,

$$< E_{(i)}, E_{(n)} >=< E_{(i)}, E_{(n+1)} >= 0$$

and

$$< E_{(i)}, E_{(j)} >= \delta_{ij}$$

for all $i, j = 1, \ldots, n - 1$. A frame with these properties is called a *null frame*.

[9]If X, Y are linearly dependent any plane passing through their common direction will do.

(1) We now calculate, in the case of the wave equation,

$$\begin{aligned}\mathbf{T}(L,L) &= \frac{1}{2}E(\phi)^2\\ \mathbf{T}(\underline{L},\underline{L}) &= \frac{1}{2}\underline{L}(\phi)^2.\end{aligned}$$

which are clearly non-negative. Now,

$$\mathbf{T}(L,\underline{L}) = \frac{1}{2}\left[L(\phi)\underline{L}(\phi) + (g^{\mu\nu}\phi_{,\mu}\phi_{,\nu} + 2V(\phi))\right]$$

and we aim to express $g^{\mu\nu}\phi_{,\mu}\phi_{,\nu}$ relative to our null frame. To do this, observe that relative to the null frame the only nonvanishing components of the metric $g_{\alpha\beta}$ are,

$$g_{n(n+1)} = -2 \quad , \quad g_{ii} = 1 \;\; i = 1, \dots, n-1$$

and those of the inverse metric $g^{\alpha\beta}$ are

$$g^{n(n+1)} = -\frac{1}{2} \quad , \quad g^{ii} = 1 \;\; i = 1, \dots, n-1$$

Therefore,

$$g^{\mu\nu}\phi_{,\mu}\phi_{,\nu} = -L(\phi)\underline{L}(\phi) + |\nabla\!\!\!/\phi|^2$$

where

$$|\nabla\!\!\!/\phi|^2 = (E_{(1)}(\phi))^2 + (E_{(2)}(\phi))^2 + \dots E_{(n-1)}(\phi)^2.$$

Therefore,

$$\mathbf{T}(L,\underline{L}) = \frac{1}{2}|\nabla\!\!\!/\phi|^2 + V(\phi).$$

(2) For wave maps we have, according to the same calculation.

$$\begin{aligned}\mathbf{T}(E,E) &= \frac{1}{2} < E(\phi), E(\phi) >\\ \mathbf{T}(\underline{E},\underline{E}) &= \frac{1}{2} < \underline{E}(\phi), \underline{E}(\phi) >\\ \mathbf{T}(E,\underline{E}) &= \frac{1}{2}\sum_{i=1}^{n-1} < E_{(i)}(\phi), E_{(i)}(\phi) > .\end{aligned}$$

The positivity of $\mathbf{T}$ is then a consequence of the Riemannian character metric h on the target manifold N.

(3) To show positivity for the energy momentum tensor of the Maxwell equations in $3+1$ dimensions we first write the tensor in the more symmetric form

$$\mathbf{T}_{\alpha\beta} = \frac{1}{2}\left(F_\alpha{}^\mu F_{\beta\mu} + {}^\star F_\alpha{}^\mu\, {}^\star F_{\beta\mu}\right) \tag{50}$$

where ${}^\star F$ is the Hodge dual of F, i.e. ${}^\star F_{\alpha\beta} = \frac{1}{2}\in_{\alpha\beta\mu\nu} F^{\mu\nu}$.

Exercise. Check formula (50).

We introduce the following null decomposition of F, at points $p \in \mathbf{M}$,

$$\begin{aligned}\alpha_A = F_{a4} \quad &, \quad \underline{\alpha}_a = F_{a3}\\ \rho = \frac{1}{2}F_{34} \quad &, \quad \sigma = \frac{1}{2}{}^\star F_{34}.\end{aligned}$$

which completely determines the tensor F. Here the indices $a = 1, 2$ correspond to the directions E_1, E_2 tangent to the sphere while the indices $3, 4$ correspond to $E_3 = \underline{L}$ and $E_4 = L$. We then calculate that for ${}^\star F$,

$$
\begin{array}{rcl}
{}^\star F_{a4} = -\,{}^\star\alpha_a = & , & {}^\star F_{a3} = {}^\star\underline{\alpha}_a \\
{}^\star F_{34} = 2\sigma & , & {}^{\star\star}F_{34} = -2\rho
\end{array}
$$

where ${}^\star\alpha_a = \in_{ab}\ \alpha_b$. Here $\in_{ab}$ is the volume form on the unit sphere, hence $\in_{ab}= \frac{1}{2} \in_{ab34}$, i.e. $\in_{11}=\in_{22}= 0$, $\in_{12}= -\epsilon_{21} = 1$. With this notation we calculate,

$$
\begin{aligned}
\mathbf{T}(E_{(4)}, E_{(4)}) &= \frac{1}{2}\sum_{a=1}^{2} (F_{4a} \cdot F_{4a} + {}^\star F_{4a} \cdot {}^\star F_{4a}) \\
&= \frac{1}{2}\sum_{a=1}^{2} (\alpha_a \cdot \alpha_a + {}^\star\alpha_a \cdot {}^\star\alpha_a) \\
&= \sum_{a=1}^{2} \alpha_a \cdot \alpha_a = |\alpha|^2 \geq 0.
\end{aligned}
$$

Similarly,

$$
\mathbf{T}(E_{(3)}, E_{(3)}) = \sum_{A=1}^{2} \underline{\alpha}_a \cdot \underline{\alpha}_a = |\underline{\alpha}|^2 \geq 0
$$

and in the same vein we find

$$
\mathbf{T}(E, \underline{E}) = \rho^2 + \sigma^2 \geq 0
$$

which proves our assertion.

■

Exercise. Consider the Maxwell equations with sources, (40). Show that the energy momentum tensor $T_{\alpha\beta}$ verifies,

$$
D^\alpha T_{\alpha\beta} = J^\lambda F_{\lambda\beta}
$$

with $J^\lambda F_{\beta\lambda}$ the Lorentz force.

3.2. Perfect fluids. A perfect fluid is a continuous distribution of matter with energy-momentum tensor,

$$
\mathbf{T}_{\mu\nu} = \rho u_\mu u_\nu + P(\mathbf{m}_{\mu\nu} + u_\mu u_\nu) \tag{51}
$$

where u^μ is a unit timelike vector-field representing the four velocity of the fluid. Also ρ and P denote, respectively, the mass-energy density and pressure of the fluid (as measured in its rest frame). The equation of motion of a perfect fluid, subject to no external forces, in a spacetime $(\mathbf{M}, \mathbf{g})$, is

$$
\mathbf{D}^\mu \mathbf{T}_{\mu\nu} = 0. \tag{52}
$$

These can be rewritten in the form,

$$\begin{aligned} u^\mu \mathbf{D}_\mu \rho + (\rho + P)\mathbf{D}^\mu u_\mu = 0, \\ (P+\rho)u^\mu \mathbf{D}_\mu u_\nu + (\mathbf{g}_{\mu\nu} + u_\mu u_\nu)\mathbf{D}^\mu P = 0 \end{aligned} \tag{53}$$

In Minkowski space, i.e. $\mathbf{g} = \mathbf{m}$, in the non relativistic limit, i.e.

$$P << \rho, \quad u^\mu = (1, v), \qquad v\frac{dP}{dt} << |\nabla P|$$

where v denotes the usual 3 velocity and ∇P the spatial gradient, we derive

$$\begin{aligned} \partial_t \rho + \nabla(\rho v) = 0, \\ \rho(\partial_t v + v \cdot \nabla v) = -\nabla P \end{aligned} \tag{54}$$

3.3. Conformal Fields. Another important property which the energy momentum tensor of a field theory may satisfy is the trace free condition, that is

$$\mathbf{g}_{\alpha\beta}\mathbf{T}^{\alpha\beta} = 0.$$

It turns out that this condition is satisfied by all field theories which are conformally invariant.

Definition. A field theory is said to be *conformally invariant* if the corresponding action integral is invariant under conformal transformations of the metric

$$\mathbf{g}_{\alpha\beta} \longrightarrow \tilde{\mathbf{g}}_{\alpha\beta} = \Omega^2 \mathbf{g}_{\alpha\beta}$$

Ω a positive smooth function on the space-time.

PROPOSITION 3.4. *The energy momentum tensor* $\mathbf{T}$ *of a conformally invariant field theory is traceless.*

Proof: Consider an arbitrary smooth function f compactly supported in $\mathcal{U} \subset \mathcal{M}$. Consider the following variation of a given metric $\mathbf{g}$,

$$\mathbf{g}_{\mu\nu}(s) = e^{sf}\mathbf{g}_{\mu\nu}.$$

Let $\mathcal{S}(s) = \mathcal{S}_{\mathcal{U}}[\psi, \mathbf{g}(s)]$. In view of the covariance of $\mathcal{S}$ we have $\mathcal{S}(s) = \mathcal{S}(0)$. Hence,

$$0 = \frac{d}{ds}\mathcal{S}(s)|_{s=0} = \int_{\mathcal{U}} \mathbf{T}^{\mu\nu}\dot{\mathbf{g}}_{\mu\nu} dv_{\mathbf{g}}$$

where

$$\dot{\mathbf{g}}_{\mu\nu} = \frac{d}{ds}\mathbf{g}_{\mu\nu}(s)\Big|_{s=0} = f\mathbf{g}_{\mu\nu}.$$

Hence, $\int_{\mathcal{U}} (\mathbf{T}^{\mu\nu}\mathbf{g}_{\mu\nu}) f dv_{\mathbf{g}} = 0$ and since f is arbitrary we infer that,

$$trT = g^{\mu\nu}\mathbf{T}_{\mu\nu} \equiv 0.$$

■

We can easily check that the Maxwell and the Yang-Mills equations are conformally invariant in $3+1$-dimensions. The wave maps field theory is conformally invariant in dimension $1+1$, i.e. if the space-time $\mathcal{M}$ is two-dimensional[10].

[10] Similarly for the linear scalar wave equation

Remark: The action integral of the Maxwell equations, $\mathbf{S} = \int_{\mathcal{U}} F_{\alpha\beta}F^{\alpha\beta}dv_{\mathbf{g}}$ is conformally invariant in any dimension provided that we also scale the electromagnetic field F. Indeed if $\tilde{\mathbf{g}}_{\alpha\beta} = \Omega^2\mathbf{g}_{\alpha\beta}$ then $dv_{\tilde{\mathbf{g}}} = \Omega^{n+1}dv_{\mathbf{g}}$ and if we also set $\tilde{F}_{\alpha\beta} = \Omega^{-\frac{n-3}{2}}F_{\alpha\beta}$ we get

$$\tilde{\mathbf{S}}[\tilde{F},\tilde{\mathbf{g}}] \quad = \quad \int \tilde{F}_{\alpha\beta}\tilde{F}_{\gamma\delta}\tilde{\mathbf{g}}^{\alpha\gamma}\tilde{\mathbf{g}}^{\beta\delta}dv_{\tilde{\mathbf{g}}} = \int F_{\alpha\beta}F_{\gamma\delta}\mathbf{g}^{\alpha\gamma}\mathbf{g}^{\beta\delta}dv_{\mathbf{g}} = \mathbf{S}[F,\mathbf{g}].$$

We finish this section with a simple observation concerning conformal field theories in $1+1$ dimensions. We specialize in fact to the Minkowski space $\mathbb{R}^{1+1}$ and consider the local conservation law, $\partial^\mu \mathbf{T}_{\nu\mu} = 0$. Setting $\nu = 0, 1$ we derive

$$\partial^0\mathbf{T}_{00} + \partial^1\mathbf{T}_{01} = 0, \quad \partial^0\mathbf{T}_{01} + \partial^1\mathbf{T}_{11} = 0 \tag{55}$$

Since the energy-momentum tensor is trace-free, we get $\mathbf{T}_{00} = \mathbf{T}_{11} = A$, say. Set $\mathbf{T}_{01} = \mathbf{T}_{10} = B$. Therefore (55) implies that both A and B satisfy the linear homogeneous wave equation;

$$\Box A \quad = \; 0 \; = \quad \Box B. \tag{56}$$

Using this observation it is is easy to prove that smooth initial data remain smooth for all time.

For example, wave maps are conformally invariant in dimension $1+1$. In this case

$$A \;=\; \mathbf{T}_{00} \;=\; \frac{1}{2}\left(<\partial_t\phi, \partial_t\phi> + <\partial_x\phi, \partial_x\phi>\right),$$

Given data in $C_0^\infty(\mathbb{R})$, (56) implies that the derivatives of ϕ remain smooth for all positive times. This proves global existence.

3.5. Conservation Laws. The energy-momentum tensor of a field theory is intimately connected with conservations laws. This connection is seen through Noether's principle,

Noether's Principle: *To any one-parameter group of transformations preserving the action there corresponds a conservation law.*

We illustrate this fundamental principle as follows: Let $\mathbf{S} = \mathbf{S}[\psi, \mathbf{g}]$ be the action integral of the fields ψ. Let χ_t be a 1-parameter group of isometries of $\mathbf{M}$, i.e., $(\chi_t)_*\mathbf{g} = \mathbf{g}$. Then

$$\begin{aligned}\mathbf{S}[(\chi_t)_*\psi, \mathbf{g}] &= \mathbf{S}[(\chi_t)_*\psi, (\chi_t)_*\mathbf{g}] \\ &= \mathbf{S}[\psi, \mathbf{g}].\end{aligned}$$

Thus the action is preserved under $\psi \to (\chi_t)_*\psi$. In view of Noether's Principle we ought to find a conservation law for the corresponding Euler-Lagrange equations[11]. We derive these laws using the Killing vectorfield X which generates χ_t.

[11] The same argument holds for conformal isometries acting on a conformally invariant field theory. We therefore also expect conservation laws in such a setting.

We begin with a general calculation involving the energy-momentum tensor $\mathbf{T}$ of ψ and an arbitrary vectorfield X. P the one-form obtained by contracting $\mathbf{T}$ with X.

$$P_\alpha = \mathbf{T}_{\alpha\beta} X^\beta$$

Since $\mathbf{T}$ is symmetric and divergence-free

$$\mathbf{D}^\alpha P_\alpha = (\mathbf{D}^\alpha \mathbf{T}_{\alpha\beta}) X^\beta + \mathbf{T}_{\alpha\beta}\left(\mathbf{D}^\alpha X^\beta\right) = \frac{1}{2}\mathbf{T}^{\alpha\beta}\ {}^{(X)}\pi_{\alpha\beta}$$

where ${}^{(X)}\pi_{\alpha\beta}$ is the deformation tensor of X.

$${}^{(X)}\pi_{\alpha\beta} = (\mathcal{L}_X \mathbf{g})_{\alpha\beta} = \mathbf{D}_\alpha X_\beta + \mathbf{D}_\beta X_\alpha$$

Notation. We denote the backward light cone with vertex $p = (\bar{t}, \bar{x}) \in \mathbb{R}^{n+1}$ by

$$\mathcal{N}^-(\bar{t}, \bar{x}) = \{(t, x)\ \big| 0 \le t \le \bar{t}; |x - \bar{x}| = \bar{t} - t\}.$$

The restriction of this set to some time interval $[t_1, t_2]$, $t_1 \le t_2 \le \bar{t}$, will be written $\mathcal{N}^-_{[t_1,t_2]}(\bar{t}, \bar{x})$. These null hypersurfaces are null boundaries of,

$$\begin{aligned}\mathcal{J}^{-1}(\bar{t}, \bar{x}) &= \{(t, x)\ \big|\ 0 \le t \le \bar{t};\ |x - \bar{x}| \le \bar{t} - t\}\\ \mathcal{J}^-_{[t_2,t_1]}(\bar{t}, \bar{x}) &= \{(t, x)\ \big| t_2 \le t \le t_1;\ |x - \bar{x}|\ \le \bar{t} - t\}\end{aligned}$$

We shall denote by $S_t = S_t(\bar{t}, \bar{x})$ and $B_t = B_t(\bar{t}, \bar{x})$ the intersection of the time slice Σ_t with $\mathcal{N}^-$, respectively $\mathcal{J}^-$.

At each point $q = (t, x)$ along $\mathcal{N}^-(p)$, we define the null pair (E_+, E_-) of future oriented null vectors

$$\underline{L} = E_+ = \partial_t + \frac{x^i - \bar{x}^i}{|x - \bar{x}|}\partial_i, \qquad L = E_- = \partial_t - \frac{x^i - \bar{x}^i}{|x - \bar{x}|}\partial_i$$

Observe that both $L, \underline{L}$ are null and $< L, \underline{L} >= -2$.

The following is a simple consequence of Stoke's theorem, in the following form.

PROPOSITION 3.6. *Let P_μ be a one-form satisfying $\partial^\mu P_\mu = F$. Then*[12]*, for all* $t_1 \le t_2 \le \bar{t}$,

$$\int_{B_{t_2}} \langle P, \partial_t \rangle + \int_{\mathcal{N}^-_{[t_1,t_2]}(p)} \langle P, E_- \rangle = \int_{B_{t_1}} \langle P, \partial_t \rangle - \int_{\mathcal{J}^-_{[t_1,t_2]}(p)} F dt dx \tag{57}$$

where,

$$\int_{\mathcal{N}^-_{[t_1,t_2]}(p)} \langle P, E_- \rangle = \int_{t_1}^{t_2} dt \int_{S_t} \langle P, E_- \rangle\, da_t.$$

Applying this proposition to Stoke's theorem to (57) we get

[12] The brackets $\langle \cdot, \cdot \rangle$ in (57) denote inner product with respect to the Minkowski metric.

THEOREM 3.7. *Let T be the energy-momentum tensor associated to a field theory and X an arbitrary vector field. Then*

$$\int_{B_{t_2}} \mathbf{T}(\partial_t, X) + \int_{\mathcal{N}^-_{[t_1,t_2]}(p)} \mathbf{T}(E_-, X) = \int_{B_{t_1}} \mathbf{T}(\partial_t, X) - \int_{\mathcal{J}^-_{[t_1,t_2]}(p)} \mathbf{T}^{\alpha\beta\ (X)}\pi_{\alpha\beta} dtdx \tag{58}$$

In the particular case when X is Killing, its deformation tensor π vanishes identically. Thus,

COROLLARY 3.8. *If X is a killing vectorfield,*

$$\int_{B_{t_2}} \mathbf{T}(\partial_t, X) + \int_{\mathcal{N}^-_{[t_1,t_2]}(p)} \mathbf{T}(L, X) = \int_{B_{t_1}} \mathbf{T}(\partial_t, X) \tag{59}$$

Moreover (59) *remains valid if* $\mathbf{T}$ *is traceless and X is conformal Killing.*

The identity (59) is usually applied to time-like future-oriented Killing vectorfields X in which case the positive energy condition for $\mathbf{T}$ insures that all integrands are positive. We know that, up to a Lorentz transformation the only Killing, future oriented timelike vectorfield is a constant multiple of ∂_t. Choosing $X = \partial_t$, (59) becomes,

$$\int_{B_{t_2}} \mathbf{T}(\partial_t, \partial_t) + \int_{\mathcal{N}^-_{[t_1,t_2]}(p)} \mathbf{T}(E_-, \partial_t) = \int_{B_{t_1}} \mathbf{T}(\partial_t, \partial_t) \tag{60}$$

In the case of a conformal field theory we can pick X to be the future timelike, conformal Killing vectorfield $X = K_0 = (t^2 + |x|^2)\partial_t + 2tx^i\partial_i$. Thus,

$$\int_{B_{t_2}} \mathbf{T}(\partial_t, K_0) + \int_{\mathcal{N}^-_{[t_1,t_2]}(p)} \mathbf{T}(L, K_0) = \int_{B_{t_1}} \mathbf{T}(\partial_t, K_0) \tag{61}$$

In (60) the term $\mathbf{T}(\partial_t, \partial_t)$ is called *energy density* while $\mathbf{T}(E_-, \partial_t)$ is called *energy flux density* . The corresponding integrals are called *energy contained in B_{t_1}, and* B_{t_2} and, respectively, *flux of energy through $\mathcal{N}^-$.* The coresponding terms in (61) are called conformal energy densities, fluxes etc.

Equation (60) can be used to derive the following fundamental properties of relativistic field theories.

(1) Finite propagation speed
(2) Uniqueness of the Cauchy problem

Proof : The first property follows from the fact that, if $\int_{B_{t_1}} \mathbf{T}(\partial_t, \partial_t)$ is zero at time $t = t_1$ then both integrals $\int_{B_{t_2}} \mathbf{T}(\partial_t, \partial_t)$ and $\int_{\mathcal{N}^-_{[t_1,t_2]}} \mathbf{T}(E_-, \partial_t)$ must vanish also. In view of the positivity properties of the $\mathbf{T}$ it follows that the corresponding integrands must also vanish. Taking into account the specific form of $\mathbf{T}$, in a particular theory, one can then show that the fields do also vanish in the domain of influence of the ball B_{t_1}. Conversely, if the initial data for the fields vanish in the complement of B_{t_1}, the the fields are identically zero in the complement of the domain of influence of of B_{t_1}.

The proof of the second property follows immediately from the first for a linear field theory. For a nonlinear theory one has to work a little more. ∎

Exercise. Formulate an initial value problem for each of the field theories we have encountered so far, scalar wave equation (SWE), Wave Maps (WM), Maxwell equations (ME). Proof uniqueness of solutions to the initial value problem, for smooth solutions.

The following is another important consequence of (60) and (61). To state the results we introduce the following quantities,

$$\mathcal{E}(t) = \int_{\mathbb{R}^n} \mathbf{T}\left(\partial_t, \partial_t\right)(t,x)dx \tag{62}$$

$$\mathcal{E}_c(t) = \int_{\mathbb{R}^n} \mathbf{T}\left(K_0, \partial_t\right)(t,x)dx \tag{63}$$

THEOREM 3.9 (Global Energy). *For an arbitrary field theory, if $\mathcal{E}(0) < \infty$, then*

$$\mathcal{E}(t) = \mathcal{E}(0) \tag{64}$$

Moreover, for a conformal field theory, if $\mathcal{E}_c(0) < \infty$,

$$\mathcal{E}_c(t) = \mathcal{E}_c(0) \tag{65}$$

Proof : Follows easily by applying (60) and (61) to past causal domains $\mathcal{J}^-(p)$ with $p = (\bar{t}, 0)$ between $t_1 = 0$ and $t_2 = t$ and letting $\bar{t} \to +\infty$. ∎

3.10. Energy dissipation. In this section we shall make use of the global conformal energy identity (65) to show how energy dissipates for a filed theories in Minkowski space. Consider a conformal field theory defined on all of $\mathbb{R}^{n+1}$. At each point of $\mathbb{R}^{n+1}$, with $t \geq 0$, define the standard null frame where

$$\begin{aligned} L = E_+ &= \partial_t + \partial_r \\ \underline{L} = E_- &= \partial_t - \partial_r. \end{aligned}$$

Observe that the conformal Killing vectorfield $K_0 = (t^2 + r^2)\partial_t + 2rt\partial_r$ can be expressed in the form,

$$K_0 = \frac{1}{2}\left[(t+r)^2 E_+ + (t-r)^2 E_-.\right]$$

Thus,

$$
\begin{aligned}
\mathcal{E}_c(t) &= \int_{\mathbb{R}^n} \frac{1}{4}(t+r)^2\mathbf{T}_{++} + \frac{1}{4}(t-r)^2\mathbf{T}_{--} + \underbrace{\left((t+r)^2+(t-r)^2\right)}_{2(t^2+r^2)}\mathbf{T}_{+-}dx. \\
&= \int_{\mathbb{R}^n} \frac{1}{4}(t+r)^2\mathbf{T}_{++} + \frac{1}{2}(t^2+r^2)\mathbf{T}_{+-} + \frac{1}{4}(t-r)^2\mathbf{T}_{--}dx \qquad (66) \\
\mathcal{E}_c(0) &= \int_{\mathbb{R}^n} \mathbf{T}(\partial_t, K_0)(0,x)dx = \int_{\mathbb{R}^n} |x|^2\mathbf{T}(\partial_t,\partial_t)dx
\end{aligned}
$$

According to (65) we have $\mathcal{E}_c(t) = \mathcal{E}_c(0)$. Assuming that $\mathcal{E}_c(0) = \int_{\mathbb{R}^n}|x|^2\mathbf{T}(\partial_t,\partial_t)dx$ is finite we conclude that,

$$
\begin{aligned}
\int_{\mathbb{R}^n} \mathbf{T}_{++}(t,\cdot)dx &\lesssim t^{-2}\mathcal{E}_c(0) \\
\int_{\mathbb{R}^n} \mathbf{T}_{+-}(t,\cdot)dx &\lesssim t^{-2}\mathcal{E}_c(0).
\end{aligned}
$$

The remaining term in (66) contains the factor $(t-r)^2$ which is constant along outgoing null directions $r = t + c$. Hence for any $0 < \epsilon < 1$

$$
\begin{aligned}
\int_{|x|>(1+\epsilon)t} \mathbf{T}_{--} &= O(t^{-2}) \\
\int_{|x|<(1-\epsilon)t} \mathbf{T}_{--} &= O(t^{-2}).
\end{aligned}
$$

We conclude that most of the energy of a conformal field is carried by the $\mathbf{T}_{--}$ component and propagates near the light cone.

CHAPTER 2

ON THE PHYSICAL CONTENT OF GENERAL RELATIVITY

1. Equivalence Principle and Derivation of the Field Equations

Einstein's point of departure was the well known experimental fact concerning the *universality of free fall* or the *weak equivalence principle.* In Newton's theory of gravitation the universality of free fall appears implicitly in the identification,

$$m_i = m_g \tag{67}$$

between the inertial mass m_i of a particle, i.e. the one which appears in $F = m_i a$ law, and the gravitational mass m_g which appears in the gravitational force between two particles separated by a distance r, $F_g = G m_g M_g r^{-2} = -m_g \nabla\Phi$ where Φ is the gravitational potential of the particle of mass $M = M_g$,

$$\Phi = -G\frac{M}{r}. \tag{68}$$

An immediate consequence of (67) is the universality of free falling test particles in a fixed gravitational field, i.e. the acceleration a of any such particle is given by,

$$a = -\nabla\Phi \tag{69}$$

Very early on in his quest for a relativistic theory of gravity Einstein realized that this mysterious equality (67) contains a deeper equivalence between inertia and gravitation. In his famous thought experiment with a freely falling elevator he observes, as consequence of the universality of free fall, that all rigid objects in the elevator appear as being at rest with respect to the freely falling reference frame attached to the elevator. Thus, relative to such a frame, the external gravitational field appears to be erased. Einstein's principle of equivalence extrapolates this impossibility of distinguishing between uniform acceleration and an external gravitational field to all physical experiments, not just free falling particles.

Thus, his *principle of equivalence* (EEP) postulates as follows.

- Any gravitational field can be locally[1] erased using an appropriate freely falling, local, reference frame. The non-gravitational physical laws (such as electromagnetism) apply in this local reference frame in the same way as they would in an inertial frame (free of gravity) in special relativity.

[1] In fact infinitesimally

- Starting with an inertial reference frame in special relativity one can create an apparent gravitational field in a local reference frame which is accelerated with respect to the first.

The passage from an inertial frame to an accelerated one is, mathematically, a change of coordinates $x^\alpha = x^\alpha(y^\beta)$ from the inertial coordinates x^α to general coordinates y^α. In the new coordinate system the metric takes the form,

$$\mathbf{g}_{\mu\nu}(y^\lambda)dy^\mu dy^\nu, \qquad \mathbf{g}_{\mu\nu} = \mathbf{m}_{\alpha\beta}\frac{\partial x^\alpha}{\partial y^\mu}\frac{\partial x^\beta}{\partial y^\nu} \tag{70}$$

This has led Einstein to consider general Lorentzian metrics. The time-like geodesics in such space-times, corresponding to the trajectory of freely falling particles,

$$\frac{d^2x^\lambda}{ds^2} + \Gamma^\lambda_{\mu\nu}\frac{dx^\mu}{ds}\frac{dx^\nu}{ds} = 0 \tag{71}$$

where

$$\Gamma^\lambda_{\mu\nu} = \frac{1}{2}\mathbf{g}^{\lambda\sigma}(\partial_\mu \mathbf{g}_{\nu\sigma} + \partial_\nu \mathbf{g}_{\mu\sigma} - \partial_\sigma \mathbf{g}_{\mu\nu})$$

Interpreting the term $\Gamma^\lambda_{\mu\nu}\frac{dx^\mu}{ds}\frac{dx^\nu}{ds}$ as corresponding to a gravitational force acting on the particle, we see that the equivalence principle here corresponds to the mathematical possibility to choose a coordinate system, in a neighborhood of a point q along the curve, such that $\Gamma(q) = 0$. In other words, we have erased the gravitational force at q, by simply making a local change of reference.

Exercise. Prove this fact. Show in fact that a coordinate system can be chosen, in a neighborhood of q along the curve, so that Γ vanishes along the curve.

Once he decided that the metric $\mathbf{g}$ describes both the geometry and spacetime and gravitation Einstein had to find which equation it satisfies. This equations has to supersede the Poisson equation for the Newton potential,

$$\Delta\Phi = 4\pi G\rho \tag{72}$$

where ρ is the mass density. The explicit form of Φ in (68) corresponds to a pointlike mass distribution. A relativistic generalization of (72) should take the form of a tensor equation. One can guess that the tensor generalization of mass density should be the energy-momentum tensor T. The following three principles have led him to

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu} \tag{73}$$

where G is Newton's gravitational constant and c the speed of light.

(1) *Principle of General Relativity* The fundamental Laws of Physics should be covariant. This is a principle of indifference, i.e. physical phenomenon do not take place in the same way (in general) in different coordinates systems (as is the case with inertial coordinates in special relativity) but no coordinate system (in general) has a privileged status relative to others.
(2) In the presence of matter the source of the gravitational field has to be the energy-momentum tensor.

(3) *Principle of correspondence* In the limit where one neglects gravitational effects, i.e. $\mathbf{g} = \mathbf{m}$, the laws of Physics should be those of special relativity. Moreover there should be a special limit with regard to which the equations being sought reduce to Newton's theory of gravity.

These three principles has led Einstein to his field equations, One can show (see [**1**]) that the three principles described above uniquely determine the Einstein field equations.

1.1. Geodesic deviation and tidal forces. Let $\gamma_s(t)$ be a smooth one - parameter family of time-like geodesics $\gamma_s : [0,1] \to \mathbb{R}$ parametrized by their arclength t. We assume that the map $(t,s) \in [0,1] \times (-\epsilon, \epsilon)$
$to(\partial_t \gamma, \partial_s \gamma)$ is smooth and has a smooth inverse so that the curves $\gamma_s(t)$ span a two dimensional surface on the original Lorentzian manifold $\mathbf{M}$. We can regard t, s as local coordinates. Let $T = \partial_t$ $X = \partial_s$ the corresponding coordinate vectorfields, i.e. for any function f on $\mathbf{M}$,

$$T(f) = \partial_t f(\gamma_s(t)), \qquad X(f) = \partial_s f(\gamma_s(t))$$

Since the curves γ_s are geodesic we have,

$$D_T T = 0.$$

We can also normalize T such that $\mathbf{g}(T,T) = -1$. We can also choose the " original" s-curve, $s \to \gamma_s(0)$, such that its tangent X is perpendicular to T. Then, since T, X commute we find,

$$T\mathbf{g}(T,X) = \mathbf{g}(T, \mathbf{D}_T X) = \mathbf{g}(T, \mathbf{D}_X T) = \frac{1}{2} X \mathbf{g}(\mathbf{T}, \mathbf{T}) = 0$$

Hence X remains orthogonal to T for all values of $t \in [0,1]$. Consider $V = \mathbf{D}_T X$, the rate of change (along a given geodesic) of the *displacement vector* X. We may interpret V as the relative velocity of an infinitesimally nearby geodesic. We now calculate the relative acceleration of an infinitesimally nearby geodesic. Since X, T commute and $D_T T = 0$ we deduce,

$$\begin{aligned} \mathbf{D}_T V &= \mathbf{D}_T \mathbf{D}_T X = \mathbf{D}_T(\mathbf{D}_X T) = \mathbf{D}_X(\mathbf{D}_T T) + [\mathbf{D}_T, \mathbf{D}_X] T \\ &= [\mathbf{D}_T, \mathbf{D}_X] T = \mathbf{R}(T,X)T \end{aligned}$$

Thus we deduce the geodesic deviation formula,

$$\mathbf{D}_T^2 X = \mathbf{R}(T,X)T \tag{74}$$

Thus, geodesics which start are parallel at $t = 0$, i.e. $V = \mathbf{D}_T X = 0$, may fail to remain so because of the non-vanishing of the curvature term R.

2. Weak Field limit

We say that the gravitational field is weak if we can decompose the metric $\mathbf{g}$ into the flat metric $\mathbf{m}$ plus a small perturbation h,

$$\mathbf{g}_{\mu\nu} = \mathbf{m}_{\mu\nu} + h_{\mu\nu}, \qquad |h_{\mu\nu}| << 1, \tag{75}$$

in an inertial system of coordinates with respect to $\mathbf{m}$. Clearly,

$$\mathbf{g}^{\mu\nu} = \mathbf{m}^{\mu\nu} - h^{\mu\nu}$$

where the indices of h are risen with respect to the minkowski metric $\mathbf{m}$. Recall that,

$$R_{\mu\nu\rho}{}^{\sigma} = \partial_\nu \Gamma^\sigma_{\mu\rho} - \partial_\mu \Gamma^\sigma_{\nu\rho} + \Gamma^\alpha_{\mu\rho}\Gamma^\sigma_{\alpha\nu} - \Gamma^\alpha_{\nu\rho}\Gamma^\sigma_{\alpha\mu}$$

Now, neglecting quadratic terms in h,

$$\Gamma^\rho_{\mu\nu} = \frac{1}{2}\mathbf{m}^{\rho\lambda}(\partial_\mu h_{\lambda\nu} + \partial_\nu h_{\mu\lambda} - \partial_\lambda h_{\mu\nu}) + O(h^2)$$

We deduce,

$$R_{\mu\nu\rho}{}^{\sigma} = \frac{1}{2}\big(\partial_\rho\partial_\nu h_{\mu\sigma} + \partial_\sigma\partial_\mu h_{\nu\rho} - \partial_\sigma\partial_\nu h_{\mu\rho} - \partial_\rho\partial_\mu h_{\nu\sigma}\big) + O(h^2 + |\partial h|^2).$$

and, with $h = \mathbf{m}^{\mu\nu} h_{\mu\nu}$,

$$\begin{aligned} R_{\mu\nu} &= \frac{1}{2}\big(\partial_\sigma\partial_\nu h^\sigma{}_\mu + \partial_\sigma\partial_\mu h^\sigma{}_\rho - \partial_\mu\partial_\nu h - \Box h_{\mu\nu}\big) + O(h^2 + |\partial h|^2). \\ R &= \partial_\mu\partial_\nu h^{\mu\nu} - \Box h \end{aligned}$$

We derive the Einstein tensor,

$$\begin{aligned} \mathbf{E}_{\mu\nu} &= \frac{1}{2}\big(\partial_\sigma\partial_\nu h^\sigma{}_\mu + \partial_\sigma\partial_\mu h^\sigma{}_\rho - \partial_\mu\partial_\nu h - \Box h_{\mu\nu} - \mathbf{m}_{\mu\nu}\partial_\alpha\partial_\beta h^{\alpha\beta} + \mathbf{m}_{\mu\nu}\Box h\big) \\ &+ O(h^2 + |\partial h|^2) \end{aligned}$$

The expression can be simplified if we introduce the trace reversed quantity,

$$\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2} m_{\mu\nu} h$$

Then,

$$E_{\mu\nu} = -\frac{1}{2}\Box\bar{h}_{\mu\nu} + \partial_\sigma\partial_{(\nu}\bar{h}^\sigma{}_{\mu)} - \frac{1}{2}m_{\mu\nu}\partial_\alpha\partial_\beta\bar{h}^{\alpha\beta} + O(h^2 + |\partial h|^2)$$

where,

$$\partial_\sigma\partial_{(\nu}\bar{h}^\sigma{}_{\mu)} = \frac{1}{2}\big(\partial_\sigma\partial_{(\nu}\bar{h}^\sigma{}_{\mu)} + \partial_\sigma\partial_\mu\bar{h}^\sigma{}_\nu\big)$$

Thus the linearized Einstein equations take the form,

$$-\frac{1}{2}\Box\bar{h}_{\mu\nu} + \partial_\sigma\partial_{(\nu}\bar{h}^\sigma{}_{\mu)} - \frac{1}{2}m_{\mu\nu}\partial_\alpha\partial_\beta\bar{h}^{\alpha\beta} = 8\pi T_{\mu\nu} \tag{76}$$

We now come to the issue of gauge invariance. Consider another copy of the same manifold $\mathbf{M}$, which we denote $\mathbf{M}_b$, endowed with the flat metric $\mathbf{m}$. We have a diffeomorphism between them which we denote

$$\Phi : \mathbf{M}_b \to \mathbf{M}$$

Note that Φ is the identity map if we consider the same coordinate system on $\mathbf{M}_b$ and $\mathbf{M}$, but we allow the possibility of considering different coordinate systems. We define the perturbation $\mathbf{h}$ by,

$$\mathbf{h} = \Phi^*\mathbf{g} - \mathbf{m} \tag{77}$$

We say that the gravitational field on $\mathbf{M}$ is weak if there exists a diffeomorphism Φ such that $|h_{\mu\nu}| << 1$ in a fixed coordinate system in $\mathbf{M}_b$. Of course, we can have many such coordinate systems and therefore we have to recognize how to pass from one *admissible* system to another. Consider an arbitrary vectorfield X on $\mathbf{M}_b$ and its induced flow $\Psi_t : \mathbf{M}_b \to \mathbf{M}_b$. We can now define the family of perturbations,

$$\begin{aligned}\mathbf{h}_t : &= (\Phi \circ \Psi_t) * \mathbf{g} - \mathbf{m} = \psi_t^* \Phi^* \mathbf{g} - \mathbf{m} = \Psi_t^* \big(\Phi^* \mathbf{g} - \mathbf{m}\big) + \Psi_t^* \mathbf{m} - \mathbf{m} \\ &= \Psi_t^* \mathbf{h} + \Psi_t^* \mathbf{m} - \mathbf{m}\end{aligned}$$

Writing $\Psi_t^* \mathbf{h} = \mathbf{h} + O(|t|)\mathbf{h}$ and,

$$\Psi_t^* \mathbf{m} - \mathbf{m} = -t \frac{\Psi_t^* \mathbf{m} - \mathbf{m}}{-t} = -t \mathcal{L}_X \mathbf{m} + O(|t|^2)$$

where, recalling the definition of Lie derivative,

$$\mathcal{L}_X \mathbf{m} = \lim_{t \to 0} \frac{\Psi_{-t} \mathbf{m} - m}{t}.$$

Hence,

$$\mathbf{h}_t \quad = \quad \mathbf{h} - t \mathcal{L}_X \mathbf{m} + O(|t|^2 + |t| \mathbf{h})$$

or, for $|t| \leq \epsilon$,

$$\mathbf{h}_t \quad = \quad \mathbf{h} - \epsilon \mathcal{L}_X \mathbf{m} + O(\epsilon^2).$$

We denote $\xi = -\epsilon X$. Thus, for small ξ, $\mathbf{h}$ and $\mathbf{h} + \mathcal{L}_\xi \mathbf{m}$ describe the same physical perturbation. This means that linear gravity has a gauge freedom given by,

$$h'_{\mu\nu} = h_{\mu\nu} + \partial_\mu \xi_\nu + \partial_\nu \xi_\mu \tag{78}$$

This is similar to the gauge freedom of the electromagnetic field, $A'_\mu = A_\mu + \partial_\mu \chi$.

Using (78) we can now simplify the linearized Einstein equations (76). Indeed we try to choose the vectorfield ξ such that,

$$\partial^\mu \bar{h}'_{\mu\nu} = 0$$

Now,

$$\bar{h}'_{\mu\nu} = \bar{h}_{\mu\nu} + \partial_\mu \xi_\nu + \partial_\nu \xi_\mu - m_{\mu\nu} \partial^\lambda \xi_\lambda$$

Thus,

$$\partial^\mu \bar{h}'_{\mu\nu} = \partial^\mu \bar{h}_{\mu\nu} + \Box \xi_\nu.$$

Hence if,

$$\Box \xi_\nu = -\partial^\mu \bar{h}_{\mu\nu} \tag{79}$$

we have,

$$\partial^\mu \bar{h}'_{\mu\nu} = 0. \tag{80}$$

In such a gauge the Einstein equations take the form,

$$\Box \bar{h}_{\mu\nu} \quad = \quad -16\pi T_{\mu\nu} \tag{81}$$

2.1. Newtonian limit. The newtonian limit is derived from the Einstein field equations, with $\mathbf{T}$ the energy-momentum tensor of a perfect fluid (see (51)) , under the following assumptions

(1) Gravity is weak, i.e. $\mathbf{g} = \mathbf{m} + \mathbf{h}$ with $\mathbf{h}$ verifying (81).
(2) The relative motion of the sources is much slower than the speed of light, i.e. $u^0 \approx 1$, $|v| << 1$ with $v^i = u^i$, and $T_{ij} << T_{0i} << T_{00}$.
(3) The space-time geometry is slowly changing, i.e. $\partial_t h_{\mu\nu}$ are small.

The assumption about sources can be reformulated as follows. There exists an inertial coordinate system such that,

$$T_{00} = \rho$$

and all other components of T are negligible. Since $\partial_0^2 \bar{h}_{\mu\nu}$ are negligible we derive from (81),

$$\begin{aligned} \Delta\bar{h}_{\mu\nu} &= 0, \quad (\mu,\nu) \neq (0,0) \\ \Delta\bar{h}_{00} &= -16\pi\rho \end{aligned}$$

Assuming that $\bar{h}_{\mu\nu}$ are well behaved at infinity we derive $\bar{h}_{0i} = \bar{h}_{ij} = 0$. Let,

$$\phi = -\frac{1}{4}\bar{h}_{00} \tag{82}$$

Thus, since,

$$h_{\mu\nu} = \bar{h}\mu\nu - \frac{1}{2}m_{\mu\nu}\bar{h}$$

$$h_{00} = -2\phi, \quad h_{ii} = 2\phi, \quad h_{0i} = 0, \qquad \Delta\phi = 4\pi\rho \tag{83}$$

This leads to the metric,

$$-(1+2\Phi)dt^2 + (1-2\Phi)\big((dx^1)^2 + (dx^2)^2 + (dx^3)^2\big) \tag{84}$$

The motion of *freely falling* test particles in this geometry is governed by the geodesic equation,

$$\frac{d^2x^\lambda}{d\tau^2} + \Gamma^\lambda_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} = 0$$

with τ proper time. For motion much slower than the speed of light we can approximate $\frac{dx^\mu}{d\tau}$ by the vector $(1,0,0,0,)$ and proper time τ by the t coordinate t. Thus,

$$\frac{d^2x^i}{dt^2} = -\Gamma^{\imath}_{00} = \frac{1}{2}\frac{\partial h_{00}}{\partial x^{\imath}} = -\partial_i\Phi$$

i.e.,

$$\frac{d^2x^i}{dt^2} = -\nabla_i\Phi \tag{85}$$

as expected.

2.2. Gravitational radiation. In linear approximation the propagation of gravitational radiation is governed by the source- free, linearized Einstein equations,

$$\partial^\mu \bar{h}_{\mu\nu} = 0, \quad \Box \bar{h}_{\mu\nu} = 0. \tag{86}$$

In deriving these equations we have used the gauge transformations (78). The remaining gauge freedom is given by transformations,

$$\Box \xi_\mu = 0. \tag{87}$$

One can use this additional freedom to obtain, in the so called *radiation gauge*,

$$h = 0, \qquad h_{0i} = 0, \qquad h_{00} = 0. \tag{88}$$

We first show how to arrange $h = 0$ by performing a gauge transformation (78) with $\xi_1 = \xi_2 = \xi_3 = 0$, i.e.

$$\begin{aligned} h' &= \mathbf{m}^{\mu\nu} h'_{\mu\nu} = \mathbf{m}^{\mu\nu}(h_{\mu\nu} + \partial_\mu \xi_\nu + \partial_\nu \xi_\mu) \\ &= h - 2\partial_t \xi_0 \end{aligned}$$

where $\Box \xi_0 = 0$. Let $f := h - 2\partial_t \xi_0$. Taking the trace of the equation

$$\Box \bar{h}_{\mu\nu} = \Box (h_{\mu\nu} - \frac{1}{2}\mathbf{m}_{\mu\nu} h) = 0$$

with respect to $\mathbf{m}$ we see that $\Box h = 0$. Thus,

$$\Box f = \Box h - 2\partial_t \Box \xi_0 = 0.$$

To show that $f = 0$ we only need to arrange the initial conditions for ξ_0 at $t = t_0$ to be such that,

$$f = 0, \quad \partial_t f = 0 \qquad \text{at } t = t_0.$$

This leads to the following choice for the initial data for ξ_0,

$$\partial_t \xi_0 = \frac{1}{2} h, \qquad \partial_t^2 \xi_0 = \frac{1}{2}\partial_t h$$

or, equivalently,

$$\Delta \xi_0 = \frac{1}{2}\partial_t h, \qquad \partial_t \xi_0 = \frac{1}{2} h. \tag{89}$$

Solving the initial value problem (89) for $\Box \xi_0$ we find in this manner a gauge transformation given by the covector $(\xi_0, 0, 0, 0)$ such that in addition to (86) we also have, $h = 0$. More precisely, there exists a gauge transformation such that,

$$\nabla^\mu h_{\mu\nu} = 0, \quad \Box h_{\mu\nu} = 0, \quad h = 0.$$

We can use the remaining degree of freedom, i.e. by solving $\Box \xi_i = 0$, $i = 1, 2, 3$, in the same manner to also obtain $h_{0i} = 0$. As a bonus we can now also show that $h_{00} = 0$. Indeed, using $\nabla^\mu h_{\mu\nu} = 0$ we must have $\partial_t h_{00} = 0$. Since also $\Box h_{00} = 0$, we infer that $h_0 0$ is a constant. By another trivial change of gauge we deduce $h_{00} = 0$. We have thus proved the following:

PROPOSITION 2.3. *In the weak field limit, we can find a gauge transformation such that the components $h_{0\mu}$, $\mu = 0, 1, 2, 3$ and the trace h vanish. Thus the linearized vacuum equations take the form,*

$$\nabla^j h_{ij} = 0, \quad \Box h_{ij} = 0, \quad \delta^{ij} h_{ij} = 0. \tag{90}$$

Observe that there only six independent components of h_{ij} verifying four differential equations. Consider now plane wave solutions to (90),

$$h_{\alpha\beta} \quad = \quad C_{\alpha\beta} e^{ik_\mu x^\mu} \tag{91}$$

with constants $C_{\alpha\beta}$ and k_μ. To verify (90) we need $C_{0\beta} = 0$ and

$$\begin{aligned} k^j C_{ij} &= 0, \quad \delta^{ij} C_{ij} = 0 \\ k_\mu k^\mu &= 0. \end{aligned}$$

Choosing $k_\mu = \omega(1, 0, 01)$ we find that the only non-vanishing components of C are $A = C_{11} = -C_{22}$ and $B = C_{12} = C_{21}$. Thus a plane wave traveling in the x^3 direction is completely characterized by A, B and ω. The solutions corresponding to $A \neq 0, B = 0$ and $A = 0, B \neq 0$ describe the two independent polarization states of plane gravitational wave.

To detect a gravitational wave one one has to study the relative acceleration of two point masses due to it. For two nearby freely falling masses, this acceleration is given by the geodesic deviation formula (74). For two bodies nearly at rest in a global inertial system x^α we have, with $T \approx \partial_t$, $\mathbf{D} \approx \partial$ the flat covariant derivative operator and X^μ the deviation vector,

$$\frac{d^2 X^\mu}{dt^2} \approx R_{\nu 00}{}^{\mu} X^\nu \tag{92}$$

In the radiation gauge,

$$R_{\nu 00\mu} \approx \frac{1}{2} \frac{\partial^2 h_{\mu\nu}}{\partial t^2} \tag{93}$$

APPENDIX A

BASIC GEOMETRIC NOTIONS

We briefly review the following topics below:

1.) Lie brackets of vectorfields. Frobenius theorem

2.) Lie derivative of a tensorfield

3.) Multilinear forms and exterior differentiation

4.) Connections and covariant derivatives

5.) Pseudo-riemannian metrics. Riemannian and Lorentzian geometry.

6.) Levi-Civita connection associated to a pseudo-riemannian metric.

7.) Parallel transport, geodesics, exponential map, completeness

8.) Curvature tensor of a pseudo-riemannian manifold. Symmetries. First and second Bianchi identities.

9.) Isometries and conformal isometries. Killing and conformal Killing vectorfields.

1. Pseudo-riemannian, Lorentzian metrics

A pseudo-riemannian manifold [1], or simply a spacetime, consist of a pair $(\mathbf{M}, \mathbf{g})$ where $\mathbf{M}$ is an orientable $p+q$-dimensional manifold and $\mathbf{g}$ is a pseudo-riemannian metric defined on it, that is a smooth, a non degenerate, 2-covariant symmetric tensor field of signature (p, q). This means that at each point $p \in \mathbf{M}$ one can choose a basis of $p+q$ vectors, $\{e_{(\alpha)}\}$, belonging to the tangent space $T\mathbf{M}_p$, such that

$$\mathbf{g}(e_{(\alpha)}, e_{(\beta)}) = \mathbf{m}_{\alpha\beta} \tag{94}$$

[1] We assume that our reader is already familiar with the basics concepts of differential geometry such as manifolds, tensor fields, covariant, Lie and exterior differentiation. For a short introduction to these concepts see Chapter 2 of Hawking and Ellis, "The large scale structure of space-time", [?]

for all $\alpha, \beta = 0, 1, ..., n$, where $\mathbf{m}$ is the diagonal matrix with -1 in the first p entries and $+1$ in the last q entries. If X is an arbitrary vector at p expressed, in terms of the basis $\{e_{(\alpha)}\}$, as $X = X^\alpha e_{(\alpha)}$, we have

$$\mathbf{g}(X, X) = -(X^1)^2 - \ldots - (X^p)^2 + (X^{p+1})^2 + ... + (X^{p+q})^2 \tag{95}$$

The case when $p = 0$ and $q = n$ corresponds to Riemannian manifolds of dimension n. The other case of interest for us is $p = 1, q = n$ which corresponds to a Lorentzian manifolds of dimension $n+1$. The primary example of Riemannian manifold is the Euclidean space $\mathbb{R}^n$. Any other Riemannian manifold looks, locally, like $\mathbb{R}^n$. Similarly, the primary example of a Lorentzian manifold is the Minkowski spacetime, the spacetime of Special Relativity. It plays the same role, in Lorentzian geometry, as the Euclidean space in Riemannian geometry. In this case the manifold $\mathbf{M}$ is diffeomorphic to $\mathbb{R}^{n+1}$ and there exists globally defined systems of coordinates, x^α, relative to which the metric takes the diagonal form $-1, 1, ..., 1$. All such systems are related through Lorentz transformations and are called inertial. We denote the Minkowski spacetime of dimension $n+1$ by $(\mathbb{R}^{n+1}, \mathbf{m})$.

Relative to a given coordinate system x^μ, the components of a pseudo-riemannian metric take the form

$$g_{\mu\nu} = \mathbf{g}(\partial_\mu, \partial_\nu)$$

where $\partial_\mu = \frac{\partial}{\partial x^\mu}$ are the associated coordinate vectorfields. We denote by $g^{\mu\nu}$ the components of the inverse metric g^{-1} relative to the same coordinates x, and by $|g|$ the determinant of the matrix $g_{\mu\nu}$. The volume element $dv_{\mathbf{M}}$ of $\mathbf{M}$ is expressed, in local coordinates, by $\sqrt{|g|}dx = \sqrt{|g|}dx^1 \ldots dx^n$. Thus the integral $\int_{\mathbf{M}} f dv_{\mathbf{M}}$ of a function f, supported in coordinate chart $U \subset \mathbf{M}$ is defined by $\int_U f(x)\sqrt{|g(x)|}dx$. The integral on $\mathbf{M}$ of an arbitrary function f is defined by making a partition of unity subordinated to a covering of $\mathbf{M}$ by coordinate charts. One can easily check that the definition is independent of the particular system of local coordinates.

In view of (95) we see that a Lorentzian metric divides the vectors in the tangent space $T\mathbf{M}_p$ at each p, into timelike, null or spacelike according to whether the quadratic form

$$\mathbf{g}(X, X) = g_{\mu\nu} X^\mu X^\nu \tag{96}$$

is, respectively, negative, zero or positive. We defined the magnitude of a vector to be $|X| = \sqrt{\mathbf{g}(X, X)}$, if X is spacelike, $|X| = \sqrt{-\mathbf{g}(X, X)}$ if X is timelike and $|X| = 0$ if X is null. Observe that the Cauchy-Schwartz inequality for timelike vectors takes the form,

$$\mathbf{g}(X, Y)^2 \geq |X|^2 |Y|^2.$$

The set of null vectors N_p forms a double cone, called the null cone of the corresponding point p. The set of timelike vectors I_p forms the interior of this cone. The vectors in the union of I_p and N_p are called causal. The set S_p of spacelike vectors is the complement of $I_p \cup N_p$. The *causal structure* of a lorentzian manifold $\mathbf{M}$ is given by specifying the null cones $N_p \subset T_p(\mathbf{M})$

PROPOSITION 1.1. *Two Lorentz metrics $\mathbf{g}_1, \mathbf{g}_2$ have the same causal structure if and only if they differ by a proportionality factor.*

Proof It suffices to show that if their causal structures are the same then $\mathbf{g}_2 = \Lambda \mathbf{g}_1$ for some non-vanishing scalar function Λ. Let X be two vectors, one spacelike the other timelike. Since the roots of the quadratic forms in $\lambda \in \mathbb{R}$, $\mathbf{g}_1(X + \lambda Y)$ and $\mathbf{g}_2(X + \lambda Y)$ coincide we deduce that the corresponding coefficients must be proportional,

$$\frac{\mathbf{g}_1(X,X)}{\mathbf{g}_1(Y,Y)} = \frac{\mathbf{g}_2(X,X)}{\mathbf{g}_2(Y,Y)}$$

Setting,

$$\Lambda = \frac{\mathbf{g}_2(X,X)}{\mathbf{g}_1(X,X)} = \frac{\mathbf{g}_2(Y,Y)}{\mathbf{g}_1(Y,Y)}$$

one can easily show that $\Lambda = \Lambda(p)$ does not depend on the particular vectors X, Y. By a simple polarization formula it then follows that, for any two non-null vectors X, Y,

$$\mathbf{g}_2(X,Y) = \Lambda \mathbf{g}_1(X,Y)$$

∎

A frame $e_{(\alpha)}$ verifying (94) is said to be orthonormal. In the case of Lorentzian manifolds it makes sense to consider, in addition to orthonormal frames, null frames. These are collections of vectorfields[2] e_α consisting of two null vectors e_{n+1}, e_n and orthonormal spacelike vectors $(e_a)_{a=1,\dots,n-1}$ which verify,

$$\begin{aligned}
\mathbf{g}(e_n, e_n) &= \mathbf{g}(e_{n+1}, e_{n+1}) = 0 \ , \ \mathbf{g}(e_n, e_{n+1}) = -2 \\
\mathbf{g}(e_n, e_a) &= \mathbf{g}(e_{n+1}, e_a) = 0 \ , \ \mathbf{g}(e_a, e_b) = \delta_{ab}
\end{aligned}$$

One-forms $A = A_\alpha dx^\alpha$ are sections of the cotangent bundle of $\mathbf{M}$. We denote by $A(X)$ the natural pairing between A and a vectorfield X. We can raise the indices of A by $A^\alpha = \mathbf{g}^{\alpha\beta}\mathbf{A}_\beta$. $A' = A^\alpha \partial_\alpha$ defines a vectorfield on $\mathbf{M}$ and we have, $A(X) = \mathbf{g}(A', X)$. Covariant tensors A of order k are k-multilinear forms on $T\mathbf{M}$.

Given a submanifold $\mathbf{N} \subset \mathbf{M}$, the restriction of $\mathbf{g}_p$ to $\mathbf{T}_p(\mathbf{N})$ defines the induced metric $\mathbf{h}$, i.e.

$$\mathbf{h}_p(X,Y) = \mathbf{g}_p(X,Y), \qquad \forall X, Y \in T_p(\mathbf{N})$$

The submanifold is said to be spacelike if its induced metric $\mathbf{h}$ is Riemannian, timelike if $\mathbf{h}$ is Lorentzian and null if $\mathbf{h}$ is degenerate, at every point $p \in \mathbf{N}$. Of particular interest are submanifolds of codimension 1, called hypersurfaces, given (locally) by the non-critical level sets of a function $f : \mathbf{M} \to \mathbb{R}$, with $df \neq 0$. We define the gradient of f to be the vector obtained by raising the indices of the 1-form df, $N_f = -g^{\mu\nu}\partial_\mu f \partial_\nu$. A level hypersurface $\mathcal{H}_f = \{f = c\}$ is spacelike, timelike or null if N_f is respectively timelike, spacelike or null. Clearly, in all cases, N_f is orthogonal to $\mathcal{H}_f$. Observe that, for any vectorfield X,

$$\mathbf{g}(\mathbf{N}_f, X) = X(f).$$

[2]We write e_α instead of $e_{(\alpha)}$ to simplify the notation, whenever there can be no confusion.

Thus X is tangent to $\mathcal{H}_f$ if and only if $\mathbf{g}(N_f, X) = 0$. In the particular case of a null hypersurface case N_f is both orthogonal and tangent to $\mathcal{H}_f$. Also,

$$0 = \mathbf{g}(N_f, N_f) = g^{\mu\nu}\partial_\mu f \partial_\nu f \tag{97}$$

i.e. f verifies the *eikonal equation* in $\mathbf{M}$.

Notation: We will use the following notational conventions: We shall use boldface characters to denote important tensors such as the metric $\mathbf{g}$, and the Riemann curvature tensor $\mathbf{R}$. Their components relative to arbitrary frames will also be denoted by boldface characters. Thus, given a frame $\{e_{(\alpha)}\}$ we write $\mathbf{g}_{\alpha\beta} = \mathbf{g}(e_\alpha, e_\beta)$, $\mathbf{R}_{\alpha\beta\gamma\delta} = \mathbf{R}(e_\alpha, e_\beta, e_\gamma, e_\delta)$ and, for an arbitrary tensor T,

$$T_{\alpha\beta\gamma\delta\ldots} \equiv T(e_\alpha, e_\beta, e_\gamma, e_\delta, \ldots)$$

We shall not use boldface characters for the components of tensors, relative to a fixed system of coordinates. Thus, for instance, in (96) $g_{\mu\nu} = \mathbf{g}(\frac{\partial}{\partial x^\mu}, \frac{\partial}{\partial x^\nu})$. In the case of a Riemannian manifold we use latin letters $i, j, k, l, \ldots$ to denote indices of coordinates $x^1, x^2, \ldots, x^n$ or tensors. For a Lorentzian manifold we use greek letters $\alpha, \beta, \gamma, \ldots$ to denote indices $0, 1, \ldots, n$. Given an arbitrary frame $E_{(\alpha)}$ in $\mathbf{M}$

2. Covariant derivatives, Lie derivatives

We recall here the three fundamental operators of the differential geometry on a Riemann or Lorentz manifold: the exterior derivative, the Lie derivative, and the Levi-Civita connection with its associated covariant derivative.

2.0.1. *The exterior derivative.* Given a scalar function f its differential df is the 1-form defined by

$$df(X) = X(f)$$

for any vector field X. This definition can be extended for all differential forms on $\mathbf{M}$ in the following way:

i) d is a linear operator defined from the space of all k-forms to that of $k+1$-forms on $\mathbf{M}$. Thus for all k-forms A,B and real numbers λ, μ

$$d(\lambda A + \mu B) = \lambda dA + \mu dB$$

ii) For any k-form A and arbitrary form B

$$d(A \wedge B) = dA \wedge B + (-1)^k A \wedge dB$$

iii) For any form A,

$$d^2 A = 0 \ .$$

We recall that, if Φ is a smooth map defined from $\mathbf{M}$ to another manifold $\mathbf{M}'$, then

$$d(\Phi^* A) = \Phi^*(dA) \ .$$

Finally if A is a one form and X, Y arbitrary vector fields, we have the equation

$$dA(X, Y) = \frac{1}{2}\Big(X(A(Y)) - Y(A(X)) - A([X, Y])\Big)$$

where $[X, Y]$ is the commutator $X(Y) - Y(X)$. This can be easily generalised to arbitrary k forms, see Spivak's book, Vol.I, Chapter 7, Theorem 13. [**?**]

2.0.2. *The Lie derivative.* Consider an arbitrary vector field X. In local coordinates x^μ, the flow of X is given by the system of differential equations

$$\frac{dx^\mu}{dt} = X^\mu(x^1(t), ..., x^{p+q}(t)) \ .$$

The corresponding curves, $x^\mu(t)$, are the integral curves of X. For each point $p \in \mathbf{M}$ there exists an open neighborhood $\mathcal{U}$, a small $\epsilon > 0$ and a family of diffeomorphism $\Phi_t : \mathcal{U} \to \mathbf{M},\ |t| \leq \epsilon$, obtained by taking each point in $\mathcal{U}$ to a parameter distance t, along the integral curves of X. We use these diffeomorphisms to construct, for any given tensor T at p, the family of tensors $(\Phi_t)_* T$ at $\Phi_t(p)$.

The Lie derivative $\mathcal{L}_X T$ of a tensor field T, with respect to X, is:

$$\mathcal{L}_X T|_p \equiv \lim_{t \to 0} \frac{1}{t} \left(T|_p - (\Phi_t)_* T|_p \right) \ .$$

It has the following properties:

i) $\mathcal{L}_X$ linearly maps (p, q)-tensor fields into tensor fields of the same type.

ii) $\mathcal{L}_X$ commutes with contractions.

iii) For any tensor fields S, T,

$$\mathcal{L}_X(S \otimes T) = \mathcal{L}_X S \otimes T + S \otimes \mathcal{L}_X T \ .$$

If X is a vector field we easily check that

$$\mathcal{L}_X Y = [X, Y]$$

by writing $(\mathcal{L}_X Y)^i = -\frac{d}{dt}((\Phi_t)_* Y)^i\Big|_{t=0}$ and expressing $(\Phi_t)_* Y)^i\Big|_p = \frac{\partial x^i(\Phi_t(q))}{\partial x^j(q)} Y^j\Big|_q$, where $q = \Phi_{-t}(p)$. (See [**?**], Hawking and Ellis, section 2.4 for details.)

If A is a k-form we have, as a consequence of the commutation formula of the exterior derivative with the pull-back Φ^*,

$$d(\mathcal{L}_X A) = \mathcal{L}_X(dA) \ .$$

For a given k-covariant tensorfield T we have,

$$\mathcal{L}_X T(Y_1, \dots, Y_k) = X T(Y_1, \dots, Y_k) - \sum_{i=1}^{k} T(Y_1, \dots, \mathcal{L}_X Y_i, \dots, Y_k)$$

We remark that the Lie bracket of two coordinate vector fields vanishes,

$$\left[\frac{\partial}{\partial x^\mu}, \frac{\partial}{\partial x^\nu} \right] = 0.$$

The converse is also true, namely, see Spivak, [**?**], Vol.I, Chapter 5, PROPOSITION 2.1. *If $X_{(0)}, ..., X_{(k)}$ are linearly independent vector fields in a neighbourhood of a point p and the Lie bracket of any two of them is zero then there exists a coordinate system x^μ, around p such that $X_{(\rho)} = \frac{\partial}{\partial x^\rho}$ for each $\rho = 0, ..., k$.*

The above proposition is the main step in the proof of Frobenius Theorem. To state the theorem we recall the definition of a k-distribution in $\mathbf{M}$. This is an arbitrary smooth assignment of a k-dimensional plane π_p at every point in a domain $\mathcal{U}$ of $\mathbf{M}$. The distribution is said to be involute if, for any vector fields X, Y on $\mathcal{U}$ with $X|_p, Y|_p \in \pi_p$, for any $p \in \mathcal{U}$, we have $[X,Y]|_p \in \pi_p$. This is clearly the case for integrable distributions[3]. Indeed if $X|_p, Y|_p \in T\mathcal{N}_p$ for all $p \in \mathcal{N}$, then X, Y are tangent to $\mathcal{N}$ and so is also their commutator $[X,Y]$. The Frobenius Theorem establishes that the converse is also true[4], that is being in involution is also a sufficient condition for the distribution to be integrable,

THEOREM 2.2. *(Frobenius Theorem) A necessary and sufficient condition for a distribution $(\pi_p)_{p\in\mathcal{U}}$ to be integrable is that it is involute.*

2.2.1. *The connection and the covariant derivative.* A connection $\mathbf{D}$ is a rule which assigns to each vectorfield X a differential operator $\mathbf{D}_X$. This operator maps vector fields Y into vector fields $\mathbf{D}_X Y$ in such a way that, with $\alpha, \beta \in \mathbb{R}$ and f, g scalar functions on $\mathbf{M}$,

$$\begin{aligned} &a)\ \ \mathbf{D}_{fX+gY}Z = f\mathbf{D}_X Z + g\mathbf{D}_Y Z \\ &b)\ \ \mathbf{D}_X(\alpha Y + \beta Z) = \alpha\mathbf{D}_X Y + \beta\mathbf{D}_X Z \\ &c)\ \ \mathbf{D}_X fY = X(f)Y + f\mathbf{D}_X Y \end{aligned} \tag{98}$$

Therefore, at a point p,

$$\mathbf{D}Y \equiv Y^\alpha_{;\beta}\,\theta^{(\beta)} \otimes e_{(\alpha)} \tag{99}$$

where the $\theta^{(\beta)}$ are the one-forms of the dual basis respect to the orthonormal frame $e_{(\beta)}$. Observe that $Y^\alpha_{;\beta} = \theta^{(\alpha)}(\mathbf{D}_{e_{(\beta)}}Y)$. On the other side, from c),

$$\mathbf{D}fY = df \otimes Y + f\mathbf{D}Y$$

so that

$$\mathbf{D}Y = \mathbf{D}(Y^\alpha e_{(\alpha)}) = dY^\alpha \otimes e_{(\alpha)} + Y^\alpha \mathbf{D}e_{(\alpha)}$$

and finally, using $df(\cdot) = e_{(\alpha)}(f)\theta^{(\alpha)}(\cdot)$,

$$\mathbf{D}Y = \Big(e_{(\beta)}(Y^\alpha) + Y^\gamma\theta^{(\alpha)}(\mathbf{D}_{e_{(\beta)}}e_{(\gamma)})\Big)\,\theta^{(\beta)} \otimes e_{(\alpha)} \tag{100}$$

Therefore

$$Y^\alpha_{;\beta} = e_{(\beta)}(Y^\alpha) + \Gamma^\alpha_{\beta\gamma}Y^\gamma$$

and the connection is, therefore, determined by its connection coefficients,

$$\Gamma^\alpha_{\beta\gamma} = \theta^{(\alpha)}(\mathbf{D}_{e_{(\beta)}}e_{(\gamma)}) \tag{101}$$

---

[3]Recall that a distribution π on $\mathcal{U}$ is said to be integrable if through every point $p \in \mathcal{U}$ there passes a unique submanifold $\mathcal{N}$, of dimension k, such that $\pi_p = T\mathcal{N}_p$.

[4]For a proof see Spivak, citeSpivak, Vol.I, Chapter 6.

which, in a coordinate basis, are the usual Christoffel symbols and have the expression

$$\Gamma^{\mu}_{\rho\nu} = dx^{\mu}(\mathbf{D}_{\frac{\partial}{\partial x^{\rho}}} \frac{\partial}{\partial x^{\nu}})$$

Finally

$$\mathbf{D}_X Y = \left(X(Y^{\alpha}) + \Gamma^{\alpha}_{\beta\gamma} X^{\beta} Y^{\gamma} \right) e_{(\alpha)} \tag{102}$$

In the particular case of a coordinate frame we have

$$\mathbf{D}_X Y = \left(X^{\mu} \frac{\partial Y^{\nu}}{\partial x^{\mu}} + \Gamma^{\nu}_{\rho\sigma} X^{\rho} Y^{\sigma} \right) \frac{\partial}{\partial x^{\nu}}$$

A connection is said to be a Levi-civita connection if $\mathbf{D}\mathbf{g} = 0$. That is, for any three vector fields X, Y, Z,

$$Z(\mathbf{g}(X, Y)) = \mathbf{g}(\mathbf{D}_Z X, Y) + \mathbf{g}(X, \mathbf{D}_Z Y) \tag{103}$$

A very simple and basic result of differential geometry asserts that for any given metric there exists a unique affine connection associated to it.

Proposition 2.3. *There exists a unique connection on* $\mathbf{M}$*, called the Levi-Civita connection, which satisfies* $\mathbf{D}\,\mathbf{g} = 0$*. The connection is torsion free, that is,*

$$\mathbf{D}_X Y - \mathbf{D}_Y X = [X, Y] \ .$$

Moreover, relative to a system of coordinates, x^{μ}*, the Christoffel symbol of the connection is given by the standard formula*

$$\Gamma^{\mu}_{\rho\nu} = \frac{1}{2} g^{\mu\tau} \left(\partial_{\rho} g_{\nu\tau} + \partial_{\nu} g_{\tau\rho} - \partial_{\tau} g_{\nu\rho} \right) \ .$$

Exercise: Prove the proposition yourself, without looking in a book.

So far we have only defined the covariant derivative of a a vector field. We can easily extend the definition to one forms $A = A_{\alpha} dx^{a}$ by the requirement that,

$$X(A(Y)) = \mathbf{D}_X A(Y) + A(D_X Y),$$

for all vectorfields X, Y. Given a k-covariant tensor field T we define its covariant derivative $\mathbf{D}_X T$ by the rule,

$$\mathbf{D}_X T(Y_1, \dots, Y_k) = XT(Y_1, \dots, Y_k) - \sum_{i=1}^{k} T(Y_1, \dots, \mathbf{D}_X Y_i, \dots, Y_k)$$

We can talk about $\mathbf{D}T$ as a covariant tensor of rank $k+1$ defined by,

$$\mathbf{D}T(X, Y_1, \dots, Y_k) = \mathbf{D}_X T(Y_1, \dots, Y_k).$$

Given a frame e_{α} we denote by $T_{\alpha_1 \dots, \alpha_k; \beta} = \mathbf{D}T(e_{\beta}, e_{a_1}, \dots, e_{\alpha_k})$ the components of $\mathbf{D}T$ relative to the frame. By repeated covariant differentiation we can define $\mathbf{D}^2 T, \dots \mathbf{D}^m \mathbf{T}$. Relative to a frame e_{α} we write,

$$\mathbf{D}_{\beta_1} \dots \mathbf{D}_{\beta_m} T_{\alpha_1 \dots \alpha_k} = T_{\alpha_1 \dots \alpha_k; \beta_1 \dots \beta_m} = \mathbf{D}^m T(e_{\beta_1} \dots, e_{\beta_m}, e_{\alpha_1}, ..., e_{\alpha_k}).$$

The fact that the Levi-Civita connection is torsion free allows us to connect covariant differentiation to the Lie derivative. Thus, if T is a k-covariant tensor we have, in a coordinate basis,

$$(\mathcal{L}_X T)_{\sigma_1 \ldots \sigma_k} = X^\mu T_{\sigma_1 \ldots \sigma_k;\mu} + X^\mu{}_{;\sigma_1} T_{\mu\sigma_2 \ldots \sigma_k} + \ldots . + X^\mu{}_{;\sigma_k} T_{\sigma_1 \ldots \sigma_{k-1}\mu} \ .$$

The covariant derivative is also connected to the exterior derivative according to the following simple formula. If A is a k-form, we have[5] $A_{[\sigma_1 \ldots \sigma_k;\mu]} = A_{[\sigma_1 \ldots \sigma_k,\mu]}$ and

$$dA = \sum A_{\sigma_1 \ldots \sigma_k;\mu} dx^\mu \wedge dx^{\sigma_1} \wedge dx^{\sigma_2} \wedge \ldots . \wedge dx^{\sigma_k} \ .$$

Given a smooth curve $\mathbf{x} : [0,1] \to \mathbf{M}$, parametrized by t, let $T = \left(\frac{\partial}{\partial t}\right)_{\mathbf{x}}$ be the corresponding tangent vector field along the curve. A vector field X, defined on the curve, is said to be parallelly transported along it if $\mathbf{D}_T X = 0$. If the curve has the parametric equations $x^\nu = x^\nu(t)$, relative to a system of coordinates, then $T^\mu = \frac{dx^\mu}{dt}$ and the components $X^\mu = X^\mu(\mathbf{x}(t))$ satisfy the ordinary differential system of equations

$$\frac{\mathbf{D}}{dt} X^\mu \equiv \frac{dX^\mu}{dt} + \Gamma^\mu_{\rho\sigma}(\mathbf{x}(t)) \frac{dx^\rho}{dt} X^\sigma = 0 \ .$$

The curve is said to be geodesic if, at every point of the curve, $\mathbf{D}_T T$ is tangent to the curve, $\mathbf{D}_T T = \lambda T$. In this case one can reparametrize the curve such that, relative to the new parameter s, the tangent vector $S = \left(\frac{\partial}{\partial s}\right)_{\mathbf{x}}$ satisfies $\mathbf{D}_S S = 0$. Such a parameter is called an "affine parameter". The affine parameter is defined up to a transformation $s = as' + b$ for a, b constants. Relative to an affine parameter s and arbitrary coordinates x^μ the geodesic curves satisfy the equations

$$\frac{d^2 x^\mu}{ds^2} + \Gamma^\mu_{\rho\sigma} \frac{dx^\rho}{ds} \frac{dx^\sigma}{ds} = 0 \ .$$

A geodesic curve parametrized by an affine parameter is simply called a geodesic. In Lorentzian geometry timelike geodesics correspond to world lines of particles freely falling in the gravitational field represented by the connection coefficients. In this case the affine parameter s is called the proper time of the particle.

Given a point $p \in \mathbf{M}$ and a vector X in the tangent space $T_p\mathbf{M}$, let $\mathbf{x}(t)$ be the unique geodesic starting at p with "velocity" X. We define the exponential map:

$$\exp_p : T_p\mathbf{M} \to \mathbf{M} \ .$$

This map may not be defined for all $X \in T_p\mathbf{M}$. The theorem of existence and uniqueness for systems of ordinary differential equations implies that the exponential map is defined in a neighbourhood of the origin in $T_p\mathbf{M}$. If the exponential map is defined for all $T_p\mathbf{M}$, for every point p the manifold $\mathbf{M}$ is said geodesically complete. In general if the connection is a C^r connection[6] there exists an open neighbourhood $\mathcal{U}_0$ of the origin in $T_p\mathbf{M}$ and an open neighbourhood of the point

[5] $[\sigma_1 \ldots \sigma_k; \mu]$ indicates the antisymmetrization with respect to all indices (i.e. $\frac{1}{k!}$(alternating sum of the tensor over all permutations of the indices)) and "$,\mu$" indicates the ordinary derivative with respect to x^μ.

[6] A C^r connection is such that if Y is a C^{r+1} vector field then $\mathbf{D}Y$ is a C^r vector field.

p in $\mathbf{M}$, $\mathcal{V}_p$, such that the map $\exp_p$ is a C^r diffeomorphism of $\mathcal{U}_0$ onto $\mathcal{V}_p$. The neighbourhood $\mathcal{V}_p$ is called a normal neighbourhood of p.

3. Riemann curvature tensor, Ricci tensor, Bianchi identities

In the flat spacetime if we parallel transport a vector along any closed curve we obtain the vector we have started with. This fails in general because the second covariant derivatives of a vector field do not commute. This lack of commutation is measured by the Riemann curvature tensor,

$$\mathbf{R}(X,Y)Z = \mathbf{D}_X(\mathbf{D}_Y Z) - \mathbf{D}_Y(\mathbf{D}_X Z) - \mathbf{D}_{[X,Y]}Z \tag{104}$$

or written in components relative to an arbitrary frame,

$$R^{\alpha}{}_{\beta\gamma\delta} = \theta^{(\alpha)}\big((\mathbf{D}_\gamma\mathbf{D}_\delta - \mathbf{D}_\delta\mathbf{D}_\gamma)e_{(\beta)}\big) \tag{105}$$

Relative to a coordinate system x^μ and written in terms of the $g_{\mu\nu}$ components, the Riemann components have the expression

$$R^{\mu}{}_{\nu\rho\sigma} = \frac{\partial\Gamma^{\mu}_{\sigma\nu}}{\partial x^\rho} - \frac{\partial\Gamma^{\mu}_{\rho\nu}}{\partial x^\sigma} + \Gamma^{\mu}_{\rho\tau}\Gamma^{\tau}_{\sigma\nu} - \Gamma^{\mu}_{\sigma\tau}\Gamma^{\tau}_{\rho\nu} \tag{106}$$

The fundamental property of the curvature tensor, first proved by Riemann, states that if $\mathbf{R}$ vanishes identically in a neighbourhood of a point p one can find families of local coordinates such that, in a neighbourhood of p, $g_{\mu\nu} = \mathbf{m}_{\mu\nu}$ [7].

The trace of the curvature tensor, relative to the metric $\mathbf{g}$, is a symmetric tensor called the Ricci tensor,

$$\mathbf{R}_{\alpha\beta} = \mathbf{g}^{\gamma\delta}\mathbf{R}_{\alpha\gamma\beta\delta}$$

The scalar curvature is the trace of the Ricci tensor

$$\mathbf{R} = \mathbf{g}^{\alpha\beta}\mathbf{R}_{\alpha\beta} \ .$$

The Riemann curvature tensor of an arbitrary spacetime $(\mathbf{M},\mathbf{g})$ has the following symmetry properties,

$$\begin{aligned}\mathbf{R}_{\alpha\beta\gamma\delta} &= -\mathbf{R}_{\beta\alpha\gamma\delta} = -\mathbf{R}_{\alpha\beta\delta\gamma} = \mathbf{R}_{\gamma\delta\alpha\beta}\\ \mathbf{R}_{\alpha\beta\gamma\delta} &+ \mathbf{R}_{\alpha\gamma\delta\beta} + \mathbf{R}_{\alpha\delta\beta\gamma} = 0\end{aligned} \tag{107}$$

The second identity in (107) is called the first Bianchi identity.

It also satisfies the second Bianchi identities, which we refer to here as the *Bianchi equations* and, in a generic frame, have the form:

$$\mathbf{D}_{[\epsilon}\mathbf{R}_{\gamma\delta]\alpha\beta} = 0 \tag{108}$$

The traceless part of the curvature tensor, $\mathbf{C}$ is called the Weyl tensor, and has the following expression in an arbitrary frame,

$$\begin{aligned}\mathbf{C}_{\alpha\beta\gamma\delta} &= \mathbf{R}_{\alpha\beta\gamma\delta} - \frac{1}{n-1}\left(\mathbf{g}_{\alpha\gamma}\mathbf{R}_{\beta\delta} + \mathbf{g}_{\beta\delta}\mathbf{R}_{\alpha\gamma} - \mathbf{g}_{\beta\gamma}\mathbf{R}_{\alpha\delta} - \mathbf{g}_{\alpha\delta}\mathbf{R}_{\beta\gamma}\right)\\ &+ \frac{1}{n(n-1)}(\mathbf{g}_{\alpha\gamma}\mathbf{g}_{\beta\delta} - \mathbf{g}_{\alpha\delta}\mathbf{g}_{\beta\gamma})\mathbf{R}\end{aligned} \tag{109}$$

[7] For a thorough discussion and proof of this fact, refer to Spivak, [**?**], Vol. II.

Observe that $\mathbf{C}$ verifies all the symmetry properties of the Riemann tensor:

$$\begin{aligned} &\mathbf{C}_{\alpha\beta\gamma\delta} = -\mathbf{C}_{\beta\alpha\gamma\delta} = -\mathbf{C}_{\alpha\beta\delta\gamma} = \mathbf{C}_{\gamma\delta\alpha\beta} \\ &\mathbf{C}_{\alpha\beta\gamma\delta} + \mathbf{C}_{\alpha\gamma\delta\beta} + \mathbf{C}_{\alpha\delta\beta\gamma} = 0 \end{aligned} \tag{110}$$

and, in addition, $\quad \mathbf{g}^{\alpha\gamma}\mathbf{C}_{\alpha\beta\gamma\delta} = 0$.

We say that two metrics $\mathbf{g}$ and $\hat{\mathbf{g}}$ are conformal if $\hat{\mathbf{g}} = \lambda^2\mathbf{g}$ for some non zero differentiable function λ. Then the following theorem holds (see Hawking- Ellis, [**?**], chapter 2, section 2.6):

THEOREM 3.1. *Let $\hat{\mathbf{g}} = \lambda^2\mathbf{g}$, $\hat{\mathbf{C}}$ the Weyl tensor relative to $\hat{\mathbf{g}}$ and $\mathbf{C}$ the Weyl tensor relative to $\mathbf{g}$. Then*

$$\hat{\mathbf{C}}^{\alpha}_{\beta\gamma\delta} = \mathbf{C}^{\alpha}_{\beta\gamma\delta} \ .$$

Thus $\mathbf{C}$ is conformally invariant.

3.2. Isometries and conformal isometries, Killing and conformal Killing vector fields.

DEFINITION 3.3. A diffeomorphism $\Phi : \mathcal{U} \subset \mathbf{M} \to \mathbf{M}$ is said to be a conformal isometry if, at every point p, $\Phi_*\mathbf{g} = \Lambda^2\mathbf{g}$, that is,

$$(\Phi^*\mathbf{g})(X,Y)|_p = \mathbf{g}(\Phi_*X, \Phi_*Y)|_{\Phi(p)} = \Lambda^2\mathbf{g}(X,Y)|_p$$

with $\Lambda \neq 0$. If $\Lambda = 1$, Φ is called an isometry of $\mathbf{M}$.

DEFINITION 3.4. A vector field K which generates a one parameter group of isometries (respectively, conformal isometries) is called a Killing (respectively, conformal Killing) vector field.

Let K be such a vector field and Φ_t the corresponding one parameter group. Since the $(\Phi_t)_*$ are conformal isometries, we infer that $\mathcal{L}_K\mathbf{g}$ must be proportional to the metric $\mathbf{g}$. Moreover $\mathcal{L}_K\mathbf{g} = 0$ if K is a Killing vector field.

DEFINITION 3.5. Given an arbitrary vector field X we denote ${}^{(X)}\pi$ the deformation tensor of X defined by the formula

$${}^{(X)}\pi_{\alpha\beta} = (\mathcal{L}_X g)_{\alpha\beta} = \mathbf{D}_\alpha X_\beta + \mathbf{D}_\beta X_\alpha \ .$$

The tensor ${}^{(X)}\pi$ measures, in a precise sense, how much the diffeomorphism generated by X differs from an isometry or a conformal isometry. The following Proposition holds, (see Hawking-Ellis, citeHawkEll, chapter 2, section 2.6):

PROPOSITION 3.6. *The vector field X is Killing if and only if ${}^{(X)}\pi = 0$. It is conformal Killing if and only if ${}^{(X)}\pi$ is proportional to $\mathbf{g}$.*

Remark: One can choose local coordinates such that $X = \frac{\partial}{\partial x^\mu}$. It then immediately follows that, relative to these coordinates the metric $\mathbf{g}$ is independent of the component x^μ.

PROPOSITION 3.7. *On any pseudo-riemannian spacetime $\mathbf{M}$, of dimension $n = p+q$, there can be no more than $\frac{1}{2}(p+q)(p+q+1)$ linearly independent Killing vector fields.*

Proof: Proposition 3.7 is an easy consequence of the following relation, valid for an arbitrary vector field X, obtained by a straightforward computation and the use of the symmetries of $\mathbf{R}$.

$$\mathbf{D}_\beta \mathbf{D}_\alpha X_\lambda = \mathbf{R}_{\lambda\alpha\beta\delta} X^\delta + {}^{(X)}\Gamma_{\alpha\beta\lambda} \tag{111}$$

where

$${}^{(X)}\Gamma_{\alpha\beta\lambda} = \frac{1}{2}\left(\mathbf{D}_\beta \pi_{\alpha\lambda} + \mathbf{D}_\alpha \pi_{\beta\lambda} - \mathbf{D}_\lambda \pi_{\alpha\beta}\right) \tag{112}$$

and $\pi \equiv {}^{(X)}\pi$ is the X deformation tensor.

If X is a Killing vector field equation (111) becomes

$$\mathbf{D}_\beta(\mathbf{D}_\alpha X_\lambda) = \mathbf{R}_{\lambda\alpha\beta\delta} X^\delta \tag{113}$$

and this implies, in view of the theorem of existence and uniqueness for ordinary differential equations, that any Killing vector field is completely determined by the $\frac{1}{2}(n+1)(n+2)$ values of X and $\mathbf{D}X$ at a given point. Indeed let p, q be two points connected by a curve $x(t)$ with tangent vector T. Let $L_{\alpha\beta} \equiv \mathbf{D}_\alpha X_\beta$, Observe that along $x(t)$, X, L verify the system of differential equations

$$\frac{\mathbf{D}}{dt} X = T \cdot L \quad , \quad \frac{\mathbf{D}}{dt} L = \mathbf{R}(\cdot, \cdot, X, T)$$

therefore the values of X, L along the curve are uniquely determined by their values at p.

The n-dimensional Riemannian manifold which possesses the maximum number of Killing vector fields is the Euclidean space $\mathbb{R}^n$. Simmilarily the Minkowski spacetime $\mathbb{R}^{n+1}$ is the Lorentzian manifold with the maximum numbers of Killing vectorfields.

3.8. Laplace-Beltrami operator. The scalar Laplace-Beltrami operator on a pseudo-riemannian manifold $\mathbf{M}$ is defined by,

$$\Delta_{\mathbf{M}} u(x) = g^{\mu\nu} \mathbf{D}_\mu \mathbf{D}_\nu u \tag{114}$$

where u is a scalar function on $\mathbf{M}$. Or, in local coordinates,

$$\Delta_{\mathbf{M}} u(x) = \frac{1}{\sqrt{|g(x)|}} \partial_\mu (g^{\mu\nu} \sqrt{|g(x)|} \partial_\nu) u(x) \tag{115}$$

The Laplace-Beltrami operator is called D'Alembertian in the particular case of a Lorentzian manifold, and is then denoted by $\Box_{\mathbf{M}}$. On any pseudo-riemannian manifold, $\Delta_{\mathbf{M}}$ is symmetric relative to the following scalar product for scalar functions u, v:

$$(u, v)_{\mathbf{M}} = \int u(x) v(x) dv_{\mathbf{M}}$$

Indeed the following identities are easily established by integration by parts, for any two smooth, compactly supported[8] functions u, v,

$$(-\Delta u, v)_{\mathbf{M}} = \int_{\mathbf{M}} \nabla u \cdot \nabla v \, dv_{\mathbf{M}} = (u, -\Delta v)_{\mathbf{M}} \tag{116}$$

where $\nabla u \cdot \nabla v = g^{ij} \partial_i u \partial_j v$. In the particular case when $u = v$ we derive, $(-\Delta u, v)_{\mathbf{M}} = \int_{\mathbf{M}} |\nabla u|^2$, with $|\nabla u|^2 = \nabla u \cdot \nabla u$. Thus, $-\Delta = -\Delta_{\mathbf{M}}$ is symmetric for functions $u \in \mathcal{C}_0^\infty(\mathbf{M})$. It is positive definite if the manifold $\mathbf{M}$ is Riemannian. This is not the case for Lorentzian manifolds: $\square_{\mathbf{M}}$ is non-definite.

4. Geometry of space-like hypersurfaces

Consider a spacelike hypersurface Σ in $\mathbf{M}, \mathbf{g}$ with unit future normal T. We define the induced metric (or first fundamental form) g and second fundamental form k,

$$g(X, Y) = \mathbf{g}(X, Y), \qquad k(X, Y) = -\mathbf{g}(D_X T, Y) \qquad \forall X, Y \in T(\Sigma) \tag{117}$$

Remark that k is symmetric. Indeed since $[X, Y] \in T(\Sigma)$,

$$k(X, Y) - k(Y, X) = -\mathbf{g}(T, [X, Y]) = 0.$$

Denoting by ∇ the induced covariant derivative operator on Σ we have, for any $X, Y \in T(\Sigma)$,

$$DXY = \nabla_X Y - k(X, Y) T \tag{118}$$

To understand the geometric significance of k we extend T to a neighborhood $\mathcal{U}$ of Σ by parallel transporting it along the geodesics perpendicular to Σ, i.e.

$$D_T T = 0$$

Clearly we continue to have $\mathbf{g}(T, T) = -1$. Also, given any vectorfield X on Σ we extend, it along the same geodesics, by solving the differential equation,

$$[T, X] = 0.$$

Observe that,

$$T\mathbf{g}(T, X) = \mathbf{g}(D_T T, X) + \mathbf{g}(T, D_T X) = \mathbf{g}(T, D_X T) = \frac{1}{2} X \mathbf{g}(T, T) = 0.$$

Since $\mathbf{g}(T, X) = 0$ on Σ we infer that the extended vectorfields X remain orthogonal to the extended T in the neighborhood $\mathcal{U}$ of Σ.

Let t be the proper time along these geodesics (i.e. $T(t) = 1$), with $t = 0$ on Σ, and let Σ_t its level hypersurfaces, Observe that,

$$TX(t) = XT(t) = X(1) = 0.$$

Hence, since X is tangent to Σ and $t = 0$ on Σ we infer that $X(t) = 0$ in $\mathcal{U}$. In other words the extended vectors X are tangent to Σ_t. Clearly, the tangent space of Σ_t is spanned by these extended vectorfields, and since they are perpendicular to

[8]This is automatically satisfied if the manifold $\mathbf{M}$ is compact.

both T and the gradient $\mathbf{g}^{\mu\nu}D_\mu t$ we deduce that T and $\mathbf{g}^{\mu\nu}D_\mu t$ are proportional. Since $T(t) = 1$ and $\mathbf{g}(T, T) = -1$ we infer that,

$$T^\mu = -\mathbf{g}^{\mu\nu}D_\nu t \tag{119}$$

In particular t verifies the equation,

$$\mathbf{g}^{\mu\nu}D_\nu t D_\mu t = -1$$

Given the extended T and two such extended X, Y of tangent vectorfields on Σ we have,

$$\mathcal{L}_T\mathbf{g}(X,Y) = T\mathbf{g}(X,Y) = \mathbf{g}(D_TX,Y) + \mathbf{g}(D_TY,X) = 2\mathbf{g}(D_XT,Y) = -2k(X,Y).$$

Denoting by g the restriction of $\mathbf{g}$ to Σ_t we deduce,

$$k(X,Y) \quad = \quad -\frac{1}{2}Tg(X,Y). \tag{120}$$

On the other hand, since $[X,T] = [Y,T] = 0$ we can compute the second variation of g in the T direction as follows

$$\begin{aligned} TK(X,Y) &= -T\mathbf{g}(D_XT,Y) = -\mathbf{g}(D_TD_XT,Y) - (D_XT,D_T,Y) \\ &= -\mathbf{g}(D_XD_TT,Y) + R(X,T,Y,T) - (D_XT,D_YT) \\ &= \mathbf{R}(X,T,Y,T) - k^2(X,Y) \end{aligned}$$

where, in an arbitrary frame $e_1, \dots e_n$ on Σ,

$$k^2(X,Y) = \sum_{i=1}^{n} k(X,e_i)k(Y,e_i).$$

We have thus derived the second variation formula,

$$TK(X,Y) \quad = \quad \mathbf{R}(X,T,Y,T) - k^2(X,Y) \tag{121}$$

Now, let $(e_i)_{i=1\dots n}$ be an orthonormal frame on Σ. The frame $e_0 = T, e_1, \dots e_n$ is a spacetime orthonormal frame along Σ. We have,

$$\begin{aligned} D_ie_j &= \nabla_ie_j - k_{ij}T \\ D_iT &= -k_{ij}e_j \end{aligned} \tag{122}$$

where D_i denotes D_{e_i}. Given a 1 form A on our manifold we recall,

$$A_{i;j} = D_jA_i = DA(e_i;e_j) = e_j(A_i) - A_{D_ie_j}$$

$$\begin{aligned} A_{i;jm} &= D_mD_jA_i = D^2A(e_i;e_j,e_m) \\ &= e_m(D_jA_i) - D_{D_me_j}A_i - D_jA_{D_me_i} \end{aligned}$$

If A is tangent to Σ (which we can extend smoothly to a neighborhood of Σ) we derive, using (122),

$$\begin{aligned} A_{i;j} &= A_{i||j} \\ A_{i;jm} &= A_{i;||jm} + k_{im}k_{js}A^s + k_{mj}D_TA_i \\ A_{i;mj} &= A_{i;||mj} + k_{ij}k_{ms}A^s + k_{jm}D_TA_i \end{aligned}$$

Therefore, subtracting,

$$A_{i;jm} - A_{i;mj} \quad = \quad A_{i;||jm} - A_{i;||mj} + (k_{im}k_{js} - k_{ij}k_{ms})A^s$$

On the other hand, in $\mathbf{M}$

$$A_{i;jm} - A_{i;mj} \quad = \quad -\mathbf{R}_{isjm}$$

and in Σ,

$$A_{i;||m} - A_{i||mj} \quad = \quad -R_{isjm}$$

We thus derive the Gauss equation,

$$\mathbf{R}_{isjm} = R_{isjm} + k_{ij}k_{ms} - k_{im}k_{js} \tag{123}$$

Now, letting T_μ be the one form obtained by lowering the indices of T we derive,

$$\begin{aligned} T_{i;mj} &= -k_{im||j} + k_{mj}D_T T \\ T_{i;jm} &= -k_{ij||m} + k_{jm}D_T T \end{aligned}$$

Hence,

$$T_{i;mj} - T_{i;jm} \quad = \quad -(k_{im||j} - k_{ij||m})$$

from which we derive,

$$\mathbf{R}_{iojm} = \nabla_m k_{ij} - \nabla_j k_{im} \tag{124}$$

We summarize the results obtained so far in the following:

Proposition 4.1. *Let Σ be a spacelike hypersurface with induced metric g and second fundamental form k.*

(1) *If X, Y, Z, W are arbitrary vectorfields tangent to Σ we have the Codazzi equations,*

$$\nabla_X k(Y,Z) - \nabla_Y k(X,Z) = \mathbf{R}(T,Z,X,Y) \tag{125}$$

and the Gauss equations,

$$R(X,Y,Z,W) + k(X,Z)k(Y,W) - k(X,W)k(Y,Z) = \mathbf{R}(X,Y,Z,W) \tag{126}$$

(2) *Extend T and $X \in T(\Sigma)$ to a neighborhood of Σ such that $D_T T = 0$ and $[T,X]$=0. We have the first and second variations of g,*

$$Tg(X,Y) \quad = \quad -2k(X,Y), \qquad Tk(X,Y) = \mathbf{R}(X,T,Y,T) \tag{127}$$

Bibliography

[1] W. Israel, *Dark Stars: The evolution of an idea* In 300 years of Graviation, edited by S. Hawking and W. Israel, Cambridge University Press, Cambridge, 1987, Chapter 7, pp 199 -276.