Relativity

Yi-Zen Chu

Contents

1 Galilean Relativity

The core tenet of relativity is that physical laws must be indistinguishable – the fundamental equations of physics must take the same form – when we switch from one inertial frame to another. The difference between Galilean Relativity underlying Newton's laws of motion and Einstein's Special Relativity is the very different notions of an inertial frame. Let us begin with the former, since it is likely already familiar.

Newton's 1st law of motion is really a *definition* of an inertial frame: if there are no forces acting on an arbitrary system and if *all* such force-free systems travel at constant velocities, then the frame in which these observations are made is an inertial one.

Newton's 2nd law In such an inertial frame, using Cartesian coordinates $\vec{x}$ to describe the location os some mass m, its acceleration is governed by Newton's second law:

$$m \frac{\mathrm{d}^2 \vec{x}}{\mathrm{d}t^2} \equiv m\ddot{\vec{x}} = \vec{F}_{\text{total}}, \tag{1.0.1}$$

where $\vec{F}_{\text{total}}$ denotes the total force acting on it. Note, however, that Newton's 2nd law does not tell you what the forces $\{\vec{F}\}$ are; they are to be determined by empirical observation of the real world.

Flat Space In writing eq. (1.0.1) using Cartesian coordinates, there is an implicit assumption that Newton's laws of motion applies in *flat space*, whose precise definition I shall delay for a while. Roughly speaking flat space is where the rules of Euclidean geometry holds: parallel lines do not cross, sum of internal angles of a triangle equals π, the Pythagorean theorem holds, etc. For example, the surface of a perfectly spherical ball is *not* a flat space; parallel lines can meet, and sum of the internal angles of a triangle is not necessarily π. Being in flat space means, we may extend a straight line from mass m_1 at $\vec{x}_1$ and mass m_2 at $\vec{x}_2$, and denote the resulting vector as

$$\vec{\Delta}_{1\to 2} \equiv \vec{x}_2 - \vec{x}_1. \tag{1.0.2}$$

In particular, if both masses experience no external forces, $\ddot{\vec{x}}_{1,2} = 0$, we must also have

$$\ddot{\vec{\Delta}}_{1\to 2} = 0, \tag{1.0.3}$$

whose solution tells us the relative displacement between them must amount to constant velocity motion:

$$\vec{x}_2(t) - \vec{x}_1(t) = \vec{\Delta}_{1\to 2} = \vec{\Delta}_0 + \vec{V} \cdot t, \tag{1.0.4}$$

for time ($t-$)independent 'initial displacement' $\vec{\Delta}_0$ and 'initial velocity' $\vec{V}$.

Problem 1.1. Force-Free Parallel Lines on 2-Sphere Consider two masses m_1 and m_2 located on the unit 2$-$sphere, with trajectories $\vec{y}_1(t)$ and $\vec{y}_2(t)$. We shall let their initial velocities at $t = 0$ be perpendicular to the equator at $\theta = \pi/2$. This means they are initially parallel. Below, we shall verify that the following trajectories are indeed force-free:

$$y^i_{1,2} = \left(\frac{\pi}{2} - v_0 t, \phi_{1,2}\right), \tag{1.0.5}$$

where $\phi_{1,2}$ are the constant azimuthal angles of the masses' motion.

Verify that the velocities of $m_{1,2}$ are both $-v_0\widehat{\theta}$. What time t do they meet at the North Pole? For $\Delta\phi \equiv \phi_2 - \phi_1$ small enough – namely, the trajectories are nearby enough – so that the local region of space containing the two masses at a given time t can be considered nearly flat space, show that the displacement vector joining m_1 to m_2 is

$$\vec{\Delta}_{1\to 2} = \cos(v_0 t)\Delta\phi \cdot \widehat{\phi}. \tag{1.0.6}$$

This problem illustrates the difference between force-free motion on a curved space versus that in flat space: not only can initially parallel trajectories eventually meet, their relative displacements are not acceleration free – unlike their counterparts in flat space in equations (1.0.3) and (1.0.4). $\square$

Recall that, if $\vec{X}$ is the relative displacement from point A to B, the Pythagorean theorem informs us that the square of the distance between them is $\vec{X}^2 \equiv \vec{X}\cdot\vec{X}$, where the $\cdot$ is the the ordinary dot product. In an infinitesimal region of space, the infinitesimal distance $\mathrm{d}\ell$ between $\vec{x}$ and $\vec{x}+\mathrm{d}\vec{x}$ is therefore

$$\mathrm{d}\ell^2 = \mathrm{d}\vec{x}\cdot\mathrm{d}\vec{x}. \tag{1.0.7}$$

A word on notation: instead of labeling the Cartesian components $\{x, y, z\}$, we shall instead call them $\{x^1, x^2, x^3\}$. Here, $\{x^i | i = 1, 2, 3\}$ does not mean x raised to the ith power; but rather the ith component of the Cartesian coordinate vector $\vec{x}$. The Pythagorean theorem reads, in $3D$ space,

$$\mathrm{d}\ell^2 = (\mathrm{d}x^1)^2 + (\mathrm{d}x^2)^2 + (\mathrm{d}x^3)^2. \tag{1.0.8}$$

This is also a good place to introduce Einstein summation notation. First, we define the Kronecker delta,

$$\delta_{ij} = 1 \qquad \text{if } i = j \tag{1.0.9}$$

$$= 0 \qquad \text{if } i \neq j. \tag{1.0.10}$$

Notice, this is simply the identity matrix in index notation. Then, decree that

> whenever a pair of indices are repeated – for e.g., A^iB^i – they are implicitly summed over; namely, $A^iB^i \equiv \sum_i A^iB^i$.

For instance, the dot product between $\vec{a}$ and $\vec{b}$ is now expressible as

$$\vec{a}\cdot\vec{b} = \sum_i a^i b^j = \delta_{ij}a^i b^j. \tag{1.0.11}$$

We may thus rephrase the Pythagorean theorem as

$$\mathrm{d}\ell^2 = \delta_{ij}\mathrm{d}x^i\mathrm{d}x^j. \tag{1.0.12}$$

If we choose instead some other (possibly curvilinear) coordinates $\{y^i\}$, we may simply compute the Jacobian $\partial x^i/\partial y^a$ in order to obtain $\mathrm{d}\ell$ in this new system:

$$\mathrm{d}\ell^2 = \delta_{ij}\frac{\partial x^i}{\partial y^a}\frac{\partial x^j}{\partial y^b}\mathrm{d}y^a\mathrm{d}y^b = \left(\frac{\partial\vec{x}}{\partial y^a}\cdot\frac{\partial\vec{x}}{\partial y^b}\right)\mathrm{d}y^a\mathrm{d}y^b \equiv g_{a'b'}(\vec{y})\mathrm{d}y^a\mathrm{d}y^b. \tag{1.0.13}$$

Much of vector calculus operations follows from this object $g_{a'b'}$, usually dubbed the 'metric tensor'.

Problem 1.2. Extremal length implies straight lines Let $\vec{z}(\lambda_0 \leq \lambda \leq \lambda_1)$ be a path in space that joins $\vec{x}'$ to $\vec{x}$:

$$\vec{z}(\lambda = \lambda_0) = \vec{x}', \qquad \vec{z}(\lambda = \lambda_1) = \vec{x}. \tag{1.0.14}$$

Total length of this path can be defined as

$$\ell(\vec{x}' \leftrightarrow \vec{x}) = \int_{\vec{x}'}^{\vec{x}} \sqrt{\mathrm{d}\vec{z} \cdot \mathrm{d}\vec{z}} = \int_{\lambda_0}^{\lambda_1} \mathrm{d}\lambda \sqrt{(\mathrm{d}\vec{z}/\mathrm{d}\lambda)^2}, \tag{1.0.15}$$

where $(\mathrm{d}\vec{z}/\mathrm{d}\lambda)^2 \equiv (\mathrm{d}\vec{z}/\mathrm{d}\lambda) \cdot (\mathrm{d}\vec{z}/\mathrm{d}\lambda)$. If ℓ is extremized, show that $\vec{z}$ are straight lines joining $\vec{x}'$ to $\vec{x}$; namely,

$$\vec{z} = \vec{x}' + f(\lambda)(\vec{x} - \vec{x}'); \tag{1.0.16}$$

where f is an arbitrary but monotonically increasing function of λ subject to the boundary conditions $f(\lambda = \lambda_0) = 0$ and $f(\lambda = \lambda_1) = 1$. $\square$

Covariance Now, even though we defined Newton's second law eq. (1.0.1) using Cartesian coordinates, we may ask how to rephrase it in *arbitrary* ones. Afterall, a car or building should function in exactly the same way no matter what coordinate system the engineer used to design them. Geometrically, the length of a curve or the area of some 2D surface ought not depend on the coordinates used to parametrize them. Coordinates are important but merely technical intermediate tools to describe Nature herself. This demand that an equation of physics be expressible in arbitrary coordinate system – that the rules of calculation remains the same – is known as *covariance*.

We may begin with a function of space $f(\vec{x})$, which returns a unique number f given a unique location $\vec{x}$ in space – temperature of a medium at some point $\vec{x}$ is an example. Imagine the trajectory of a point mass $\vec{x}(t)$ passing through this medium, so that $f(\vec{x}(t))$ is the value f measured by it as a function of time. The time derivative is, by the chain rule,

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \frac{\mathrm{d}x^i}{\mathrm{d}t} \frac{\partial f}{\partial x^i}. \tag{1.0.17}$$

If we change the coordinate systems,

$$\vec{x} = \vec{x}(\vec{y}), \tag{1.0.18}$$

$$\vec{x}(t) \equiv \vec{x}(\vec{y}(t)); \tag{1.0.19}$$

then we had better get back the same f as long as $\vec{x} = \vec{x}(\vec{y})$ remains the same point:

$$f'(\vec{y}) \equiv f(\vec{x}(\vec{y})), \tag{1.0.20}$$

where the prime does not denote a derivative, but rather $f'(\vec{y})$ is the function f but now written in the $\vec{u}$ coordinate system. This, in fact, is the definition of a scalar function. Moreover, we may consider the time derivative

$$\frac{\mathrm{d}f'(\vec{y}(t))}{\mathrm{d}t} = \frac{\mathrm{d}y^a}{\mathrm{d}t} \frac{\partial f(\vec{x}(\vec{y}(t)))}{\partial y^a} = \frac{\mathrm{d}y^a}{\mathrm{d}t} \frac{\partial x^i}{\partial y^a} \frac{\partial f(\vec{x}(t))}{\partial x^i} = \frac{\mathrm{d}x^i}{\mathrm{d}t} \frac{\partial f(\vec{x}(t))}{\partial x^i}. \tag{1.0.21}$$

Since f itself was arbitrary, we may therefore identify the *tangent vector* along the trajectory to be

$$\frac{\mathrm{d}}{\mathrm{d}t} = \frac{\mathrm{d}y^a}{\mathrm{d}t}\frac{\partial}{\partial y^a} = \frac{\mathrm{d}y^a}{\mathrm{d}t}\frac{\partial x^i}{\partial y^a}\frac{\partial}{\partial x^i} = \frac{\mathrm{d}x^i}{\mathrm{d}t}\frac{\partial}{\partial x^i}. \tag{1.0.22}$$

Notice, from the second and last equality, that this notion of a tangent vector – i.e., the velocity tangent to some prescribed path – takes the same form no matter the coordinate system used. The two expressions in the $\vec{x}-$ and $\vec{y}-$system are in fact related by a contraction with the relevant Jacobian. Moreover, notice $\mathrm{d}/\mathrm{d}t$ for an arbitrary trajectory is really a superposition of the partial derivatives with respect to the coordinates employed; hence the collection of all such $\mathrm{d}/\mathrm{d}t$ is the vector space at a given point in space spanned by these $\{\partial_i\}$.

Let's turn to the second derivative version of $\mathrm{d}/\mathrm{d}t$, which we shall denote as $D^2/\mathrm{d}t^2$. We demand, like $\mathrm{d}/\mathrm{d}t$, that it takes the same form no matter the coordinate system used. The answer is

$$\frac{D^2}{\mathrm{d}t^2} = \left(\frac{\mathrm{d}^2 y^i}{\mathrm{d}t^2} + \Gamma^i{}_{ab}(\vec{y})\frac{\mathrm{d}y^a}{\mathrm{d}t}\frac{\mathrm{d}y^b}{\mathrm{d}t}\right)\frac{\partial}{\partial y^i}, \tag{1.0.23}$$

where the $\Gamma^i{}_{ab}$ are known as Christoffel symbols. (Note: $(\mathrm{d}/\mathrm{d}t)^2 \neq D^2/\mathrm{d}t^2$.) For any spatial metric $\mathrm{d}\ell^2 = g_{ab}\mathrm{d}y^a\mathrm{d}y^b$, it can be computed as

$$\Gamma^i{}_{ab}(\vec{y}) = \frac{1}{2}(g^{-1})^{ic}(\vec{y})\left(\partial_{y^a} g_{bc} + \partial_{y^b} g_{ac} - \partial_{y^c} g_{ab}\right). \tag{1.0.24}$$

By viewing g_{ab} as a matrix (with a and b being the row and column number) we have defined $(g^{-1})^{ab}$ as the (a,b)-component of its inverse. For the flat metric at hand, if $\vec{x}$ are Cartesian coordinates and $\vec{y}$ are some other (possibly curvilinear) ones, so that $g_{ab} = \partial_{y^a}\vec{x}\cdot\partial_{y^b}\vec{x}$,

$$\begin{aligned}\Gamma^i{}_{ab}(\vec{y}) &= \frac{1}{2}(g^{-1})^{ic}(\vec{y})\left(\partial_{y^a}(\partial_{y^b}\vec{x}\cdot\partial_{y^c}\vec{x}) + \partial_{y^b}(\partial_{y^a}\vec{x}\cdot\partial_{y^c}\vec{x}) - \partial_{y^c}(\partial_{y^a}\vec{x}\cdot\partial_{y^b}\vec{x})\right) && (1.0.25)\\ &= \frac{1}{2}(g^{-1})^{ic}(\vec{y})\left(2\partial_{y^a y^b}\vec{x}\cdot\partial_{y^c}\vec{x} + \partial_{y^a}\vec{x}\cdot\partial_{y^b y^c}\vec{x} + \partial_{y^b}\vec{x}\cdot\partial_{y^a y^c}\vec{x} - \partial_{y^c y^a}\vec{x}\cdot\partial_{y^b}\vec{x} - \partial_{y^a}\vec{x}\cdot\partial_{y^c y^b}\vec{x}\right)\\ &= (g^{-1})^{ic}(\vec{y})\frac{\partial\vec{x}}{\partial y^a\partial y^b}\cdot\frac{\partial\vec{x}}{\partial y^c}. && (1.0.26)\end{aligned}$$

The inverse of the metric is

$$(g^{-1})^{ab}(\vec{y}) = \delta^{ij}\frac{\partial y^a}{\partial x^i}\frac{\partial y^b}{\partial x^j} \tag{1.0.27}$$

because

$$\begin{aligned}(g^{-1}g)^a{}_b = (g^{-1})^{ac}g_{cb} &= \frac{\partial y^a}{\partial x^i}\frac{\partial y^c}{\partial x^i}\frac{\partial x^l}{\partial y^c}\frac{\partial x^l}{\partial y^b} && (1.0.28)\\ &= \frac{\partial y^a}{\partial x^i}\frac{\partial x^l}{\partial x^i}\frac{\partial x^l}{\partial y^b} = \frac{\partial y^a}{\partial x^i}\delta^l_i\frac{\partial x^l}{\partial y^b} && (1.0.29)\\ &= \frac{\partial y^a}{\partial x^l}\frac{\partial x^l}{\partial y^b} = \frac{\partial y^a}{\partial y^b} = \delta^a_b. && (1.0.30)\end{aligned}$$

In other words,

$$\frac{D^2}{\mathrm{d}t^2} = \left(\frac{\mathrm{d}^2 y^i}{\mathrm{d}t^2} + (g^{-1})^{ic}(\vec{y})\frac{\partial \vec{x}}{\partial y^a \partial y^b} \cdot \frac{\partial \vec{x}}{\partial y^c}\frac{\mathrm{d}y^a}{\mathrm{d}t}\frac{\mathrm{d}y^b}{\mathrm{d}t}\right)\frac{\partial}{\partial y^i} \tag{1.0.31}$$

$$= \left(\frac{\mathrm{d}^2 y^i}{\mathrm{d}t^2} + \frac{\partial y^i}{\partial x^l}\frac{\partial y^c}{\partial x^l}\frac{\partial x^k}{\partial y^a \partial y^b}\frac{\partial x^k}{\partial y^c}\frac{\mathrm{d}y^a}{\mathrm{d}t}\frac{\mathrm{d}y^b}{\mathrm{d}t}\right)\frac{\partial}{\partial y^i} \tag{1.0.32}$$

$$= \left(\frac{\mathrm{d}^2 y^i}{\mathrm{d}t^2} + \frac{\partial y^i}{\partial x^l}\frac{\partial x^l}{\partial y^a \partial y^b}\frac{\mathrm{d}y^a}{\mathrm{d}t}\frac{\mathrm{d}y^b}{\mathrm{d}t}\right)\frac{\partial}{\partial y^i}. \tag{1.0.33}$$

Problem 1.3. Coordinate Transformation Consider changing coordinates $\vec{y} = \vec{y}(\vec{z})$, so that, for instance,

$$\frac{\partial}{\partial y^a} = \frac{\partial z^k}{\partial y^a}\frac{\partial}{\partial z^k} \qquad \text{and} \qquad \frac{\mathrm{d}y^i}{\mathrm{d}t} = \frac{\partial y^i}{\partial z^a}\frac{\mathrm{d}z^a}{\mathrm{d}t} \tag{1.0.34}$$

– show that $D^2/\mathrm{d}t^2$ does indeed take the same form:

$$\frac{D^2}{\mathrm{d}t^2} = \left(\frac{\mathrm{d}^2 z^i}{\mathrm{d}t^2} + \frac{\partial z^i}{\partial x^l}\frac{\partial x^l}{\partial z^a \partial z^b}\frac{\mathrm{d}z^a}{\mathrm{d}t}\frac{\mathrm{d}z^b}{\mathrm{d}t}\right)\frac{\partial}{\partial z^i}. \tag{1.0.35}$$

Observe that, we recover the ordinary acceleration $\mathrm{d}^2 x^i/\mathrm{d}t^2$ when $\vec{z} = \vec{x}$; in fact, one approach to this problem is to show that

$$\ddot{x}^i \partial_{x^i} = \frac{D^2 z^i}{\mathrm{d}t^2}\partial_{z^i} \tag{1.0.36}$$

for arbitrary but invertible $\vec{x}(\vec{z})$. Moreover, since this definition takes the same form under arbitrary coordinate systems, we may take it to denote the fully covariant form of acceleration $a^i \equiv \ddot{z}^i + (\partial z^i/\partial x^l)(\partial^2 x^l/\partial z^a \partial z^b)\dot{z}^a \dot{z}^b$. □

Problem 1.4. Classical Mechanics on 2-Sphere We may exploit the first line in eq. (1.0.31) to describe acceleration $D^2 z^i/\mathrm{d}t^2$ on the 2−sphere. Simply view $\vec{x}$ as the unit-length Cartesian displacement vector parametrized in spherical coordinates:

$$\vec{x} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta). \tag{1.0.37}$$

Next, define

$$y^i \equiv (\theta, \phi). \tag{1.0.38}$$

Show that

$$g_{ab}(\vec{y}) = \frac{\partial \vec{x}}{\partial y^a} \cdot \frac{\partial \vec{x}}{\partial y^b} = \begin{bmatrix} 1 & 0 \\ 0 & \sin^2\theta \end{bmatrix} \qquad \text{and} \qquad (g^{-1})^{ab}(\vec{y}) = \begin{bmatrix} 1 & 0 \\ 0 & 1/\sin^2\theta \end{bmatrix}. \tag{1.0.39}$$

By computing $\partial y/\partial x$ and $\partial^2 x/\partial x \partial x$, show that the trajectories in eq. (1.0.5) are indeed acceleration-free: $D^2\theta/\mathrm{d}t^2 = 0 = D^2\phi/\mathrm{d}t^2$. □

Euclidean symmetry We record here without proof that the most general coordinate transformation $\vec{x} = \vec{x}(\vec{x}')$ that preserves the form of the metric in equations (1.0.7) and (1.0.12) – namely,

$$\mathrm{d}\vec{x} \cdot \mathrm{d}\vec{x} = \left(\frac{\partial \vec{x}}{\partial x'^a} \cdot \frac{\partial \vec{x}}{\partial x'^b}\right) \mathrm{d}x'^a \mathrm{d}x'^b = \mathrm{d}\vec{x}' \cdot \mathrm{d}\vec{x}' \tag{1.0.40}$$

is given by

$$\vec{x} = \widehat{R} \cdot \vec{x}' + \vec{a}, \tag{1.0.41}$$

where $\widehat{R}$ is an orthogonal matrix obeying $\widehat{R}^{\mathrm{T}} \widehat{R} = \mathbb{I}$ and $\vec{a}$ is a constant vector. **YZ: Switch to Analytical Methods.**

Galiean Transformation and Newtonian Gravity An important example is that of Newtonian gravity of N point masses. In an inertial frame, Newton's second law for the A-th mass reads

$$m_A \ddot{\vec{x}}_A = -\sum_{B \neq A} \frac{G_{\mathrm{N}} m_A m_B (\vec{x}_A - \vec{x}_B)}{|\vec{x}_A - \vec{x}_B|^3}. \tag{1.0.42}$$

That we are in flat space is what allows us to write the displacement vector between m_A and m_B as $\vec{x}_A - \vec{x}_B$ and the associated distance as $|\vec{x}_A - \vec{x}_B|$. Additionally, let us perform the Galilean transformation

$$\vec{x} = \widehat{R} \cdot \vec{x}' + \vec{a} + \vec{V} \cdot t' \qquad \text{and} \qquad t = t', \tag{1.0.43}$$

where $\widehat{R}$, $\vec{V}$, and $\vec{a}$ are constant; moreover $\widehat{R}^{\mathrm{T}} \widehat{R} = \mathbb{I}$. This relates two inertial frames by a constant velocity displacement as well as spatial rotation and/or parity flips. Note that eq. (1.0.41) is a subset of eq. (1.0.43); i.e., where $\vec{V} = \vec{0}$.

For Newtonian gravity, eq. (1.0.43) leads to

$$\vec{x}_A - \vec{x}_B = \widehat{R} \cdot (\vec{x}'_A - \vec{x}'_B), \tag{1.0.44}$$

$$|\vec{x}_A - \vec{x}_B| = |\vec{x}'_A - \vec{x}'_B|, \tag{1.0.45}$$

whereas

$$\ddot{\vec{x}} = \widehat{R} \cdot \ddot{\vec{x}}'. \tag{1.0.46}$$

Altogether, Newtonian gravity now reads

$$m_A \widehat{R} \cdot \ddot{\vec{x}}'_A = -\sum_{B \neq A} \frac{G_{\mathrm{N}} m_A m_B \widehat{R} \cdot (\vec{x}'_A - \vec{x}'_B)}{|\vec{x}'_A - \vec{x}'_B|^3}. \tag{1.0.47}$$

In index notation,

$$m_A \widehat{R}^i{}_j \cdot \ddot{x}'^j_A = -\widehat{R}^i{}_j \sum_{B \neq A} \frac{G_{\mathrm{N}} m_A m_B (x'^j_A - x'^j_B)}{|\vec{x}'_A - \vec{x}'_B|^3}. \tag{1.0.48}$$

Multiplying both sides by $\widehat{R}^{\mathrm{T}}$, we see that Newtonian gravity is in fact invariant under the transformations in eq. (1.0.43):

$$m_A \ddot{\vec{x}}'_A = -\sum_{B \neq A} \frac{G_{\mathrm{N}} m_A m_B \cdot (\vec{x}'_A - \vec{x}'_B)}{|\vec{x}'_A - \vec{x}'_B|^3}. \tag{1.0.49}$$

To ensure that Newtonian gravity may be expressed in arbitrary coordinate systems, we contract both sides with the partial derivatives:

$$m_A \ddot{x}^i_A \partial_{x^i} = -\sum_{B \neq A} \frac{G_{\mathrm{N}} m_A m_B \cdot (x^i_A - x^i_B)}{|\vec{x}_A - \vec{x}_B|^3} \partial_{x^i}. \tag{1.0.50}$$

Previously, we have already seen that if $\vec{x}(\vec{z})$ is given the coordinate invariant version of the LHS is $\ddot{x}^i \partial_{x^i} = (D^2 z^i / \mathrm{d}t^2) \partial_{z^i}$. On the other hand, $\partial_{x^i} = (\partial z^a / \partial x^i) \partial_{z^a}$. We therefore arrive at

$$m_A \frac{D^2 z^a_A}{\mathrm{d}t^2} \partial_{z^a} = -\sum_{B \neq A} \frac{G_{\mathrm{N}} m_A m_B \cdot (x^i_A - x^i_B)}{|\vec{x}_A - \vec{x}_B|^3} \frac{\partial z^a}{\partial x^i} \partial_{z^a}. \tag{1.0.51}$$

Problem 1.5. 2-Body Newtonian Gravity: Spherical Coordinates Suppose a small mass m is orbiting a much heavier one M, i.e., $m \lll M$, so that Newton's law of gravity reduces for the small mass' trajectory $\vec{x}$ to

$$\ddot{x}^i = -\frac{G_{\mathrm{N}} M}{|\vec{x}|^2} \frac{\vec{x}}{|\vec{x}|}, \tag{1.0.52}$$

where $\vec{x}$ are Cartesian coordinates. Use eq. (1.0.35) to show that, in spherical coordinates,

$$\ddot{r} - r \cdot \dot{\theta}^2 - r \cdot \sin^2\theta \dot{\phi}^2 = -\frac{G_{\mathrm{N}} M}{r^2}, \tag{1.0.53}$$

$$\ddot{\theta} + \frac{2}{r} \dot{r} \dot{\theta} - \sin\theta \cos\theta \dot{\phi}^2 = 0, \tag{1.0.54}$$

$$\ddot{\phi} + \frac{2}{r} \dot{r} \dot{\phi} + 2 \cot\theta \dot{\theta} \dot{\phi} = 0. \tag{1.0.55}$$

For practical purposes, it is useful to choose the coordinate system such that the orbit takes place on the $\theta = \pi/2$ plane. □

Classical Mechanics & Galilean Symmetry For slowing moving classical systems ($v/c \ll 1$), the Galilean transformation in eq. (1.0.43) are expected to preserve the form of all fundamental physical laws.

Mathematically, we may package it as the following matrix relation:

$$\begin{bmatrix} t \\ \vec{x} \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & \vec{0}^{\mathrm{T}} & 0 \\ \vec{V} & \widehat{R} & \vec{a} \\ 0 & \vec{0}^{\mathrm{T}} & 1 \end{bmatrix} \begin{bmatrix} t' \\ \vec{x}' \\ 1 \end{bmatrix}. \tag{1.0.56}$$

(The final row does not contain physical information; it is inserted just to make the matrix multiplication work out properly.) We see that a Galilean transformation can be encoded with a $(D+2)\times(D+2)$ matrix, containing the constant velocity $\vec{V}$, the rotation and/or parity flip $\widehat{R}$, and the constant spatial displacement $\vec{a}$. Denoting

$$\Pi\left(\vec{V},\widehat{R},\vec{a}\right) \equiv \begin{bmatrix} 1 & \vec{0}^{\mathrm{T}} & 0 \\ \vec{V} & \widehat{R} & \vec{a} \\ 0 & \vec{0}^{\mathrm{T}} & 1 \end{bmatrix}, \tag{1.0.57}$$

we multiply two such matrices to uncover

$$\Pi\left(\vec{V}_1,\widehat{R}_1,\vec{a}_1\right)\cdot\Pi\left(\vec{V}_2,\widehat{R}_2,\vec{a}_2\right) = \begin{bmatrix} 1 & \vec{0}^{\mathrm{T}} & 0 \\ \vec{V}_1 & \widehat{R}_1 & \vec{a}_1 \\ 0 & \vec{0}^{\mathrm{T}} & 1 \end{bmatrix}\begin{bmatrix} 1 & \vec{0}^{\mathrm{T}} & 0 \\ \vec{V}_2 & \widehat{R}_2 & \vec{a}_2 \\ 0 & \vec{0}^{\mathrm{T}} & 1 \end{bmatrix} \tag{1.0.58}$$

$$= \begin{bmatrix} 1 & \vec{0}^{\mathrm{T}} & 0 \\ \vec{V}_1+\widehat{R}_1\vec{V}_2 & \widehat{R}_1\widehat{R}_2 & \widehat{R}_1\vec{a}_2+\vec{a}_1 \\ 0 & \vec{0}^{\mathrm{T}} & 1 \end{bmatrix} \tag{1.0.59}$$

$$= \Pi\left(\vec{V}_1+\widehat{R}_1\vec{V}_2,\widehat{R}_1\widehat{R}_2,\widehat{R}_1\vec{a}_2+\vec{a}_1\right). \tag{1.0.60}$$

The identity transformation is

$$\mathbb{I}_{(D+2)\times(D+2)} = \Pi\left(\vec{0},\mathbb{I}_{D\times D},\vec{0}\right). \tag{1.0.61}$$

and therefore

$$\Pi\left(\vec{V},\widehat{R},\vec{a}\right)^{-1} = \Pi\left(-\widehat{R}^{\mathrm{T}}\vec{V},\widehat{R}^{\mathrm{T}},-\widehat{R}^{\mathrm{T}}\vec{a}\right). \tag{1.0.62}$$

These relations verify that $\{\Pi(\vec{V},\widehat{R},\vec{a})\}$ forms a group.

Problem 1.6. Derivatives Explain why

$$\begin{bmatrix} \frac{\partial t}{\partial t'} & \frac{\partial t}{\partial x'^b} \\ \frac{\partial x^a}{\partial t'} & \frac{\partial x^a}{\partial x'^b} \end{bmatrix}^{-1} = \begin{bmatrix} \frac{\partial t'}{\partial t} & \frac{\partial t'}{\partial x^b} \\ \frac{\partial x'^a}{\partial t} & \frac{\partial x'^a}{\partial x^b} \end{bmatrix}. \tag{1.0.63}$$

Use this result or otherwise to deduce from eq. (1.0.43) the relations

$$\partial_t = \partial_{t'} - V^a\widehat{R}^{ab}\partial_{x'^b} \tag{1.0.64}$$

and

$$\partial_{x^i} = \widehat{R}^{ij}\partial_{x'^j}. \tag{1.0.65}$$

These results are important in determining if certain partial differential equations of physics are in fact invariant under the Galilean transformations of eq. (1.0.43). □

Problem 1.7. Covariant Acceleration If the Cartesian $\vec{x}$ are given a transformation into $\vec{z}-$coordinates, i.e., $\vec{x}(\vec{z})$ is given, show that eq. (1.0.43) then implies

$$\frac{\partial x^a}{\partial z^i} = \widehat{R}^{ab}\frac{\partial x'^b}{\partial z^i} \tag{1.0.66}$$

and

$$\frac{\partial z^i}{\partial x^a} = \frac{\partial z^i}{\partial x'^b} \widehat{R}^{ab}. \tag{1.0.67}$$

Next, prove that the acceleration in eq. (1.0.35) is in fact invariant under Galilean transformations. □

Surface Waves: Toy Model Let x^3 be the height of the 2D surface of some substance made out of many point particles – say, a rubber sheet. Let there be a wave propagating along the positive 1−direction, so that

$$x^3 = A \sin(x^1 - vt), \tag{1.0.68}$$

where A is the amplitude of the wave and v is its (constant) speed. These $\vec{x} = (x^1, x^2, x^3)$ are defined with respect to the rest frame of this substance. Now, if Galilean symmetry holds (cf. (1.0.43)), then in the inertial $\vec{x}'$−frame moving at velocity $\vec{V}$ parallel to the rubber sheet, namely

$$(x^1, x^2, x^3) = (x'^1 + a^1 + V^1 \cdot t, x'^2 + a^2 + V^2 \cdot t, x'^3), \tag{1.0.69}$$

we have

$$x'^3 = A \sin(x'^1 + a^1 + (V^1 - v)t). \tag{1.0.70}$$

In other words, the velocity of the wave in this new $\vec{x}'$−frame is now $V^1 - v$. We shall see the electromagnetic waves *do not* transform in such a manner.

2 Local Conservation Laws

Non-relativistic You would be rightly shocked if you had stored a sealed tank of water on your rooftop only to find its contents gradually disappearing over time – the total mass of water ought to be a constant. Assuming a flat space geometry, if you had instead connected the tank to two pipes, one that pumps water into the tank and the other pumping water out of it, the rate of change of the total mass of the water

$$M \equiv \int_{\text{tank}} \rho(t, \vec{x}) \mathrm{d}^3\vec{x} \tag{2.0.1}$$

in the tank – where t is time, $\vec{x}$ are Cartesian coordinates, and $\rho(t, \vec{x})$ is the water's mass density – is

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{\text{tank}} \rho \mathrm{d}^3\vec{x} = -\left(\int_{\text{cross section of 'in' pipe}} + \int_{\text{cross section of 'out' pipe}} \right) \mathrm{d}^2\vec{\Sigma} \cdot (\rho \vec{v}). \tag{2.0.2}$$

Note that $\mathrm{d}^2\vec{\Sigma}$ points *outwards* from the tank, so at the 'in' pipe-tank interface, if the water were indeed following into the pipe, $-\mathrm{d}^2\vec{\Sigma} \cdot (\rho\vec{v}) > 0$ and its contribution to the rate of increase is positive. At the 'out' pipe-tank interface, if the water were indeed following out of the pipe, $-\mathrm{d}^2\vec{\Sigma}\cdot(\rho\vec{v}) < 0$. If we apply Gauss' theorem,

$$\int_{\text{tank}} \dot{\rho}\mathrm{d}^3\vec{x} = -\int_{\text{tank}} \mathrm{d}^3\vec{x}\vec{\nabla}\cdot(\rho\vec{v}). \tag{2.0.3}$$

If we applied the same sort of reasoning to any infinitesimal packet of fluid, with some local mass density ρ, we would find the following local conservation law

$$\dot{\rho} = -\partial_i\left(\rho\cdot v^i\right). \tag{2.0.4}$$

This is a "local" conservation law in the sense that mass cannot simply vanish from one location and re-appear a finite distance away, without first flowing to a neighboring location.

Relativistic We have implicitly assumed a non-relativistic system, where $|\vec{v}| \ll 1$. This is an excellent approximation for most hydrodynamics problems. Strictly speaking, however, relativistic effects – length contraction, in particular – imply that mass density is not a Lorentz scalar. If we define $\rho(t,\vec{x})$ to be the mass density at $(t,\vec{x})$ in a frame instantaneously at rest (aka 'co-moving') with the fluid packet, then the mass density current that is a locally conserved Lorentz vector is given by

$$J^\mu(t,\vec{x}) \equiv \rho(t,\vec{x})v^\mu(t,\vec{x}). \tag{2.0.5}$$

Along its integral curve v^μ should be viewed as the proper velocity $\mathrm{d}(t,\vec{x})^\mu/\mathrm{d}\tau$ of the fluid packet, where τ is the latter's proper time. Moreover, as long as the velocity v^μ is timelike, which is certainly true for fluids, let us recall it is always possible to find a (local) Lorentz transformation $\Lambda^\mu{}_\nu(t,\vec{x})$ such that

$$(1,\vec{0})^\mu \equiv v'^\mu = \Lambda^\mu{}_\nu(t,\vec{x})v^\nu(t,\vec{x}). \tag{2.0.6}$$

and the mass density-current is now

$$J'^\mu = \rho(t',\vec{x}')v'^\mu = \rho(x')\cdot\delta^\mu_0. \tag{2.0.7}$$

The local conservation law obeyed by this relativistically covariant current J^μ is now (in Cartesian coordinates)

$$\partial_\mu J^\mu = 0; \tag{2.0.8}$$

which in turn is a Lorentz invariant statement. Total mass M in a given global inertial frame at a fixed time t is

$$M \equiv \int_{\mathbb{R}^3}\mathrm{d}^3\vec{x}J^0. \tag{2.0.9}$$

To show it is a constant, we take the time derivative, and employ eq. (6.2.8):

$$\dot{M} = \int_{\mathbb{R}^3}\mathrm{d}^3\vec{x}\partial_0 J^0 = -\int_{\mathbb{R}^3}\mathrm{d}^3\vec{x}\partial_i J^i. \tag{2.0.10}$$

The divergence theorem tells us that this is equal to the flux of J^i at spatial infinity. But there is no J^i at spatial infinity for physically realistic – i.e., isolated – systems.

Perfect Fluids

3 Scalar Fields in Minkowski Spacetime

Field theory in Minkowski spacetime indicates we wish to construct partial differential equations obeyed by fields such that they take the same form in all inertial frames – i.e., the PDEs are Lorentz covariant. As a warm-up, we shall in this section study the case of scalar fields.

A scalar field $\varphi(x)$ is an object that transforms, under Poincaré transformations

$$x^\mu = \Lambda^\mu{}_\nu x'^\nu + a^\mu \tag{3.0.1}$$

as simply

$$\varphi(x) = \varphi\left(x^\mu = \Lambda^\mu{}_\nu x'^\nu + a^\mu\right). \tag{3.0.2}$$

To ensure that this is the case, we would like the PDE it obeys to take the same form in the two inertial frames $\{x^\mu\}$ and $\{x'^\mu\}$ related by eq. (6.0.1). The simplest example is the wave equation with some external scalar source $J(x)$. Let's first write it in the x^μ coordinate system.

$$\eta^{\mu\nu}\partial_\mu\partial_\nu\varphi(x) = J(x), \qquad \partial_\mu \equiv \partial/\partial x^\mu. \tag{3.0.3}$$

If putting a prime on the index denotes derivative with respect to x'^μ, namely $\partial_{\mu'} \equiv \partial/\partial x'^\mu$, then by the chain rule,

$$\partial_{\mu'} = \frac{\partial x^\sigma}{\partial x'^\mu}\frac{\partial}{\partial x^\sigma} = \partial_{\mu'}\left(\Lambda^\sigma{}_\rho x'^\rho + a^\sigma\right)\partial_\sigma \tag{3.0.4}$$

$$= \Lambda^\sigma{}_\mu \partial_\sigma. \tag{3.0.5}$$

Therefore the wave operator indeed takes the same form in both coordinate systems:

$$\eta^{\mu\nu}\partial_{\mu'}\partial_{\nu'} = \eta^{\mu\nu}\Lambda^\sigma{}_\mu\Lambda^\rho{}_\nu\partial_\sigma\partial_\rho \tag{3.0.6}$$

$$= \eta^{\sigma\rho}\partial_\sigma\partial_\rho. \tag{3.0.7}$$

because of Lorentz invariance

$$\eta^{\mu\nu}\Lambda^\sigma{}_\mu\Lambda^\rho{}_\nu = \eta^{\sigma\rho}. \tag{3.0.8}$$

A generalization of the wave equation in eq. (6.0.3) is to add a potential $V(\varphi)$:

$$\partial^2\varphi + V'(\varphi) = J, \tag{3.0.9}$$

where $\partial^2 \equiv \eta^{\mu\nu}\partial_\mu\partial_\nu$ and the prime is a derivative with respect to the argument.

Problem 3.1. Yukawa potential in $(3+1)$D Let the potential in eq. (6.0.9) be that of a mass term

$$V(\varphi) = \frac{m^2}{2}\varphi^2. \tag{3.0.10}$$

Consider a static point mass resting at $\vec{x} = 0$ in the $\{x^\mu\}$ inertial frame, namely

$$J(\vec{x}) = J_0\delta^{(3)}(\vec{x}), \qquad J_0 \text{ constant}. \tag{3.0.11}$$

Solve φ. Hint: You may assume the time derivatives in eq. (6.0.9) can be neglected. Then go to Fourier $\vec{k}-$space. You should find

$$\widetilde{\varphi}(\vec{k}) = \frac{J_0}{\vec{k}^2 + m^2}. \tag{3.0.12}$$

You should find a short-range force that, when $m \to 0$, recovers the Coulomb/Newtonian $1/r$ potential.

Next, consider an inertial frame $\{x'^\mu\}$ that is moving relative to the $\{x^\mu\}$ frame at velocity v along the positive x^3 axis. What is $\varphi(x')$ in the new frame? $\square$

4 Electromagnetism in Minkowski Spacetime

In this section we will discuss in some detail Minkowski spacetime electromagnetism to illustrate both its Lorentz and gauge symmetries. It will also provide us the opportunity to introduce the action principle, which is key formulating both classical and quantum field theories.

Maxwell & Lorentz We begin with Maxwell's equations in the following Lorentz covariant form, written in Cartesian coordinates $\{x^\mu\}$ so that $g_{\mu\nu} = \eta_{\mu\nu}$:

$$\partial_\mu F^{\mu\nu} = J^\nu, \qquad \partial_{[\mu} F_{\alpha\beta]} = 0, \qquad F_{\mu\nu} = -F_{\nu\mu}. \tag{4.0.1}$$

The $J^\mu \equiv \rho v^\mu$ is the electromagnetic current. Assuming J^μ is timelike, v^μ is its $d-$proper velocity with $v^2 \equiv v^\mu v_\mu = 1$; and $\rho \equiv J^\mu v_\mu$ is the electric charge in the (local) rest frame where $v^\mu = \delta^\mu_0$. Defined this way, ρ is a Lorentz scalar and J^μ is a Lorentz vector since v^μ is a Lorentz vector. It is then reasonable to suppose $F_{\mu\nu}$ is a rank-2 Lorentz tensor. Specifically, let two inertial frames $\{x^\mu\}$ and $\{x'^\mu\}$ be related via the Lorentz transformation

$$x^\mu = \Lambda^\mu{}_\alpha x'^\alpha, \qquad \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta \eta_{\mu\nu} = \eta_{\alpha\beta}. \tag{4.0.2}$$

Then the Faraday tensor transforms as

$$F_{\alpha'\beta'}(x') = F_{\mu\nu}\left(x(x') = \Lambda \cdot x'\right) \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta \tag{4.0.3}$$

Its derivatives are also Lorentz covariant, for keeping in mind eq. (7.1.2),

$$\partial_{\lambda'} F_{\alpha'\beta'}(x') = \frac{\partial x^\sigma}{\partial x'^\lambda} \partial_\sigma F_{\mu\nu}\left(x(x') = \Lambda \cdot x'\right) \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta \tag{4.0.4}$$

$$= \Lambda^\sigma{}_\lambda \partial_\sigma F_{\mu\nu}\left(x(x')\right) \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta. \tag{4.0.5}$$

This immediately tells us $\partial_\mu F^{\mu\nu} = \eta^{\mu\alpha} \partial_\mu F_{\alpha\beta} \eta^{\beta\nu}$ in eq. (7.1.1) is a Lorentz vector.

Problem 4.1. 4D Maxwell's Equations in term of $(\vec{E}, \vec{B})$ Let us check that eq. (7.1.1) does in fact reproduce Maxwell's equations in terms of electric E^i and magnetic B^i fields in 4D. Given a Lorentzian inertial frame, define

$$F^{i0} \equiv E^i \qquad \text{and} \qquad F^{ij} \equiv \epsilon^{ijk} B^k; \tag{4.0.6}$$

with $\epsilon^{123} \equiv -1$. Show that the $\partial_\mu F^{\mu\nu} = J^\nu$ from eq. (7.1.1) translates to

$$\vec{\nabla} \cdot \vec{E} = J^0 \qquad \text{and} \qquad \vec{\nabla} \times \vec{B} - \partial_t \vec{E} = \vec{J}. \tag{4.0.7}$$

(The over-arrow refers to the spatial components; for instance $\vec{B} = (B^1, B^2, B^3)$.) The $\partial_{[\alpha} F_{\mu\nu]} = 0$ from eq. (7.1.1) translates to

$$\partial_t \vec{B} + \vec{\nabla} \times \vec{E} = 0 \qquad \text{and} \qquad \vec{\nabla} \cdot \vec{B} = 0. \tag{4.0.8}$$

Hint: Note that $(\vec{\nabla} \times \vec{A})^i = -\epsilon^{ijk} \partial_j A^k$, for any Cartesian vector $\vec{A}$. Also, when you compute $\partial_{[i} F_{jk]}$, you simply need to set $\{i, j, k\}$ to be any distinct permutation of $\{1, 2, 3\}$. (Why?)

Next, verify the Lorentz invariant relations, with $\epsilon^{0123} \equiv -1$:

$$F_{\mu\nu} F^{\mu\nu} = -2\left(\vec{E}^2 - \vec{B}^2\right), \qquad \vec{E}^2 \equiv E^i E^i, \ \vec{B}^2 \equiv B^i B^i, \tag{4.0.9}$$

$$\epsilon^{\mu\nu\alpha\beta} F_{\mu\nu} F_{\alpha\beta} = 4\partial_\mu \left(\epsilon^{\mu\nu\alpha\beta} A_\nu \partial_\alpha A_\beta\right) = 8\vec{E} \cdot \vec{B}. \tag{4.0.10}$$

How does $F_{\mu\nu} F^{\mu\nu}$ transform under time reversal, $t \equiv x^0 \to -t$? How does it transform under parity flips, $x^i \to -x^i$ (for a fixed i)? Answer the same questions for $\widetilde{F}^{\mu\nu} F_{\mu\nu}$, where the dual of $F_{\mu\nu}$ is

$$\widetilde{F}^{\mu\nu} \equiv \frac{1}{2} \widetilde{\epsilon}^{\mu\nu\alpha\beta} F_{\alpha\beta}. \tag{4.0.11}$$

$d \neq 4$ Can you comment what the analog of the magnetic field ought to be in spacetime dimensions different from 4 – is it still a 'vector'? – and what is the lowest dimension that the magnetic field still exists? How many components does the electric field have in 1+1 dimensions? □

Current conservation Taking the divergence of $\partial_\mu F^{\mu\nu} = J^\nu$ yields the conservation of the electric current as a consistency condition. For, by the antisymmetry $F_{\mu\nu} = -F_{\nu\mu}$, $\partial_\nu \partial_\mu F^{\mu\nu} = (1/2)\partial_\nu \partial_\mu F^{\mu\nu} - (1/2)\partial_\mu \partial_\nu F^{\nu\mu} = 0$.

$$\partial_\mu J^\mu = 0. \tag{4.0.12}$$

Problem 4.2. Total charge is constant in all inertial frames Even though we defined ρ in the $J^\mu = \rho v^\mu$ as the charge density in the local rest frame of the electric current itself, we may also define the charge density $\widehat{J^0} \equiv u_\mu J^\mu$ in the rest frame of an arbitrary family of inertial time-like observers whose worldlines' tangent vector is $u^\mu \partial_\mu = \partial_\tau$. (In other words, in their frame, the spacetime metric is $\mathrm{d}s^2 = (\mathrm{d}\tau)^2 - \mathrm{d}\vec{x} \cdot \mathrm{d}\vec{x}$.) Show that total charge is independent of the Lorentz frame by demonstrating that

$$Q \equiv \int_{\mathbb{R}^D} \mathrm{d}^D \Sigma_\mu J^\mu, \qquad \mathrm{d}^D \Sigma_\mu \equiv \mathrm{d}^D \vec{x} u_\mu, \quad D \equiv d - 1, \tag{4.0.13}$$

is a constant. □

Vector Potential & Gauge Symmetry The other Maxwell equation (cf eq. (7.1.1)) leads us to introduce a vector potential A_μ. For $\partial_{[\mu} F_{\alpha\beta]} = 0 \Leftrightarrow \mathrm{d}F = 0$ tells us, by the Poincaré lemma, that

$$F = \mathrm{d}A \qquad \Leftrightarrow \qquad F_{\mu\nu} = \partial_{[\mu} A_{\nu]} = \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{4.0.14}$$

Notice the dynamics in eq. (7.1.1) is not altered if we add to A_μ any object L_μ that obeys $\mathrm{d}L = 0$, because that does not alter the Faraday tensor: $F = \mathrm{d}(A + L) = F + \mathrm{d}L = F$. Now, $\mathrm{d}L = 0$ means, again by the Poincaré lemma, that $L_\mu = \partial_\mu L$, where L on the right hand side is a scalar. *Gauge symmetry*, in the context of electromagnetism, is the statement that the following replacement involving the gauge potential

$$A_\mu(x) \to A_\mu(x) + \partial_\mu L(x) \tag{4.0.15}$$

leaves the dynamics encoded in Maxwell's equations (7.1.1) unchanged.

The use of the gauge potential A_μ makes the $\mathrm{d}F = 0$ portion of the dynamics in eq. (7.1.1) redundant; and what remains is the vector equation

$$\partial_\mu F^{\mu\nu} = \partial_\mu \left(\partial^\mu A^\nu - \partial^\nu A^\mu\right) = J^\nu. \tag{4.0.16}$$

The symmetry under the gauge transformation of eq. (4.0.15) means that solutions to eq. (4.0.16) cannot be unique – in particular, since A_μ and $A_\mu + \partial_\mu L$ are simultaneously solutions, there really is an infinity of solutions parametrized by the arbitrary function L. In this same vein, by going to Fourier space, namely

$$A_\mu(x) \equiv \int_{\mathbb{R}^d} \frac{\mathrm{d}^d k}{(2\pi)^d} \widetilde{A}_\mu(k) e^{-ik_\mu x^\mu} \qquad \text{and} \qquad J_\mu(x) \equiv \int_{\mathbb{R}^d} \frac{\mathrm{d}^d k}{(2\pi)^d} \widetilde{J}_\mu(k) e^{-ik_\mu x^\mu}, \tag{4.0.17}$$

we may see that the differential operator in eq. (4.0.16) cannot be inverted because it has a zero eigenvalue. Firstly, the Fourier version of eq. (4.0.16) reads

$$-K^{\mu\nu} \widetilde{A}_\mu = \widetilde{J}^\nu, \tag{4.0.18}$$

$$K^{\mu\nu} \equiv k_\sigma k^\sigma \eta^{\mu\nu} - k^\nu k^\mu. \tag{4.0.19}$$

If K^{-1} exists, the solution in Fourier space would be (schematically) $\widetilde{A} = -K^{-1}\widetilde{J}$. However, since $K^{\mu\nu} = K^{\nu\mu}$ is a real symmetric matrix, it must be diagonalizable via an orthogonal transformation, with $\det K^{\mu\nu}$ equal to the product of its eigenvalues. That $\det K^{\mu\nu} = 0$ and therefore K^{-1} does not exist can now be seen by observing that k_μ is in fact its null eigenvector:

$$K^{\mu\nu} k_\mu = (k_\sigma k^\sigma) k^\nu - k^\nu k^\mu k_\mu = 0. \tag{4.0.20}$$

Problem 4.3. Can you explain why eq. (4.0.20) amounts to the statement that $F_{\mu\nu}$ is invariant under the gauge transformation of eq. (4.0.15)? Hint: Consider eq. (4.0.15) in Fourier space.

Lorenz gauge To make $K^{\mu\nu}$ invertible, one *fixes a gauge*. A common choice is the Lorenz gauge; in Fourier spacetime:

$$k^\mu \widetilde{A}_\mu = 0. \tag{4.0.21}$$

In 'position'/real spacetime, this reads instead

$$\partial^\mu A_\mu = 0 \qquad \text{(Lorenz gauge)}. \tag{4.0.22}$$

With the constraint in eq. (4.0.21), Maxwell's equations in eq. (4.0.18) becomes

$$-\left(k_\sigma k^\sigma \widetilde{A}^\nu - k^\nu (k^\mu \widetilde{A}_\mu)\right) = -k_\sigma k^\sigma \widetilde{A}^\nu = \widetilde{J}^\nu. \tag{4.0.23}$$

Now, Maxwell's equations have become invertible:

$$\widetilde{A}_\mu(k) = \frac{\widetilde{J}_\mu(k)}{-k^2}, \qquad k^2 \equiv k_\sigma k^\sigma, \qquad \text{(Lorenz gauge)}. \tag{4.0.24}$$

In position/real spacetime, eq. (4.0.23) is equivalent to

$$\partial^2 A^\nu(x) = J^\nu(x) \qquad \partial^2 \equiv \eta^{\mu\nu}\partial_\mu\partial_\nu. \tag{4.0.25}$$

[1]In the Lorenz gauge, we have d Minkowski scalar wave equations, one for each Cartesian component. We may express its position spacetime solution by inverting the Fourier transform in eq. (4.0.24):

$$A_\mu(x) = \int_{\mathbb{R}^{d-1,1}} \mathrm{d}^d x' G_d^+(x-x') J_{\mu'}(x'), \tag{4.0.26}$$

$$G_d^+(x-x') \equiv \int_{\mathbb{R}^d} \frac{\mathrm{d}^d k}{(2\pi)^d} \frac{e^{-ik\cdot(x-x')}}{-k^2}. \tag{4.0.27}$$

Because A_μ is not gauge-invariant, its physical interpretation can be ambiguous. Classically it is the electromagnetic fields $F_{\mu\nu}$ that exert forces on charges/currents, so we need its solution. In fact, we may take the curl of eq. (4.0.25) to see that

$$\partial^2 F_{\mu\nu} = \partial_{[\mu} J_{\nu]}; \tag{4.0.28}$$

this means, using the same Green's function in eq. (4.0.27):

$$F_{\mu\nu}(x) = \int_{\mathbb{R}^{d-1,1}} \mathrm{d}^d x' G_d^+(x-x') \partial_{[\mu'} J_{\nu']}(x'). \tag{4.0.29}$$

We may verify that equations (4.0.26) and (4.0.27) solve eq. (4.0.25) readily:

$$\begin{aligned}
\partial_x^2 G_d^+(x-x') &= \int_{\mathbb{R}^d} \frac{\mathrm{d}^d k}{(2\pi)^d} \frac{\partial_\sigma \partial^\sigma e^{-ik\cdot(x-x')}}{-k^2} \\
&= \int_{\mathbb{R}^d} \frac{\mathrm{d}^d k}{(2\pi)^d} \frac{\partial_\sigma(-ik^\rho \delta_\rho^\sigma e^{-ik\cdot(x-x')})}{-k^2} \\
&= \int_{\mathbb{R}^d} \frac{\mathrm{d}^d k}{(2\pi)^d} \frac{\partial_\sigma(-ik^\sigma e^{-ik\cdot(x-x')})}{-k^2}
\end{aligned}$$

[1]Eq. (4.0.25) is valid in any dimension $d \geq 3$. In 2D, the $\mathrm{d}F = 0$ portion of Maxwell's equations is trivial – i.e., *any* F would satisfy it – because there cannot be three distinct indices in $\partial_{[\mu} F_{\alpha\beta]} = 0$.

$$= \int_{\mathbb{R}^d} \frac{\mathrm{d}^d k}{(2\pi)^d} \frac{(-ik_\sigma)(-ik^\sigma) e^{-ik\cdot(x-x')}}{-k^2}$$
$$= \int_{\mathbb{R}^d} \frac{\mathrm{d}^d k}{(2\pi)^d} e^{-ik\cdot(x-x')} = \delta^{(d)}(x-x'); \tag{4.0.30}$$

with a similar calculation showing $\partial^2_{x'} G^+_d(x-x') = \delta^{(d)}(x-x')$. To sum,

$$\partial^2_x G^+_d(x-x') = \partial^2_{x'} G^+_d(x-x') = \delta^{(d)}(x-x'); \tag{4.0.31}$$

Moreover, comparing each Cartesian component of the wave equation in eq. (4.0.25) with the one obeyed by the Green's function in eq. (4.0.31), we may identify the source J of the Green's function itself to be a unit strength spacetime point source at some fixed location x'. It is often useful to think of x as the spacetime location of some observer; so $x^0 - x'^0 \equiv t - t'$ is the time elapsed while $|\vec{x} - \vec{x}'|$ is the observer-source spatial distance. Altogether, we may now view the solution in eq. (4.0.26) as the sum of the field generated by all spacetime point sources, weighted by the physical electric current $J_\mu(x')$.

We now may verify directly that eq. (4.0.26) is indeed a solution to eq. (4.0.25).

$$\partial^2_x A_\mu(x) = \partial^2_x \left(\int_{\mathbb{R}^{d-1,1}} \mathrm{d}^d x' G^+_d(x-x') J_{\mu'}(x') \right) = \int_{\mathbb{R}^{d-1,1}} \mathrm{d}^d x' \delta^{(d)}(x-x') J_{\mu'}(x')$$
$$= J_\mu(x). \tag{4.0.32}$$

Lorenz gauge: Existence That we have managed to solve Maxwell's equations using the Lorenz gauge, likely convinces the practical physicist that the Lorenz gauge itself surely exists. However, it is certainly possible to provide a general argument. For suppose $\partial^\mu A_\mu$ were not zero, then all one has to show is that we may perform a gauge transformation (cf. (4.0.15)) that would render the new gauge potential $A'_\mu \equiv A_\mu - \partial_\mu L$ satisfy

$$\partial^\mu A'_\mu = \partial^\mu A_\mu - \partial^2 L = 0. \tag{4.0.33}$$

But all that means is, we have to solve $\partial^2 L = \partial^\mu A_\mu$; and since the Green's function $1/\partial^2$ exists, we have proved the assertion.

Lorenz gauge and current conservation You may have noticed, by taking the divergence of both sides of eq. (4.0.25),

$$\partial^2 \left(\partial^\sigma A_\sigma\right) = \partial^\sigma J_\sigma. \tag{4.0.34}$$

This teaches us the consistency of the Lorenz gauge is intimately tied to the conservation of the electric current $\partial^\sigma J_\sigma = 0$. Another way to see this, is to take the time derivative of the divergence of the vector potential, followed by subtracting and adding the spatial Laplacian of A_0 so that $\partial^2 A_0 = J_0$ may be employed:

$$\partial^\sigma \dot{A}_\sigma = \ddot{A}_0 + \partial^i \dot{A}_i = \partial^0 \partial_0 A_0 + \partial^i \partial_i A_0 + \partial^i \partial_0 A_i - \partial^i \partial_i A_0$$
$$= \partial^2 A_0 - \partial^i \left(\partial_i A_0 - \partial_0 A_i\right)$$
$$\partial_0 \left(\partial^\sigma A_\sigma\right) = J_0 - \partial^i F_{i0}. \tag{4.0.35}$$

Notice the right hand side of the last line is zero if the $\nu = 0$ component of $\partial_\mu F^{\mu\nu} = J^\nu$ is obeyed – and if the latter is obeyed the 'left-hand-side' of Lorenz gauge condition $\partial_\mu A^\mu$ is a time independent quantity.

4.1 4 dimensions

4D Maxwell We now focus on the physically most relevant case of $(3+1)D$. In 4D, the wave operator ∂^2 has the following inverse – i.e., retarded Green's function – that obeys causality:

$$G_4^+(x-x') \equiv \frac{\delta(t-t'-|\vec{x}-\vec{x}'|)}{4\pi|\vec{x}-\vec{x}'|}, \qquad x^\mu = (t,\vec{x}),\ x'^\mu = (t',\vec{x}'), \tag{4.1.1}$$

$$\partial_x^2 G_4^+(x-x') = \partial_{x'}^2 G_4^+(x-x') = \delta^{(4)}(x-x'), \tag{4.1.2}$$

$$\partial_x^2 \equiv \eta^{\mu\nu}\frac{\partial}{\partial x^\mu}\frac{\partial}{\partial x^\nu} \qquad \partial_{x'}^2 \equiv \eta^{\mu\nu}\frac{\partial}{\partial x'^\mu}\frac{\partial}{\partial x'^\nu}. \tag{4.1.3}$$

To see that G_4^+ obeys causality, that it respects the principle that cause precedes effect, one merely needs to focus on the δ-function in eq. (4.1.1). It is non-zero only when the time elapsed $t-t'$ is precisely equal to the observer-source distance $|\vec{x}-\vec{x}'|$. That is, if the source is located at a spatial distance $R = |\vec{x}-\vec{x}'|$ away from the observer, and if the source emitted an instantaneous flash at time t', then the observer would see a signal at time R later (i.e., at $t = t'+R$). In other words, the retarded Green's function propagates signals on the *forward* light cone of the source.[2]

Problem 4.4. Lorentz covariance Suppose $\Lambda^\alpha{}_\mu$ is a Lorentz transformation; let two inertial frames $\{x^\mu\}$ and $\{x'^\mu\}$ be related via

$$x^\mu = \Lambda^\mu{}_\alpha x'^\alpha. \tag{4.1.4}$$

Suppose we solved the Lorenz gauge Maxwell's equations in the $\{x^\mu\}$ frame, namely

$$\frac{\partial A^\mu(x)}{\partial x^\mu} = 0, \qquad \eta^{\mu\nu}\frac{\partial}{\partial x^\mu}\frac{\partial}{\partial x^\nu}A_\alpha(x) = J_\alpha(x). \tag{4.1.5}$$

Explain how to solve $A_{\alpha'}(x')$, the solution in the $\{x'^\mu\}$ frame. □

Problem 4.5. Analogy: Driven Simple Harmonic Oscillator Suppose we only Fourier-transformed the spatial coordinates in the Lorenz gauge Maxwell eq. (4.0.25). Show that this leads to

$$\ddot{\widetilde{A}}_\mu(t,\vec{k}) + k^2\widetilde{A}_\mu(t,\vec{k}) = \widetilde{J}_\mu(t,\vec{k}), \qquad k \equiv |\vec{k}|. \tag{4.1.6}$$

[3]Compare this to the simple harmonic oscillator (in flat space), with Cartesian coordinate vector $\vec{q}(t)$, mass m, spring constant σ, and driven by an external force $\vec{f}$:

$$m\ddot{\vec{q}} + \sigma\vec{q} = \vec{f}, \tag{4.1.7}$$

where each over-dot corresponds to a time derivative. Identify k^2 and $\widetilde{J}$ in eq. (4.1.6) with the appropriate quantities in eq. (4.1.7). Even though the Lorenz gauge Maxwell equations are fully

[2]The advanced Green's function $G_4^-(x-x') = \delta(t-t'+|\vec{x}-\vec{x}'|)/(4\pi|\vec{x}-\vec{x}'|)$ also solves eq. (4.0.31), but propagates signals on the past light cone: $t = t'-R$.

[3]This equation actually holds in all dimensions $d \geq 3$.

relativistic, notice the analogy with the non-relativistic driven harmonic oscillator! In particular, when the electric current is not present (i.e., $J_\mu = 0$), the 'mixed-space' equations of (4.1.6) are in fact a collection of free simple harmonic oscillators.

Now, how does one solve eq. (4.1.7)? Explain why the inverse of $(\mathrm{d}/\mathrm{d}t)^2 + k^2$ is

$$G_{\text{SHO}}(t-t', k) = -\int_{\mathbb{R}} \frac{\mathrm{d}\omega}{2\pi} \frac{e^{-i\omega(t-t')}}{\omega^2 - k^2}. \tag{4.1.8}$$

That is, verify that this equation satisfies

$$\left(\frac{\mathrm{d}^2}{\mathrm{d}t^2} + k^2\right) G_{\text{SHO}}(t-t', k) = \left(\frac{\mathrm{d}^2}{\mathrm{d}t'^2} + k^2\right) G_{\text{SHO}}(t-t', k) = \delta(t-t'). \tag{4.1.9}$$

If one tries to integrate ω over the real line in eq. (4.1.8), one runs into trouble – explain the issue. In other words, eq. (4.1.8) is actually ambiguous as it stands.

Now evaluate the Green's function G^+_{SHO} in eq. (4.1.8) using the contour running just slightly above the real line, i.e., $\omega \in (-\infty + i0^+, +\infty + i0^+)$. You should find

$$G^+_{\text{SHO}}(t-t', k) = \Theta(t-t') \frac{\sin(k(t-t'))}{k}. \tag{4.1.10}$$

Here, Θ is the step function

$$\Theta(x) = 1, \qquad \text{if } x > 0, \tag{4.1.11}$$
$$= 0, \qquad \text{if } x < 0. \tag{4.1.12}$$

Hence, the mixed-space Maxwell's equations have the solution

$$\widetilde{A}_\mu(t, \vec{k}) = \int_{-\infty}^{t} \mathrm{d}t' G^+_{\text{SHO}}(t-t', k) \widetilde{J}_\mu(t', \vec{k}). \tag{4.1.13}$$

By performing an inverse-Fourier transform, namely

$$A_\mu(x) = \int_{\mathbb{R}^{3,1}} \mathrm{d}^4 x' G^+_4(x-x') J_{\mu'}(x'), \tag{4.1.14}$$

arrive at the expression in eq. (4.1.1) $\square$

Vacuum solution & Massless Spin-1 (Helicity-1) Let us examine the simplest situation in 4D flat spacetime, where there are no electric charges nor currents present: $J_\nu = 0$. In Fourier space, setting $\widetilde{J} = 0$ in eq. (4.1.6) leads us to

$$\ddot{\widetilde{A}}_\mu(t, \vec{k}) + k^2 \widetilde{A}_\mu(t, \vec{k}) = 0, \qquad k \equiv |\vec{k}|. \tag{4.1.15}$$

These are the free simple harmonic oscillators alluded to earlier. The solutions are $\widetilde{A}_\mu(t, \vec{k}) = \exp(\pm ikt)$ for $k \equiv |\vec{k}| \geq 0$. Hence, the general solution is the superposition

$$\begin{aligned} A_\mu &= \int_{\mathbb{R}^3} \frac{\mathrm{d}^3\vec{k}}{(2\pi)^3} \left(a_\mu(\vec{k}) \exp(-ikt + i\vec{k}\cdot\vec{x}) + b_\mu(\vec{k}) \exp(ikt + i\vec{k}\cdot\vec{x})\right) \\ &= \int_{\mathbb{R}^3} \frac{\mathrm{d}^3\vec{k}}{(2\pi)^3} \left(a_\mu(\vec{k}) \exp(-ikt + i\vec{k}\cdot\vec{x}) + b_\mu(-\vec{k}) \exp(ikt - i\vec{k}\cdot\vec{x})\right). \end{aligned} \tag{4.1.16}$$

Referring to eq. (4.0.25), since J_μ is real, so is A_μ. Thus it must be that $a_\mu(\vec{k})^* = b_\mu(-\vec{k})$:

$$A_\mu = \int_{\mathbb{R}^3} \frac{\mathrm{d}^3\vec{k}}{(2\pi)^3} \left(a_\mu(\vec{k}) e^{-ik\cdot x} + a_\mu(\vec{k})^* e^{ik\cdot x} \right). \tag{4.1.17}$$

Since a_μ has been arbitrary thus far, we may write a single plane wave solution to eq. (4.1.6) as

$$\begin{aligned} \mathrm{Re}\left\{ \widetilde{A}_\mu(t,\vec{k}) e^{i\vec{k}\cdot\vec{x}} \right\} = \mathrm{Re}\left\{ \epsilon_\mu(\vec{k}) e^{-ik\cdot t} e^{i\vec{k}\cdot\vec{x}} \right\} = \mathrm{Re}\left\{ \epsilon_\mu(\vec{k}) e^{-ik\cdot x} \right\}, \\ k_\mu \equiv (k, k_i), \qquad\qquad k \equiv |\vec{k}| = \sqrt{\delta^{ab} k_a k_b}. \end{aligned} \tag{4.1.18}$$

The Lorenz gauge says $k^\mu \widetilde{A}_\mu = 0$. Since the $\exp(-ik_\mu x^\mu)$ are basis functions, it must be that the polarization vector ϵ_μ itself is orthogonal to the momentum vector k^μ:

$$k^\mu \epsilon_\mu(\vec{k}) = 0. \tag{4.1.19}$$

Let us suppose k_i points in the positive 3−axis, so that

$$k_\mu = k(1,0,0,-1) \qquad \text{and} \qquad k^\mu = k(1,0,0,1). \tag{4.1.20}$$

This means the plane wave itself becomes

$$\exp(-ik_\mu x^\mu) = \exp(-ik(t - x^3)); \tag{4.1.21}$$

i.e., it indeed describes propagation in the positive 3−direction. The polarization vector may then be decomposed as follows:

$$\epsilon_\mu = \kappa_+ \ell^+{}_\mu + \kappa_- \ell^-{}_\mu + a_+ \epsilon^+{}_\mu + a_- \epsilon^-{}_\mu; \tag{4.1.22}$$

where the κ and a's are (scalar) complex amplitudes; the null basis vectors $\ell^\pm$ are

$$\ell^\pm{}_\mu \equiv \frac{1}{\sqrt{2}} (1,0,0,\pm 1)^{\mathrm{T}}; \tag{4.1.23}$$

wheres the spatial basis vectors $\epsilon^\pm$ are

$$\epsilon^\pm{}_\mu \equiv \frac{1}{\sqrt{2}} (0,\mp 1, i, 0)^{\mathrm{T}}. \tag{4.1.24}$$

Now, under the following rotation on the $(1,2)$-plane orthogonal to $\vec{k}$, namely

$$\widehat{R}(\theta)^\mu{}_\nu \doteq \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) & 0 \\ 0 & \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{4.1.25}$$

the null polarization vectors in eq. (4.1.23) remain unchanged ($\widehat{R}(\theta)^\mu{}_\nu \ell^{\pm\nu} = \ell^{\pm\mu}$) – they are the spin-0 modes – while the spatial polarizations in eq. (4.1.24) transform as

$$\epsilon^\pm{}_\mu \widehat{R}(\theta)^\mu{}_\nu = e^{-i(\pm 1)\theta} \epsilon^\pm{}_\nu. \tag{4.1.26}$$

These $\epsilon^\pm{}_\nu$ are the helicity-1 modes.

Problem 4.6. Verify eq. (4.1.26).

We now turn to imposing the Lorenz gauge condition $k^\mu \widetilde{A}_\mu = 0$.

$$k\epsilon_0 + k\epsilon_3 = 0 \qquad \Rightarrow \qquad \epsilon_3 = -\epsilon_0. \tag{4.1.27}$$

Since the 0th component has to be negative the 3rd, the ℓ^+ cannot occur in the decomposition of eq. (4.1.22). But since ℓ^- is proportional to k_μ (cf. eq. (4.1.23)) and $k^2 \equiv k_\nu k^\nu = 0$, we see this remaining spin-0 piece of the polarization tensor simultaneously satisfies the Lorenz gauge and is a gradient term – and hence 'pure gauge' (cf. the $\partial_\mu L$ terms of eq. (4.0.15)) – in position spacetime:

$$\kappa_- \ell^-{}_\mu = \frac{\kappa_-}{\sqrt{2}} \frac{k_\mu}{k}. \tag{4.1.28}$$

Since this term will not contribute to the electromagnetic fields $F_{\mu\nu}$, we may perform a Lorenz-gauge-preserving gauge transformation to cancel this term:

$$\text{Re } \widetilde{A}'_\mu(t,\vec{k}) e^{i\vec{k}\cdot\vec{x}} \equiv \text{Re}\left\{ \epsilon_\mu(\vec{k}) e^{-ik\cdot x} - \frac{\kappa_-}{\sqrt{2}} \frac{k_\mu}{k} e^{-ik\cdot x} \right\}. \tag{4.1.29}$$

And now that we have canceled the 0th and 3rd component of the polarization vector in eq. (4.1.22),

$$\text{Re } \widetilde{A}'_\nu(t,\vec{k}) e^{i\vec{k}\cdot\vec{x}} = \text{Re}\left\{ \left(a_+ \epsilon^+{}_\mu + a_- \epsilon^-{}_\mu \right) e^{-ik\cdot x} \right\}. \tag{4.1.30}$$

General Case When k_i is not necessarily (anti)parallel to the 3−axis, we may continue to define $\ell^-{}_\mu$ to be the normalized version of k_μ, i.e.,

$$\ell^-{}_\mu \equiv \frac{k_\mu}{\sqrt{2}k}, \qquad k \equiv k_0. \tag{4.1.31}$$

The ℓ^+, on the other hand, is the solution to the constraints

$$\ell^+ \cdot \ell^- \equiv \eta^{\mu\nu} \ell^+{}_\mu \ell^-{}_\nu = +1, \tag{4.1.32}$$

$$\epsilon^{(1)} \cdot \ell^+ = \epsilon^{(2)} \cdot \ell^+ = 0; \tag{4.1.33}$$

where the $\epsilon^{(1)}$ and $\epsilon^{(2)}$ are themselves mutually orthogonal spatial basis vectors perpendicular to ℓ^+ – namely

$$\epsilon^{(\mathrm{I})} \cdot \epsilon^{(\mathrm{J})} = -\delta^{\mathrm{IJ}}, \qquad \mathrm{I}, \mathrm{J} \in \{1,2\} \tag{4.1.34}$$

$$\epsilon^{(1)} \cdot \ell^- = \epsilon^{(2)} \cdot \ell^- = 0. \tag{4.1.35}$$

The spin−1 basis vectors can be constructed from the $\epsilon^{(\mathrm{I})}$ via the definitions

$$\epsilon^\pm \equiv \frac{\mp 1}{\sqrt{2}} \epsilon^{(1)} + \frac{i}{\sqrt{2}} \epsilon^{(2)}. \tag{4.1.36}$$

Altogether, the Minkowski metric would obey the following completeness relation

$$\eta_{\mu\nu} = \ell^+{}_{\{\mu} \ell^-{}_{\nu\}} - \epsilon^{(1)}{}_\mu \epsilon^{(1)}{}_\nu - \epsilon^{(2)}{}_\mu \epsilon^{(2)}{}_\nu \tag{4.1.37}$$

$$= \ell^+{}_{\{\mu} \ell^-{}_{\nu\}} + \epsilon^+{}_{\{\mu} \epsilon^-{}_{\nu\}}. \tag{4.1.38}$$

(3+1)D Spin-1 Waves To sum, given an inertial frame, the electromagnetic vector potential A_μ in vacuum is given by the following superposition of spin-1 waves:

$$A_\mu(x) = \text{Re} \int_{\mathbb{R}^3} \frac{\text{d}^3\vec{k}}{(2\pi)^3} \left(a_+ {\epsilon^+}_\mu(\vec{k}) + a_- {\epsilon^-}_\mu(\vec{k}) \right) e^{-ik\cdot x}, \tag{4.1.39}$$

where ${\epsilon^\pm}_\mu$ are purely spatial polarization tensors orthogonal to the k_i; and, under a rotation by an angle θ around the plane perpendicular to k_i transforms as $\epsilon^\pm \to \exp(-i(\pm 1)\theta)\epsilon^\pm$.[4]

Problem 4.7. Circularly Polarized Light from 4D Spin-1 Consider a single spin-1 plane wave (cf. (4.1.24)) propagating along the 3−axis, with $k_\mu = k(1,0,0,-1)$:

$$A_\mu^\pm(t,x,y,z) \equiv \text{Re}\left\{ a_\pm {\epsilon^\pm}_\mu e^{-ik(t-z)} \right\}, \qquad a_\pm \in \mathbb{R}. \tag{4.1.40}$$

Compute the electric field ${}_\pm E^i = F^{i0}$ and show that these plane waves give rise to circularly polarized light, i.e., for either a fixed time t or spatial location z – the electric field direction rotates in a circular fashion:

$${}_\pm E^i = \frac{k a_\pm}{\sqrt{2}} \left(\pm \sin(k(t-z))\widehat{x}^i + \cos(k(t-z))\widehat{y}^i \right), \tag{4.1.41}$$

where $\widehat{x}$ and $\widehat{y}$ are the unit vectors in the 1− and 2−directions:

$$\widehat{x}^i \doteq (1,0,0) \qquad \text{and} \qquad \widehat{y}^i \doteq (0,1,0). \tag{4.1.42}$$

□

Redshift For each Lorenz-gauge plane wave in an inertial frame $\{x^\mu = (t,\vec{x})\}$,

$${\epsilon^\pm}_\mu(k)\exp(-ik\cdot x) = {\epsilon^\pm}_\mu(k)\exp(-ik_j x^j)\exp(-i\omega t), \qquad \omega \equiv |\vec{k}|, \tag{4.1.43}$$

we may read off its frequency ω as the coefficient of the time coordinate t. Quantum mechanics tells us ω is also the energy of the associated photon. Suppose a different Lorentz inertial frame $\{x'\}$ is related to the previous through the Lorentz transformation ${\Lambda^\alpha}_\mu$: $x^\alpha = {\Lambda^\alpha}_\mu x'^\mu$. Because the phase in the plane wave solution of eq. (4.1.43) is a scalar, in the $\{x'\}$ Lorentz frame

$$-ik_\alpha x^\alpha = -ik_\alpha {\Lambda^\alpha}_\mu x'^\mu = -i(k_\alpha {\Lambda^\alpha}_0)t' - i(k_\alpha {\Lambda^\alpha}_i)x'^i. \tag{4.1.44}$$

The frequency ω' and hence the photon's energy in this $\{x'\}$ frame is therefore

$$\omega' = k_\alpha {\Lambda^\alpha}_0 = \omega\left({\Lambda^0}_0 + \widehat{k}_i {\Lambda^i}_0 \right) \tag{4.1.45}$$

$$\widehat{k}_i \equiv k_i/|\vec{k}| = k_i/\omega. \tag{4.1.46}$$

[4]For a given inertial frame and within the Lorenz gauge, we have been able to get rid of the 'pure gauge' spin-0 mode by a gauge transformation, leaving only the massless spin-1 (simple-harmonic) waves. Note however, these waves in eq. (4.1.39) would no longer be an admixture of pure spin-1 modes – simply by viewing them in a different reference frame, i.e., upon a Lorentz boost.

There is a slightly different way to express this redshift result that would help us generalize the analysis to curved spacetime, at least in the high frequency 'JWKB' limit. To extract the frequency directly from the phase $S \equiv k \cdot x$, we may take its time derivative using the unit norm vector $u \equiv \partial_t = \partial_0$ that we may associate with the worldlines of observers at rest in the $\{x\}$ frame:

$$u^\mu \partial_\mu S = \partial_0 (k_\alpha x^\alpha) = \omega. \tag{4.1.47}$$

The observers at rest in the $\{x'\}$ frame have $u' \equiv \partial_{t'} = \partial_{0'}$ as their timelike unit norm tangent vector. (Note: $x^\alpha = \Lambda^\alpha{}_\mu x'^\mu \Leftrightarrow \partial_{\mu'} = \Lambda^\alpha{}_\mu \partial_\alpha$.) The energy of the photon is then

$$\begin{aligned} u'^\alpha \partial_{\alpha'} S = \partial_{t'} S &= \Lambda^\alpha{}_0 \partial_\alpha (k \cdot x) \\ &= \Lambda^\alpha{}_0 k_\alpha = \omega \left(\Lambda^0{}_0 + \widehat{k}_i \Lambda^i{}_0 \right). \end{aligned} \tag{4.1.48}$$

Problem 4.8. Consider a single photon with wave vector $k_\mu = \omega(1, \widehat{n}_i)$ (where $\widehat{n}_i \widehat{n}_j \delta^{ij} = 1$) in some inertial frame $\{x^\mu\}$. Let a family of inertial observers be moving with constant velocity $v^\mu \equiv (1, v^i)$ with respect to the frame $\{x^\mu\}$. What is the photon's frequency ω' in their frame? Compute the redshift formula for ω'/ω. Comment on the redshift result when v^i is (anti)parallel to $\widehat{n}_i$ and when v^i is perpendicular to $\widehat{n}_i$.

Problem 4.9. Dispersion relations Consider the *massive* Klein-Gordon equation in Minkowski spacetime:

$$\left(\partial^2 + m^2\right) \varphi(t, \vec{x}) = 0, \tag{4.1.49}$$

where φ is a real scalar field. Find the general solution for φ in terms of plane waves $\exp(-ik \cdot x)$ and obtain the dispersion relation:

$$k^2 = m^2 \qquad \Leftrightarrow \qquad E^2 = \vec{p}^2 + m^2, \tag{4.1.50}$$

$$E \equiv k^0, \qquad \qquad \vec{p} \equiv \vec{k}. \tag{4.1.51}$$

If each plane wave is associated with a particle of $d-$momentum k_μ, this states that it has mass m. The photon, which obeys $k^2 = 0$, has zero mass.

Bonus: Can you restore the factors of $\hbar$ and c in eq. (6.5.24)? $\square$

4.2 Gauge Invariant Variables for Vector Potential

Although the vector potential A_μ itself is not a gauge invariant object, we will now exploit the spatial translation symmetry of Minkowski spacetime to seek a gauge-invariant set of partial differential equations involving a "scalar-vector" decomposition of A_μ. There are at least two reasons for doing so.

- This will be a warm-up to an analogous analysis for gravitation linearized about a Minkowski "background" spacetime.

- We will witness how, for a given inertial frame, the only portion of the vector potential A_μ that obeys a wave equation is its gauge-invariant "transverse" spatial portion. (Even though every component of A_μ in the Lorenz gauge (cf. eq. (4.0.25)) obeys the wave equation, remember such a statement is not gauge-invariant.) We shall also identify a gauge-invariant scalar potential sourced by charge density.

Scalar-Vector Decomposition The scalar-vector decomposition is the statement that the spatial components of the vector potential may be expressed as a gradient of a scalar α plus a transverse vector α_i:

$$A_i = \partial_i \alpha + \alpha_i, \tag{4.2.1}$$

where by "transverse" we mean

$$\partial_i \alpha_i = 0. \tag{4.2.2}$$

To demonstrate the generality of eq. (4.2.1) we shall first write A_i in Fourier space

$$A_i(t, \vec{x}) = \int_{\mathbb{R}^D} \frac{\mathrm{d}^D \vec{k}}{(2\pi)^D} \widetilde{A}_i(t, \vec{k}) e^{i\vec{k}\cdot\vec{x}}; \tag{4.2.3}$$

where $\vec{k} \cdot \vec{x} \equiv \delta_{ij} k^i x^j = -k_j x^j$. Every spatial derivative ∂_j acting on $A_i(t, \vec{x})$ becomes in Fourier space a $-ik_j$, since

$$\begin{aligned}
\partial_j A_i &= \int_{\mathbb{R}^D} \frac{\mathrm{d}^D \vec{k}}{(2\pi)^D} \partial_j \left(i\delta_{ab} k^a x^b \right) \widetilde{A}_i(t, \vec{k}) e^{i\vec{k}\cdot\vec{x}} \\
&= \int_{\mathbb{R}^D} \frac{\mathrm{d}^D \vec{k}}{(2\pi)^D} \left(i\delta_{ab} k^a \delta^b_j \right) \widetilde{A}_i(t, \vec{k}) e^{i\vec{k}\cdot\vec{x}} \\
&= \int_{\mathbb{R}^D} \frac{\mathrm{d}^D \vec{k}}{(2\pi)^D} ik^j \widetilde{A}_i(t, \vec{k}) e^{i\vec{k}\cdot\vec{x}} \\
&= \int_{\mathbb{R}^D} \frac{\mathrm{d}^D \vec{k}}{(2\pi)^D} (-ik_j) \widetilde{A}_i(t, \vec{k}) e^{i\vec{k}\cdot\vec{x}}.
\end{aligned} \tag{4.2.4}$$

As such, the transverse property of $\alpha_i(t, \vec{x})$ would in Fourier space become

$$-ik_i \widetilde{\alpha}_i(t, \vec{k}) = 0. \tag{4.2.5}$$

At this point we simply write down

$$\widetilde{A}_i(t, \vec{k}) = \left(\delta_{ij} - \frac{k_i k_j}{\vec{k}^2} \right) \widetilde{A}_j(t, \vec{k}) + \frac{k_i k_j}{\vec{k}^2} \widetilde{A}_j(t, \vec{k}). \tag{4.2.6}$$

This is mere tautology, of course. However, we may now check that the first term on the left hand side of eq. (4.2.6) is transverse:

$$-ik_i \left(\delta_{ij} - \frac{k_i k_j}{\vec{k}^2} \right) \widetilde{A}_j(t, \vec{k}) = -i \left(k_j - \frac{\vec{k}^2 k_j}{\vec{k}^2} \right) \widetilde{A}_j(t, \vec{k}) = 0. \tag{4.2.7}$$

The second term on the right hand side of eq. (4.2.6) is a gradient because it is

$$-ik_i\left(\frac{ik_j}{\vec{k}^2}\widetilde{A}_j\right). \tag{4.2.8}$$

To sum, we have identified the α and α_i terms of eq. (4.2.1) as

$$\alpha(t,\vec{x}) = \int_{\mathbb{R}^D}\frac{\mathrm{d}^D\vec{k}}{(2\pi)^D}\frac{ik_j}{\vec{k}^2}\widetilde{A}_j(t,\vec{k})e^{i\vec{k}\cdot\vec{x}}; \tag{4.2.9}$$

and the transverse portion as

$$\begin{aligned}\alpha_i(t,\vec{x}) &= \int_{\mathbb{R}^D}\frac{\mathrm{d}^D\vec{k}}{(2\pi)^D}P_{ij}(\vec{k})\widetilde{A}_j(t,\vec{k})e^{i\vec{k}\cdot\vec{x}}, \\ P_{ij}(\vec{k}) &\equiv \delta_{ij} - \frac{k_i k_j}{\vec{k}^2}.\end{aligned} \tag{4.2.10}$$

Notice it is really the projector P_{ij} that is "transverse"; i.e.

$$k_i P_{ij}(\vec{k}) = 0. \tag{4.2.11}$$

Let us also note that this scalar-vector decomposition is unique, in that – if we have the Fourier-space equation

$$-ik_i\widetilde{\alpha} + \widetilde{\alpha}_i = -ik_i\widetilde{\beta} + \widetilde{\beta}_i, \tag{4.2.12}$$

where $k_i\widetilde{\alpha}_i = k_i\widetilde{\beta}_i = 0$, then

$$\widetilde{\alpha} = \widetilde{\beta} \qquad \text{and} \qquad \widetilde{\alpha}_i = \widetilde{\beta}_i. \tag{4.2.13}$$

For, we may first "dot" both sides of eq. (4.2.12) with $\vec{k}$ and see that – for $\vec{k} \neq \vec{0}$,

$$\vec{k}^2\widetilde{\alpha} = \vec{k}^2\widetilde{\beta} \qquad \Leftrightarrow \qquad \widetilde{\alpha} = \widetilde{\beta}. \tag{4.2.14}$$

Plugging this result back into eq. (4.2.12), we also conclude $\widetilde{\alpha}_i = \widetilde{\beta}_i$.

Now, this scalar-vector decomposition is really just a mathematical fact, and may even be performed in a curved space – as long as the latter is infinite – since it depends on the existence of the Fourier transform and not on the metric structure. (A finite space would call for a discrete Fourier-like series of sorts.) However, to determine its usefulness, we would need to insert it into the partial differential equations obeyed by A_i, where the metric structure does matter. As we now turn to examine, because of the spatial translation symmetry of Minkowski spacetime, Maxwell's equations themselves admit a scalar-vector decomposition. This, in turn, would lead to PDEs for the gauge-invariant portions of A_μ.

Gauge transformations We first examine how the gauge transformation of eq. (4.0.15) is implemented on a scalar-vector decomposed A_μ.

$$A_0 \to A_0 + \dot{L} \tag{4.2.15}$$

$$\begin{aligned}A_i = \partial_i\alpha + \alpha_i &\to \partial_i\alpha + \alpha_i + \partial_i L \\ &= \partial_i(\alpha + L) + \alpha_i.\end{aligned} \tag{4.2.16}$$

From the uniqueness discussion above, we may thus identify the gauge-transformed "scalar" portion of A_i

$$\alpha \to \alpha' \equiv \alpha + L \tag{4.2.17}$$

and the "transverse-vector" portion of A_i to be gauge-invariant:

$$\alpha_i \to \alpha_i . \tag{4.2.18}$$

Let us now identify

$$\Phi \equiv A_0 - \dot{\alpha} \tag{4.2.19}$$

because it is gauge-invariant; for, according to equations (4.2.15) and (4.2.17)

$$\Phi \to A_0 + \dot{L} - \partial_0(\alpha + L) = A_0 - \dot{\alpha}. \tag{4.2.20}$$

In terms of Φ and α_i, the components of the gauge-invariant electromagnetic tensor read

$$F_{0i} \equiv \dot{A}_i - \partial_i A_0 = \dot{\alpha}_i + \partial_i \dot{\alpha} - \partial_i A_0 \tag{4.2.21}$$

$$= \dot{\alpha}_i - \partial_i \Phi \tag{4.2.22}$$

$$F_{ij} = \partial_{[i} A_{j]} = \partial_{[i} \alpha_{j]}. \tag{4.2.23}$$

Electric current We also need to perform a scalar-vector decomposition of the electric current

$$J_\mu \equiv \left(\rho_{\mathrm{E}}, \partial_i \mathcal{J} + \mathcal{J}_i\right). \tag{4.2.24}$$

Its conservation $\partial^\mu J_\mu = 0$ now reads

$$\dot{\rho}_{\mathrm{E}} - \partial_i \left(\partial_i \mathcal{J} + \mathcal{J}_i\right) = 0 \tag{4.2.25}$$

$$\dot{\rho}_{\mathrm{E}} = \vec{\nabla}^2 \mathcal{J}. \tag{4.2.26}$$

Maxwell's Equations At this point, we are ready to write down Maxwell's equations $\partial^\mu F_{\mu\nu} = J_\nu$. From eq. (4.2.22), the $\nu = 0$ component is

$$-\partial_i F_{i0} = \partial_i(\dot{\alpha}_i - \partial_i \Phi) = -\vec{\nabla}^2 \Phi = \rho_{\mathrm{E}}. \tag{4.2.27}$$

The $\nu = i$ component of $\partial^\mu F_{\mu\nu} = J_\nu$, according to eq. (4.2.22) and (4.2.23),

$$\partial_0 F_{0i} - \partial_j F_{ji} = \partial_i \mathcal{J} + \mathcal{J}_i \tag{4.2.28}$$

$$\ddot{\alpha}_i - \partial_i \dot{\Phi} - \partial_j \left(\partial_j \alpha_i - \partial_i \alpha_j\right) = \partial_i \mathcal{J} + \mathcal{J}_i \tag{4.2.29}$$

$$\partial^2 \alpha_i - \partial_i \dot{\Phi} = \mathcal{J}_i + \partial_i \mathcal{J}. \tag{4.2.30}$$

As already advertised, we see that the spatial components of Maxwell's equations does admit a scalar-vector decomposition. By the uniqueness argument above, we may read off the "transverse-vector" portion to be

$$\partial^2 \alpha_i = \mathcal{J}_i. \tag{4.2.31}$$

and the "scalar" portion to be

$$-\dot{\Phi} = \mathcal{J}. \tag{4.2.32}$$

We have gotten 3 (groups of) equations – (4.2.27), (4.2.31), (4.2.32) – for 2 sets of variables (Φ, α_i). Let us argue that eq. (4.2.32) is actually redundant. Taking into account eq. (4.2.26), we may take a time derivative of both sides of eq. (4.2.27),

$$-\vec{\nabla}^2\dot{\Phi} = \dot{\rho}_{\mathrm{E}} = \vec{\nabla}^2\mathcal{J}. \tag{4.2.33}$$

For the physically realistic case of isolated electric currents, where we may assume implies both $\dot{\Phi} \to 0$ and $\mathcal{J} \to 0$ as the observer-J_i distance goes to infinity, the solution to this above Poisson equation is then unique. This hands us eq. (4.2.32).

> **Gauge-Invariant Formalism** To sum: for physically realistic situations in Minkowski spacetime, if we perform a scalar-vector decomposition of the photon vector potential A_μ through eq. (4.2.1) and that of the current J_μ through eq. (4.2.24), we find a gauge-invariant Poisson equation
>
> $$-\vec{\nabla}^2\Phi = \rho_{\mathrm{E}}, \qquad \Phi \equiv A_0 - \dot{\alpha}; \tag{4.2.34}$$
>
> as well as a gauge-invariant wave equation
>
> $$\partial^2\alpha_i = \mathcal{J}_i. \tag{4.2.35}$$

These illuminate the theoretical structure of electromagnetism. As you may recall, our explicit discussions in 4D leading up to the spin-1 modes of eq. (4.1.39) led us to conclude that the non-trivial homogeneous wave solutions of Maxwell's equations are in fact of the "transverse-vector" type. The gauge-invariant formalism for this section thus allows us to identify the source of these spin-1 waves – they are the "transverse-vector" portion of the spatial electric current.

Remark: (1+1)D The one constraint $\partial_i\alpha_i = 0$ obeyed by the spin-1 photon α_i means it has really $D - 1 = d - 2$ independent components, since in Fourier space $k_i\widetilde{\alpha}_i = 0$ implies (for $\vec{k} \neq 0$) the $\{\widetilde{\alpha}_i\}$ are linearly dependent. In particular, in $(1+1)$D $k_1\widetilde{\alpha}_1 = 0$ and as long as $k_1 \neq 0$, the spin-1 photon itself is trivial: $\widetilde{\alpha}_1 = 0$.

5 Action Principles & Classical Field Theory

In §(7.1), we elucidated the Lorentz and gauge symmetries enjoyed by Maxwell's equations. There is in fact an efficient means to define a theory such that it would enjoy the symmetries one desires. This is the action principle. You may encountered it in (non-relativistic) Classical Mechanics, where Newton's second law emerges from demanding the integral

$$S \equiv \int_{t_\mathrm{i}}^{t_\mathrm{f}} L \mathrm{d}t, \tag{5.0.1}$$

$$L \equiv \frac{1}{2} m \dot{\vec{x}}(t)^2 - V(\vec{x}(t)). \tag{5.0.2}$$

Here, L is called the Lagrangian, and in this context is the difference between the particle's kinetic and potential energy. The action of a field theory also plays a central role in its quantum theory when phrased in the path integral formulation; roughly speaking, $\exp(iS)$ is related to the infinitesimal quantum transition amplitude. For these reasons, we shall study the classical field theories – leading up to General Relativity itself – through the principle of stationary action.

General covariance In field theory one defines an object similar to the one in eq. (6.1.1), except the integrand $\mathcal{L}$ is now a Lagrangian *density* (per unit spacetime volume). To obtain generally covariant equations, we now demand that the Lagrangian density is, possibly up to a total divergence, a scalar under coordinate transformations and other symmetry transformations relevant to the problem at hand.

$$S \equiv \int_{t_\mathrm{i}}^{t_\mathrm{f}} \mathrm{d}^d x \sqrt{|g|} \mathcal{L} \tag{5.0.3}$$

One then demands that the action is extremized under the boundary conditions that the field configurations at some initial t_i and final time t_f are fixed. If the spatial boundaries of the spacetime are a finite distance away, one would also have to impose appropriate boundary conditions there; otherwise, if space is infinite, the fields are usually assumed to fall off to zero sufficiently quickly at spatial infinity – below, we will assume the latter for technical simplicity. (In particle mechanics, the action principle also assumes the initial and final positions of the particle are specified.) In curved spacetime, note that the time coordinate x^0 need not correspond to same variable defining the initial t_i and final t_f times; the latter are really shorthand for any appropriately defined spacelike 'constant-time' hyper-surfaces.

5.1 Scalar Fields

Let us begin with a scalar field φ. For concreteness, we shall form its Lagrangian density $\mathcal{L}(\varphi, \nabla_\alpha \varphi)$ out of φ and its first covariant derivatives $\nabla_\alpha \varphi$. Demanding the resulting action be extremized means its first order variation need to vanish. That is, we shall replace $\varphi \to \varphi + \delta\varphi$ (which also means $\nabla_\alpha \varphi \to \nabla_\alpha \varphi + \nabla_\alpha \delta\varphi$) and demand that the portion of the action linear in $\delta\varphi$ be zero.

$$\begin{aligned} \delta_\varphi S &= \int_{t_\mathrm{i}}^{t_\mathrm{f}} \mathrm{d}^d x \sqrt{|g|} \left(\frac{\partial \mathcal{L}}{\partial \varphi} \delta\varphi + \frac{\partial \mathcal{L}}{\partial(\nabla_\alpha \varphi)} \nabla_\alpha \delta\varphi \right) \\ &= \left[\int \mathrm{d}^{d-1} \Sigma_\alpha \frac{\partial \mathcal{L}}{\partial(\nabla_\alpha \varphi)} \delta\varphi \right]_{t_\mathrm{i}}^{t_\mathrm{f}} + \int_{t_\mathrm{i}}^{t_\mathrm{f}} \mathrm{d}^d x \sqrt{|g|} \delta\varphi \left(\frac{\partial \mathcal{L}}{\partial \varphi} - \nabla_\alpha \frac{\partial \mathcal{L}}{\partial(\nabla_\alpha \varphi)} \right) \end{aligned} \tag{5.1.1}$$

Because the initial and final field configurations $\varphi(t_{\rm i})$ and $\varphi(t_{\rm f})$ are assumed fixed, their respective variations are zero by definition: $\delta\varphi(t_{\rm i}) = \delta\varphi(t_{\rm f}) = 0$. This sets to zero the first term on the second equality. At this point, the requirement that the action be stationary means $\delta_\varphi S$ be zero for any small but arbitrary $\delta\varphi$, which in turn implies the coefficient of $\delta\varphi$ must be zero. That leaves us with the Euler-Lagrangian equations

$$\frac{\partial\mathcal{L}}{\partial\varphi} = \nabla_\alpha \frac{\partial\mathcal{L}}{\partial(\nabla_\alpha\varphi)}. \tag{5.1.2}$$

We may now consider a coordinate transformation $x(x')$. Assuming $\mathcal{L}$ is a coordinate scalar, this means the only ingredient that is not a scalar is the derivative with respect to $\nabla_\alpha\varphi$. Since

$$\frac{\partial x^\alpha}{\partial x'^\mu}\nabla_\alpha\varphi(x) = \nabla_{\mu'}\varphi(x') \equiv \nabla_{\mu'}\varphi\left(x(x')\right), \tag{5.1.3}$$

we have

$$\frac{\partial\mathcal{L}}{\partial(\nabla_\alpha\varphi(x))} = \frac{\partial(\nabla_{\mu'}\varphi(x'))}{\partial(\nabla_\alpha\varphi(x))}\frac{\partial\mathcal{L}}{\partial(\nabla_{\mu'}\varphi(x'))} = \frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial\mathcal{L}}{\partial(\nabla_{\mu'}\varphi(x'))}. \tag{5.1.4}$$

That is, $\partial\mathcal{L}/\partial(\nabla_\alpha\varphi(x))$ transforms as a rank-1 vector; and $\nabla_\alpha\{\partial\mathcal{L}/\partial(\nabla_\alpha\varphi(x))\}$ is its divergence, i.e., a scalar. Altogether, we have thus demonstrated that the Euler-Lagrange equations in eq. (6.1.5), for a scalar field φ, is itself a scalar. This is a direct consequence of the fact that $\mathcal{L}$ is a coordinate scalar by construction. A common example of such a scalar action is

$$S[\varphi] \equiv \int \mathrm{d}^d x\sqrt{|g|}\left(\frac{1}{2}g^{\mu\nu}\nabla_\mu\varphi\nabla_\nu\varphi - V(\varphi)\right), \tag{5.1.5}$$

where V is its scalar potential.

Internal Global O_N Symmetry To provide an example of a symmetry other than the invariance under coordinate transformations, let us consider the following action involving $N > 1$ scalar fields $\{\varphi^{\rm I} | {\rm I} = 1, 2, 3, \dots, N\}$:

$$S \equiv \int \mathrm{d}^d x\sqrt{|g|}\mathcal{L}\left(g^{\mu\nu}\nabla_\mu\varphi^{\rm I}\nabla_\nu\varphi^{\rm I}, \varphi^{\rm I}\varphi^{\rm I}\right). \tag{5.1.6}$$

With summation covention in force, we see that the sum over the scalar field label 'I' is simply a dot product in 'field space'. This in turn leads us to observe that the action is invariant under a global rotation:

$$\varphi^{\rm I} \equiv \widehat{R}^{\rm I}{}_{\rm J}\varphi'^{\rm J}, \tag{5.1.7}$$

where $\widehat{R}^{\rm I}{}_{\rm A}\widehat{R}^{\rm J}{}_{\rm B}\delta_{\rm IJ} = \delta_{\rm AB}$. (By 'global' rotation, we mean the rotation matrices $\{\widehat{R}^{\rm I}{}_{\rm J}\}$ do not depend on spacetime.) Explicitly,

$$\int \mathrm{d}^d x\sqrt{|g|}\mathcal{L}\left(g^{\mu\nu}\nabla_\mu\varphi^{\rm I}\nabla_\nu\varphi^{\rm I}, \varphi^{\rm I}\varphi^{\rm I}\right) = \int \mathrm{d}^d x\sqrt{|g|}\mathcal{L}\left(g^{\mu\nu}\nabla_\mu\varphi'^{\rm I}\nabla_\nu\varphi'^{\rm I}, \varphi'^{\rm I}\varphi'^{\rm I}\right). \tag{5.1.8}$$

Let us now witness, because we have constructed a Lagrangian density that is invariant under such an internal O_N symmetry, the resulting equations of motion transform covariantly under rotations. Firstly, the I-th Euler-Lagrange equation, gotten by varying eq. (6.1.9) with respect to φ^{I}, reads

$$\frac{\partial \mathcal{L}}{\partial \varphi^{\mathrm{I}}} = \nabla_\alpha \frac{\partial \mathcal{L}}{\partial(\nabla_\alpha \varphi^{\mathrm{I}})}. \tag{5.1.9}$$

Under rotation, eq. (6.1.10) is equivalent to

$$\left(\widehat{R}^{-1}\right)^{\mathrm{J}}{}_{\mathrm{I}} \varphi^{\mathrm{I}} = \varphi'^{\mathrm{J}}, \tag{5.1.10}$$

which in turn tells us

$$\left(\widehat{R}^{-1}\right)^{\mathrm{J}}{}_{\mathrm{I}} \nabla_\alpha \varphi^{\mathrm{I}} = \nabla_\alpha \varphi'^{\mathrm{J}}. \tag{5.1.11}$$

Therefore eq. (6.1.12) becomes

$$\frac{\partial \varphi'^{\mathrm{J}}}{\partial \varphi^{\mathrm{I}}} \frac{\partial \mathcal{L}}{\partial \varphi'^{\mathrm{J}}} = \frac{\partial \nabla_\alpha \varphi'^{\mathrm{J}}}{\partial \nabla_\alpha \varphi^{\mathrm{I}}} \nabla_\alpha \frac{\partial \mathcal{L}}{\partial(\nabla_\alpha \varphi^{\mathrm{J}})}, \tag{5.1.12}$$

$$\left(\widehat{R}^{-1}\right)^{\mathrm{J}}{}_{\mathrm{I}} \frac{\partial \mathcal{L}}{\partial \varphi'^{\mathrm{J}}} = \left(\widehat{R}^{-1}\right)^{\mathrm{J}}{}_{\mathrm{I}} \nabla_\alpha \frac{\partial \mathcal{L}}{\partial(\nabla_\alpha \varphi^{\mathrm{J}})}. \tag{5.1.13}$$

The PDEs for our O_N-invariant scalar field theory transforms covariantly as a vector under global rotation of the fields $\{\varphi^{\mathrm{I}}\}$.

5.2 Maxwell's Electromagnetism

We have seen how Maxwell's equations are both gauge-invariant and Lorentz invariant. (In fact, the former meant we had to gauge fix the equations before we could solve them.) In curved spacetime, the gauge-invariant electromagnetic tensor

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu = \nabla_\mu A_\nu - \nabla_\nu A_\mu \tag{5.2.1}$$

does not actually contain any metric because the Christoffel symbols cancel out. We may form a coordinate scalar as follows:

$$\mathcal{L}_{\text{Maxwell}} \equiv -\frac{1}{4} F_{\mu\nu} F^{\mu\nu}. \tag{5.2.2}$$

We now claim that, given an externally prescribed electric current J^μ, Maxwell's equations arise from the following action:

$$S_{\text{Maxwell}} \equiv \int_{t_{\mathrm{i}}}^{t_{\mathrm{f}}} \mathrm{d}^d x \sqrt{|g|} \left(\mathcal{L}_{\text{Maxwell}} - A_\mu J^\mu\right). \tag{5.2.3}$$

The $A_\mu J^\mu$ term, under gauge transformation $A_\mu \to A_\mu + \partial_\mu L$, is altered as

$$\begin{aligned} A_\mu J^\mu &\to A_\mu J^\mu + \nabla_\mu L \cdot J^\mu \\ &= A_\mu J^\mu + \nabla_\mu (L \cdot J^\mu) - L \nabla_\mu J^\mu \end{aligned} \tag{5.2.4}$$

If we require that $L(t_\mathrm{i}) = L(t_\mathrm{f}) = 0$, then we see that such a gauge transformation changes the Maxwell action in eq. (5.2.3) as

$$S_\mathrm{Maxwell} \to S_\mathrm{Maxwell} - \left[\int \mathrm{d}^{d-1}\Sigma_\mu (L \cdot J^\mu) \right]_{t_\mathrm{i}}^{t_\mathrm{f}} + \int_{t_\mathrm{i}}^{t_\mathrm{f}} \mathrm{d}^d x \sqrt{|g|} L \nabla_\mu J^\mu \tag{5.2.5}$$

$$\to S_\mathrm{Maxwell} + \int_{t_\mathrm{i}}^{t_\mathrm{f}} \mathrm{d}^d x \sqrt{|g|} L \nabla_\mu J^\mu . \tag{5.2.6}$$

We have already witnessed in §(7.1) how gauge-invariance is intimately tied to current conservation: here we see that the Maxwell action would not be invariant under gauge transformations unless J^μ is conserved.

Let us now proceed to vary the Maxwell action, and see how Maxwell's equations emerge. Consider

$$A_\mu \to A_\mu + \delta A_\mu \tag{5.2.7}$$

and read off the first order in δA_μ terms in the resulting action:

$$\begin{aligned} \delta_A S_\mathrm{Maxwell} &= \int_{t_\mathrm{i}}^{t_\mathrm{f}} \mathrm{d}^d x \sqrt{|g|} \left(-\frac{1}{4} \partial_{[\mu} \delta A_{\nu]} F^{\mu\nu} - \frac{1}{4} F^{\mu\nu} \partial_{[\mu} \delta A_{\nu]} - \delta A_\mu J^\mu \right) \\ &= \int_{t_\mathrm{i}}^{t_\mathrm{f}} \mathrm{d}^d x \sqrt{|g|} \left(-\frac{1}{2} \nabla_\mu \delta A_\nu F^{[\mu\nu]} - \delta A_\mu J^\mu \right) \\ &= \left[- \int \mathrm{d}^{d-1}\Sigma_\mu \delta A_\nu F^{\mu\nu} \right]_{t_\mathrm{i}}^{t_\mathrm{f}} + \int_{t_\mathrm{i}}^{t_\mathrm{f}} \mathrm{d}^d x \sqrt{|g|} \delta A_\nu \left(\nabla_\mu F^{\mu\nu} - J^\nu \right) . \end{aligned} \tag{5.2.8}$$

Let us notice not all the components of the vector potential A_μ need to be fixed at t_i and t_f for the first term of the last equality to vanish. As a simple example, suppose we focus on the Minkowski case and let t_i and t_f correspond to constant x^0-surfaces, then we have $\int \mathrm{d}^{d-1}\Sigma_\mu \delta A_\nu F^{\mu\nu} = \int_{\mathbb{R}^{d-1}} \mathrm{d}^{d-1}\vec{x} \delta A_i F^{0i}$ because, by the antisymmetry of $F^{\mu\nu}$, $F^{00} = 0$. In any case, once the boundary field configurations are fixed, $\delta A_i(t_\mathrm{i}) = \delta A_i(t_\mathrm{f}) = 0$, we have to demand the remaining coefficient of δA_ν be zero.

$$\nabla_\mu F^{\mu\nu} = J^\nu \tag{5.2.9}$$

This is Maxwell's equations in curved spacetime. Here, we are using A_μ as our fundamental field variable; but if we were instead working with $F_{\mu\nu}$, we need to impose $F_{\mu\nu} = -F_{\nu\mu}$ and $\mathrm{d}F = 0$:

$$\nabla_{[\alpha} F_{\mu\nu]} = \partial_{[\alpha} F_{\mu\nu]} = 0. \tag{5.2.10}$$

Of course, by the Poincaré lemma, this does imply $F = \mathrm{d}A$, i.e., $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$.

Problem 5.1. Conservation of electric current Explain why

$$\nabla_\nu \nabla_\mu F^{\mu\nu} = 0 \tag{5.2.11}$$

is an identity. Therefore, by taking the divergence on both sides of eq. (5.2.9), we see that Maxwell's equations in curved spacetime continue to require the conservation of its electric current. □

Electromagnetism of point charges As a non-trivial application of the action principle for electromagnetism, we will now demonstrate why the following action describes the electromagnetism of N point electric charges.

$$S' \equiv -\frac{1}{4}\int_{t_{\rm i}}^{t_{\rm f}} \mathrm{d}^d x\sqrt{|g|}F_{\mu\nu}F^{\mu\nu} - \sum_{a=1}^{N}\left(m_a \int_{t_{\rm i}}^{t_{\rm f}} \mathrm{d}\lambda_a \sqrt{g_{\mu\nu}\dot{z}_a^\mu \dot{z}_a^\nu} + q_a \int_{t_{\rm i}}^{t_{\rm f}} A_\mu(z_a)\mathrm{d}x^\mu\right), \tag{5.2.12}$$

where m_a and q_a are respectively the mass and electric charge of the ath point particle; and $\dot{z}_a^\mu \equiv \mathrm{d}z_a^\mu/\mathrm{d}\lambda_a$. In particular, this action leads to Maxwell's equations sourced by point charges

$$\nabla_\mu F^{\mu\nu} = J^\nu \tag{5.2.13}$$

where the electric current here is

$$J^\nu = \sum_a q_a \int \mathrm{d}\lambda_a \frac{\mathrm{d}z_a^\nu}{\mathrm{d}\lambda_a}\frac{\delta^{(d)}\left(x - z(\lambda_a)\right)}{\sqrt[4]{g(x)g(z)}}; \tag{5.2.14}$$

as well as the covariant Lorentz force law

$$m_a \frac{D^2 z_a^\mu}{\mathrm{d}\tau_a^2} = q_a F^\mu{}_\nu \frac{\mathrm{d}z_a^\nu}{\mathrm{d}\tau_a}, \tag{5.2.15}$$

where τ_a is the proper time of the ath point charge and covariant acceleration on the left-hand-side is

$$\frac{D^2 z_a^\mu}{\mathrm{d}\tau_a^2} \equiv \frac{\mathrm{d}^2 z_a^\mu}{\mathrm{d}\tau_a^2} + \Gamma^\mu{}_{\alpha\beta}\frac{\mathrm{d}z_a^\alpha}{\mathrm{d}\tau_a}\frac{\mathrm{d}z_a^\beta}{\mathrm{d}\tau_a}. \tag{5.2.16}$$

Gauge symmetry Let us observe S' in eq. (5.2.12) is gauge invariant. We already know the $F_{\mu\nu}F^{\mu\nu}$ is gauge invariant, so we only need to check the $A_\mu \mathrm{d}x^\mu$ term. Upon the replacement $A_\mu \mathrm{d}x^\mu \to A_\mu \mathrm{d}x^\mu + \partial_\mu L \mathrm{d}x^\mu = A_\mu \mathrm{d}x^\mu + \mathrm{d}L$, and as long as L is chosen to vanish at the initial $t_{\rm i}$ and final $t_{\rm f}$ times of the trajectories

$$\begin{aligned}\sum_a \int A_\mu(z_a)\mathrm{d}x^\mu \to \sum_a &\left(\int A_\mu(z_a)\mathrm{d}x^\mu + \int \mathrm{d}L(z_a)\right) \\ &= \sum_a \left(\int A_\mu(z_a)\mathrm{d}x^\mu + L(t_{\rm f}, \vec{z}_a(t_{\rm f})) - L(t_{\rm i}, \vec{z}_a(t_{\rm i}))\right) \\ &= \sum_a \int A_\mu(z_a)\mathrm{d}x^\mu. \end{aligned} \tag{5.2.17}$$

Variational calculation We demand the action be stationary under the variation of the gauge field as well as the individual trajectories $\{z_a^\mu\}$. By re-writing the $A_\mu \mathrm{d}x^\mu$ terms as

$$-\sum_a q_a \int A_\mu \mathrm{d}x^\mu = -\int_{t_{\rm i}}^{t_{\rm f}} \mathrm{d}^d x\sqrt{|g(x)|}A_\mu(x)\sum_a q_a \int \mathrm{d}\lambda_a \frac{\mathrm{d}z_a^\mu}{\mathrm{d}\lambda_a}\frac{\delta^{(d)}\left(x - z_a\right)}{\sqrt[4]{|g(x)g(z_a)|}}, \tag{5.2.18}$$

the δ-functions tell us we are dealing with point charges. If we compare this expression against the $A_\mu J^\mu$ term in eq. (5.2.3) – namely, by reading off the coefficient of A_μ – this allows us to identify the electric current in eq. (5.2.14). We have thus arrived at eq. (5.2.13).

Next, we vary the action with respect to the ath trajectory z_a, namely

$$z_a^\mu \to z_a^\mu + \delta z_a^\mu; \tag{5.2.19}$$

assuming $\delta z_a(t_{\rm i}) = \delta z_a(t_{\rm f}) = 0$. The terms linear in the small perturbation δz_a^μ are

$$\begin{aligned} \delta_{z_a} S' &= \int_{t_{\rm i}}^{t_{\rm f}} \mathrm{d}\tau_a \delta z_a^\alpha g_{\alpha\beta} m_a \frac{D^2 z_a^\beta}{\mathrm{d}\tau_a^2} - q_a \int_{t_{\rm i}}^{t_{\rm f}} \mathrm{d}\tau_a \left(\delta z_a^\alpha \partial_\alpha A_\beta \dot{z}_a^\beta + A_\beta \delta \dot{z}_a^\beta \right) \\ &= \int_{t_{\rm i}}^{t_{\rm f}} \mathrm{d}\tau_a \delta z_a^\alpha \left\{ g_{\alpha\beta} \cdot m_a \frac{D^2 z_a^\beta}{\mathrm{d}\tau_a^2} - q_a \left(\partial_\alpha A_\beta - \partial_\beta A_\alpha \right) \dot{z}_a^\beta \right\} \\ &= \int_{t_{\rm i}}^{t_{\rm f}} \mathrm{d}\tau_a \delta z_a^\alpha g_{\alpha\beta} \left\{ m_a \frac{D^2 z_a^\beta}{\mathrm{d}\tau_a^2} - q_a F^\beta{}_\gamma \dot{z}_a^\gamma \right\}. \end{aligned} \tag{5.2.20}$$

We have thus recovered eq. (5.2.15). Note that, in the first equality above, we have changed variables from λ_a to proper time τ_a *after variation*.

Problem 5.2. Lorentz force law in 4D flat spacetime Express the 4D Minkowski spacetime version of the Lorentz force law in eq. (5.2.15) in terms of electric $E^i \equiv F^{i0}$ and magnetic fields $F^{ij} = \epsilon^{0ijk} B^k$. (The ϵ^{0ijk} is the Levi-Civita tensor in flat Minkowski spacetime, with $\epsilon_{0123} \equiv 1$.) Express your time derivatives with respect to coordinate time. You should find the zeroth component of eq. (5.2.15) to be redundant; to arrive at this conclusion more rapidly you may want to start with the action in eq. (5.2.12) but written in coordinate time t:

$$S'_{\rm pp} \equiv - \sum_{a=1}^{N} \left(m_a \int_{t_{\rm i}}^{t_{\rm f}} \mathrm{d}t \sqrt{\eta_{\mu\nu} \dot{z}_a^\mu \dot{z}_a^\nu} + q_a \int_{t_{\rm i}}^{t_{\rm f}} A_\mu(z_a) \dot{z}_a^\mu \mathrm{d}t \right), \tag{5.2.21}$$

where $\dot{z}_a^\mu \equiv \mathrm{d}z_a^\mu / \mathrm{d}t$. $\square$

Problem 5.3. Lorenz gauge Let us impose the Lorenz gauge condition in curved spacetime,

$$\nabla^\mu A_\mu = 0. \tag{5.2.22}$$

Show that Maxwell's equation in eq. (5.2.9) then reads

$$\Box A_\nu - R_\nu{}^\sigma A_\sigma = J_\nu, \qquad \Box \equiv g^{\mu\nu} \nabla_\mu \nabla_\nu = \nabla_\mu \nabla^\mu. \tag{5.2.23}$$

Hint: You may need to 'commute' the covariant derivatives in the term $\nabla^\mu \nabla_\nu A_\mu$. $\square$

Problem 5.4. Weyl invariance in (3+1)D Consider replacing the metric $g_{\mu\nu}$ by multiplying it with an overall scalar function $\Omega(x)^2$, i.e.,

$$g_{\mu\nu}(x) \to \Omega(x)^2 g_{\mu\nu}(x). \tag{5.2.24}$$

Show that the Maxwell action in eq. (5.2.3) is invariant in 4 spacetime dimensions if we simultaneously make the replacement in eq. (5.2.24) and

$$J^\mu(x) \to \Omega(x)^p J^\mu(x) \tag{5.2.25}$$

for an appropriate p – what is p? – as well as

$$A_\mu(x) \to A_\mu(x). \tag{5.2.26}$$

□

Problem 5.5. Null Geodesics & Weyl Transformations Suppose two geometries $g_{\mu\nu}$ and $\bar{g}_{\mu\nu}$ are related via a Weyl transformation

$$g_{\mu\nu}(x) = \Omega(x)^2 \bar{g}_{\mu\nu}(x). \tag{5.2.27}$$

Consider the null geodesic equation in the geometry $g_{\mu\nu}(x)$,

$$k'^\sigma \nabla_\sigma k'^\mu = 0, \qquad g_{\mu\nu} k'^\mu k'^\nu = 0 \tag{5.2.28}$$

where ∇ is the covariant derivative with respect to $g_{\mu\nu}$; as well as the null geodesic equation in $\bar{g}_{\mu\nu}(x)$,

$$k^\sigma \overline{\nabla}_\sigma k^\mu = 0, \qquad \bar{g}_{\mu\nu} k^\mu k^\nu = 0; \tag{5.2.29}$$

where $\overline{\nabla}$ is the covariant derivative with respect to $\bar{g}_{\mu\nu}$. Show that

$$k^\mu = \Omega^2 \cdot k'^\mu. \tag{5.2.30}$$

Hint: First show that the Christoffel symbol $\overline{\Gamma}^\mu{}_{\alpha\beta}[\bar{g}]$ built solely out of $\bar{g}_{\mu\nu}$ is related to $\Gamma^\mu{}_{\alpha\beta}[g]$ built out of $g_{\mu\nu}$ through the relation

$$\Gamma^\mu{}_{\alpha\beta}[g] = \bar{\Gamma}^\mu{}_{\alpha\beta}[\bar{g}] + \delta^\mu_{\{\beta} \overline{\nabla}_{\alpha\}} \ln \Omega - \bar{g}_{\alpha\beta} \overline{\nabla}^\mu \ln \Omega. \tag{5.2.31}$$

Remember to use the constraint $g_{\mu\nu} k'^\mu k'^\nu = 0 = \bar{g}_{\mu\nu} k^\mu k^\nu$.

A spacetime is said to be conformally flat if it takes the form

$$g_{\mu\nu}(x) = \Omega(x)^2 \eta_{\mu\nu}. \tag{5.2.32}$$

Solve the null geodesic equation explicitly in such a spacetime. □

Problem 5.6. Non-Minimal Electromagnetic-Gravitational Interactions in 4D Consider adding the following action to the Maxwell one of eq. (5.2.3):

$$S_{\text{EM-Gravity I}} \equiv \int_{t_i}^{t_f} \mathrm{d}^4 x \sqrt{|g|} C_{6,1} R^{\mu\nu\alpha\beta} F_{\mu\nu} F_{\alpha\beta}. \tag{5.2.33}$$

By ensuring the dimension of the Maxwell action in eq. (5.2.3) is the same as that of $S_{\text{EM-Gravity I}}$, determine the dimension of $C_{6,1}$. That is, $[C_{6,1}] = \text{Mass}^p$ – what is p? Write down as many such actions as you can with coefficients that share the same mass dimension. (Hint: Evaluate the $\sqrt{|g|}$ and Lagrangian density in a Fermi Normal Coordinate System.)

Note that quantum corrections to electromagnetism in 4D curved spacetimes does in fact generate such terms – and infinitely many more! – in addition to the Maxwell action of eq. (5.2.3). For low energy processes, photons interact with gravity in increasingly complicated ways through the exchange of virtual electron-positrons propagating in spacetime, and the coefficient $C_{6,1}$ and its analogs would scale as some power of $1/m_e$, where m_e is the electron mass. To see such interactions are indeed quantum in nature, put back the factors of $\hbar$; i.e., write down $S_{\text{EM-Gravity I}}$ as

$$S_{\text{EM-Gravity I}} = \int_{t_\text{i}}^{t_\text{f}} \mathrm{d}^4x\sqrt{|g|}C'_{6,1}\hbar^q m_e^p R^{\mu\nu\alpha\beta}F_{\mu\nu}F_{\alpha\beta}, \tag{5.2.34}$$

with the appropriate powers of q and p and some dimensionless $C'_{6,1}$. Finally, by comparing the length scales involved, i.e., $F_{\mu\nu}F^{\mu\nu}$ versus $C_{6,1}R^{\mu\nu\alpha\beta}F_{\mu\nu}F_{\alpha\beta}$, describe qualitatively the relative importance of such quantum effects encoded with this $C_{6,1}$ term and its cousins. $\square$

Problem 5.7. Hodge dual formulation of Maxwell's equations Define the dual of the Faraday tensor as

$$\widetilde{F}^{\mu_1\dots\mu_{d-2}} \equiv \frac{1}{2}\widetilde{\epsilon}^{\mu_1\dots\mu_{d-2}\alpha\beta}F_{\alpha\beta}. \tag{5.2.35}$$

Verify the Hodge dual formulation of Maxwell's equations (5.2.9) and (5.2.10):

$$\chi\cdot\widetilde{\epsilon}^{\mu_1\dots\mu_{d-1}\nu}\nabla_{\mu_1}\widetilde{F}_{\mu_2\dots\mu_{d-1}} = J^\nu, \tag{5.2.36}$$

$$\nabla_\sigma\widetilde{F}^{\sigma\mu_3\dots\mu_{d-1}} = 0; \tag{5.2.37}$$

and work out the numerical constant χ. $\square$

Problem 5.8. Second order form of Maxwell's Equations By taking the divergence of the Bianchi portion of Maxwell's equations, $\mathrm{d}F = 0$, followed by using the geometric Bianchi identity $R_{\mu[\nu\alpha\beta]} = 0$, show that

$$\Box F_{\alpha\beta} + R_{\mu\nu\alpha\beta}F^{\mu\nu} + R^{\sigma}{}_{[\alpha}F_{\beta]\sigma} = -\nabla_{[\alpha}\nabla^\sigma F_{\beta]\sigma}, \qquad \Box\equiv\nabla_\sigma\nabla^\sigma. \tag{5.2.38}$$

Next, use the other Maxwell's equation, $\mathrm{div}F = J$, to obtain the second order form of Maxwell's equations:

$$\Box F_{\alpha\beta} + R_{\mu\nu\alpha\beta}F^{\mu\nu} + R^{\sigma}{}_{[\alpha}F_{\beta]\sigma} = \nabla_{[\alpha}J_{\beta]}. \tag{5.2.39}$$

This indicates electromagnetic fields $F_{\mu\nu}$ obey a wave equation in curved spacetimes. $\square$

5.3 JWKB Approximation and Gravitational Redshift

In this section we will apply the JWKB (more commonly dubbed WKB) approximation to study the vacuum (i.e., $J_\mu = 0$ limit of) Maxwell's equations in eq. (5.2.23). At leading orders in perturbation theory, we will argue – in the limit where the wavelength of the photons are much shorter than that of the background geometric curvature – that photons propagate on the light cone and their polarization tensors are largely parallel transported along their null geodesics. We will also see that the photon's phase S would allow us to define its frequency as the number density of constant-S surfaces piercing the timelike worldline of the observer. This also leads us to recognize that, not only is $k^\mu \equiv \nabla^\mu S$ null it obeys the geodesic equation $k^\sigma \nabla_\sigma k^\mu = 0$.

Eikonal/Geometric Optics/JWKB Ansatz We will begin by postulating that the vector potential can be modeled as the (real part of) a slowly varying amplitude a_μ multiplied by a rapidly oscillating phase $\exp(iS)$:

$$A_\mu = \mathrm{Re}\left\{ a_\mu \exp(iS/\epsilon) \right\}. \tag{5.3.1}$$

[5]The $\{a_\mu\}$ can be complex but S is real. We shall also allow the amplitude itself to be a power series in ϵ:

$$a_\mu = \sum_{\ell=0}^{\infty} \epsilon^\ell \, {}_\ell a_\mu. \tag{5.3.2}$$

The $0 < \epsilon \ll 1$ is a fictitious parameter that reminds us of the hierarchy of length scales in the problem – specifically, ϵ should be viewed as the ratio between the short wavelength of the photon to the long wavelength of the background geometric curvature. To this end, we shall re-write the vacuum version of the Lorenz-gauge Maxwell's equation (5.2.23) with ϵ^2 multiplying the wave operator $\Box$:

$$\Box A_\mu - \epsilon^2 R_\mu{}^\sigma A_\sigma = 0. \tag{5.3.3}$$

In a locally freely-falling frame (i.e., flat coordinate system), this equation takes the schematic form

$$\partial^2 A - \epsilon^2 (\partial^2 g) A = 0. \tag{5.3.4}$$

The first term from the left goes as $A/(\text{wavelength of } A)^2$ while the second as $A/(\text{wavelength of } g)^2$, and as thus already advertised ϵ^2 is a power counting parameter reminding us of the relative strength of the two terms.

Wave Equation Plugging the ansatz of eq. (5.3.1) into eq. (5.3.3):

$$\begin{aligned} 0 &= (\Box a_\mu - \epsilon^2 R_\mu{}^\sigma a_\sigma) e^{iS/\epsilon} + 2\nabla_\sigma a_\mu \frac{i}{\epsilon} \nabla^\sigma S \cdot e^{iS/\epsilon} + a_\mu \nabla_\sigma \left(i(\nabla^\sigma S/\epsilon) e^{iS/\epsilon} \right) \\ &= (\Box a_\mu - \epsilon^2 R_\mu{}^\sigma a_\sigma) e^{iS/\epsilon} + 2\nabla_\sigma a_\mu \frac{i}{\epsilon} (\nabla^\sigma S) \cdot e^{iS/\epsilon} + a_\mu \left(i(\Box S/\epsilon) e^{iS/\epsilon} + (i\nabla S/\epsilon)^2 e^{iS/\epsilon} \right). \end{aligned} \tag{5.3.5}$$

Employing the power series of eq. (5.3.2),

$$\begin{aligned} 0 &= \Box\, {}_0 a_\mu + \epsilon \Box\, {}_1 a_\mu + \epsilon^2 \Box\, {}_2 a_\mu + \dots \\ &\quad - R_\mu{}^\sigma \left(\epsilon^2 \, {}_0 a_\sigma + \epsilon^3 \, {}_1 a_\sigma + \dots \right) \\ &\quad + 2i\epsilon^{-1} (\nabla\, {}_0 a_\mu \cdot \nabla S) + 2i\epsilon^0 (\nabla\, {}_1 a_\mu \cdot \nabla S) + 2i\epsilon (\nabla\, {}_2 a_\mu \cdot \nabla S) + \dots \end{aligned}$$

[5]Recall that this ansatz becomes an exact solution in Minkowski spacetime, where $S = \pm k_\mu x^\mu$ and both k_μ and a_μ are constant.

$$
\begin{aligned}
&+ i\epsilon^{-1}\, {}_0a_\mu \Box S + i\epsilon^0\, {}_1a_\mu \Box S + i\epsilon\, {}_2a_\mu \Box S + \dots \\
&- (\nabla S)^2 \left(\epsilon^{-2}\, {}_0a_\mu + \epsilon^{-1}\, {}_1a_\mu + \epsilon^0\, {}_2a_\mu + \epsilon\, {}_3a_\mu + \dots\right).
\end{aligned} \tag{5.3.6}
$$

Negative Two Setting the coefficient of ϵ^{-2} to zero

$$
k_\mu k^\mu = 0, \qquad k_\mu \equiv \nabla_\mu S. \tag{5.3.7}
$$

Because S is a scalar, $\nabla_\nu k_\mu = \nabla_\nu \nabla_\mu S = \nabla_\mu \nabla_\nu S = \nabla_\mu k_\nu$ and hence

$$
0 = \nabla_\nu (k^2) = 2k^\mu \nabla_\nu k_\mu = 2k^\mu \nabla_\mu k_\nu. \tag{5.3.8}
$$

That is, the gradient of the phase S sweeps out null geodesics in spacetime:

$$
(k \cdot \nabla) k^\mu = 0. \tag{5.3.9}
$$

Negative One Setting the coefficient of ϵ^{-1} to zero:

$$
0 = {}_0a_\mu \Box S + 2\nabla^\sigma S \nabla_\sigma\, {}_0a_\mu, \tag{5.3.10}
$$
$$
0 = \overline{{}_0a_\mu} \Box S + 2\nabla^\sigma S \nabla_\sigma \overline{{}_0a_\mu}, \tag{5.3.11}
$$

where the second line is simply the complex conjugate of the first. Note that $\nabla|a|^2 = (\nabla a)\bar{a} + a(\nabla \bar{a})$. Guided by this, we may multiply the first equation by $\overline{{}_0a^\mu}$ and the second equation by ${}_0a^\mu$, followed by adding them.

$$
0 = |{}_0a|^2 \Box S + 2\overline{{}_0a^\mu} \nabla^\sigma S \nabla_\sigma\, {}_0a_\mu \tag{5.3.12}
$$
$$
0 = |{}_0a|^2 \Box S + 2\, {}_0a^\mu \nabla^\sigma S \nabla_\sigma \overline{{}_0a_\mu} \tag{5.3.13}
$$
$$
0 = 2|{}_0a|^2 \Box S + 2\nabla^\sigma S \nabla_\sigma |{}_0a|^2, \qquad |{}_0a|^2 \equiv {}_0a_\mu \overline{{}_0a^\mu}. \tag{5.3.14}
$$

The right hand side of the final equation can be expressed as a divergence.

$$
0 = \nabla_\sigma \left(|{}_0a|^2 \nabla^\sigma S\right) = \nabla_\sigma \left(|{}_0a|^2 k^\sigma\right) \tag{5.3.15}
$$

Up to an overall normalization constant, we may interpret $n^\sigma \equiv |{}_0a|^2 k^\sigma$ as a photon number current, and this equation as its conservation law.

We turn to examining the derivative along $k \equiv \nabla S$ the normalized leading order photon amplitude ${}_0a_\mu/\sqrt{|{}_0a|^2}$:

$$
\nabla^\sigma S \nabla_\sigma \left(\frac{{}_0a_\mu}{\sqrt{|{}_0a|^2}}\right) = \frac{\nabla^\sigma S \nabla_\sigma\, {}_0a_\mu}{\sqrt{|{}_0a|^2}} - \frac{{}_0a_\mu}{2(|{}_0a|^2)^{3/2}} \nabla^\sigma S \nabla_\sigma |{}_0a|^2. \tag{5.3.16}
$$

Eq. (5.3.15) says $\nabla S \cdot \nabla |{}_0a|^2 = -|{}_0a|^2 \Box S$, while eq. (5.3.10), in turn, states ${}_0a_\mu \Box S = -2\nabla^\sigma S \nabla_\sigma\, {}_0a_\mu$.

$$
\begin{aligned}
\nabla^\sigma S \nabla_\sigma \left(\frac{{}_0a_\mu}{\sqrt{|{}_0a|^2}}\right) &= \frac{\nabla^\sigma S \nabla_\sigma\, {}_0a_\mu}{\sqrt{|{}_0a|^2}} + \frac{|{}_0a|^2}{2(|{}_0a|^2)^{3/2}}\, {}_0a_\mu \Box S \\
&= \frac{\nabla^\sigma S \nabla_\sigma\, {}_0a_\mu}{\sqrt{|{}_0a|^2}} - \frac{\nabla^\sigma S \nabla_\sigma\, {}_0a_\mu}{\sqrt{|{}_0a|^2}} = 0.
\end{aligned} \tag{5.3.17}
$$

Lorenz gauge Let us not forget the Lorenz gauge condition: $0 = \nabla^\mu A_\mu = ((\nabla^\mu a_\mu) + (i/\epsilon)\nabla^\mu S a_\mu)e^{iS/\epsilon}$.

$$\begin{aligned} 0 &= \nabla^\mu {}_0a_\mu + \epsilon\nabla^\mu {}_1a_\mu + \epsilon^2\nabla^\mu {}_2a_\mu + \dots \\ &\quad + i\epsilon^{-1}\nabla S \cdot {}_0a + i\epsilon^0\nabla S \cdot {}_1a + i\epsilon\nabla S \cdot {}_2a + \dots \end{aligned} \tag{5.3.18}$$

Negative One Setting the coefficient of ϵ^{-1} to zero, we find the leading order polarization vector must be orthogonal to the wave vector:

$$k^\mu {}_0a_\mu = 0. \tag{5.3.19}$$

Zero Setting the coefficient of ϵ^0 to zero,

$$k^\mu {}_1a_\mu = i\nabla^\mu {}_0a_\mu. \tag{5.3.20}$$

This is telling us that the polarization vector does not remain perpendicular to k^μ at the next order.

To summarize, we have worked out the first two orders of the Lorenz gauge vacuum Maxwell's equations in the JWKB/eikonal/geometric optics limit. Up to this level of accuracy, perturbation theory teaches us:

- The gradient of the phase of the photon field $k^\mu \equiv \nabla^\mu S$ – which we may interpret as its dominant direction of propagation – follows null geodesics in the curved spacetime.
- The photon number current is covariantly conserved.
- The normalized polarization vector is parallel transported along k^μ.
- This same wave vector is orthogonal to the polarization of the photon at leading order; and the first deviation to non-orthogonality occurring at the next order is proportional to the divergence of the polarization vector itself.

Problem 5.9. Electromagnetic Fields In classical theory, it is the electromagnetic fields $F_{\mu\nu}$ that are directly observable, as opposed to the gauge-dependent vector potential A_μ. In the leading order of the JWKB approximation, argue that

$$F_{\mu\nu} \approx \mathrm{Re}\left\{\frac{i}{\epsilon}k_{[\mu}a_{\nu]}\exp\left(iS/\epsilon\right)\right\}, \qquad k_\mu \equiv \nabla_\mu S. \tag{5.3.21}$$

Why is $k^\mu F_{\mu\nu} \approx 0$? This might appear at first sight to depend on the Lorenz gauge condition in eq. (5.3.19), but argue that the Lorenz gauge condition continues to hold – at the leading JWKB approximation – upon any gauge transformation of the form

$$A_\mu \to A_\mu + \mathrm{Re}\left\{\nabla_\mu\left(\ell \cdot \exp(iS/\epsilon)\right)\right\}, \tag{5.3.22}$$

where ℓ is a slowly varying function of spacetime compared to the phase $\exp(iS/\epsilon)$. □

Gravitational Redshift As alluded to at the beginning of this section, the frequency of light according to a timelike observer may be defined as the number density of constant phase surfaces piercing its worldline. This, in turn, may be formalized using the unit normal vector $u^\mu \partial_\mu$ tangent to the said worldline:

$$\omega \equiv |u \cdot \nabla S| = \left| \frac{\mathrm{d}S}{\mathrm{d}\tau} \right| = k^{\widehat{0}} = k_{\widehat{0}}, \tag{5.3.23}$$

where τ is the observer's proper time. In other words, the frequency *is* the zeroth component of the wave vector ($\equiv$ momentum) in an orthonormal basis in the observer's frame.

Static Spherically Symmetric Metrics Near the surface of the Earth, we may model its geometry – at least as a first pass! – as a static spherically symmetric one, given by

$$\mathrm{d}s^2 = (A(r)\mathrm{d}t)^2 - (B(r)\mathrm{d}r)^2 - r^2 \mathrm{d}\Omega^2, \tag{5.3.24}$$

$$A(r) = \sqrt{1 - \frac{r_{s,\mathrm{E}}}{r}}, \qquad B(r) = \frac{1}{\sqrt{1 - r_{s,\mathrm{E}}/r}}, \qquad r_{s,\mathrm{E}} \equiv 2G_\mathrm{N} M_\mathrm{E}. \tag{5.3.25}$$

The associated Lagrangian for the geodesic equation is

$$L = \frac{1}{2}\left((Ak^0)^2 - (Bk^r)^2 - (rk^\theta)^2 - (r\sin(\theta)k^\phi)^2 \right); \tag{5.3.26}$$

where $k^\mu \equiv \mathrm{d}(t, r, \theta, \phi)^\mu / \mathrm{d}\lambda$. (Remember this Lagrangian yields $k^\mu \nabla_\mu k^\nu = 0$.) The static assumption allows us to immediately identify 'energy' E as the conserved quantity

$$E = \frac{\partial L}{\partial k^0} = A^2 k^0. \tag{5.3.27}$$

An observer at a fixed position ($\mathrm{d}r = \mathrm{d}\theta = 0$) has proper time

$$\mathrm{d}\tau = A(r)\mathrm{d}t. \tag{5.3.28}$$

This allows us to identify the 0th vierbein $\varepsilon^{\widehat{0}}{}_\mu \mathrm{d}x^\mu = A\mathrm{d}t$ as the observer worldline's unit timelike tangent vector. That, in turn, inform us eq. (5.3.27) is now

$$E = A(r)k^{\widehat{0}} = A(r)\omega(r) \qquad \Rightarrow \qquad \omega(r) = \frac{E}{A(r)}. \tag{5.3.29}$$

Here, we have recalled from our JWKB discussion above that $k^{\widehat{0}}$ *is* the frequency of the photon measured by our observer. We may now send electromagnetic waves between observers at different radii r_1 and r_2 (say, the bottom and top ends of the Pound-Rebka experiment) near the surface of the Earth:

$$\frac{\omega(r_2)}{\omega(r_1)} = \frac{A(r_1)}{A(r_2)}. \tag{5.3.30}$$

Compare this result to the time dilation result we worked out in Problem (**??**). See also the Wikipedia article on the Pound-Rebka experiment, the first verification of the gravitational time dilation effect. □

Problem 5.10. Co-Moving Redshift in Cosmology At large scales, we live in a universe well described by a spatially flat Friedmann-Lemaître–Robertson–Walker (FLRW) universe:

$$\mathrm{d}s^2 = \mathrm{d}t^2 - a(t)^2 \mathrm{d}\vec{x} \cdot \mathrm{d}\vec{x}, \qquad a(t) > 0. \tag{5.3.31}$$

The observers at rest in this geometry – the ones that witness a perfectly isotropic Cosmic Microwave Background sky (i.e., with a zero dipole) – have trajectories described by

$$Z^\mu = (t, \vec{Z}_0), \qquad \vec{Z}_0 \text{ constant}. \tag{5.3.32}$$

Exploiting the spatial translation symmetry of this geometry, we may postulate the following ansatz for the Lorenz gauge vector potential:

$$A_\mu = \mathrm{Re}\left\{ a_\mu(kt) e^{i\Sigma(t)} e^{i\vec{k}\cdot\vec{x}} \right\}, \qquad k \equiv |\vec{k}|; \tag{5.3.33}$$

where the slowly-varying amplitude a_μ does not depend on the spatial coordinates $\{x^i\}$.

Show that, to leading order, the frequency of the photon according to a co-moving observer redshifts as $1/a$. Specifically, demonstrate that

$$\frac{\omega(t_{\text{observer}})}{\omega(t_{\text{emission}})} = \frac{a(t_{\text{emission}})}{a(t_{\text{observer}})} \equiv (1+z)^{-1}; \tag{5.3.34}$$

where z is dubbed the redshift parameter. Since z is observable[6] – oftentimes inferred using atomic spectral lines – we may employ it to determine the epoch of emission of a given signal. □

[6]Example: the Cosmic Microwave Background radiation is roughly 2.7 Kelvins, emitted from 'the last scattering surface' roughly 380,000 years after the Big Bang and as observed from our vantage point has $z \approx 1,100$.

6 Classical Scalar Fields in Minkowski Spacetime

Field theory in Minkowski spacetime indicates we wish to construct partial differential equations obeyed by fields such that they take the same form in all inertial frames – i.e., the PDEs are Lorentz covariant. As a warm-up, we shall in this section study the case of scalar fields.

A scalar field $\varphi(x)$ is an object that transforms, under Poincaré transformations

$$x^\mu = \Lambda^\mu{}_\nu x'^\nu + a^\mu \tag{6.0.1}$$

as simply

$$\varphi(x(x')) = \varphi\left(x^\mu = \Lambda^\mu{}_\nu x'^\nu + a^\mu\right) \equiv \varphi(x'). \tag{6.0.2}$$

To ensure that this is the case, we would like the PDE it obeys to take the same form in the two inertial frames $\{x^\mu\}$ and $\{x'^\mu\}$ related by eq. (6.0.1). The simplest example is the wave equation with some external scalar source $J(x)$. Let's first write it in the x^μ coordinate system.

$$\eta^{\mu\nu}\partial_\mu\partial_\nu\varphi(x) = J(x), \qquad \partial_\mu \equiv \partial/\partial x^\mu. \tag{6.0.3}$$

If putting a prime on the index denotes derivative with respect to x'^μ, namely $\partial_{\mu'} \equiv \partial/\partial x'^\mu$, then by the chain rule,

$$\partial_{\mu'} = \frac{\partial x^\sigma}{\partial x'^\mu}\frac{\partial}{\partial x^\sigma} = \partial_{\mu'}\left(\Lambda^\sigma{}_\rho x'^\rho + a^\sigma\right)\partial_\sigma \tag{6.0.4}$$

$$= \Lambda^\sigma{}_\mu \partial_\sigma. \tag{6.0.5}$$

Therefore the wave operator indeed takes the same form in both coordinate systems:

$$\eta^{\mu\nu}\partial_{\mu'}\partial_{\nu'} = \eta^{\mu\nu}\Lambda^\sigma{}_\mu\Lambda^\rho{}_\nu\partial_\sigma\partial_\rho \tag{6.0.6}$$

$$= \eta^{\sigma\rho}\partial_\sigma\partial_\rho. \tag{6.0.7}$$

because of Lorentz invariance

$$\eta^{\mu\nu}\Lambda^\sigma{}_\mu\Lambda^\rho{}_\nu = \eta^{\sigma\rho}. \tag{6.0.8}$$

A generalization of the wave equation in eq. (6.0.3) is to add a potential $V(\varphi)$:

$$\partial^2\varphi + V'(\varphi) = J, \tag{6.0.9}$$

where $\partial^2 \equiv \eta^{\mu\nu}\partial_\mu\partial_\nu$ and the prime is a derivative with respect to the argument.

6.1 Action Principle and Symmetries

There is in fact an efficient means to define a theory such that it would enjoy the symmetries one desires. This is the action principle. You may encountered it in (non-relativistic) Classical Mechanics, where Newton's second law emerges from demanding the integral

$$S \equiv \int_{t_\text{i}}^{t_\text{f}} L\text{d}t, \tag{6.1.1}$$

$$L \equiv \frac{1}{2}m\dot{\vec{x}}(t)^2 - V(\vec{x}(t)). \tag{6.1.2}$$

Here, L is called the Lagrangian, and in this context is the difference between the particle's kinetic and potential energy. The action of a field theory also plays a central role in its quantum theory when phrased in the path integral formulation; roughly speaking, $\exp(iS)$ is related to the infinitesimal quantum transition amplitude. For these reasons, we shall study the classical field theories – leading up to General Relativity itself – through the principle of stationary action.

Lorentz covariance In field theory one defines an object similar to the one in eq. (6.1.1), except the integrand $\mathcal{L}$ is now a Lagrangian *density* (per unit spacetime volume). To obtain Lorentz covariant equations, we now demand that the Lagrangian density is, possibly up to a total divergence, a scalar under spacetime Lorentz transformations and other symmetry transformations relevant to the problem at hand.

$$S \equiv \int_{t_{\mathrm{i}}}^{t_{\mathrm{f}}} \mathcal{L} \mathrm{d}^d x \tag{6.1.3}$$

One then demands that the action is extremized under the boundary conditions that the field configurations at some initial t_{i} and final time t_{f} are fixed. If the spatial boundaries of the spacetime are a finite distance away, one would also have to impose appropriate boundary conditions there; otherwise, if space is infinite, the fields are usually assumed to fall off to zero sufficiently quickly at spatial infinity – below, we will assume the latter for technical simplicity. (In particle mechanics, the action principle also assumes the initial and final positions of the particle are specified.)

Let us begin with a scalar field φ. For concreteness, we shall form its Lagrangian density $\mathcal{L}(\varphi, \partial_\alpha \varphi)$ out of φ and its first derivatives $\partial_\alpha \varphi$. Demanding the resulting action be extremized means its first order variation need to vanish. That is, we shall replace $\varphi \to \varphi + \delta\varphi$ (which also means $\partial_\alpha \varphi \to \partial_\alpha \varphi + \partial_\alpha \delta\varphi$) and demand that the portion of the action linear in $\delta\varphi$ be zero.

$$\begin{aligned} \delta_\varphi S &= \int_{t_{\mathrm{i}}}^{t_{\mathrm{f}}} \mathrm{d}^d x \left(\frac{\partial \mathcal{L}}{\partial \varphi} \delta\varphi + \frac{\partial \mathcal{L}}{\partial(\partial_\alpha \varphi)} \partial_\alpha \delta\varphi \right) \\ &= \left[\int \mathrm{d}^{d-1} \Sigma_\alpha \frac{\partial \mathcal{L}}{\partial(\partial_\alpha \varphi)} \delta\varphi \right]_{t_{\mathrm{i}}}^{t_{\mathrm{f}}} + \int_{t_{\mathrm{i}}}^{t_{\mathrm{f}}} \mathrm{d}^d x \delta\varphi \left(\frac{\partial \mathcal{L}}{\partial \varphi} - \partial_\alpha \frac{\partial \mathcal{L}}{\partial(\partial_\alpha \varphi)} \right) \end{aligned} \tag{6.1.4}$$

Because the initial and final field configurations $\varphi(t_{\mathrm{i}})$ and $\varphi(t_{\mathrm{f}})$ are assumed fixed, their respective variations are zero by definition: $\delta\varphi(t_{\mathrm{i}}) = \delta\varphi(t_{\mathrm{f}}) = 0$. This sets to zero the first term on the second equality. At this point, the requirement that the action be stationary means $\delta_\varphi S$ be zero for any small but arbitrary $\delta\varphi$, which in turn implies the coefficient of $\delta\varphi$ must be zero. That leaves us with the Euler-Lagrangian equations

$$\frac{\partial \mathcal{L}}{\partial \varphi} = \partial_\alpha \frac{\partial \mathcal{L}}{\partial(\partial_\alpha \varphi)}. \tag{6.1.5}$$

We may now consider a coordinate transformation $x(x')$. Assuming $\mathcal{L}$ is a coordinate scalar, this means the only ingredient that is not a scalar is the derivative with respect to $\partial_\alpha \varphi$. Since

$$\frac{\partial x^\alpha}{\partial x'^\mu} \partial_\alpha \varphi(x) = \partial_{\mu'} \varphi(x') \equiv \partial_{\mu'} \varphi\left(x(x')\right), \tag{6.1.6}$$

we have

$$\frac{\partial\mathcal{L}}{\partial(\partial_\alpha\varphi(x))} = \frac{\partial(\partial_{\mu'}\varphi(x'))}{\partial(\partial_\alpha\varphi(x))}\frac{\partial\mathcal{L}}{\partial(\partial_{\mu'}\varphi(x'))} = \frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial\mathcal{L}}{\partial(\partial_{\mu'}\varphi(x'))}. \tag{6.1.7}$$

That is, $\partial\mathcal{L}/\partial(\partial_\alpha\varphi(x))$ transforms as a rank-1 vector; and $\partial_\alpha\{\partial\mathcal{L}/\partial(\partial_\alpha\varphi(x))\}$ is its divergence, i.e., a scalar. Altogether, we have thus demonstrated that the Euler-Lagrange equations in eq. (6.1.5), for a scalar field φ, is itself a scalar. This is a direct consequence of the fact that $\mathcal{L}$ is a coordinate scalar by construction. A common example of such a scalar action is

$$S[\varphi] \equiv \int \mathrm{d}^d x \left(\frac{1}{2}\eta^{\mu\nu}\partial_\mu\varphi\partial_\nu\varphi - V(\varphi)\right), \tag{6.1.8}$$

where V is its scalar potential.

Problem 6.1. Show from eq. (6.1.5) that the equations derived from the action in eq. (6.1.8) is $\partial^2\varphi = -V'(\phi)$. □

Internal Global O_N Symmetry To provide an example of a symmetry other than the invariance under coordinate transformations, let us consider the following action involving $N > 1$ scalar fields $\{\varphi^{\mathrm{I}}|\mathrm{I} = 1, 2, 3, \dots, N\}$:

$$S \equiv \int \mathrm{d}^d x \mathcal{L}\left(\eta^{\mu\nu}\partial_\mu\varphi^{\mathrm{I}}\partial_\nu\varphi^{\mathrm{I}}, \varphi^{\mathrm{I}}\varphi^{\mathrm{I}}\right). \tag{6.1.9}$$

With summation covention in force, we see that the sum over the scalar field label 'I' is simply a dot product in 'field space'. This in turn leads us to observe that the action is invariant under a global rotation:

$$\varphi^{\mathrm{I}} \equiv \widehat{R}^{\mathrm{I}}{}_{\mathrm{J}}\varphi'^{\mathrm{J}}, \tag{6.1.10}$$

where $\widehat{R}^{\mathrm{I}}{}_{\mathrm{A}}\widehat{R}^{\mathrm{J}}{}_{\mathrm{B}}\delta_{\mathrm{IJ}} = \delta_{\mathrm{AB}}$. (By 'global' rotation, we mean the rotation matrices $\{\widehat{R}^{\mathrm{I}}{}_{\mathrm{J}}\}$ do not depend on spacetime.) Explicitly,

$$\int \mathrm{d}^d x \mathcal{L}\left(\eta^{\mu\nu}\partial_\mu\varphi^{\mathrm{I}}\partial_\nu\varphi^{\mathrm{I}}, \varphi^{\mathrm{I}}\varphi^{\mathrm{I}}\right) = \int \mathrm{d}^d x \mathcal{L}\left(\eta^{\mu\nu}\partial_\mu\varphi'^{\mathrm{I}}\partial_\nu\varphi'^{\mathrm{I}}, \varphi'^{\mathrm{I}}\varphi'^{\mathrm{I}}\right). \tag{6.1.11}$$

Let us now witness, because we have constructed a Lagrangian density that is invariant under such an internal O_N symmetry, the resulting equations of motion transform covariantly under rotations. Firstly, the I-th Euler-Lagrange equation, gotten by varying eq. (6.1.9) with respect to φ^{I}, reads

$$\frac{\partial\mathcal{L}}{\partial\varphi^{\mathrm{I}}} = \partial_\alpha\frac{\partial\mathcal{L}}{\partial(\partial_\alpha\varphi^{\mathrm{I}})}. \tag{6.1.12}$$

Under rotation, eq. (6.1.10) is equivalent to

$$\left(\widehat{R}^{-1}\right)^{\mathrm{J}}{}_{\mathrm{I}}\varphi^{\mathrm{I}} = \varphi'^{\mathrm{J}}, \tag{6.1.13}$$

which in turn tells us

$$\left(\widehat{R}^{-1}\right)^{\mathrm{J}}{}_{\mathrm{I}}\partial_\alpha\varphi^{\mathrm{I}} = \partial_\alpha\varphi'^{\mathrm{J}}. \tag{6.1.14}$$

Therefore eq. (6.1.12) becomes

$$\frac{\partial\varphi'^{\mathrm{J}}}{\partial\varphi^{\mathrm{I}}}\frac{\partial\mathcal{L}}{\partial\varphi'^{\mathrm{J}}} = \frac{\partial\partial_\alpha\varphi'^{\mathrm{J}}}{\partial\partial_\alpha\varphi^{\mathrm{I}}}\partial_\alpha\frac{\partial\mathcal{L}}{\partial(\partial_\alpha\varphi^{\mathrm{J}})}, \tag{6.1.15}$$

$$\left(\widehat{R}^{-1}\right)^{\mathrm{J}}{}_{\mathrm{I}}\frac{\partial\mathcal{L}}{\partial\varphi'^{\mathrm{J}}} = \left(\widehat{R}^{-1}\right)^{\mathrm{J}}{}_{\mathrm{I}}\partial_\alpha\frac{\partial\mathcal{L}}{\partial(\partial_\alpha\varphi^{\mathrm{J}})}. \tag{6.1.16}$$

The PDEs for our O_N-invariant scalar field theory transforms covariantly as a vector under global rotation of the fields $\{\varphi^{\mathrm{I}}\}$.

6.2 Local Conservation Laws

Non-relativistic You would be rightly shocked if you had stored a sealed tank of water on your rooftop only to find its contents gradually disappearing over time – the total mass of water ought to be a constant. Assuming a flat space geometry, if you had instead connected the tank to two pipes, one that pumps water into the tank and the other pumping water out of it, the rate of change of the total mass of the water

$$M \equiv \int_{\text{tank}} \rho(t,\vec{x})\mathrm{d}^3\vec{x} \tag{6.2.1}$$

in the tank – where t is time, $\vec{x}$ are Cartesian coordinates, and $\rho(t,\vec{x})$ is the water's mass density – is

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\text{tank}}\rho\mathrm{d}^3\vec{x} = -\left(\int_{\text{cross section of 'in' pipe}} + \int_{\text{cross section of 'out' pipe}}\right)\mathrm{d}^2\vec{\Sigma}\cdot(\rho\vec{v}). \tag{6.2.2}$$

Note that $\mathrm{d}^2\vec{\Sigma}$ points *outwards* from the tank, so at the 'in' pipe-tank interface, if the water were indeed following into the pipe, $-\mathrm{d}^2\vec{\Sigma}\cdot(\rho\vec{v}) > 0$ and its contribution to the rate of increase is positive. At the 'out' pipe-tank interface, if the water were indeed following out of the pipe, $-\mathrm{d}^2\vec{\Sigma}\cdot(\rho\vec{v}) < 0$. If we apply Gauss' theorem,

$$\int_{\text{tank}}\dot{\rho}\mathrm{d}^3\vec{x} = -\int_{\text{tank}}\mathrm{d}^3\vec{x}\vec{\nabla}\cdot(\rho\vec{v}). \tag{6.2.3}$$

If we applied the same sort of reasoning to any infinitesimal packet of fluid, with some local mass density ρ, we would find the following local conservation law

$$\dot{\rho} = -\partial_i\left(\rho\cdot v^i\right). \tag{6.2.4}$$

This is a "local" conservation law in the sense that mass cannot simply vanish from one location and re-appear a finite distance away, without first flowing to a neighboring location.

Relativistic We have implicitly assumed a non-relativistic system, where $|\vec{v}| \ll 1$. This is an excellent approximation for most hydrodynamics problems. Strictly speaking, however, relativistic effects – length contraction, in particular – imply that mass density is not a Lorentz scalar. If we define $\rho(t,\vec{x})$ to be the mass density at $(t,\vec{x})$ in a frame instantaneously at rest (aka 'co-moving') with the fluid packet, then the mass density current that is a locally conserved Lorentz vector is given by

$$J^\mu(t,\vec{x}) \equiv \rho(t,\vec{x})v^\mu(t,\vec{x}). \tag{6.2.5}$$

Along its integral curve v^μ should be viewed as the proper velocity $\mathrm{d}(t,\vec{x})^\mu/\mathrm{d}\tau$ of the fluid packet, where τ is the latter's proper time. Moreover, as long as the velocity v^μ is timelike, which is certainly true for fluids, let us recall it is always possible to find a (local) Lorentz transformation $\Lambda^\mu{}_\nu(t,\vec{x})$ such that

$$(1,\vec{0})^\mu \equiv v'^\mu = \Lambda^\mu{}_\nu(t,\vec{x})v^\nu(t,\vec{x}). \tag{6.2.6}$$

and the mass density-current is now

$$J'^\mu = \rho(t',\vec{x}')v'^\mu = \rho(x')\cdot\delta^\mu_0. \tag{6.2.7}$$

The local conservation law obeyed by this relativistically covariant current J^μ is now (in Cartesian coordinates)

$$\partial_\mu J^\mu = 0; \tag{6.2.8}$$

which in turn is a Lorentz invariant statement. Total mass M in a given global inertial frame at a fixed time t is

$$M \equiv \int_{\mathbb{R}^3} \mathrm{d}^3\vec{x} J^0. \tag{6.2.9}$$

To show it is a constant, we take the time derivative, and employ eq. (6.2.8):

$$\dot{M} = \int_{\mathbb{R}^3} \mathrm{d}^3\vec{x}\partial_0 J^0 = -\int_{\mathbb{R}^3} \mathrm{d}^3\vec{x}\partial_i J^i. \tag{6.2.10}$$

The divergence theorem tells us that this is equal to the flux of J^i at spatial infinity. But there is no J^i at spatial infinity for physically realistic – i.e., isolated – systems.

6.3 Noether: Continuous Symmetries and Conserved Currents

A field is a substance permeating spacetime. In this section, we shall attempt to associate with it energy-momentum at every location in spacetime, by identifying the Noether's currents associated with the symmetries of Minkowski spacetime. Specifically, the conservation of energy is due to the time translation symmetry of the system at hand. The conservation of linear momentum is due to its spatial translation symmetry; whereas the conservation of angular momentum is due to rotational symmetry. Throughout this discussion, we will assume the dynamics of the field theory is governed by some Lorentz invariant Lagrangian density that depends on the field, and on its first derivatives – but no higher.

Spacetime Translations and Stress-Energy Tensor The physical interpretation delineated here for the components of $T^{\widehat{\mu}\widehat{\nu}}$ is really an assertion. Let us attempt to justify it partially, by appealing to the flat spacetime limit, where the momentum of a classical field theory may be viewed as the conserved Noether current of spacetime translation symmetry. Specifically, let us analyze the canonical scalar field theory of eq. (6.1.8) but with $g_{\mu\nu} = \eta_{\mu\nu}$.

$$\mathcal{L}(x) = \frac{1}{2}\eta^{\mu\nu}\partial_\mu\varphi(x)\partial_\nu\varphi(x) - V\left(\varphi(x)\right). \tag{6.3.1}$$

Since $\mathcal{L}$ is Lorentz invariant, we may consider an infinitesimal spacetime displacement,

$$x^\mu = x'^\mu + a^\mu, \tag{6.3.2}$$

for constant but 'small' a^μ.

$$\mathcal{L}(x) = \mathcal{L}(x') + a^\mu \partial_{\mu'}\mathcal{L}(x') + \mathcal{O}\left(a^2\right). \tag{6.3.3}$$

On the other hand, $\partial/\partial x^\mu = \partial_\mu = \partial_{\mu'} = \partial/\partial x'^\mu$ and

$$\begin{aligned}\mathcal{L}(x'+a) &= \mathcal{L}(x') + \frac{\partial\mathcal{L}}{\partial\varphi(x')}a^\nu\partial_{\nu'}\varphi(x') + \frac{\partial\mathcal{L}}{\partial\partial_{\mu'}\varphi(x')}a^\nu\partial_{\nu'}\partial_{\mu'}\varphi(x') + \mathcal{O}\left(a^2\right) &(6.3.4)\\ &= \mathcal{L}(x') + a^\nu\partial_{\nu'}\varphi(x')\left\{\frac{\partial\mathcal{L}}{\partial\varphi(x')} - \partial_{\mu'}\frac{\partial\mathcal{L}}{\partial\partial_{\mu'}\varphi(x')}\right\} + a^\nu\partial_{\mu'}\left(\partial_{\nu'}\varphi(x')\frac{\partial\mathcal{L}}{\partial\partial_{\mu'}\varphi(x')}\right) + \mathcal{O}\left(a^2\right). &(6.3.5)\end{aligned}$$

Using the equations-of-motion for the scalar field

$$\frac{\partial\mathcal{L}}{\partial\varphi(x')} - \partial_{\mu'}\frac{\partial\mathcal{L}}{\partial\partial_{\mu'}\varphi(x')} = 0, \tag{6.3.6}$$

eq. (6.3.5) becomes

$$\mathcal{L}(x'+a) = \mathcal{L}(x') + a^\nu\partial_{\mu'}\left(\partial_{\nu'}\varphi(x')\frac{\partial\mathcal{L}}{\partial\partial_{\mu'}\varphi(x')}\right). \tag{6.3.7}$$

We may now equate the linear-in-a^ν terms on the right hand sides of equations (6.3.3) and (6.3.7), and find the following conservation law:

$$\partial_{\mu'}\left\{a^\gamma\left(\partial_{\gamma'}\varphi(x')\frac{\partial\mathcal{L}}{\partial\partial_{\mu'}\varphi(x')} - \delta^\mu_\gamma\mathcal{L}(x')\right)\right\} = 0. \tag{6.3.8}$$

By setting $a^\gamma = \delta^\gamma_\nu$, for a fixed ν, we may identify the conserved quantity inside the $\{\dots\}$ as the Noether momentum p_ν due to translation symmetry along the ν-th direction.[7] Doing so now allows us to identify the conserved stress tensor

$$T^\mu{}_\nu = \partial_{\nu'}\varphi(x')\frac{\partial\mathcal{L}}{\partial\partial_{\mu'}\varphi(x')} - \delta^\mu_\nu\mathcal{L}(x'). \tag{6.3.9}$$

[7] As a simple parallel to the situation here: in classical mechanics, because the free Lagrangian $L = (1/2)\dot{\vec{x}}^2$ is space-translation invariant, $\partial L/\partial x^i = 0$, we may identify the momentum $p_i \equiv \partial L/\partial \dot{x}^i$ as the corresponding Noether charge.

Applying this to eq. (6.3.1), we obtain

$$T^{\mu}{}_{\nu} = \partial_\nu\varphi\partial^\mu\varphi - \delta^\mu_\nu\left(\frac{1}{2}(\partial\varphi)^2 - V(\varphi)\right). \tag{6.3.10}$$

It is possible to obtain the same result by first writing the scalar action in curved spacetime, and reading off $T^{\mu\nu}$ as the coefficient of $-(1/2)\delta g_{\mu\nu}$ upon perturbing $g_{\mu\nu} \to g_{\mu\nu}+\delta g_{\mu\nu}$. Unfortunately, this procedure does not yield a unique $T^{\mu\nu}$, let alone a necessarily gauge-invariant one. (This issue has a long history, starting from at least [**?**, **?**].)

Interpretation of $J^\mu{}_\nu a^\nu$. Just as the local conservation of mass or electric charge leads to their appropriate currents J^μ and their divergence free properties $\partial_\mu J^\mu$, we shall see here that the conservation of energy and momentum leads us to the divergence-less energy-momentum-shear-stress tensor – or, more commonly, the energy-momentum or stress-energy tensor.

In a given inertial frame, we associate time translation symmetry with the conservation of energy. Hence, by choosing $a^\nu\partial_\nu = \partial_t$, we may associate $J^\mu{}_0$ as the energy current. Whenever it is timelike, the zeroth component $J^0{}_0$ may be associated with the co-moving energy density; whereas J^{i0} is the momentum density (i.e., energy per time per area perpendicular to J^{i0}).

In the same inertial frame, we associate translation symmetry in the ith spatial direction with the conservation of the ith component of momentum. By choosing $a^\nu\partial_\nu = \partial_i$, we may associate $J^\mu{}_i$ with the current associated with the ith component of the (spatial) momentum. The zeroth component $J^0{}_i$ is the density of the ith component of momentum; this tells us $J^{0i} \sim J^{i0}$ (up to an overall sign). Whereas $J^k{}_i$ is the ith component of momentum per unit time across the spatial surface perpendicular to the kth spatial direction. In particular, when $k = i$, this would be the momentum per unit time through the surface perpendicular to the $i = k$th direction – but this is simply the pressure (force per unit time) acting on an infinitesimal slab between x^i and $x^i + \mathrm{d}x^i$ in the $i = k$ direction. For $i \neq k$, the $J^k{}_i$ is shear: force in the ith direction per unit area perpendicular to the kth direction. Now, if the force in the kth direction per unit area perpendicular to the ith direction were not equal to the $J^k{}_i$, there will be a torque generated on the (i,k) plane on an infinitesimal area. Hence, we expect $J^{ki} = J^{ik}$.

All these considerations allow us to identify the components $J^{\alpha\beta} \equiv T^{\alpha\beta}$ as those of the energy-momentum-shear-stress (note: stress $\equiv$ pressure) tensor, the flux of the μth component of energy-momentum across the hypersurface orthogonal to the νth direction.

- T^{00} is the energy density ($\equiv$ energy per unit spatial volume).
- $T^{0i} = T^{i0}$ is the linear momentum density ($\equiv$ energy per unit time per unit area perpendicular to the ith direction).
- $T^{ij} = T^{ji}$ for $i \neq j$ is the shear density; the flow of the ith component of momentum per unit time per unit area perpendicular to the jth surface.
- T^{ii} is the pressure/stress ($\equiv$ force per unit area) in the ith direction.

Noether: General Case Let us suppose that a small change is induced on some real scalar field $\delta\varphi$ that leaves the Lagrangian invariant up to a total derivative.

$$\varphi \to \varphi + \delta\varphi, \tag{6.3.11}$$
$$\mathcal{L} \to \mathcal{L} + \partial_\mu K^\mu. \tag{6.3.12}$$

On the other hand, we may expand

$$\partial_\mu K^\mu = \frac{\partial \mathcal{L}}{\partial \varphi}\delta\varphi + \frac{\partial \mathcal{L}}{\partial \partial_\mu \varphi}\partial_\mu \delta\varphi \tag{6.3.13}$$

$$= \left(\frac{\partial \mathcal{L}}{\partial \varphi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial \partial_\mu \varphi}\right)\delta\varphi + \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial \partial_\mu \varphi}\delta\varphi\right) \tag{6.3.14}$$

The first group of terms on the right hand side of the second equality is the Euler-Lagrange operation on the Lagrangian. In particular, we see that – if the EoM of the scalar is satisfied – then we may identify

$$\partial_\mu J^\mu = 0 \tag{6.3.15}$$

$$J^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)}\delta\varphi - K^\mu. \tag{6.3.16}$$

Noether's Currents: Ambiguities Notice we read off the Noether current from a divergence equation of the form $\partial_\mu J^\mu = 0$. That means we may add an identically conserved current to the LHS and, hence, yield a different Noether J^μ. It turns out, the most general identically conserved current takes the form $\partial_\mu \partial_\nu K^{\mu\nu}$ for arbitrary but anti-symmetric $K^{\mu\nu} = -K^{\nu\mu}$. Hence, Noether currents are always ambiguous up to this additive $\partial_\nu K^{\mu\nu}$ term. Furthermore, if J^μ is conserved, so is a constant A times it. To sum,

> If, due to some continuous symmetry, J^μ is conserved when evaluated on the solutions to the Euler-Lagrange equations, so is
>
> $$J'^\mu \equiv A \cdot (J^\mu + \partial_\nu K^{\mu\nu}) \tag{6.3.17}$$
>
> for arbitrary constant A and anti-symmetric $K^{\mu\nu} = -K^{\nu\mu}$.

This means the stress energy tensor we 'derived' earlier is, likewise, ambiguous in the same manner. In a given physical situation, therefore, we need additional criteria to pin down the precise physical meanings of the components of the Noether currents.

Internal SO$_D$ Example If a Lagrangian involves 3 scalar fields φ^{I} such that the former is invariant under global SO$_D$ rotations of the latter:

$$\mathcal{L} = \frac{1}{2}\partial_\alpha \varphi^{\mathrm{I}} \partial^\alpha \varphi^{\mathrm{I}} - V\left(\varphi^{\mathrm{I}}\varphi^{\mathrm{I}}\right), \tag{6.3.18}$$

$$\varphi^{\mathrm{I}} \to \widehat{R}^{\mathrm{I}}{}_{\mathrm{J}}\varphi^{\mathrm{J}}, \tag{6.3.19}$$

$$\mathcal{L} \to \mathcal{L}. \tag{6.3.20}$$

Under infinitesimal rotations, we may rotate the pairs $(1,2)$, $(1,3)$ and $(2,3)$.

$$\varphi^{\mathrm{I}} \to \varphi^{\mathrm{I}} - i\theta \left(\widehat{J}^{\mathrm{AB}}\right)_{\mathrm{IJ}} \varphi^{\mathrm{J}} \tag{6.3.21}$$

Recalling that $i(\widehat{J}^{\mathrm{AB}})_{\mathrm{IJ}} = \delta^{\mathrm{A}}_{[\mathrm{I}}\delta^{\mathrm{B}}_{\mathrm{J}]}$, we may recognize

$$\delta\varphi^{\mathrm{I}} = \delta^{\mathrm{I}[\mathrm{A}}\varphi^{\mathrm{B}]}. \tag{6.3.22}$$

The Noether currents – one for each rotation generator – are

$$J^{\mu}_{\mathrm{K}} = \frac{\partial \mathcal{L}}{\partial \partial_\mu \varphi^{\mathrm{I}}} \delta^{\mathrm{I}[\mathrm{A}} \varphi^{\mathrm{B}]} \tag{6.3.23}$$

$$= \partial^\mu \varphi^{[\mathrm{A}} \cdot \varphi^{\mathrm{B}]}. \tag{6.3.24}$$

We may check explicitly that this is conserved. Firstly, the EoMs are

$$\partial^2 \varphi^{\mathrm{I}} = -2\varphi^{\mathrm{I}} V'(\vec{\varphi}^2). \tag{6.3.25}$$

Hence,

$$\partial_\mu J^\mu = \partial^2 \varphi^{[\mathrm{A}} \cdot \varphi^{\mathrm{B}]} + \partial^\mu \varphi^{[\mathrm{I}} \cdot \partial_\mu \varphi^{\mathrm{B}]} \tag{6.3.26}$$

$$= -2V' \cdot \varphi^{[\mathrm{A}} \cdot \varphi^{\mathrm{B}]} = 0. \tag{6.3.27}$$

Problem 6.2. Noether, Lorentz and Angular Momentum Above, we consider the Noether current $T^{\mu}{}_{\alpha}$, obeying $\partial_\mu T^{\mu}{}_{\alpha} = 0$. corresponding to spacetime translation symmetry.

Let us now consider the Noether current $J^{\mu}{}_{\alpha\beta}$, obeying $\partial_\mu J^{\mu}{}_{\alpha\beta} = 0$, from the Lorentz transformation $x^\alpha \to \Lambda^{\alpha}{}_{\beta} x^\beta$. (Why are there two extra indices $(\alpha\beta)$ on the Noether current? Hint: How are the Lorentz generators labeled?) Show that it is possible to obtain

$$J^{\mu\alpha\beta} = T^{\mu[\alpha} x^{\beta]}, \tag{6.3.28}$$

where $T^{\mu\nu}$ is the Noether current of spacetime translations in eq. (6.3.9). Interpret the components of $J^{\mu\alpha\beta}$; i.e., what is the Noether current of spatial rotations? And of boosts? □

Problem 6.3. Symmetric Noether Currents Explain why, if $\partial_\mu J^{\mu\alpha\beta} = 0$, where $J^{\mu\alpha\beta}$ is given by eq. (6.3.28), then $T^{\mu\nu} = T^{\nu\mu}$.

In other words, if we can obtain the Noether current of Lorentz transformations to be related to that of spacetime translations in the form of eq. (6.3.28), then the Noether current of spacetime translations must be a symmetric tensor. This symmetry property is important for interpreting $T^{\mu\nu}$ as the energy-momentum-stress tensor. □

6.4 Hamiltonian Formulation

1D Particle Mechanics: Review In particle mechanics, from the Lagrangian $L(q, \dot{q})$, we may define the momentum conjugate to the (generalized) position q as

$$p \equiv \left(\frac{\partial L(q, \dot{q})}{\partial \dot{q}} \right)_q . \tag{6.4.1}$$

This relation between p and $(q, \dot{q})$ usually allows us to invert $\dot{q}$ for the pair (q, p); so that every variable can now be expressed in terms of this pair – this allows the interpretation of (q, p) as *independent* "phase space" variables in what follows. The Hamiltonian itself is

$$H(q, p) \equiv p \cdot \dot{q}(q, p) - L(q, p), \tag{6.4.2}$$

$$L(q, p) \equiv L(q, \dot{q}(q, p)). \tag{6.4.3}$$

Hamilton's equations now reads

$$\dot{q} = \frac{\partial H}{\partial p} \qquad \text{and} \qquad \dot{p} = -\frac{\partial H}{\partial q}. \tag{6.4.4}$$

Field Theory For field theory, an analogous discussion follows. The Lagrangian $\mathcal{L}(\varphi, \partial\varphi) = \mathcal{L}(\varphi, \dot{\varphi}, \vec{\nabla}\varphi)$ depends on the field (which is analogous to the position q) and its partial derivatives. The time derivative $\dot{\varphi} \equiv \partial_0\varphi$ is analogous to $\dot{q}$; this, in turn, allows us to define the momentum conjugate to φ as

$$\Pi(x) \equiv \frac{\partial\mathcal{L}}{\partial(\partial_0\varphi(x))}. \tag{6.4.5}$$

Like the particle mechanics case, we shall assume it is possible to solve $\dot{\varphi}$ in terms of φ and Π. We then define the Hamiltonian density via the Legendre transform:

$$\mathcal{H} \equiv \Pi \cdot \partial_0\varphi - \mathcal{L}. \tag{6.4.6}$$

We may vary this Legendre transform,

$$\delta\mathcal{H} = \delta\Pi \cdot \partial_0\varphi + \Pi \cdot \partial_0\delta\varphi - \frac{\partial\mathcal{L}}{\partial\varphi}\delta\varphi - \frac{\partial\mathcal{L}}{\partial\partial_0\varphi}\partial_0\delta\varphi - \frac{\partial\mathcal{L}}{\partial\partial_i\varphi}\partial_i\delta\varphi \tag{6.4.7}$$

$$= \delta\Pi \cdot \partial_0\varphi + \Pi \cdot \partial_0\delta\varphi - \frac{\partial\mathcal{L}}{\partial\varphi}\delta\varphi - \Pi \cdot \partial_0\delta\varphi - \frac{\partial\mathcal{L}}{\partial\partial_i\varphi}\partial_i\delta\varphi \tag{6.4.8}$$

$$= \delta\Pi \cdot \partial_0\varphi - \frac{\partial\mathcal{L}}{\partial\varphi}\delta\varphi + \partial_i\frac{\partial\mathcal{L}}{\partial\partial_i\varphi}\delta\varphi - \partial_i\left(\frac{\partial\mathcal{L}}{\partial\partial_i\varphi}\delta\varphi\right). \tag{6.4.9}$$

Applying the Euler-Lagrange equations

$$\partial_0\frac{\partial\mathcal{L}}{\partial\partial_0\varphi} + \partial_i\frac{\partial\mathcal{L}}{\partial\partial_i\varphi} = \frac{\partial\mathcal{L}}{\partial\varphi}, \tag{6.4.10}$$

$$\partial_i\frac{\partial\mathcal{L}}{\partial\partial_i\varphi} = \frac{\partial\mathcal{L}}{\partial\varphi} - \partial_0\Pi. \tag{6.4.11}$$

to eq. (6.4.9), we find that

$$\delta\mathcal{H} = \delta\Pi \cdot \partial_0\varphi - \frac{\partial\mathcal{L}}{\partial\varphi}\delta\varphi + \left(\frac{\partial\mathcal{L}}{\partial\varphi} - \dot{\Pi}\right)\delta\varphi - \partial_i\left(\frac{\partial\mathcal{L}}{\partial\partial_i\varphi}\delta\varphi\right) \tag{6.4.12}$$

$$= \delta\Pi \cdot \dot{\varphi} - \dot{\Pi} \cdot \delta\varphi - \partial_i\left(\frac{\partial\mathcal{L}}{\partial\partial_i\varphi}\delta\varphi\right). \tag{6.4.13}$$

On the other hand, we may vary the Hamiltonian as a function of the field, its spatial gradient, and the conjugate momentum:

$$\varphi \to \varphi + \delta\varphi, \tag{6.4.14}$$

$$\Pi \to \Pi + \delta\Pi; \tag{6.4.15}$$

and discover

$$\delta\mathcal{H} = \frac{\partial\mathcal{H}}{\partial\Pi}\cdot\delta\Pi + \frac{\partial\mathcal{H}}{\partial\varphi}\cdot\delta\varphi + \frac{\partial\mathcal{H}}{\partial\partial_i\varphi}\cdot\partial_i\delta\varphi \tag{6.4.16}$$

$$= \frac{\partial\mathcal{H}}{\partial\Pi}\cdot\delta\Pi + \left(\frac{\partial\mathcal{H}}{\partial\varphi} - \partial_i\frac{\partial\mathcal{H}}{\partial\partial_i\varphi}\right)\cdot\delta\varphi + \partial_i\left(\frac{\partial\mathcal{H}}{\partial\partial_i\varphi}\cdot\delta\varphi\right). \tag{6.4.17}$$

Comparing the two variation results,

$$\delta\Pi\cdot\left(\dot{\varphi} - \frac{\partial\mathcal{H}}{\partial\Pi}\right) - \left(\dot{\Pi} + \frac{\partial\mathcal{H}}{\partial\varphi} - \partial_i\frac{\partial\mathcal{H}}{\partial\partial_i\varphi}\right)\delta\varphi = \partial_i\left\{\left(\frac{\partial\mathcal{H}}{\partial\partial_i\varphi} + \frac{\partial\mathcal{L}}{\partial\partial_i\varphi}\right)\delta\varphi\right\}. \tag{6.4.18}$$

If we integrate both sides over space, the right hand side will be converted into a surface integral at spatial infinity, which we may argue should fall off to zero as long as $\delta\varphi$ does.

$$\int_{\mathbb{R}^D} \mathrm{d}^D\vec{x}\left\{\delta\Pi\cdot\left(\dot{\varphi} - \frac{\partial\mathcal{H}}{\partial\Pi}\right) - \left(\dot{\Pi} + \frac{\partial\mathcal{H}}{\partial\varphi} - \partial_i\frac{\partial\mathcal{H}}{\partial\partial_i\varphi}\right)\delta\varphi\right\} = 0 \tag{6.4.19}$$

By viewing Π and φ as independent variables, the coefficients of their variations on the left hand side must therefore be individually zero because $\delta\Pi$ and $\delta\varphi$ are arbitrary at every point in space.

$$\dot{\varphi} = \frac{\partial\mathcal{H}}{\partial\Pi} \tag{6.4.20}$$

$$\dot{\Pi} = \partial_i\frac{\partial\mathcal{H}}{\partial\partial_i\varphi} - \frac{\partial\mathcal{H}}{\partial\varphi} \tag{6.4.21}$$

This in turn implies, the right hand side of eq. (6.4.18) must be zero too. And since $\delta\varphi$ was arbitrary,

$$\frac{\partial\mathcal{H}}{\partial(\partial_i\varphi)} = -\frac{\partial\mathcal{L}}{\partial(\partial_i\varphi)}. \tag{6.4.22}$$

Example Let us work out Hamilton's equations for the canonical scalar field in eq. (6.3.1).

$$\Pi = \frac{\partial\mathcal{L}}{\partial\dot{\varphi}} = \dot{\varphi} \tag{6.4.23}$$

The Lagrangian is thus

$$\mathcal{L} = \frac{1}{2}\Pi^2 - \frac{1}{2}(\vec{\nabla}\varphi)^2 - V. \tag{6.4.24}$$

The Legendre transform now reads

$$\mathcal{H} = \Pi^2 - \left(\frac{1}{2}\Pi^2 - \frac{1}{2}(\vec{\nabla}\varphi)^2 - V\right) \tag{6.4.25}$$

$$= \frac{1}{2}\Pi^2 + \frac{1}{2}(\vec{\nabla}\varphi)^2 + V. \tag{6.4.26}$$

Hamilton's equations are

$$\dot{\varphi} = \Pi, \tag{6.4.27}$$
$$\dot{\Pi} = \vec{\nabla}^2\varphi - V'(\varphi). \tag{6.4.28}$$

When combined, they simply yield $\ddot{\varphi} - \vec{\nabla}^2\varphi = -V'$ as before. We may also readily verify

$$\frac{\partial\mathcal{H}}{\partial(\partial_i\varphi)} = \partial_i\varphi = -\frac{\partial\mathcal{L}}{\partial(\partial_i\varphi)}. \tag{6.4.29}$$

Problem 6.4. Lagrangians from Hamiltonians Starting from the relationship between the Hamiltonian and the Lagrangian in eq. (6.4.6) – but in terms of Π, $\partial_i\varphi$ and φ – show that Hamilton's equations in equations (6.4.20) and (6.4.21) imply the Euler-Lagrange equation $\partial_\mu(\partial\mathcal{L}/\partial(\partial_\mu\varphi)) = \partial\mathcal{L}/\partial\varphi$. □

Problem 6.5. Hamilton's Equations for O_N model Work out Hamilton's equations from the Lagrangian in eq. (6.3.18). □

Problem 6.6. Poisson Brackets and Time Evolution Define the 'equal-time' functional derivatives

$$\frac{\delta\varphi(t,\vec{x})}{\delta\varphi(t,\vec{y})} = \delta^{(D)}(\vec{x}-\vec{y}), \tag{6.4.30}$$
$$\frac{\delta\Pi(t,\vec{x})}{\delta\Pi(t,\vec{y})} = \delta^{(D)}(\vec{x}-\vec{y}), \tag{6.4.31}$$
$$\frac{\delta\varphi(t,\vec{x})}{\delta\Pi(t,\vec{y})} = 0 = \frac{\delta\Pi(t,\vec{x})}{\delta\varphi(t,\vec{y})}; \tag{6.4.32}$$

as well as the 'equal-time' Poisson bracket

$$\begin{aligned} &\{f(\varphi(t,\vec{x}),\Pi(t,\vec{x})), g(\varphi(t,\vec{x}'),\Pi(t,\vec{x}'))\} \\ &\equiv \int_{\mathbb{R}^D} \mathrm{d}^D\vec{y}\left(\frac{\delta f(t,\vec{x})}{\delta\varphi(t,\vec{y})}\frac{\delta g(t,\vec{x}')}{\delta\Pi(t,\vec{y})} - \frac{\delta g(t,\vec{x})}{\delta\varphi(t,\vec{y})}\frac{\delta f(t,\vec{x})}{\delta\Pi(t,\vec{y})}\right). \end{aligned} \tag{6.4.33}$$

Show that, for the canonical Hamiltonian in eq. (6.4.26), the Poisson bracket equations

$$\partial_t\varphi(t,\vec{x}) = \{\varphi(t,\vec{x}), H(t)\}, \tag{6.4.34}$$
$$\partial_t\Pi(t,\vec{x}) = \{\Pi(t,\vec{x}), H(t)\} \tag{6.4.35}$$

involving the total Hamiltonian

$$H(t) \equiv \int_{\mathbb{R}^D} \mathrm{d}^D\vec{x}''\mathcal{H}(t,\vec{x}''); \tag{6.4.36}$$

is in fact equivalent to Hamilton's equations. □

6.5 Fourier Space(time)

One of the key insights that we need for doing perturbative field theory in Minkowski spacetime, is that the linear massive wave equation is really an infinite collection of simple harmonic oscillators (SHOs) in Fourier space. That is, if we decompose

$$\varphi(t, \vec{x}) = \int_{\mathbb{R}^D} \frac{\mathrm{d}^D \vec{k}}{(2\pi)^D} \widetilde{\varphi}(t, \vec{k}) e^{i\vec{k}\cdot\vec{x}} = \int_{\mathbb{R}^D} \frac{\mathrm{d}^D \vec{k}}{(2\pi)^D} \widetilde{\varphi}(t, \vec{k}) e^{-ik_j x^j}, \tag{6.5.1}$$

we see that

$$\partial^2 \varphi(x) = \int \frac{\mathrm{d}^3 \vec{k}}{(2\pi)^3} \left(\ddot{\widetilde{\varphi}} + \vec{k}^2 \widetilde{\varphi}\right) e^{i\vec{k}\cdot\vec{x}} \tag{6.5.2}$$

because each spatial derivative acting on the exponential amounts to the replacement rule $\partial_j \to -ik_j$ because

$$\partial_j e^{i\vec{k}\cdot\vec{x}} = \partial_j \left(-ik_l x^l\right) e^{i\vec{k}\cdot\vec{x}} \tag{6.5.3}$$

$$= \left(-ik_l \delta^l_j\right) e^{i\vec{k}\cdot\vec{x}} \tag{6.5.4}$$

$$= \left(-ik_j\right) e^{i\vec{k}\cdot\vec{x}}. \tag{6.5.5}$$

The wave equation sourced by some external J,

$$\ddot{\varphi}(t, \vec{x}) - \vec{\nabla}^2 \varphi(t, \vec{x}) + m^2 \varphi(t, \vec{x}) = J(t, \vec{x}) \tag{6.5.6}$$

becomes the driven SHO equation

$$\ddot{\widetilde{\varphi}}(t, \vec{k}) + E_{\vec{k}}^2 \widetilde{\varphi}(t, \vec{k}) = \widetilde{J}(t, \vec{k}), \tag{6.5.7}$$

where the SHO frequency is the energy

$$E_{\vec{k}} \equiv \sqrt{\vec{k}^2 + m^2}. \tag{6.5.8}$$

This insight is particularly important when we later quantize φ, because we may then view each Fourier mode of the scalar field as a *quantum* SHO. But let's first turn to two key physical features of our scalar field. First, for the linear wave equation sourced by J, let us witness that the presence or absence of the mass m determines whether φ itself is short- or long-ranged.

Classical Linear Solutions Remember that the retarded Green's function of the SHO with frequency Ω is

$$G_{\text{SHO}}(T) = \Theta(T) \frac{\sin(\Omega T)}{\Omega}, \tag{6.5.9}$$

where

$$\left(\frac{\mathrm{d}^2}{\mathrm{d}T^2} + \Omega^2\right) G_{\text{SHO}}(T) = \delta(T). \tag{6.5.10}$$

The classical causality-obeying solution to eq. (6.5.7) is thus

$$\widetilde{\varphi}(t,\vec{k}) = \int_{-\infty}^{t} \frac{\sin(E_{\vec{k}}(t-t'))}{E_{\mathrm{k}}} \cdot \widetilde{J}\left(t',\vec{k}\right) \mathrm{d}t', \tag{6.5.11}$$

$$\varphi(x) = \int_{\mathbb{R}^D} \mathrm{d}^D\vec{x}' \int_{\mathbb{R}} \mathrm{d}t' \left(\Theta(t-t') \int_{\mathbb{R}^D} \frac{\mathrm{d}^D\vec{k}}{(2\pi)^D} e^{i\vec{k}\cdot(\vec{x}-\vec{x}')} \frac{\sin(E_{\vec{k}}(t-t'))}{E_k}\right) J\left(t',\vec{x}'\right) \tag{6.5.12}$$

$$\equiv \int_{\mathbb{R}^D} \mathrm{d}^D\vec{x}' \int_{\mathbb{R}} \mathrm{d}t' G_{\mathrm{ret}}(x-x') J\left(x'\right). \tag{6.5.13}$$

Problem 6.7. Long or Short Range? Massive versus Massless Consider a static point mass resting at $\vec{x}=0$ in the $\{x^\mu\}$ inertial frame, namely

$$J(t,\vec{x}) = J_0 \delta^{(3)}\left(\vec{x}\right), \qquad J_0 \text{ constant}. \tag{6.5.14}$$

Solve eq. (6.5.6). Hint: You may assume the time derivatives in eq. (6.0.9) can be neglected. Then go to Fourier $\vec{k}-$space. You should find

$$\widetilde{\varphi}(\vec{k}) = \frac{J_0}{\vec{k}^2 + m^2}. \tag{6.5.15}$$

You should find

$$\varphi(t, r\equiv|\vec{x}|, \theta, \phi) = J_0 \frac{\exp(-mr)}{4\pi r}. \tag{6.5.16}$$

That is, φ describes a short-range force (with range $1/m$) that, when $m\to 0$, recovers the long-range Coulomb/Newtonian $1/r$ potential. (Hint: The 3D Fourier integral can be reduced to a 1D integral, which can then be tackled by closing the contour on the complex plane.)

Next, consider an inertial frame $\{x'^\mu\}$ that is moving relative to the $\{x^\mu\}$ frame at velocity v along the positive x^3 axis. What is $\varphi(x')$ in the new frame? $\square$

Problem 6.8. Nonlinearities as Self-Coupled SHOs By going to Fourier space, consider a potential that is a polynomial of degree n in the field φ, with minimum at $\varphi=0$,

$$V(\varphi) = \sum_{\ell=2}^{n} \frac{p_\ell}{\ell}\varphi^\ell. \tag{6.5.17}$$

Show that the Fourier space version of eq. (6.0.9) is:

$$\ddot{\widetilde{\varphi}}(t,\vec{k}) + (k^2+p_2)\widetilde{\varphi}(t,\vec{k}) = \widetilde{J}(t,\vec{k}) - \sum_{\ell=2}^{n-1} p_{\ell+1} \prod_{s=1}^{\ell} \left(\int \frac{\mathrm{d}^3\vec{k}_s}{(2\pi)^3}\widetilde{\varphi}(t,\vec{k}_s)\right)(2\pi)^3\delta^{(3)}\left(\vec{k}-\vec{k}_1-\dots-\vec{k}_\ell\right). \tag{6.5.18}$$

Observe, for a given $\vec{k}$, the non-linearities of the potential $V(\phi)$ give rise to a driving force – the second line on the right hand side – due to the field itself but from superposing over a range of Fourier modes.

Hint: Let's work out the $\ell = 2$ contribution as an example. This is comes from the cubic p_3 term in the potential:

$$V'(\varphi) = p_3\varphi(t,\vec{x})^2 + \dots. \tag{6.5.19}$$

The Fourier decomposition of φ^2 is

$$\widetilde{\varphi^2}(t,\vec{k}) = \int \mathrm{d}^3\vec{x}\varphi(x)^2 e^{-i\vec{k}\cdot\vec{x}} \tag{6.5.20}$$

$$= \int \mathrm{d}^3\vec{x} \int \frac{\mathrm{d}^3\vec{k}_1}{(2\pi)^3}\widetilde{\varphi}(t,\vec{k}_1)e^{i\vec{k}_1\cdot\vec{x}} \int \frac{\mathrm{d}^3\vec{k}_2}{(2\pi)^3}\widetilde{\varphi}(t,\vec{k}_2)e^{i\vec{k}_2\cdot\vec{x}}e^{-i\vec{k}\cdot\vec{x}} \tag{6.5.21}$$

$$= \int \frac{\mathrm{d}^3\vec{k}_1}{(2\pi)^3}\widetilde{\varphi}(t,\vec{k}_1) \int \frac{\mathrm{d}^3\vec{k}_2}{(2\pi)^3}\widetilde{\varphi}(t,\vec{k}_2) \int \mathrm{d}^3\vec{x} e^{i(\vec{k}_1+\vec{k}_2-\vec{k})\cdot\vec{x}} \tag{6.5.22}$$

$$= \int \frac{\mathrm{d}^3\vec{k}_1}{(2\pi)^3}\widetilde{\varphi}(t,\vec{k}_1) \int \frac{\mathrm{d}^3\vec{k}_2}{(2\pi)^3}\widetilde{\varphi}(t,\vec{k}_2)(2\pi)^3\delta^{(3)}\left(\vec{k}-\vec{k}_1-\vec{k}_2\right). \tag{6.5.23}$$

□

Problem 6.9. Dispersion relations Consider the *massive* Klein-Gordon equation in Minkowski spacetime:

$$\left(\partial^2 + m^2\right)\varphi(t,\vec{x}) = 0, \tag{6.5.24}$$

where φ is a real scalar field. Find the general solution for φ in terms of plane waves $\exp(-ik\cdot x)$ and obtain the dispersion relation:

$$k^2 = m^2 \qquad \Leftrightarrow \qquad E^2 = \vec{p}^2 + m^2, \tag{6.5.25}$$

$$E \equiv k^0, \qquad \vec{p} \equiv \vec{k}. \tag{6.5.26}$$

If each plane wave is associated with a particle of $d-$momentum k_μ, this states that it has mass m. The photon, which obeys $k^2 = 0$, has zero mass. □

6.6 ⋆Uniqueness of Lagrangians

In this section, let us ask the following question. Suppose two Lagrangian densities $\mathcal{L}(\varphi,\partial\varphi)$ and $\mathcal{L}'(\varphi,\partial\varphi)$, which we shall assume only depends on φ and its first derivatives, yield the same EoM:

$$\frac{\partial\mathcal{L}}{\partial\varphi} - \partial_\mu\frac{\partial\mathcal{L}}{\partial\partial_\mu\varphi} = \frac{\partial\mathcal{L}'}{\partial\varphi} - \partial_\mu\frac{\partial\mathcal{L}'}{\partial\partial_\mu\varphi}. \tag{6.6.1}$$

What is the most general $\Delta \equiv \mathcal{L} - \mathcal{L}'$? In particular, this means Δ solves the Euler-Lagrange equation identically – namely, its form should not depend on the specific solution of φ.

$$0 = \frac{\partial\Delta}{\partial\varphi} - \partial_\mu\frac{\partial\Delta}{\partial\partial_\mu\varphi} \tag{6.6.2}$$

$$= \frac{\partial\Delta}{\partial\varphi} - \partial_\mu\varphi\frac{\partial^2\Delta}{\partial\varphi\partial\partial_\mu\varphi} - \partial_\mu\partial_\gamma\varphi\frac{\partial^2\Delta}{\partial\partial_\gamma\varphi\partial\partial_\mu\varphi} \tag{6.6.3}$$

Since this holds for any φ, the only way the second derivative terms vanish is

$$\frac{\partial^2 \Delta}{\partial \partial_\gamma \varphi \partial \partial_\mu \varphi} = 0, \tag{6.6.4}$$

i.e., Δ can only be linear in $\partial\varphi$.

$$\Delta = \Delta_{\mathrm{I}}(\varphi) + \partial_\alpha \varphi \Delta_{\mathrm{II}}^\alpha(\varphi) \tag{6.6.5}$$

Euler-Lagrange then reduces to

$$\Delta_{\mathrm{I}}'(\varphi) + \partial_\alpha \varphi \Delta_{\mathrm{II}}^{\alpha\prime}(\varphi) - \partial_\alpha \Delta_{\mathrm{II}}^\alpha(\varphi) = 0 \tag{6.6.6}$$

$$\Delta_{\mathrm{I}}'(\varphi) + \partial_\alpha \varphi \Delta_{\mathrm{II}}^{\alpha\prime}(\varphi) - \partial_\alpha \varphi \Delta_{\mathrm{II}}^{\alpha\prime}(\varphi) = 0 \tag{6.6.7}$$

$$\Delta_{\mathrm{I}}'(\varphi) = 0. \tag{6.6.8}$$

Hence, if we define

$$\Delta_0^\alpha(z) \equiv \int^z \Delta_{\mathrm{II}}^\alpha(z')\mathrm{d}z', \tag{6.6.9}$$

most general Δ is a total divergence (plus an irrelevant constant):

$$\Delta = \partial_\alpha \Delta_0^\alpha(\varphi) = \partial_\alpha \varphi \Delta_{\mathrm{II}}^\alpha(\varphi). \tag{6.6.10}$$

> If two Lagrangian densities depending on φ, $\partial\varphi$ but no higher derivatives yield the same Euler-Lagrange equations-of-motion, they must differ only up to a total divergence.

Actually, if the single scalar theory is Lorentz invariant, $\Delta_{\mathrm{II}}^\alpha$ would have to be proportional to $\partial^\alpha\varphi$, since there are no other vectors in the problem at hand. But $\Delta_{\mathrm{II}}^\alpha$ does not depend on the derivatives of φ. Hence, the result is likely stronger: a field theory involving a *single scalar* that has a Lorentz invariant Lagrangian density $\mathcal{L}(\varphi, \partial\varphi)$ must be unique up to an additive constant.

7 Last update: March 6, 2025

References