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). □

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., $A^i B^i$ – they are implicitly summed over; namely, $A^i B^i \equiv \sum_i A^i B^i$.

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 $f(\lambda = \lambda_1) = 1$. □

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

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

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. $\square$

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

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

we multiply two such matrices to uncover

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

$$=\left[\begin{array}{ccc} 1 & \vec{0}^{\mathrm{T}} & 0 \\ \vec{V}_1+\widehat{R}_1\vec{V}_2 & \widehat{R}_1\widehat{R}_2 & \widehat{R}_1\vec{a}_2+\vec{a}_1 \\ 0 & \vec{0}^{\mathrm{T}} & 1 \end{array}\right] \tag{1.0.59}$$

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

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

$$\left[\begin{array}{cc} \frac{\partial t}{\partial t'} & \frac{\partial t}{\partial x'^b} \\ \frac{\partial x^a}{\partial t'} & \frac{\partial x^a}{\partial x'^b} \end{array}\right]^{-1}=\left[\begin{array}{cc} \frac{\partial t'}{\partial t} & \frac{\partial t'}{\partial x^b} \\ \frac{\partial x'^a}{\partial t} & \frac{\partial x'^a}{\partial x^b} \end{array}\right]. \tag{1.0.63}$$

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). $\Box$

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.

Problem 1.8. Schrödinger Equation If V is Galilean invariant, show that the Schrödinger equation

$$i\partial_t \psi = \left(-\frac{\vec{\nabla}^2_{\vec{x}}}{2m} + V(\vec{x})\right)\psi \tag{1.0.71}$$

is Galilean invariant – namely,

$$i\partial_{t'} \psi' = \left(-\frac{\vec{\nabla}^2_{\vec{x}'}}{2m} + V(\vec{x}')\right)\psi' \tag{1.0.72}$$

– provided the wave function itself transforms as $\psi = \psi' e^{i\vartheta}$. Work out the phase ϑ. □

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. (3.2.8):

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

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 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

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

as simply

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

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. (3.0.1). 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. (3.0.3) 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.

3.1 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

$$S \equiv \int_{t_\mathrm{i}}^{t_\mathrm{f}} L \mathrm{d}t, \tag{3.1.1}$$

$$L \equiv \frac{1}{2}m\dot{\vec{x}}(t)^2 - V(\vec{x}(t)). \tag{3.1.2}$$

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. (3.1.1), 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.

$$S \equiv \int_{t_\mathrm{i}}^{t_\mathrm{f}} \mathcal{L} \mathrm{d}^d x \tag{3.1.3}$$

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.

$$\begin{aligned} \delta_\varphi S &= \int_{t_\mathrm{i}}^{t_\mathrm{f}} \mathrm{d}^d x \left(\frac{\partial \mathcal{L}}{\partial \varphi} \delta\varphi + \frac{\partial \mathcal{L}}{\partial(\partial_\alpha \varphi)} \partial_\alpha \delta\varphi \right) \\ &= \left[\int \mathrm{d}^{d-1} \Sigma_\alpha \frac{\partial \mathcal{L}}{\partial(\partial_\alpha \varphi)} \delta\varphi \right]_{t_\mathrm{i}}^{t_\mathrm{f}} + \int_{t_\mathrm{i}}^{t_\mathrm{f}} \mathrm{d}^d x \delta\varphi \left(\frac{\partial \mathcal{L}}{\partial \varphi} - \partial_\alpha \frac{\partial \mathcal{L}}{\partial(\partial_\alpha \varphi)} \right) \end{aligned} \tag{3.1.4}$$

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

$$\frac{\partial \mathcal{L}}{\partial \varphi} = \partial_\alpha \frac{\partial \mathcal{L}}{\partial(\partial_\alpha \varphi)}. \tag{3.1.5}$$

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

$$\frac{\partial x^\alpha}{\partial x'^\mu} \partial_\alpha \varphi(x) = \partial_{\mu'} \varphi(x') \equiv \partial_{\mu'} \varphi\left(x(x')\right), \tag{3.1.6}$$

we have

$$\frac{\partial\mathcal{L}}{\partial(\partial_\alpha\varphi(x))} = \frac{\partial(\partial_{\mu'}\varphi(x'))}{\partial(\partial_\alpha\varphi(x))}\frac{\partial\mathcal{L}}{\partial(\partial_{\mu'}\varphi(x'))} = \frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial\mathcal{L}}{\partial(\partial_{\mu'}\varphi(x'))}. \tag{3.1.7}$$

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. (3.1.5), 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

$$S[\varphi] \equiv \int \mathrm{d}^d x \left(\frac{1}{2}\eta^{\mu\nu}\partial_\mu\varphi\partial_\nu\varphi - V(\varphi)\right), \tag{3.1.8}$$

where V is its scalar potential.

Problem 3.1. Show from eq. (3.1.5) that the equations derived from the action in eq. (3.1.8) is $\partial^2\varphi = -V'(\phi)$. $\square$

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 $\{\varphi^{\mathrm{I}}|\mathrm{I} = 1, 2, 3, \dots, N\}$:

$$S \equiv \int \mathrm{d}^d x \mathcal{L}\left(\eta^{\mu\nu}\partial_\mu\varphi^{\mathrm{I}}\partial_\nu\varphi^{\mathrm{I}}, \varphi^{\mathrm{I}}\varphi^{\mathrm{I}}\right). \tag{3.1.9}$$

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:

$$\varphi^{\mathrm{I}} \equiv \widehat{R}^{\mathrm{I}}{}_{\mathrm{J}}\varphi'^{\mathrm{J}}, \tag{3.1.10}$$

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,

$$\int \mathrm{d}^d x \mathcal{L}\left(\eta^{\mu\nu}\partial_\mu\varphi^{\mathrm{I}}\partial_\nu\varphi^{\mathrm{I}}, \varphi^{\mathrm{I}}\varphi^{\mathrm{I}}\right) = \int \mathrm{d}^d x \mathcal{L}\left(\eta^{\mu\nu}\partial_\mu\varphi'^{\mathrm{I}}\partial_\nu\varphi'^{\mathrm{I}}, \varphi'^{\mathrm{I}}\varphi'^{\mathrm{I}}\right). \tag{3.1.11}$$

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. (3.1.9) with respect to φ^{I}, reads

$$\frac{\partial\mathcal{L}}{\partial\varphi^{\mathrm{I}}} = \partial_\alpha\frac{\partial\mathcal{L}}{\partial(\partial_\alpha\varphi^{\mathrm{I}})}. \tag{3.1.12}$$

Under rotation, eq. (3.1.10) is equivalent to

$$\left(\widehat{R}^{-1}\right)^{\mathrm{J}}{}_{\mathrm{I}}\varphi^{\mathrm{I}} = \varphi'^{\mathrm{J}}, \tag{3.1.13}$$

which in turn tells us

$$\left(\widehat{R}^{-1}\right)^{\mathrm{J}}{}_{\mathrm{I}} \partial_\alpha \varphi^{\mathrm{I}} = \partial_\alpha \varphi'^{\mathrm{J}}. \tag{3.1.14}$$

Therefore eq. (3.1.12) becomes

$$\frac{\partial \varphi'^{\mathrm{J}}}{\partial \varphi^{\mathrm{I}}} \frac{\partial \mathcal{L}}{\partial \varphi'^{\mathrm{J}}} = \frac{\partial \partial_\alpha \varphi'^{\mathrm{J}}}{\partial \partial_\alpha \varphi^{\mathrm{I}}} \partial_\alpha \frac{\partial \mathcal{L}}{\partial(\partial_\alpha \varphi^{\mathrm{J}})}, \tag{3.1.15}$$

$$\left(\widehat{R}^{-1}\right)^{\mathrm{J}}{}_{\mathrm{I}} \frac{\partial \mathcal{L}}{\partial \varphi'^{\mathrm{J}}} = \left(\widehat{R}^{-1}\right)^{\mathrm{J}}{}_{\mathrm{I}} \partial_\alpha \frac{\partial \mathcal{L}}{\partial(\partial_\alpha \varphi^{\mathrm{J}})}. \tag{3.1.16}$$

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

3.2 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

$$M \equiv \int_{\text{tank}} \rho(t, \vec{x}) \mathrm{d}^3\vec{x} \tag{3.2.1}$$

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

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{\text{tank}} \rho \mathrm{d}^3\vec{x} = -\left(\int_{\text{cross section of 'in' pipe}} + \int_{\text{cross section of 'out' pipe}}\right) \mathrm{d}^2\vec{\Sigma} \cdot (\rho \vec{v}). \tag{3.2.2}$$

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,

$$\int_{\text{tank}} \dot{\rho} \mathrm{d}^3\vec{x} = -\int_{\text{tank}} \mathrm{d}^3\vec{x} \vec{\nabla} \cdot (\rho \vec{v}). \tag{3.2.3}$$

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

$$\dot{\rho} = -\partial_i \left(\rho \cdot v^i\right). \tag{3.2.4}$$

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

$$J^\mu(t,\vec{x}) \equiv \rho(t,\vec{x}) v^\mu(t,\vec{x}). \tag{3.2.5}$$

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

$$(1,\vec{0})^\mu \equiv v'^\mu = {\Lambda^\mu}_\nu(t,\vec{x}) v^\nu(t,\vec{x}). \tag{3.2.6}$$

and the mass density-current is now

$$J'^\mu = \rho(t',\vec{x}') v'^\mu = \rho(x') \cdot \delta^\mu_0. \tag{3.2.7}$$

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

$$\partial_\mu J^\mu = 0; \tag{3.2.8}$$

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

$$M \equiv \int_{\mathbb{R}^3} \mathrm{d}^3\vec{x} J^0. \tag{3.2.9}$$

To show it is a constant, we take the time derivative, and employ eq. (3.2.8):

$$\dot{M} = \int_{\mathbb{R}^3} \mathrm{d}^3\vec{x} \partial_0 J^0 = -\int_{\mathbb{R}^3} \mathrm{d}^3\vec{x} \partial_i J^i. \tag{3.2.10}$$

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.

3.3 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. (3.1.8) but with $g_{\mu\nu} = \eta_{\mu\nu}$.

$$\mathcal{L}(x) = \frac{1}{2}\eta^{\mu\nu}\partial_\mu\varphi(x)\partial_\nu\varphi(x) - V\left(\varphi(x)\right). \tag{3.3.1}$$

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

$$x^\mu = x'^\mu + a^\mu, \tag{3.3.2}$$

for constant but 'small' a^μ.

$$\mathcal{L}(x) = \mathcal{L}(x') + a^\mu\partial_{\mu'}\mathcal{L}(x') + \mathcal{O}\left(a^2\right). \tag{3.3.3}$$

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

$$\mathcal{L}(x'+a) = \mathcal{L}(x') + \frac{\partial\mathcal{L}}{\partial\varphi(x')}a^\nu\partial_{\nu'}\varphi(x') + \frac{\partial\mathcal{L}}{\partial\partial_{\mu'}\varphi(x')}a^\nu\partial_{\nu'}\partial_{\mu'}\varphi(x') + \mathcal{O}\left(a^2\right) \tag{3.3.4}$$

$$= \mathcal{L}(x') + a^\nu\partial_{\nu'}\varphi(x')\left\{\frac{\partial\mathcal{L}}{\partial\varphi(x')} - \partial_{\mu'}\frac{\partial\mathcal{L}}{\partial\partial_{\mu'}\varphi(x')}\right\} + a^\nu\partial_{\mu'}\left(\partial_{\nu'}\varphi(x')\frac{\partial\mathcal{L}}{\partial\partial_{\mu'}\varphi(x')}\right) + \mathcal{O}\left(a^2\right). \tag{3.3.5}$$

Using the equations-of-motion for the scalar field

$$\frac{\partial\mathcal{L}}{\partial\varphi(x')} - \partial_{\mu'}\frac{\partial\mathcal{L}}{\partial\partial_{\mu'}\varphi(x')} = 0, \tag{3.3.6}$$

eq. (3.3.5) becomes

$$\mathcal{L}(x'+a) = \mathcal{L}(x') + a^\nu\partial_{\mu'}\left(\partial_{\nu'}\varphi(x')\frac{\partial\mathcal{L}}{\partial\partial_{\mu'}\varphi(x')}\right). \tag{3.3.7}$$

We may now equate the linear-in-a^ν terms on the right hand sides of equations (3.3.3) and (3.3.7), and find the following conservation law:

$$\partial_{\mu'}\left\{a^\gamma\left(\partial_{\gamma'}\varphi(x')\frac{\partial\mathcal{L}}{\partial\partial_{\mu'}\varphi(x')} - \delta^\mu_\gamma\mathcal{L}(x')\right)\right\} = 0. \tag{3.3.8}$$

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.[1] Doing so now allows us to identify the conserved stress tensor

$$T^\mu{}_\nu = \partial_{\nu'}\varphi(x')\frac{\partial\mathcal{L}}{\partial\partial_{\mu'}\varphi(x')} - \delta^\mu_\nu\mathcal{L}(x'). \tag{3.3.9}$$

[1]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. (3.3.1), we obtain

$$T^{\mu}{}_{\nu} = \partial_\nu \varphi \partial^\mu \varphi - \delta^\mu_\nu \left(\frac{1}{2} (\partial \varphi)^2 - V(\varphi) \right). \tag{3.3.10}$$

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.

$$\varphi \to \varphi + \delta\varphi, \tag{3.3.11}$$
$$\mathcal{L} \to \mathcal{L} + \partial_\mu K^\mu. \tag{3.3.12}$$

On the other hand, we may expand

$$\partial_\mu K^\mu = \frac{\partial \mathcal{L}}{\partial \varphi}\delta\varphi + \frac{\partial \mathcal{L}}{\partial \partial_\mu \varphi}\partial_\mu \delta\varphi \tag{3.3.13}$$

$$= \left(\frac{\partial \mathcal{L}}{\partial \varphi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial \partial_\mu \varphi}\right)\delta\varphi + \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial \partial_\mu \varphi}\delta\varphi\right) \tag{3.3.14}$$

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

$$\partial_\mu J^\mu = 0 \tag{3.3.15}$$

$$J^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)}\delta\varphi - K^\mu. \tag{3.3.16}$$

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
>
> $$J'^\mu \equiv A \cdot (J^\mu + \partial_\nu K^{\mu\nu}) \tag{3.3.17}$$
>
> 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:

$$\mathcal{L} = \frac{1}{2}\partial_\alpha \varphi^{\mathrm{I}} \partial^\alpha \varphi^{\mathrm{I}} - V\left(\varphi^{\mathrm{I}}\varphi^{\mathrm{I}}\right), \tag{3.3.18}$$

$$\varphi^{\mathrm{I}} \to \widehat{R}^{\mathrm{I}}{}_{\mathrm{J}}\varphi^{\mathrm{J}}, \tag{3.3.19}$$

$$\mathcal{L} \to \mathcal{L}. \tag{3.3.20}$$

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

$$\varphi^{\mathrm{I}} \to \varphi^{\mathrm{I}} - i\theta\left(\widehat{J}^{\mathrm{AB}}\right)_{\mathrm{IJ}}\varphi^{\mathrm{J}} \tag{3.3.21}$$

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

$$\delta\varphi^{\mathrm{I}} = \delta^{\mathrm{I}[\mathrm{A}}\varphi^{\mathrm{B}]}. \tag{3.3.22}$$

The Noether currents – one for each rotation generator – are

$$J^\mu_{\rm K} = \frac{\partial \mathcal{L}}{\partial\partial_\mu\varphi^{\rm I}}\delta^{\rm I[A}\varphi^{\rm B]} \tag{3.3.23}$$

$$= \partial^\mu\varphi^{\rm [A}\cdot\varphi^{\rm B]}. \tag{3.3.24}$$

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

$$\partial^2\varphi^{\rm I} = -2\varphi^{\rm I}V'(\vec{\varphi}^2). \tag{3.3.25}$$

Hence,

$$\partial_\mu J^\mu = \partial^2\varphi^{\rm [A}\cdot\varphi^{\rm B]} + \partial^\mu\varphi^{\rm [I}\cdot\partial_\mu\varphi^{\rm B]} \tag{3.3.26}$$

$$= -2V'\cdot\varphi^{\rm [A}\cdot\varphi^{\rm B]} = 0. \tag{3.3.27}$$

Problem 3.2. 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

$$J^{\mu\alpha\beta} = T^{\mu[\alpha}x^{\beta]}, \tag{3.3.28}$$

where $T^{\mu\nu}$ is the Noether current of spacetime translations in eq. (3.3.9). Interpret the components of $J^{\mu\alpha\beta}$; i.e., what is the Noether current of spatial rotations? And of boosts? □

Problem 3.3. Symmetric Noether Currents Explain why, if $\partial_\mu J^{\mu\alpha\beta} = 0$, where $J^{\mu\alpha\beta}$ is given by eq. (3.3.28), 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. (3.3.28), 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. □

3.4 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

$$p \equiv \left(\frac{\partial L(q,\dot{q})}{\partial \dot{q}}\right)_q. \tag{3.4.1}$$

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

$$H(q,p) \equiv p\cdot\dot{q}(q,p) - L(q,p), \tag{3.4.2}$$

$$L(q,p) \equiv L(q,\dot{q}(q,p)). \tag{3.4.3}$$

Hamilton's equations now reads

$$\dot{q} = \frac{\partial H}{\partial p} \qquad \text{and} \qquad \dot{p} = -\frac{\partial H}{\partial q}. \tag{3.4.4}$$

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

$$\Pi(x) \equiv \frac{\partial \mathcal{L}}{\partial(\partial_0\varphi(x))}. \tag{3.4.5}$$

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:

$$\mathcal{H} \equiv \Pi \cdot \partial_0\varphi - \mathcal{L}. \tag{3.4.6}$$

We may vary this Legendre transform,

$$\delta\mathcal{H} = \delta\Pi \cdot \partial_0\varphi + \Pi \cdot \partial_0\delta\varphi - \frac{\partial\mathcal{L}}{\partial\varphi}\delta\varphi - \frac{\partial\mathcal{L}}{\partial\partial_0\varphi}\partial_0\delta\varphi - \frac{\partial\mathcal{L}}{\partial\partial_i\varphi}\partial_i\delta\varphi \tag{3.4.7}$$

$$= \delta\Pi \cdot \partial_0\varphi + \Pi \cdot \partial_0\delta\varphi - \frac{\partial\mathcal{L}}{\partial\varphi}\delta\varphi - \Pi \cdot \partial_0\delta\varphi - \frac{\partial\mathcal{L}}{\partial\partial_i\varphi}\partial_i\delta\varphi \tag{3.4.8}$$

$$= \delta\Pi \cdot \partial_0\varphi - \frac{\partial\mathcal{L}}{\partial\varphi}\delta\varphi + \partial_i\frac{\partial\mathcal{L}}{\partial\partial_i\varphi}\delta\varphi - \partial_i\left(\frac{\partial\mathcal{L}}{\partial\partial_i\varphi}\delta\varphi\right). \tag{3.4.9}$$

Applying the Euler-Lagrange equations

$$\partial_0\frac{\partial\mathcal{L}}{\partial\partial_0\varphi} + \partial_i\frac{\partial\mathcal{L}}{\partial\partial_i\varphi} = \frac{\partial\mathcal{L}}{\partial\varphi}, \tag{3.4.10}$$

$$\partial_i\frac{\partial\mathcal{L}}{\partial\partial_i\varphi} = \frac{\partial\mathcal{L}}{\partial\varphi} - \partial_0\Pi. \tag{3.4.11}$$

to eq. (3.4.9), we find that

$$\delta\mathcal{H} = \delta\Pi \cdot \partial_0\varphi - \frac{\partial\mathcal{L}}{\partial\varphi}\delta\varphi + \left(\frac{\partial\mathcal{L}}{\partial\varphi} - \dot{\Pi}\right)\delta\varphi - \partial_i\left(\frac{\partial\mathcal{L}}{\partial\partial_i\varphi}\delta\varphi\right) \tag{3.4.12}$$

$$= \delta\Pi \cdot \dot{\varphi} - \dot{\Pi} \cdot \delta\varphi - \partial_i\left(\frac{\partial\mathcal{L}}{\partial\partial_i\varphi}\delta\varphi\right). \tag{3.4.13}$$

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

$$\varphi \to \varphi + \delta\varphi, \tag{3.4.14}$$

$$\Pi \to \Pi + \delta\Pi; \tag{3.4.15}$$

and discover

$$\delta\mathcal{H} = \frac{\partial\mathcal{H}}{\partial\Pi}\cdot\delta\Pi + \frac{\partial\mathcal{H}}{\partial\varphi}\cdot\delta\varphi + \frac{\partial\mathcal{H}}{\partial\partial_i\varphi}\cdot\partial_i\delta\varphi \tag{3.4.16}$$

$$= \frac{\partial\mathcal{H}}{\partial\Pi}\cdot\delta\Pi + \left(\frac{\partial\mathcal{H}}{\partial\varphi} - \partial_i\frac{\partial\mathcal{H}}{\partial\partial_i\varphi}\right)\cdot\delta\varphi + \partial_i\left(\frac{\partial\mathcal{H}}{\partial\partial_i\varphi}\cdot\delta\varphi\right). \tag{3.4.17}$$

Comparing the two variation results,

$$\delta\Pi\cdot\left(\dot{\varphi} - \frac{\partial\mathcal{H}}{\partial\Pi}\right) - \left(\dot{\Pi} + \frac{\partial\mathcal{H}}{\partial\varphi} - \partial_i\frac{\partial\mathcal{H}}{\partial\partial_i\varphi}\right)\delta\varphi = \partial_i\left\{\left(\frac{\partial\mathcal{H}}{\partial\partial_i\varphi} + \frac{\partial\mathcal{L}}{\partial\partial_i\varphi}\right)\delta\varphi\right\}. \tag{3.4.18}$$

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.

$$\int_{\mathbb{R}^D} \mathrm{d}^D\vec{x}\left\{\delta\Pi\cdot\left(\dot{\varphi} - \frac{\partial\mathcal{H}}{\partial\Pi}\right) - \left(\dot{\Pi} + \frac{\partial\mathcal{H}}{\partial\varphi} - \partial_i\frac{\partial\mathcal{H}}{\partial\partial_i\varphi}\right)\delta\varphi\right\} = 0 \tag{3.4.19}$$

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.

$$\dot{\varphi} = \frac{\partial\mathcal{H}}{\partial\Pi} \tag{3.4.20}$$

$$\dot{\Pi} = \partial_i\frac{\partial\mathcal{H}}{\partial\partial_i\varphi} - \frac{\partial\mathcal{H}}{\partial\varphi} \tag{3.4.21}$$

This in turn implies, the right hand side of eq. (3.4.18) must be zero too. And since $\delta\varphi$ was arbitrary,

$$\frac{\partial\mathcal{H}}{\partial(\partial_i\varphi)} = -\frac{\partial\mathcal{L}}{\partial(\partial_i\varphi)}. \tag{3.4.22}$$

Example Let us work out Hamilton's equations for the canonical scalar field in eq. (3.3.1).

$$\Pi = \frac{\partial\mathcal{L}}{\partial\dot{\varphi}} = \dot{\varphi} \tag{3.4.23}$$

The Lagrangian is thus

$$\mathcal{L} = \frac{1}{2}\Pi^2 - \frac{1}{2}(\vec{\nabla}\varphi)^2 - V. \tag{3.4.24}$$

The Legendre transform now reads

$$\mathcal{H} = \Pi^2 - \left(\frac{1}{2}\Pi^2 - \frac{1}{2}(\vec{\nabla}\varphi)^2 - V\right) \tag{3.4.25}$$

$$= \frac{1}{2}\Pi^2 + \frac{1}{2}(\vec{\nabla}\varphi)^2 + V. \tag{3.4.26}$$

Hamilton's equations are

$$\dot{\varphi} = \Pi, \tag{3.4.27}$$

$$\dot{\Pi} = \vec{\nabla}^2 \varphi - V'(\varphi). \tag{3.4.28}$$

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

$$\frac{\partial \mathcal{H}}{\partial(\partial_i \varphi)} = \partial_i \varphi = -\frac{\partial \mathcal{L}}{\partial(\partial_i \varphi)}. \tag{3.4.29}$$

Problem 3.4. Lagrangians from Hamiltonians Starting from the relationship between the Hamiltonian and the Lagrangian in eq. (3.4.6) – but in terms of Π, $\partial_i\varphi$ and φ – show that Hamilton's equations in equations (3.4.20) and (3.4.21) imply the Euler-Lagrange equation $\partial_\mu(\partial\mathcal{L}/\partial(\partial_\mu\varphi)) = \partial\mathcal{L}/\partial\varphi$. □

Problem 3.5. Hamilton's Equations for O_N model Work out Hamilton's equations from the Lagrangian in eq. (3.3.18). □

Problem 3.6. Poisson Brackets and Time Evolution Define the 'equal-time' functional derivatives

$$\frac{\delta\varphi(t,\vec{x})}{\delta\varphi(t,\vec{y})} = \delta^{(D)}(\vec{x}-\vec{y}), \tag{3.4.30}$$

$$\frac{\delta\Pi(t,\vec{x})}{\delta\Pi(t,\vec{y})} = \delta^{(D)}(\vec{x}-\vec{y}), \tag{3.4.31}$$

$$\frac{\delta\varphi(t,\vec{x})}{\delta\Pi(t,\vec{y})} = 0 = \frac{\delta\Pi(t,\vec{x})}{\delta\varphi(t,\vec{y})}; \tag{3.4.32}$$

as well as the 'equal-time' Poisson bracket

$$\begin{aligned}
&\{f(\varphi(t,\vec{x}),\Pi(t,\vec{x})), g(\varphi(t,\vec{x}'),\Pi(t,\vec{x}'))\} \\
&\equiv \int_{\mathbb{R}^D} \mathrm{d}^D\vec{y} \left(\frac{\delta f(t,\vec{x})}{\delta\varphi(t,\vec{y})}\frac{\delta g(t,\vec{x}')}{\delta\Pi(t,\vec{y})} - \frac{\delta g(t,\vec{x})}{\delta\varphi(t,\vec{y})}\frac{\delta f(t,\vec{x})}{\delta\Pi(t,\vec{y})} \right).
\end{aligned} \tag{3.4.33}$$

Show that, for the canonical Hamiltonian in eq. (3.4.26), the Poisson bracket equations

$$\partial_t\varphi(t,\vec{x}) = \{\varphi(t,\vec{x}), H(t)\}, \tag{3.4.34}$$

$$\partial_t\Pi(t,\vec{x}) = \{\Pi(t,\vec{x}), H(t)\} \tag{3.4.35}$$

involving the total Hamiltonian

$$H(t) \equiv \int_{\mathbb{R}^D} \mathrm{d}^D\vec{x}''\mathcal{H}(t,\vec{x}''); \tag{3.4.36}$$

is in fact equivalent to Hamilton's equations. □

3.5 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

$$\varphi(t,\vec{x}) = \int_{\mathbb{R}^D} \frac{\mathrm{d}^D\vec{k}}{(2\pi)^D}\widetilde{\varphi}(t,\vec{k})e^{i\vec{k}\cdot\vec{x}} = \int_{\mathbb{R}^D} \frac{\mathrm{d}^D\vec{k}}{(2\pi)^D}\widetilde{\varphi}(t,\vec{k})e^{-ik_jx^j}, \tag{3.5.1}$$

we see that

$$\partial^2\varphi(x) = \int \frac{\mathrm{d}^3\vec{k}}{(2\pi)^3}\left(\ddot{\widetilde{\varphi}} + \vec{k}^2\widetilde{\varphi}\right)e^{i\vec{k}\cdot\vec{x}} \tag{3.5.2}$$

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

$$\partial_j e^{i\vec{k}\cdot\vec{x}} = \partial_j\left(-ik_lx^l\right)e^{i\vec{k}\cdot\vec{x}} \tag{3.5.3}$$

$$= \left(-ik_l\delta^l_j\right)e^{i\vec{k}\cdot\vec{x}} \tag{3.5.4}$$

$$= \left(-ik_j\right)e^{i\vec{k}\cdot\vec{x}}. \tag{3.5.5}$$

The wave equation sourced by some external J,

$$\ddot{\varphi}(t,\vec{x}) - \vec{\nabla}^2\varphi(t,\vec{x}) + m^2\varphi(t,\vec{x}) = J(t,\vec{x}) \tag{3.5.6}$$

becomes the driven SHO equation

$$\ddot{\widetilde{\varphi}}(t,\vec{k}) + E_{\vec{k}}^2\widetilde{\varphi}(t,\vec{k}) = \widetilde{J}(t,\vec{k}), \tag{3.5.7}$$

where the SHO frequency is the energy

$$E_{\vec{k}} \equiv \sqrt{\vec{k}^2 + m^2}. \tag{3.5.8}$$

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

$$G_{\text{SHO}}(T) = \Theta(T)\frac{\sin(\Omega T)}{\Omega}, \tag{3.5.9}$$

where

$$\left(\frac{\mathrm{d}^2}{\mathrm{d}T^2} + \Omega^2\right)G_{\text{SHO}}(T) = \delta(T). \tag{3.5.10}$$

The classical causality-obeying solution to eq. (3.5.7) is thus

$$\widetilde{\varphi}(t,\vec{k}) = \int_{-\infty}^{t} \frac{\sin(E_{\vec{k}}(t-t'))}{E_{\mathrm{k}}} \cdot \widetilde{J}\left(t',\vec{k}\right) \mathrm{d}t', \tag{3.5.11}$$

$$\varphi(x) = \int_{\mathbb{R}^D} \mathrm{d}^D\vec{x}' \int_{\mathbb{R}} \mathrm{d}t' \left(\Theta(t-t') \int_{\mathbb{R}^D} \frac{\mathrm{d}^D\vec{k}}{(2\pi)^D} e^{i\vec{k}\cdot(\vec{x}-\vec{x}')} \frac{\sin(E_{\vec{k}}(t-t'))}{E_k}\right) J\left(t',\vec{x}'\right) \tag{3.5.12}$$

$$\equiv \int_{\mathbb{R}^D} \mathrm{d}^D\vec{x}' \int_{\mathbb{R}} \mathrm{d}t' G_{\mathrm{ret}}(x-x') J\left(x'\right). \tag{3.5.13}$$

Problem 3.7. Long or Short Range? Massive versus Massless Consider a static point mass resting at $\vec{x}=0$ in the $\{x^\mu\}$ inertial frame, namely

$$J(t,\vec{x}) = J_0 \delta^{(3)}\left(\vec{x}\right), \qquad J_0 \text{ constant}. \tag{3.5.14}$$

Solve eq. (3.5.6). Hint: You may assume the time derivatives in eq. (3.0.9) can be neglected. Then go to Fourier $\vec{k}-$space. You should find

$$\widetilde{\varphi}(\vec{k}) = \frac{J_0}{\vec{k}^2 + m^2}. \tag{3.5.15}$$

You should find

$$\varphi(t, r \equiv |\vec{x}|, \theta, \phi) = J_0 \frac{\exp(-mr)}{4\pi r}. \tag{3.5.16}$$

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? □

Problem 3.8. 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 $\varphi=0$,

$$V(\varphi) = \sum_{\ell=2}^{n} \frac{p_\ell}{\ell} \varphi^\ell. \tag{3.5.17}$$

Show that the Fourier space version of eq. (3.0.9) is:

$$\begin{aligned} \ddot{\widetilde{\varphi}}(t,\vec{k}) + (k^2 + p_2)\widetilde{\varphi}(t,\vec{k}) &= \widetilde{J}(t,\vec{k}) \\ &- \sum_{\ell=2}^{n-1} p_{\ell+1} \prod_{s=1}^{\ell} \left(\int \frac{\mathrm{d}^3\vec{k}_s}{(2\pi)^3} \widetilde{\varphi}(t,\vec{k}_s) \right) (2\pi)^3 \delta^{(3)}\left(\vec{k} - \vec{k}_1 - \dots - \vec{k}_\ell\right). \end{aligned} \tag{3.5.18}$$

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:

$$V'(\varphi) = p_3 \varphi(t, \vec{x})^2 + \dots. \tag{3.5.19}$$

The Fourier decomposition of φ^2 is

$$\widetilde{\varphi^2}(t, \vec{k}) = \int \mathrm{d}^3\vec{x} \varphi(x)^2 e^{-i\vec{k}\cdot\vec{x}} \tag{3.5.20}$$

$$= \int \mathrm{d}^3\vec{x} \int \frac{\mathrm{d}^3\vec{k}_1}{(2\pi)^3} \widetilde{\varphi}(t, \vec{k}_1) e^{i\vec{k}_1\cdot\vec{x}} \int \frac{\mathrm{d}^3\vec{k}_2}{(2\pi)^3} \widetilde{\varphi}(t, \vec{k}_2) e^{i\vec{k}_2\cdot\vec{x}} e^{-i\vec{k}\cdot\vec{x}} \tag{3.5.21}$$

$$= \int \frac{\mathrm{d}^3\vec{k}_1}{(2\pi)^3} \widetilde{\varphi}(t, \vec{k}_1) \int \frac{\mathrm{d}^3\vec{k}_2}{(2\pi)^3} \widetilde{\varphi}(t, \vec{k}_2) \int \mathrm{d}^3\vec{x} e^{i(\vec{k}_1+\vec{k}_2-\vec{k})\cdot\vec{x}} \tag{3.5.22}$$

$$= \int \frac{\mathrm{d}^3\vec{k}_1}{(2\pi)^3} \widetilde{\varphi}(t, \vec{k}_1) \int \frac{\mathrm{d}^3\vec{k}_2}{(2\pi)^3} \widetilde{\varphi}(t, \vec{k}_2) (2\pi)^3 \delta^{(3)}\left(\vec{k} - \vec{k}_1 - \vec{k}_2\right). \tag{3.5.23}$$

$\square$

Problem 3.9. Dispersion relations Consider the *massive* Klein-Gordon equation in Minkowski spacetime:

$$\left(\partial^2 + m^2\right) \varphi(t, \vec{x}) = 0, \tag{3.5.24}$$

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:

$$k^2 = m^2 \qquad \Leftrightarrow \qquad E^2 = \vec{p}^2 + m^2, \tag{3.5.25}$$

$$E \equiv k^0, \qquad \vec{p} \equiv \vec{k}. \tag{3.5.26}$$

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. $\square$

3.6 *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:

$$\frac{\partial \mathcal{L}}{\partial \varphi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial \partial_\mu \varphi} = \frac{\partial \mathcal{L}'}{\partial \varphi} - \partial_\mu \frac{\partial \mathcal{L}'}{\partial \partial_\mu \varphi}. \tag{3.6.1}$$

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.

$$0 = \frac{\partial \Delta}{\partial \varphi} - \partial_\mu \frac{\partial \Delta}{\partial \partial_\mu \varphi} \tag{3.6.2}$$

$$= \frac{\partial \Delta}{\partial \varphi} - \partial_\mu \varphi \frac{\partial^2 \Delta}{\partial \varphi \partial \partial_\mu \varphi} - \partial_\mu \partial_\gamma \varphi \frac{\partial^2 \Delta}{\partial \partial_\gamma \varphi \partial \partial_\mu \varphi} \tag{3.6.3}$$

Since this holds for any φ, the only way the second derivative terms vanish is

$$\frac{\partial^2 \Delta}{\partial \partial_\gamma \varphi \partial \partial_\mu \varphi} = 0, \tag{3.6.4}$$

i.e., Δ can only be linear in $\partial\varphi$.

$$\Delta = \Delta_{\mathrm{I}}(\varphi) + \partial_\alpha \varphi \Delta_{\mathrm{II}}^\alpha(\varphi) \tag{3.6.5}$$

Euler-Lagrange then reduces to

$$\Delta_{\mathrm{I}}'(\varphi) + \partial_\alpha \varphi \Delta_{\mathrm{II}}^{\alpha}{}'(\varphi) - \partial_\alpha \Delta_{\mathrm{II}}^\alpha(\varphi) = 0 \tag{3.6.6}$$

$$\Delta_{\mathrm{I}}'(\varphi) + \partial_\alpha \varphi \Delta_{\mathrm{II}}^{\alpha}{}'(\varphi) - \partial_\alpha \varphi \Delta_{\mathrm{II}}^{\alpha}{}'(\varphi) = 0 \tag{3.6.7}$$

$$\Delta_{\mathrm{I}}'(\varphi) = 0. \tag{3.6.8}$$

Hence, if we define

$$\Delta_0^\alpha(z) \equiv \int^z \Delta_{\mathrm{II}}^\alpha(z')\mathrm{d}z', \tag{3.6.9}$$

most general Δ is a total divergence (plus an irrelevant constant):

$$\Delta = \partial_\alpha \Delta_0^\alpha(\varphi) = \partial_\alpha \varphi \Delta_{\mathrm{II}}^\alpha(\varphi). \tag{3.6.10}$$

> 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.

4 Maxwell's Equations

4.1 Relativistic Formulation

This section is not meant to be an introduction to Maxwell's equations; but rather, to its *relativistic* form. I shall assume the reader is familiar with them in the following first order form, involving electric $\vec{E}$ and magnetic $\vec{B}$ fields, which are in turn sourced by the electric charge density ρ and spatial current density $\vec{J}$. In an inertial frame residing within (3+1)D Minkowski spacetime, with Cartesian coordinates $x^\mu \equiv (t, \vec{x})$,

$$\partial_i E^i = \rho, \tag{4.1.1}$$

$$\epsilon^{ijk}\partial_j B^k - \partial_t E^i = J^i, \tag{4.1.2}$$

$$\epsilon^{ijk}\partial_j E^k + \partial_t B^i = 0, \tag{4.1.3}$$

$$\partial_i B^i = 0. \tag{4.1.4}$$

We have employed index notation, where English alphabets run over the 3 space dimensions. The divergence of the electric and magnetic fields are

$$\vec{\nabla} \cdot \vec{E} = \partial_i E^i \qquad \text{and} \qquad \vec{\nabla} \cdot \vec{B} = \partial_i B^i; \tag{4.1.5}$$

whereas the ith component of their curls are

$$(\vec{\nabla} \times \vec{E})^i = \epsilon^{ijk} \partial_j E^k \qquad \text{and} \qquad (\vec{\nabla} \times \vec{B})^i = \epsilon^{ijk} \partial_j B^k. \tag{4.1.6}$$

A point mass m with spatial trajectory $\vec{z}(t)$ and electric charge q will, in the non-relativistic limit, experience the Lorentz force law

$$m \frac{\mathrm{d}^2 \vec{z}}{\mathrm{d}t^2} = q \left(\vec{E} + \frac{\mathrm{d}\vec{z}}{\mathrm{d}t} \times \vec{B} \right). \tag{4.1.7}$$

Problem 4.1. Current Conservation Taking the time derivative of eq. (4.1.1) and the spatial divergence of eq. (4.1.2), show that Maxwell's equations are consistent only if

$$\partial_t \rho = -\vec{\nabla} \cdot \vec{J}. \tag{4.1.8}$$

This is a statement of *local* electric charge conservation – any net change of total electric charge in a small spatial volume must be due to its flow out of the boundaries of the *same* volume. □

Euclidean Symmetry Let us consider the transformation

$$x^i = \widehat{R}^{ij} x'^j + a^i, \tag{4.1.9}$$

where $\widehat{R}$ and $\vec{a}$ are constant; and the former obeys $\widehat{R}^{\mathrm{T}} \widehat{R} = \mathbb{I}$. By a direction calculation and the orthogonal nature of $\widehat{R}$,

$$\frac{\partial x^i}{\partial x'^j} = \widehat{R}^{ij} = \frac{\partial x'^j}{\partial x^i}. \tag{4.1.10}$$

This means

$$\partial_{x^i} = \widehat{R}^{ij} \partial_{x'^j}. \tag{4.1.11}$$

and therefore, if we postulate that the electric and magnetic fields transform as

$$E^i(\vec{x} = \widehat{R} \cdot \vec{x}' + \vec{a}) = \widehat{R}^{ij} E^{j'}(\vec{x}') \qquad \text{and} \qquad B^i(\vec{x} = \widehat{R} \cdot \vec{x}' + \vec{a}) = \widehat{R}^{ij} B^{j'}(\vec{x}'); \tag{4.1.12}$$

then the divergence is in fact a rotation scalar, as

$$\partial_{x^i} E^i(\vec{x}) = \widehat{R}^{ik} \widehat{R}^{ij} \partial_{x'^j} E^{k'}(\vec{x}') = \delta^j_k \partial_{x'^j} E^{k'}(\vec{x}') = \partial_{x'^i} E^{i'} \tag{4.1.13}$$

and

$$\partial_{x^i} B^i(\vec{x}) = \partial_{x'^i} B^{i'}(\vec{x}'). \tag{4.1.14}$$

Problem 4.2. Euclidean Covariance Show that the curl transforms as

$$\epsilon^{ijk}\partial_{x^j}E^k(\vec{x}) = (\det\widehat{R})\widehat{R}^{il}\epsilon^{ljk}\partial_{x'^j}E^{k'}(\vec{x}'), \tag{4.1.15}$$

$$\epsilon^{ijk}\partial_{x^j}B^k(\vec{x}) = (\det\widehat{R})\widehat{R}^{il}\epsilon^{ljk}\partial_{x'^j}B^{k'}(\vec{x}'). \tag{4.1.16}$$

Proceed to demonstrate that Maxwell's equations transform covariantly under the Euclidean transformations if (I) $\widehat{R}$ is restricted to pure rotations and (II) $\vec{J}$ also transforms covariantly as $J^a(\vec{x}) = \widehat{R}^{ab}J^{b'}(\vec{x}')$.

Similarly, explain why the Lorentz force law transforms covariantly under Euclidean transformations if $\widehat{R}$ is restricted to pure rotations. □

Problem 4.3. Spatial Parity Let us define spatial parity flips as

$$x^i = -x'^i. \tag{4.1.17}$$

Argue that solutions to Maxwell's equations transform as

$$\vec{E}(\vec{x}) = -\vec{E}(-\vec{x}) \qquad \text{and} \qquad \vec{B}(\vec{x}) = \vec{B}(-\vec{x}) \tag{4.1.18}$$

– i.e., $\vec{E}$ is parity odd but $\vec{B}$ is parity even. □

Time Reversal We turn to time reversal

$$t = -t'. \tag{4.1.19}$$

Because $\vec{J}$ is supposed to describe the flow of electric charge, $\vec{J} \sim \rho \mathrm{d}\vec{z}/\mathrm{d}\tau$, we suppose that reversing time would reverse its flow:

$$J^i(t,\vec{x}) = -J^i(t',\vec{x}). \tag{4.1.20}$$

This means the time-derivative terms transform as

$$\epsilon^{ijk}\partial_j B^k(t,\vec{x}) - \partial_{t'}(-\vec{E}(t,\vec{x})) = -\vec{J}(t',\vec{x}), \tag{4.1.21}$$

$$\epsilon^{ijk}\partial_j(-B^k(t,\vec{x})) - \partial_{t'}\vec{E}(t,\vec{x}) = \vec{J}(t',\vec{x}). \tag{4.1.22}$$

Similarly,

$$\epsilon^{ijk}\partial_j E^k(t,\vec{x}) + \partial_{t'}(-\vec{B}(t,\vec{x})) = \vec{0}. \tag{4.1.23}$$

We therefore recover the same form of the equations upon the identification

$$\vec{E}(t,\vec{x}) = \vec{E}(t',\vec{x}) \qquad \text{and} \qquad \vec{B}(t,\vec{x}) = -\vec{B}(t',\vec{x}). \tag{4.1.24}$$

The electric field is time reversal even; the magnetic field is time-reversal odd.

Scalar and Vector Potentials Vector calculus tells us, whenever the divergence of a vector is zero, the vector is the curl of another vector. Recalling eq. (4.1.4), we deduce the magnetic field must be the curl of an object we shall dub the (spatial) vector potential $\vec{A}$; namely,

$$\vec{B} = \vec{\nabla} \times \vec{A}. \tag{4.1.25}$$

In turn, eq. (4.1.3) now becomes

$$\vec{\nabla} \times \left(\vec{E} + \partial_t \vec{A}\right) = 0. \tag{4.1.26}$$

Whenever the curl of a vector is zero, vector calculus says it must be the (negative) gradient of a scalar; in our case, we shall dub it the scalar potential ψ.

$$\vec{E} = -\vec{\nabla}\psi - \partial_t \vec{A}. \tag{4.1.27}$$

We have thus managed to reduce Maxwell's equations to the remaining pair in equations (4.1.1) and (4.1.2) involving the electric charge and spatial current, by first exploiting the other two to re-express the electric and magnetic fields in terms of the potentials $(\psi, \vec{A})$.

Wave Equation for Electromagnetism Taking the curl of the curl equations (4.1.2) and (4.1.3), and utilizing the identity

$$\left(\vec{\nabla} \times \left(\vec{\nabla} \times \vec{A}\right)\right)^i = \nabla^i \left(\vec{\nabla} \cdot \vec{A}\right) - \vec{\nabla}^2 A^i, \tag{4.1.28}$$

we deduce

$$\partial_i \partial_k B^k - \vec{\nabla}^2 B^i - \epsilon^{ijk}\partial_j \dot{E}^k = \epsilon^{ijk}\partial_j J^k, \tag{4.1.29}$$

$$\partial_i \partial_k E^k - \vec{\nabla}^2 E^i + \epsilon^{ijk}\partial_j \dot{B}^k = 0. \tag{4.1.30}$$

Using the divergence equations and time derivatives of the curl ones, we arrive at the conclusion that the electric and magnetic fields obey wave equations sourced by derivatives of the electric charge and current densities:

$$(\partial_t^2 - \vec{\nabla}^2)B^i = \epsilon^{ijk}\partial_j J^k, \tag{4.1.31}$$

$$(\partial_t^2 - \vec{\nabla}^2)E^i = \partial_t J^i - \partial_i \rho. \tag{4.1.32}$$

In vacuum $(\rho, \vec{J}) = (0, \vec{0})$ the Cartesian components of the electromagnetic fields obey homogeneous scalar wave equations $(\partial_t^2 - \vec{\nabla}^2)\vec{E} = \vec{0} = (\partial_t^2 - \vec{\nabla}^2)\vec{B}$. In Fourier $\vec{k}-$space, a 3D vector can be decomposed into its complex $\vec{k}-$ dependent components along the wave vector $\widehat{k} \equiv \vec{k}/|\vec{k}|$ and the two (real) unit vectors $\{\widehat{e}_{\mathrm{I}=1,2}(\vec{k}) | \widehat{e}_{\mathrm{I}} \cdot \widehat{k} = 0\}$. The solutions are thus

$$E^i(t, \vec{x}) = \int_{\mathbb{R}^3} \frac{\mathrm{d}^3\vec{k}}{(2\pi)^3} \left\{ \left(\widetilde{E}_{\parallel}(\vec{k})\widehat{k}^i + \widetilde{E}^{\mathrm{J}} \widehat{e}^i_{\mathrm{J}}(\vec{k}) \right) e^{i|\vec{k}|\left(\widehat{k}\cdot\vec{x} - t\right)} + \mathrm{c.c.} \right\}, \tag{4.1.33}$$

$$B^i(t, \vec{x}) = \int_{\mathbb{R}^3} \frac{\mathrm{d}^3\vec{k}}{(2\pi)^3} \left\{ \left(\widetilde{B}_{\parallel}(\vec{k})\widehat{k}^i + \widetilde{B}^{\mathrm{J}} \widehat{e}^i_{\mathrm{J}}(\vec{k}) \right) e^{i|\vec{k}|\left(\widehat{k}\cdot\vec{x} - t\right)} + \mathrm{c.c.} \right\}. \tag{4.1.34}$$

Problem 4.4. Orthogonality in vacuum Insert the above solutions into the vacuum Maxwell equations

$$\partial_i E^i = 0, \tag{4.1.35}$$

$$\epsilon^{ijk}\partial_j B^k - \partial_t E^i = 0, \tag{4.1.36}$$

$$\epsilon^{ijk}\partial_j E^k + \partial_t B^i = 0, \tag{4.1.37}$$

$$\partial_i B^i = 0; \tag{4.1.38}$$

to deduce the electric and magnetic fields in Fourier space must be orthogonal to $\vec{k}$

$$\widetilde{E}_\parallel = 0 = \widetilde{B}_\parallel \tag{4.1.39}$$

and

$$\vec{E}(t,\vec{x}) = \int_{\mathbb{R}^3} \frac{\mathrm{d}^3\vec{k}}{(2\pi)^3} \left\{ \widetilde{E}^{\mathrm{J}} \widehat{e}_{\mathrm{J}}(\vec{k}) e^{i|\vec{k}|\left(\widehat{k}\cdot\vec{x}-t\right)} + \text{c.c.} \right\}, \tag{4.1.40}$$

$$\vec{B}(t,\vec{x}) = \int_{\mathbb{R}^3} \frac{\mathrm{d}^3\vec{k}}{(2\pi)^3} \left\{ \widetilde{E}^{\mathrm{J}} \left(\widehat{e}_{\mathrm{J}}(\vec{k}) \times \widehat{k} \right) e^{i|\vec{k}|\left(\widehat{k}\cdot\vec{x}-t\right)} + \text{c.c.} \right\}. \tag{4.1.41}$$

Vacuum electric and magnetic fields are not only mutually perpendicular, they are transverse to the wave vector and really contains two degrees of freedom $\{\widetilde{E}^{\mathrm{J}}\}$ only . $\square$

Poincaré Covariance Why do the electric and magnetic fields travel at unit speed (i.e., at c)? In which inertial reference frame do they do so? These questions puzzled physicists in the late 1800's. We shall address this issue by way of Einstein: Maxwell's equations are Poincaré covariant – they take the same form in all inertial frames defined by Minkowski spacetime $\mathrm{d}s^2 = \mathrm{d}t^2 - \mathrm{d}\vec{x}^2 = \eta_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu$ – and this is the reason why the wavefront[2] of electromagnetic radiation travel at c in all inertial reference frames.

We first postulate that, in such a Cartesian system $\{x^\mu\}$, the electric charge and spatial current must transform as a Lorentz 4−vector,

$$J^\mu(t,\vec{x}) \equiv J^\mu(x) \equiv \left(\rho(x), \vec{J}(x)\right)^{\mathrm{T}}, \tag{4.1.42}$$

$$J^\alpha\left(x^\beta = \Lambda^\beta{}_\delta x'^\delta + a^\beta\right)\Lambda_\alpha{}^\mu = J^{\mu'}(x'). \tag{4.1.43}$$

Then, Maxwell's equations may be re-packaged into the following Lorentz-vector equations

$$\partial_\mu F^{\mu\nu} = J^\nu, \tag{4.1.44}$$

$$\epsilon^{\mu\nu\alpha\beta}\partial_\nu F_{\alpha\beta} = 0; \tag{4.1.45}$$

where $F_{\mu\nu}$ is an anti-symmetric rank−2 Lorentz tensor; i.e.,

$$F_{\mu\nu} = -F_{\nu\mu}. \tag{4.1.46}$$

Because the Levi-Civita $\epsilon^{\mu\nu\alpha\beta}$ is fully anti-symmetric, note that there is an implicit antisymmetrization of the indices on the partial derivatives acting on $F_{\mu\nu}$ in eq. (4.1.45). Furthermore, we may multiply both sides of eq. (4.1.45) with, say, $\epsilon_{\mu\delta\sigma\lambda}$ to obtain its alternate – but completely equivalent – rank−3 form

$$\partial_{[\delta}F_{\sigma\rho]} = 0, \tag{4.1.47}$$

where $[\dots]$ denotes the full-anti-symmetrization operation.

[2]In (3+1)D Minkowski electromagnetic waves travel strictly on the null cone; however, in (2+1)D and all higher odd dimensional Minkowski spacetimes, electromagnetic waves propagate *inside* the null cone. Hence, it is more accurate to assert it is the wavefront that travels at unit speed in all inertial frames.

Problem 4.5. Bianchi Explain why eq. (4.1.47) is equivalent to

$$\partial_\alpha F_{\mu\nu} = \partial_\mu F_{\alpha\nu} + \partial_\nu F_{\mu\alpha}. \tag{4.1.48}$$

Hint: Expand out the anti-symmetrization and exploit the anti-symmetric character of $F_{\mu\nu}$. □

Lorentz Force Law In order to match Maxwell's equations in (4.1.1)-(4.1.4), we first attempt to recover the Lorentz force law in eq. (4.1.7). As we shall soon see, this will allow us to identify the electric and magnetic components of $F_{\mu\nu}$. The mass times acceleration side of the relativistic version of Newton's second law must be $m\mathrm{d}^2 z^\mu/\mathrm{d}\tau^2$, where z^μ is the particle's trajectory, τ is its proper time, and q electric charge. On the force side, we need to construct a 4−vector. The only tensors in the problem at hand are $F_{\mu\nu}$, the particle's velocity $U^\mu \equiv \mathrm{d}z^\mu/\mathrm{d}\tau$, and acceleration. (We shall assume the force law, like its Newtonian counterpart, is at most second order in time.) Since the right hand side of eq. (4.1.7) does not contain accelerations, we therefore postulate

$$m\frac{\mathrm{d}^2 z^\mu}{\mathrm{d}\tau^2} = qF^\mu{}_\nu \frac{\mathrm{d}z^\nu}{\mathrm{d}\tau}. \tag{4.1.49}$$

Utilizing $\mathrm{d}\tau = \mathrm{d}t\sqrt{1-(\mathrm{d}\vec{z}/\mathrm{d}t)^2}$, and expanding the indices into its time and space components,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{m}{\sqrt{1-(\mathrm{d}\vec{z}/\mathrm{d}t)^2}}\right) = qF^0{}_j\frac{\mathrm{d}z^j}{\mathrm{d}t} = qF^{j0}\frac{\mathrm{d}z^j}{\mathrm{d}t}, \tag{4.1.50}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{m\mathrm{d}z^i/\mathrm{d}t}{\sqrt{1-(\mathrm{d}\vec{z}/\mathrm{d}t)^2}}\right) = q\left(F^i{}_0 + F^i{}_j\frac{\mathrm{d}z^j}{\mathrm{d}t}\right) = q\left(F^{i0} + \frac{\mathrm{d}z^j}{\mathrm{d}t}F^{ji}\right); \tag{4.1.51}$$

where in the first line we have recognized $F^0{}_0 = F^{00} = F_{00} = 0$ because of its anti-symmetry. In the non-relativistic limit $(\mathrm{d}\vec{z}/\mathrm{d}t)^2 \ll 1$, the left hand side of eq. (4.1.51) does reduce to $m\mathrm{d}^2\vec{z}/\mathrm{d}t^2$. The right hand side would match the right hand side of eq. (4.1.7) if we identify the velocity-independent term directly with the electric field,

$$E^i = F^{i0} = -F^{0i} = F_{0i} = -F_{i0}; \tag{4.1.52}$$

and the velocity-dependent term with $(\mathrm{d}\vec{z}/\mathrm{d}t)\times\vec{B}$,

$$\frac{\mathrm{d}z^j}{\mathrm{d}t}F^{ji} = \epsilon^{ijk}\frac{\mathrm{d}z^j}{\mathrm{d}t}B^k, \tag{4.1.53}$$

with $\epsilon^{123} \equiv +1$. How can a rank−2 F^{ji} be related to a rank−1 B^k? The answer lies in taking the spatial Hodge dual of F^{ij}. Let us declare it to be

$$F^{ij} \equiv \epsilon^{ijk}\widetilde{F}^k. \tag{4.1.54}$$

Then

$$-\epsilon^{ijk}\frac{\mathrm{d}z^j}{\mathrm{d}t}\widetilde{F}^k = \epsilon^{ijk}\frac{\mathrm{d}z^j}{\mathrm{d}t}B^k, \tag{4.1.55}$$

and we may identify

$$-\epsilon^{ijk}B^k = F_{ij} = F^{ij} = -F_{ji} = -F^{ji} \qquad \Leftrightarrow \qquad B^i = -\frac{1}{2}\epsilon^{ijk}F^{jk}. \tag{4.1.56}$$

This identification of $\vec{B}$ (a spatial vector) with F_{ij} only works in 3D space. Upon taking the spatial Hodge dual of F_{ij}, one may witness that in 2D space the magnetic field F_{ij} becomes a spatial scalar while in higher dimensions D it returns a $D-2$ form.

Before examining Maxwell's equations, let us recall that

$$P^\mu \equiv m \cdot U^\mu = m\frac{\mathrm{d}z^\mu/\mathrm{d}t}{\sqrt{1-(\mathrm{d}\vec{z}/\mathrm{d}t)^2}} \tag{4.1.57}$$

is the energy-momentum vector of the point mass m. With the identification in eq. (4.1.52), we gather from the zeroth component of the Lorentz force law in eq. (4.1.50): the power imparted to the point particle is driven by the electric field via

$$\frac{\mathrm{d}P^0}{\mathrm{d}t} = \vec{E} \cdot \frac{\mathrm{d}\vec{z}}{\mathrm{d}t}. \tag{4.1.58}$$

Problem 4.6. Show that eq. (4.1.51) implies eq. (4.1.50). In other words, the zeroth component of the Lorentz force law does not constitute an independent equation. □

We turn to expanding eq. (4.1.44) into its time and space components. Recognizing $F^{00} = 0$, the zeroth component is

$$\partial_i F^{i0} = \vec{\nabla} \cdot \vec{E} = J^0 = \rho; \tag{4.1.59}$$

whereas the ith components are

$$\partial_0 F^{0i} + \partial_j F^{ji} = J^i, \tag{4.1.60}$$

$$\epsilon^{ijk}\partial_j B^k - \partial_t E^i = J^i. \tag{4.1.61}$$

These are equations (4.1.1) and (4.1.2).

Problem 4.7. Curl E and Div B Use the identifications in equations (4.1.52) and (4.1.56) to recover equations (4.1.3) and (4.1.4) from eq. (4.1.45). □

Problem 4.8. Four-Current Conservation By taking the 4−divergence of eq. (4.1.44), and exploiting $F^{\mu\nu} = -F^{\nu\mu}$, show that Maxwell's equations are consistent only if

$$\partial_\mu J^\mu = 0. \tag{4.1.62}$$

Be sure to compare equations (4.1.8) and (4.1.62). □

Problem 4.9. 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 $J^{\widehat{0}} \equiv u_\mu J^\mu$ 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

$$Q \equiv \int_{\mathbb{R}^D} \mathrm{d}^D\Sigma_\mu J^\mu, \qquad \mathrm{d}^D\Sigma_\mu \equiv \mathrm{d}^D\vec{x}u_\mu, \quad D \equiv d-1, \tag{4.1.63}$$

is a constant. □

Problem 4.10. Four Vector Potential The Poincaré lemma states, a fully antisymmetric $N-$form $T_{\mu_1 \dots \mu_N}$ is the exterior derivative (aka generalized curl) of a fully anti-symmetric $(N-1)-$form $A_{\mu_1 \dots \mu_{N-1}}$ in a simply connected region of spacetime if and only if the exterior derivative of T is zero.

$$\partial_{[\mu_0} T_{\mu_1 \dots \mu_N]} = 0 \qquad \text{iff} \qquad T_{\mu_1 \dots \mu_N} = \partial_{[\mu_1} A_{\mu_2 \dots \mu_N]} \tag{4.1.64}$$

Explain why eq. (4.1.47) – which, recall, is equivalent to equations (4.1.3) and (4.1.4) – therefore implies

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu, \tag{4.1.65}$$

for some A_μ which we shall name the four vector potential. How is A^μ related to $(\psi, \vec{A})$? □

Gauge Invariance At the classical level, notice Maxwell's equations and the Lorentz force laws only involve $F_{\mu\nu}$ and not A_μ, even though the former is built out of the latter. We may ask: what is the most general A_μ such that $F_{\mu\nu}[A] = \partial_{[\mu} A_{\nu]} = 0$? The answer, according to the Poincaré lemma, is a pure gradient $A_\mu = \partial_\mu \Lambda$. Hence, at least for linear shifts, the most general transformation of the form $A_\mu \to A_\mu + \delta A_\mu$ that leaves the electromagnetic fields invariant, $F_{\mu\nu}[A + \delta A] = F_{\mu\nu}[A]$, is given by

$$A_\mu \to A_\mu + \partial_\mu \Lambda, \tag{4.1.66}$$

for arbitrary spacetime scalar Λ. This is known as a gauge transformation; and $F_{\mu\nu}$ is gauge invariant.

Vacuum Electromagnetic Waves We may now re-derive the wave equations of the electromagnetic fields $F_{\mu\nu}$ but in a fully Lorentz covariant manner. Starting from the vacuum form of the equations

$$\partial_\mu F^{\mu\nu} = 0 = \partial_{[\alpha} F_{\mu\nu]}, \tag{4.1.67}$$

we may proceed by taking the divergence of the right equality.

$$0 = \partial^\alpha \left(\partial_\alpha F_{\mu\nu} - \partial_\mu F_{\alpha\nu} - \partial_\nu F_{\mu\alpha} \right) \tag{4.1.68}$$

$$= \partial^\alpha \partial_\alpha F_{\mu\nu} - \partial_\mu \partial^\alpha F_{\alpha\nu} + \partial_\nu \partial^\alpha F_{\alpha\mu}. \tag{4.1.69}$$

But the second and third terms are zero by the first set of Maxwell's equations. Hence, in such a Cartesian system, the electromagnetic fields obey the following Lorentz-covariant wave equation:

$$\partial^2 F_{\mu\nu} = 0 = \partial^2 F^{\mu\nu}. \tag{4.1.70}$$

These are equivalent to $(\partial_t^2 - \vec{\nabla}^2)\vec{E} = \vec{0} = (\partial_t^2 - \vec{\nabla}^2)\vec{B}$, through the identifications in equations (4.1.52) and (4.1.56).

Problem 4.11. Second Order Form When $J^\mu \neq 0$, derive the more general second order form of Maxwell's equations:

$$\partial^2 F_{\mu\nu} = \partial_{[\mu} J_{\nu]}. \tag{4.1.71}$$

(Hint: This follows from the analysis above.) Compare this manifestly Lorentz covariant result with equations (4.1.31) and (4.1.32). □

Lorentz Covariant Flux The electromagnetic field strength tensor (aka Faraday tensor) $F_{\mu\nu}$ ought to be regarded as the components of a $2-$form:

$$F = \frac{1}{2} F_{\mu\nu} \mathrm{d}x^\mu \wedge \mathrm{d}x^\nu, \tag{4.1.72}$$

where the basis $2-$forms $\{\mathrm{d}x^\mu \wedge \mathrm{d}x^\nu = -\mathrm{d}x^\nu \wedge \mathrm{d}x^\mu\}$ are anti-symmetric. For a fixed $\mu\nu$, they describe the infinitesimal area spanned by the infinitesimal displacements in the μth and νth directions. For instance, $\mathrm{d}t \wedge \mathrm{d}x^1$ is the infinitesimal area on the time-x^1 plane; $\mathrm{d}x^2 \wedge \mathrm{d}x^3$ is that on the spatial $(2,3)-$plane; etc. Hence, we should view F as a *flux* – the 'flow' of some field through the corresponding area – and the electromagnetic components $F_{\mu\nu}$ as flux density. This is the relativistic generalization of the flux $\vec{E} \cdot \mathrm{d}(\vec{\mathrm{Area}})$ in, say, Gauss' law.

Since $\mathrm{d}x^\mu = \Lambda^\mu{}_\nu \mathrm{d}x'^\nu$ when we switch from inertial frame $\{x^\mu\}$ to frame $\{x'^\mu\}$, we see that

$$F_{\mu\nu}(x) \mathrm{d}x^\mu \wedge \mathrm{d}x^\nu = F_{\mu\nu}(x = \Lambda \cdot x' + a) \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta \mathrm{d}x'^\alpha \wedge \mathrm{d}x'^\beta \tag{4.1.73}$$

and therefore

$$F_{\alpha'\beta'}(x') = F_{\mu\nu}(x = \Lambda \cdot x' + a) \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta. \tag{4.1.74}$$

Electromagnetic fields transform covariantly as Lorentz rank-2 tensors. The Euclidean transformations for the $\vec{E}$ and $\vec{B}$ fields above are merely a subset of this larger group of symmetry transformations. Moreover, the Lorentz index structure allows the discrete time-reversal and parity flips to be easily worked out. The electric F_{0i} is both odd ($F_{0i} \to -F_{0i}$) under time reversal $t \to -t$ and parity flips $\vec{x} \to -\vec{x}$; it is a vector under pure rotations, $F_{0i} \to \widehat{R}_{ij} F_{0j}$. The magnetic F_{ij}, having no time indices, is even under both time reversal and parity flips ($F_{ij} \to F_{ij}$); its spatial Hodge dual is a vector under pure rotations, $\epsilon^{ijk} F_{jk} \to \widehat{R}^{il} \epsilon^{ljk} F_{jk}$.

Problem 4.12. Poincaré transformation of Maxwell Assuming the (external) electric $4-$current J^μ transforms covariantly under Poincaré transformations – namely, it obeys eq. (4.1.43) – prove that Maxwell's equations in (4.1.44) and (4.1.45) implies $F_{\mu\nu}$ necessarily transforms as a rank-2 Poincaré tensor; i.e., it obeys eq. (4.1.74). $\square$

Problem 4.13. Lorentz boosts Under a Lorentz boost $x^\mu = \Lambda^\mu{}_\nu x'^\nu$ given by

$$\Lambda = \begin{bmatrix} \gamma(\vec{v}) & -\gamma(\vec{v}) \cdot \vec{v}^{\mathrm{T}} \\ -\gamma(\vec{v}) \cdot \vec{v} & \gamma(\vec{v}) \widehat{v}\widehat{v}^{\mathrm{T}} + \mathbb{I}_{D\times D} - \widehat{v}\widehat{v}^{\mathrm{T}} \end{bmatrix}, \tag{4.1.75}$$

which boosts $U^\mu = (1, \vec{v})/\sqrt{1 - \vec{v}^2}$ to its rest frame $(1, \vec{0}) = \Lambda \cdot U$, show that the electric and magnetic fields $\{E^{i'}(x'), B^{i'}(x')\}$ in the $x'-$frame are related to the electric and magnetic fields $\{E^i(x), B^i(x)\}$ in the $x-$frame by

$$E^{i'} = E^i_\parallel + \gamma \left(E^i_\perp - (\vec{v} \times \vec{B})^i \right), \tag{4.1.76}$$

$$E^i_\parallel \equiv \widehat{v}^i \left(\widehat{v} \cdot \vec{E} \right) \qquad E^i_\perp \equiv E^i - \widehat{v}^i \left(\widehat{v} \cdot \vec{E} \right); \tag{4.1.77}$$

and

$$B^{k'} = B^k_{\parallel} + \gamma \left(B^k_{\perp} + \left(\vec{v} \times \vec{E} \right)^k \right), \tag{4.1.78}$$

$$B^i_{\parallel} = \left(\widehat{v} \cdot \vec{B} \right) \widehat{v}^i \qquad\qquad B^i_{\perp} = B^i - \left(\widehat{v} \cdot \vec{B} \right) \widehat{v}^i. \tag{4.1.79}$$

The subscripts $\parallel$ and $\perp$ denote, respectively, the components of the associated spatial vector field parallel and perpendicular to the boost direction $\widehat{v}$. $\square$

Hodge Dual In (3+1)D electromagnetism, we may define another rank-2 tensor by taking the Hodge dual of $F_{\mu\nu}$:

$$\widetilde{F}^{\mu\nu} \equiv \frac{1}{2} \epsilon^{\mu\nu\alpha\beta} F_{\alpha\beta}. \tag{4.1.80}$$

Its $0i$ components are in fact magnetic,

$$\widetilde{F}^{0i} = -\frac{1}{2} \epsilon_{0ijk} F_{jk} = B^i; \tag{4.1.81}$$

while it is the ij components that are electric,

$$\widetilde{F}^{ij} = -\epsilon_{ij0k} F_{0k} = -\epsilon^{ijk} E^k. \tag{4.1.82}$$

Maxwell's equations in (4.1.44) and (4.1.45) may now be written solely in terms of divergences:

$$\partial_\mu F^{\mu\nu} = J^\nu \qquad \text{and} \qquad \partial_\mu \widetilde{F}^{\mu\nu} = 0. \tag{4.1.83}$$

Problem 4.14. Time Reversals, Parity Flips Explain why $\widetilde{F}^{\mu\nu}$ *does not* transform as a tensor under time-reversal $t \to -t$ or parity flips $\vec{x} \to -\vec{x}$. This is the primary reason for defining it; because it allows us to consider electromagnetic forces that 'violate' time-reversal and/or parity symmetry. $\square$

On the other hand, if the Lorentz force law in eq. (4.1.49) were to transform as a vector under both time reversal and parity flips, this is why we did not include an additional $\widetilde{F}^{\mu}{}_{\nu} \mathrm{d}z^\nu / \mathrm{d}\tau$ term on its right hand side. Moreover, because the right hand side of the second equation in eq. (4.1.83) is zero, Maxwell's equations are in fact time-reversal and parity invariant – as long as the J^μ transforms accordingly – recall, too, its alternate form in eq. (4.1.47). For instance, if instead

$$\partial_\mu F^{\mu\nu} = J^\nu_{\mathrm{E}} \qquad \text{and} \qquad \partial_\mu \widetilde{F}^{\mu\nu} = J^\nu_{\mathrm{B}} \tag{4.1.84}$$

– where J^ν_{E} is the usual electric current and J^ν_{B} is a hypothetical 'magnetic current' – then these equations would transform as spacetime vectors under continuous Lorentz transformations but not under the discrete ones of time-reversal and parity inversions.

Problem 4.15. 'Local' electric and magnetic fields If $U^\mu(x)$ is a unit length timelike vector, and therefore may be associated with the 4$-$velocity of an observer at x, explain why

$$E^\mu \equiv F^{\mu\nu} U_\nu \qquad \text{and} \qquad B^\mu \equiv \widetilde{F}^{\mu\nu} U_\nu \tag{4.1.85}$$

may be viewed respectively as the electric and magnetic fields in the observer's instantaneous rest frame. $\square$

Problem 4.16. Lorentz Scalars Show that $\vec{E}^2 - \vec{B}^2$ is a Lorentz invariant (i.e., a spacetime scalar) by demonstrating that

$$F_{\mu\nu} F^{\mu\nu} = -2\left(\vec{E}^2 - \vec{B}^2\right). \tag{4.1.86}$$

Show that

$$\widetilde{F}^{\mu\nu} F_{\mu\nu} = 2\partial_\mu\left(\epsilon^{\mu\nu\alpha\beta} A_\nu \partial_\alpha A_\beta\right) = 4\vec{E}\cdot\vec{B}. \tag{4.1.87}$$

Therefore $\vec{E}\cdot\vec{B}$ is also a Lorentz invariant. Note that $\widetilde{F}^{\mu\nu}\widetilde{F}_{\mu\nu}$ does not yield any new invariants – can you see why? □

Variational Principles For a prescribed electric current J^μ, and the recognition that $F_{\mu\nu} = \partial_{[\mu} A_{\nu]}$, we may write down the following action for Maxwell's equations (4.1.44).

$$S[A] \equiv \int_{t_1}^{t_2} \mathrm{d}t \int_{\mathbb{R}^3} \mathrm{d}^3\vec{x} \left(-\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - A_\mu J^\mu\right). \tag{4.1.88}$$

We shall assume the spatial portion of the vector potential $A_i(t = t_{1,2}, \vec{x})$ are specified at the initial and final time surfaces; all the components A_μ fall off to zero sufficiently quickly at spatial infinity; and proceed to extremize this action – the result shall be Maxwell's equations (4.1.44). Before we do so, however, let us recognize the above Lagrangian density to be Lorentz invariant. The total action is also invariant under the gauge transformation $A_\mu \to A_\mu + \partial_\mu \Lambda$, provided the current is conserved and Λ falls off quickly enough at spatial infinity and is zero at the constant $t = t_0$ and $t = t_1$ surfaces. For, we may examine

$$S[A + \partial\Lambda] = S[A] + \delta S \tag{4.1.89}$$

$$\delta S \equiv -\int_{t_1}^{t_2} \mathrm{d}t \int_{\mathbb{R}^3} \mathrm{d}^3\vec{x} \left(\partial_\mu \Lambda \cdot J^\mu\right). \tag{4.1.90}$$

If $\partial^\mu J_\mu = 0$, the integrand may be expressed as the total divergence $\partial_\mu \Lambda \cdot J^\mu = \partial_\mu(\Lambda \cdot J^\mu)$. Gauss' theorem then turns this shift in the action δS into a surface integral at spatial infinity and at the $t = t_0$ and $t = t_1$ surfaces. By assuming $\Lambda(t_1 \leq t \leq t_2, |\vec{x}| \to \infty) \to 0$ sufficiently quickly and $\Lambda(t = t_{1,2}, \vec{x} \in \mathbb{R}^3) = 0$, we conclude $\delta S = 0$. On the other hand, if we are given that $\delta S = 0$, then we may integrate by parts to obtain

$$\delta S = \int_{t_1}^{t_2} \mathrm{d}t \int_{\mathbb{R}^3} \mathrm{d}^3\vec{x} \left(\Lambda \cdot \partial_\mu J^\mu\right) = 0. \tag{4.1.91}$$

Since Λ was otherwise arbitrary, we conclude $\partial^\mu J_\mu = 0$ throughout this spacetime region.

Let us now examine the first order perturbation of the action induced by the variation $A_\mu \to A_\mu + \delta A_\mu$:

$$\delta S[A] = \int_{t_1}^{t_2} \mathrm{d}t \int_{\mathbb{R}^3} \mathrm{d}^3\vec{x} \left(-(\partial_0 \delta A_i - \partial_i \delta A_0) F^{0i} - \frac{1}{2}\partial_{[i}\delta A_{j]} F^{ij} - \delta A_\mu J^\mu\right) \tag{4.1.92}$$

$$\begin{aligned} &= \int_{\mathbb{R}^3} \mathrm{d}^3\vec{x} \left[\delta A_i F^{i0}\right]_{t=t_1}^{t=t_2} \\ &\quad + \int_{t_1}^{t_2} \mathrm{d}t \int_{\mathbb{R}^3} \mathrm{d}^3\vec{x} \left(\delta A_i \partial_0 F^{0i} + \delta A_0 \partial_i F^{i0} + \delta A_j \partial_i F^{ij} - \delta A_\mu J^\mu\right). \end{aligned} \tag{4.1.93}$$

Since A_i is specified at the $t = t_{1,2}$ surfaces, its variation must be zero there, thereby eliminating the boundary terms. Moreover, the sum over spatial indices in the $\delta A_i \partial_0 F^{0i}$ can be extended to include the time index, because $F^{00} = 0$. This allows all the indices to be restored to the spacetime covariant form

$$\delta S[A] = \int_{t_1}^{t_2} \mathrm{d}t \int_{\mathbb{R}^3} \mathrm{d}^3\vec{x} \delta A_\nu \left(\partial_\mu F^{\mu\nu} - J^\nu\right). \tag{4.1.94}$$

Problem 4.17. Lorentz Force Law Consider the following Lorentz-invariant action involving the electromagnetic field strength tensor $F_{\mu\nu}$ and the trajectories $\{z^\mu_{\mathrm{I}} | \mathrm{I} = 1, 2, \dots, N\}$ of the corresponding point charges $\{q_{\mathrm{I}}\}$ with associated masses $\{m_{\mathrm{I}}\}$:

$$S \equiv \int_{t_1}^{t_2} \mathrm{d}t \int_{\mathbb{R}^3} \mathrm{d}^3\vec{x} \left(-\frac{1}{4} F_{\mu\nu} F^{\mu\nu}\right) - \sum_{\mathrm{I}} \left(q_{\mathrm{I}} \int_{z^0_{\mathrm{I}}=t_1}^{z^0_{\mathrm{I}}=t_2} A_\mu \mathrm{d}z^\mu_{\mathrm{I}} + m_{\mathrm{I}} \int_{\tau_{\mathrm{I}}(t_1)}^{\tau_{\mathrm{I}}(t_2)} \mathrm{d}\tau_{\mathrm{I}}\right), \tag{4.1.95}$$

where the proper time τ_{I} of the I-th particle is related to the inertial frame time t via $\mathrm{d}\tau_{\mathrm{I}} = \sqrt{1 - (\mathrm{d}\vec{z}_{\mathrm{I}}/\mathrm{d}t)^2}\mathrm{d}t$. The worldline integral $\int A_\mu \mathrm{d}z^\mu_{\mathrm{I}}$ may be parametrized using proper time or inertial frame time t,

$$\int A_\mu \mathrm{d}z^\mu_{\mathrm{I}} = \int A_\mu(z) \frac{\mathrm{d}z^\mu_{\mathrm{I}}}{\mathrm{d}\tau_{\mathrm{I}}} \mathrm{d}\tau_{\mathrm{I}} = \int A_\mu(z) \frac{\mathrm{d}z^\mu_{\mathrm{I}}}{\mathrm{d}t} \mathrm{d}t. \tag{4.1.96}$$

Show that extremizing S yields the Lorentz force law in (4.1.51) if one uses inertial time t. Extremize the action with respect to A_μ and read off the (total) electric current to be:

$$J^\mu(x) = \sum_{\mathrm{I}} q_{\mathrm{I}} \int \mathrm{d}z_{\mathrm{I}} \delta^{(4)}(x - z(\tau)) = \sum_{\mathrm{I}} q_{\mathrm{I}} \int \mathrm{d}\tau_{\mathrm{I}} \frac{\mathrm{d}z^\mu_{\mathrm{I}}}{\mathrm{d}\tau} \delta^{(4)}(x - z(\tau)). \tag{4.1.97}$$

Explain why this object is a Lorentz vector. Hint: How does the $\delta-$function transform under $x = \Lambda \cdot x' + a$ and $z = \Lambda \cdot z' + a$? $\square$

Problem 4.18. Theta term Consider the alternate action

$$S[A] \equiv \int_{t_1}^{t_2} \mathrm{d}t \int_{\mathbb{R}^3} \mathrm{d}^3\vec{x} \left(-\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \frac{\theta}{2} \widetilde{F}^{\mu\nu} F_{\mu\nu} - A_\mu J^\mu\right), \tag{4.1.98}$$

where θ is a constant. Explain why extremizing this action with respect to A_μ yields the same Maxwell's equations in eq. (4.1.44). Hint: Exploit the second equality in eq. (4.1.87). $\square$

4.2 Gauge Potentials

Vector Potential & Gauge Symmetry The other Maxwell equation (cf eq. (**??**)) 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

$$F = \mathrm{d}A \qquad \Leftrightarrow \qquad F_{\mu\nu} = \partial_{[\mu} A_{\nu]} = \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{4.2.1}$$

Notice the dynamics in eq. (**??**) 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

$$A_\mu(x) \to A_\mu(x) + \partial_\mu L(x) \tag{4.2.2}$$

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

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

$$\partial_\mu F^{\mu\nu} = \partial_\mu \left(\partial^\mu A^\nu - \partial^\nu A^\mu\right) = J^\nu. \tag{4.2.3}$$

The symmetry under the gauge transformation of eq. (4.2.2) means that solutions to eq. (4.2.3) 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

$$A_\mu(x) \equiv \int_{\mathbb{R}^d} \frac{\mathrm{d}^d k}{(2\pi)^d} \widetilde{A}_\mu(k) e^{-ik_\mu x^\mu} \qquad \text{and} \qquad J_\mu(x) \equiv \int_{\mathbb{R}^d} \frac{\mathrm{d}^d k}{(2\pi)^d} \widetilde{J}_\mu(k) e^{-ik_\mu x^\mu}, \tag{4.2.4}$$

we may see that the differential operator in eq. (4.2.3) cannot be inverted because it has a zero eigenvalue. Firstly, the Fourier version of eq. (4.2.3) reads

$$-K^{\mu\nu} \widetilde{A}_\mu = \widetilde{J}^\nu, \tag{4.2.5}$$

$$K^{\mu\nu} \equiv k_\sigma k^\sigma \eta^{\mu\nu} - k^\nu k^\mu. \tag{4.2.6}$$

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:

$$K^{\mu\nu} k_\mu = (k_\sigma k^\sigma) k^\nu - k^\nu k^\mu k_\mu = 0. \tag{4.2.7}$$

Problem 4.19. Can you explain why eq. (4.2.7) amounts to the statement that $F_{\mu\nu}$ is invariant under the gauge transformation of eq. (4.2.2)? Hint: Consider eq. (4.2.2) 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:

$$k^\mu \widetilde{A}_\mu = 0. \tag{4.2.8}$$

In 'position'/real spacetime, this reads instead

$$\partial^\mu A_\mu = 0 \qquad\qquad \text{(Lorenz gauge)}. \tag{4.2.9}$$

With the constraint in eq. (4.2.8), Maxwell's equations in eq. (4.2.5) becomes

$$-\left(k_\sigma k^\sigma \widetilde{A}^\nu - k^\nu (k^\mu \widetilde{A}_\mu)\right) = -k_\sigma k^\sigma \widetilde{A}^\nu = \widetilde{J}^\nu. \tag{4.2.10}$$

Now, Maxwell's equations have become invertible:

$$\widetilde{A}_\mu(k) = \frac{\widetilde{J}_\mu(k)}{-k^2}, \qquad k^2 \equiv k_\sigma k^\sigma, \qquad \text{(Lorenz gauge)}. \tag{4.2.11}$$

In position/real spacetime, eq. (4.2.10) is equivalent to

$$\partial^2 A^\nu(x) = J^\nu(x) \qquad \partial^2 \equiv \eta^{\mu\nu}\partial_\mu\partial_\nu. \tag{4.2.12}$$

Problem 4.20. Lorentz covariance Suppose $\Lambda^\alpha{}_\mu$ is a Lorentz transformation; let two inertial frames $\{x^\mu\}$ and $\{x'^\mu\}$ be related via

$$x^\mu = \Lambda^\mu{}_\alpha x'^\alpha + a^\mu, \tag{4.2.13}$$

where a^μ is a constant vector. Suppose we solved the Lorenz gauge Maxwell's equations in the $\{x^\mu\}$ frame, namely

$$\frac{\partial A^\mu(x)}{\partial x^\mu} = 0, \qquad \eta^{\mu\nu}\frac{\partial}{\partial x^\mu}\frac{\partial}{\partial x^\nu} A_\alpha(x) = J_\alpha(x). \tag{4.2.14}$$

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

Problem 4.21. Poincaré transformations in Fourier spacetime Given a 1−form $A_\mu(x)$ in Minkowski spacetime – not necessarily the vector potential – let us define its Fourier transform as

$$A_\mu(x) = \int_{\mathbb{R}^{d-1,1}} \frac{\mathrm{d}^d k}{(2\pi)^d} \widetilde{A}_\mu(k) e^{-ik\cdot x}, \qquad k\cdot x \equiv k_\sigma x^\sigma. \tag{4.2.15}$$

Show that, under a Poincaré transformation,

$$x^\mu = \Lambda^\mu{}_\nu x'^\nu + a^\mu, \tag{4.2.16}$$

the Fourier coefficient in the $\{x'\}$ frame

$$A_{\mu'}(x') = \int_{\mathbb{R}^{d-1,1}} \frac{\mathrm{d}^d k'}{(2\pi)^d} \widetilde{A}_{\mu'}(k') e^{-ik'\cdot x'}, \qquad k'\cdot x' \equiv k'_\sigma x'^\sigma, \tag{4.2.17}$$

is related to the one in the $\{x\}$ frame (cf. (4.2.15)) through

$$\widetilde{A}_{\mu'}(k') = \widetilde{A}_\alpha(k) e^{-ik\cdot a} \Lambda^\alpha{}_\mu \Big|_{k_\alpha = k'_\mu (\Lambda^{-1})^\mu{}_\alpha}. \tag{4.2.18}$$

Because the equation obeyed by the Lorenz gauge vector potential, namely eq. (4.2.12), is Lorentz covariant – its Fourier coefficients must transform according to eq. (4.2.18). □

[3]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

[3]Eq. (4.2.12) 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 $\partial_{[\mu} F_{\alpha\beta]} = 0$.

in eq. (4.2.11):

$$A_\mu(x) = \int_{\mathbb{R}^{d-1,1}} \mathrm{d}^d x' G_d^+(x-x') J_{\mu'}(x'), \tag{4.2.19}$$

$$G_d^+(x-x') \equiv \int_{\mathbb{R}^d} \frac{\mathrm{d}^d k}{(2\pi)^d} \frac{e^{-ik\cdot(x-x')}}{-k^2}. \tag{4.2.20}$$

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. (4.2.12) to see that

$$\partial^2 F_{\mu\nu} = \partial_{[\mu} J_{\nu]}; \tag{4.2.21}$$

this means, using the same Green's function in eq. (4.2.20):

$$F_{\mu\nu}(x) = \int_{\mathbb{R}^{d-1,1}} \mathrm{d}^d x' G_d^+(x-x') \partial_{[\mu'} J_{\nu']}(x'). \tag{4.2.22}$$

We may verify that equations (4.2.19) and (4.2.20) solve eq. (4.2.12) readily:

$$\begin{aligned}
\partial_x^2 G_d^+(x-x') &= \int_{\mathbb{R}^d} \frac{\mathrm{d}^d k}{(2\pi)^d} \frac{\partial_\sigma \partial^\sigma e^{-ik\cdot(x-x')}}{-k^2} \\
&= \int_{\mathbb{R}^d} \frac{\mathrm{d}^d k}{(2\pi)^d} \frac{\partial_\sigma(-ik^\rho \delta_\rho^\sigma e^{-ik\cdot(x-x')})}{-k^2} \\
&= \int_{\mathbb{R}^d} \frac{\mathrm{d}^d k}{(2\pi)^d} \frac{\partial_\sigma(-ik^\sigma e^{-ik\cdot(x-x')})}{-k^2} \\
&= \int_{\mathbb{R}^d} \frac{\mathrm{d}^d k}{(2\pi)^d} \frac{(-ik_\sigma)(-ik^\sigma) e^{-ik\cdot(x-x')}}{-k^2} \\
&= \int_{\mathbb{R}^d} \frac{\mathrm{d}^d k}{(2\pi)^d} e^{-ik\cdot(x-x')} = \delta^{(d)}(x-x');
\end{aligned} \tag{4.2.23}$$

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

$$\partial_x^2 G_d^+(x-x') = \partial_{x'}^2 G_d^+(x-x') = \delta^{(d)}(x-x'); \tag{4.2.24}$$

Moreover, comparing each Cartesian component of the wave equation in eq. (4.2.12) with the one obeyed by the Green's function in eq. (4.2.24), 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. (4.2.19) 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. (4.2.19) is indeed a solution to eq. (4.2.12).

$$\begin{aligned}
\partial_x^2 A_\mu(x) = \partial_x^2 \left(\int_{\mathbb{R}^{d-1,1}} \mathrm{d}^d x' G_d^+(x-x') J_{\mu'}(x') \right) &= \int_{\mathbb{R}^{d-1,1}} \mathrm{d}^d x' \delta^{(d)}(x-x') J_{\mu'}(x') \\
&= J_\mu(x).
\end{aligned} \tag{4.2.25}$$

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. (4.2.2)) that would render the new gauge potential $A'_\mu \equiv A_\mu - \partial_\mu L$ satisfy

$$\partial^\mu A'_\mu = \partial^\mu A_\mu - \partial^2 L = 0. \tag{4.2.26}$$

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. (4.2.12),

$$\partial^2 \left(\partial^\sigma A_\sigma\right) = \partial^\sigma J_\sigma. \tag{4.2.27}$$

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:

$$\begin{aligned}\partial^\sigma \dot{A}_\sigma = \ddot{A}_0 + \partial^i \dot{A}_i &= \partial^0\partial_0 A_0 + \partial^i\partial_i A_0 + \partial^i\partial_0 A_i - \partial^i\partial_i A_0 \\ &= \partial^2 A_0 - \partial^i \left(\partial_i A_0 - \partial_0 A_i\right) \\ \partial_0 \left(\partial^\sigma A_\sigma\right) &= J_0 - \partial^i F_{i0}. \end{aligned} \tag{4.2.28}$$

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.

4.3 Gauge Invariant Variables for Electromagnetic 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.

- 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. (4.2.12)) 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.
- This will be a warm-up to an analogous analysis for gravitation linearized about a Minkowski "background" spacetime.

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:

$$A_i = \partial_i \alpha + \alpha_i, \tag{4.3.1}$$

where by "transverse" we mean

$$\partial_i \alpha_i = 0. \tag{4.3.2}$$

To demonstrate the generality of eq. (4.3.1) we shall first write A_i in Fourier space

$$A_i(t,\vec{x}) = \int_{\mathbb{R}^D} \frac{\mathrm{d}^D\vec{k}}{(2\pi)^D} \widetilde{A}_i(t,\vec{k}) e^{i\vec{k}\cdot\vec{x}}; \tag{4.3.3}$$

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

$$\begin{aligned}
\partial_j A_i &= \int_{\mathbb{R}^D} \frac{\mathrm{d}^D\vec{k}}{(2\pi)^D} \partial_j \left(i\delta_{ab}k^a x^b\right) \widetilde{A}_i(t,\vec{k}) e^{i\vec{k}\cdot\vec{x}} \\
&= \int_{\mathbb{R}^D} \frac{\mathrm{d}^D\vec{k}}{(2\pi)^D} \left(i\delta_{ab}k^a \delta^b_j\right) \widetilde{A}_i(t,\vec{k}) e^{i\vec{k}\cdot\vec{x}} \\
&= \int_{\mathbb{R}^D} \frac{\mathrm{d}^D\vec{k}}{(2\pi)^D} ik^j \widetilde{A}_i(t,\vec{k}) e^{i\vec{k}\cdot\vec{x}} \\
&= \int_{\mathbb{R}^D} \frac{\mathrm{d}^D\vec{k}}{(2\pi)^D} (-ik_j) \widetilde{A}_i(t,\vec{k}) e^{i\vec{k}\cdot\vec{x}}.
\end{aligned} \tag{4.3.4}$$

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

$$-ik_i \widetilde{\alpha}_i(t,\vec{k}) = 0. \tag{4.3.5}$$

At this point we simply write down

$$\widetilde{A}_i(t,\vec{k}) = \left(\delta_{ij} - \frac{k_i k_j}{\vec{k}^2}\right) \widetilde{A}_j(t,\vec{k}) + \frac{k_i k_j}{\vec{k}^2} \widetilde{A}_j(t,\vec{k}). \tag{4.3.6}$$

This is mere tautology, of course. However, we may now check that the first term on the left hand side of eq. (4.3.6) is transverse:

$$-ik_i \left(\delta_{ij} - \frac{k_i k_j}{\vec{k}^2}\right) \widetilde{A}_j(t,\vec{k}) = -i\left(k_j - \frac{\vec{k}^2 k_j}{\vec{k}^2}\right) \widetilde{A}_j(t,\vec{k}) = 0. \tag{4.3.7}$$

The second term on the right hand side of eq. (4.3.6) is a gradient because it is

$$-ik_i \left(\frac{ik_j}{\vec{k}^2} \widetilde{A}_j\right). \tag{4.3.8}$$

To sum, we have identified the α and α_i terms of eq. (4.3.1) as

$$\alpha(t,\vec{x}) = \int_{\mathbb{R}^D} \frac{\mathrm{d}^D\vec{k}}{(2\pi)^D} \frac{ik_j}{\vec{k}^2} \widetilde{A}_j(t,\vec{k}) e^{i\vec{k}\cdot\vec{x}}; \tag{4.3.9}$$

and the transverse portion as

$$\alpha_i(t,\vec{x}) = \int_{\mathbb{R}^D} \frac{\mathrm{d}^D\vec{k}}{(2\pi)^D} P_{ij}(\vec{k})\widetilde{A}_j(t,\vec{k})e^{i\vec{k}\cdot\vec{x}},$$
$$P_{ij}(\vec{k}) \equiv \delta_{ij} - \frac{k_i k_j}{\vec{k}^2}. \tag{4.3.10}$$

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

$$k_i P_{ij}(\vec{k}) = 0. \tag{4.3.11}$$

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

$$-ik_i\widetilde{\alpha} + \widetilde{\alpha}_i = -ik_i\widetilde{\beta} + \widetilde{\beta}_i, \tag{4.3.12}$$

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

$$\widetilde{\alpha} = \widetilde{\beta} \qquad \text{and} \qquad \widetilde{\alpha}_i = \widetilde{\beta}_i. \tag{4.3.13}$$

For, we may first "dot" both sides of eq. (4.3.12) with $\vec{k}$ and see that – for $\vec{k} \neq \vec{0}$,

$$\vec{k}^2\widetilde{\alpha} = \vec{k}^2\widetilde{\beta} \qquad \Leftrightarrow \qquad \widetilde{\alpha} = \widetilde{\beta}. \tag{4.3.14}$$

Plugging this result back into eq. (4.3.12), 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.

Problem 4.22. Gauge transformations We first define

$$\Phi \equiv A_0 - \dot{\alpha}, \tag{4.3.15}$$
$$A_i^{\mathrm{T}} \equiv \alpha_i. \tag{4.3.16}$$

By first doing a scalar-vector decomposition of the gauge transformation rules $A_\mu \to A_\mu + \partial_\mu\Lambda$, show that Φ and A_i^{T} are gauge-invariant. (Recall the uniqueness discussion above.) Proceed to show that

$$F_{0i} = \dot{\alpha}_i - \partial_i\Phi \tag{4.3.17}$$
$$F_{ij} = \partial_{[i}\alpha_{j]}. \tag{4.3.18}$$

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

$$J_\mu \equiv (\rho_{\rm E}, \partial_i \mathcal{J} + \mathcal{J}_i) \,. \tag{4.3.19}$$

Show that its conservation, namely $\partial^\mu J_\mu$, leads to

$$\dot{\rho}_{\rm E} = \vec{\nabla}^2 \mathcal{J}. \tag{4.3.20}$$

Notice the transverse portion of the current does not appear in the conservation law. □

Maxwell's Equations At this point, we are ready to write down Maxwell's equations $\partial^\mu F_{\mu\nu} = J_\nu$. From eq. (4.3.17), the $\nu = 0$ component is

$$-\partial_i F_{i0} = \partial_i(\dot{\alpha}_i - \partial_i \Phi) = -\vec{\nabla}^2 \Phi = \rho_{\rm E}. \tag{4.3.21}$$

The $\nu = i$ component of $\partial^\mu F_{\mu\nu} = J_\nu$, according to eq. (4.3.17) and (4.3.18),

$$\partial_0 F_{0i} - \partial_j F_{ji} = \partial_i \mathcal{J} + \mathcal{J}_i \tag{4.3.22}$$

$$\ddot{\alpha}_i - \partial_i \dot{\Phi} - \partial_j \left(\partial_j \alpha_i - \partial_i \alpha_j\right) = \partial_i \mathcal{J} + \mathcal{J}_i \tag{4.3.23}$$

$$\partial^2 \alpha_i - \partial_i \dot{\Phi} = \mathcal{J}_i + \partial_i \mathcal{J}. \tag{4.3.24}$$

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

$$\partial^2 \alpha_i = \mathcal{J}_i. \tag{4.3.25}$$

and the "scalar" portion to be

$$-\dot{\Phi} = \mathcal{J}. \tag{4.3.26}$$

We have gotten 3 (groups of) equations – (4.3.21), (4.3.25), (4.3.26) – for 2 sets of variables (Φ, α_i). Let us argue that eq. (4.3.26) is actually redundant. Taking into account eq. (4.3.20), we may take a time derivative of both sides of eq. (4.3.21),

$$-\vec{\nabla}^2 \dot{\Phi} = \dot{\rho}_{\rm E} = \vec{\nabla}^2 \mathcal{J}. \tag{4.3.27}$$

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. (4.3.26).

> **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. (4.3.1) and that of the current J_μ through eq. (4.3.19), we find a gauge-invariant Poisson equation
>
> $$-\vec{\nabla}^2 \Phi = \rho_{\rm E} \tag{4.3.28}$$

for the scalar potential Φ sourced by the electric charge density ρ_{E}; as well as a gauge-invariant wave equation

$$\partial^2 \alpha_i = \mathcal{J}_i; \tag{4.3.29}$$

for the transverse photon α_i sourced by the transverse portion of the electric current $\mathcal{J}_i$. The gauge-invariant scalar Φ and photon $A_i^{\mathrm{T}} \equiv \alpha_i$ are defined in equations (4.3.15) and (4.3.16).

These illuminate the theoretical structure of electromagnetism.[4]

Vacuum solution & Spin/Helicity-1 As we now turn to study, the vacuum solutions to the transverse-vector portion of A_μ may be identified with massless spin-1 photons in (3+1)-dimensions. By vacuum we mean the absence of any electric current, namely $J_\mu = 0$. For a given inertial frame, eq. (4.3.28) tells us $\Phi = 0$ (assuming $\Phi \to 0$ at spatial infinity); and eq. (4.3.29) translates to the vacuum wave equation

$$\partial^2 \alpha_i = 0. \tag{4.3.30}$$

In Fourier space we may immediately write down

$$\alpha_i(t, \vec{x}) = \int_{\mathbb{R}^3} \frac{\mathrm{d}^3\vec{k}}{(2\pi)^{d-1}} \left(\delta_{ij} - \frac{k_i k_j}{\vec{k}^2}\right) \left\{\epsilon_j(\vec{k}) e^{-ik\cdot x} + \text{c.c.}\right\}, \tag{4.3.31}$$

$$e^{-ik\cdot x} = e^{-i|\vec{k}|t + i\vec{k}\cdot\vec{x}}. \tag{4.3.32}$$

Here, the ϵ_j is an arbitrary $\vec{k}-$dependent $(d-1)$-component object and 'c.c.' is the complex conjugate of the preceding term.

Let us examine a single $\vec{k}-$mode and suppose k_i points in the positive 3−axis, so that

$$k_\mu = k(1, 0, 0, -1) \qquad \text{and} \qquad k^\mu = k(1, 0, 0, 1). \tag{4.3.33}$$

This means the plane wave itself becomes

$$\exp(-ik_\mu x^\mu) = \exp(-ik(t - x^3)); \tag{4.3.34}$$

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

$$\epsilon_j = \kappa \cdot \epsilon^{\parallel}{}_j + a_+ \epsilon^+{}_j + a_- \epsilon^-{}_j; \tag{4.3.35}$$

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

$$\epsilon^{\parallel}{}_j \equiv (0, 0, 1)^T, \tag{4.3.36}$$

$$\epsilon^\pm{}_\mu \equiv \frac{1}{\sqrt{2}} (0, \mp 1, i, 0)^T. \tag{4.3.37}$$

[4] *(1+1)D remark:* 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: $\widetilde{\alpha}_1 = 0$.

Problem 4.23. Show that for $k_\mu = k(1,0,0,-1)$ and referring to eq. (4.3.35),

$$\left(\delta_{ij} - \frac{k_i k_j}{\vec{k}^2}\right) \epsilon_j = a_+ {\epsilon^+}_i + a_- {\epsilon^-}_i. \tag{4.3.38}$$

That is, in a given inertial frame, the projector in eq. (4.3.31) selects only the 2D space of polarization vectors perpendicular to $\vec{k}$. $\square$

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

$$\widehat{R}(\theta)_{ij} \doteq \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{4.3.39}$$

the spatial polarizations in eq. (4.3.37) transform as

$$\widehat{R}(\theta)_{ij} {\epsilon^\pm}_j = \left(e^{-i\theta J}\right)_{ij} {\epsilon^\pm}_j = e^{\pm i\theta} {\epsilon^\pm}_j. \tag{4.3.40}$$

These $\epsilon^\pm$ are the spin-1 modes because they are eigenvectors of the (Hermitian) generator of rotations on the 2D (x^1, x^2) plane.

Problem 4.24. Verify eq. (4.3.40). $\square$

> **(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:
>
> $$\alpha_j(t, \vec{x}) = \text{Re} \int_{\mathbb{R}^3} \frac{\mathrm{d}^3\vec{k}}{(2\pi)^3} \left(a_+ {\epsilon^+}_j(\vec{k}) + a_- {\epsilon^-}_j(\vec{k})\right) e^{-ik\cdot x}, \tag{4.3.41}$$
>
> where ${\epsilon^\pm}_\mu$ are 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 $\epsilon^\pm \to \exp(\pm i\theta)\epsilon^\pm$.

Problem 4.25. Circularly polarized light from 4D spin-1 Consider a single spin-1 vacuum plane wave (cf. (4.3.37)) propagating along the 3−axis, with $k_\mu = k(1,0,0,-1)$:

$$\alpha_j^\pm(t,x,y,z) \equiv \text{Re}\left\{a_\pm {\epsilon^\pm}_j e^{-ik(t-z)}\right\}, \qquad a_\pm \in \mathbb{R}. \tag{4.3.42}$$

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:

$${}_\pm E^i = -\frac{k a_\pm}{\sqrt{2}} \left(\pm \sin(k(t-z))\widehat{x}^i + \cos(k(t-z))\widehat{y}^i\right), \tag{4.3.43}$$

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

$$\widehat{x}^i \doteq (1,0,0) \qquad \text{and} \qquad \widehat{y}^i \doteq (0,1,0). \tag{4.3.44}$$

$\square$

4.4 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:

$$G_4^+(x-x') \equiv \frac{\delta(t-t'-|\vec{x}-\vec{x}'|)}{4\pi|\vec{x}-\vec{x}'|}, \qquad x^\mu = (t,\vec{x}),\ x'^\mu = (t',\vec{x}'), \tag{4.4.1}$$

$$\partial_x^2 G_4^+(x-x') = \partial_{x'}^2 G_4^+(x-x') = \delta^{(4)}(x-x'), \tag{4.4.2}$$

$$\partial_x^2 \equiv \eta^{\mu\nu}\frac{\partial}{\partial x^\mu}\frac{\partial}{\partial x^\nu} \qquad \partial_{x'}^2 \equiv \eta^{\mu\nu}\frac{\partial}{\partial x'^\mu}\frac{\partial}{\partial x'^\nu}. \tag{4.4.3}$$

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. (4.4.1). 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.[5]

Problem 4.26. Analogy: Driven Simple Harmonic Oscillator Suppose we only Fourier-transformed the spatial coordinates in the Lorenz gauge Maxwell eq. (4.2.12). Show that this leads to

$$\ddot{\widetilde{A}}_\mu(t,\vec{k}) + k^2\widetilde{A}_\mu(t,\vec{k}) = \widetilde{J}_\mu(t,\vec{k}), \qquad k \equiv |\vec{k}|. \tag{4.4.4}$$

[6]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 $\vec{f}$:

$$m\ddot{\vec{q}} + \sigma\vec{q} = \vec{f}, \tag{4.4.5}$$

where each over-dot corresponds to a time derivative. Identify k^2 and $\widetilde{J}$ in eq. (4.4.4) with the appropriate quantities in eq. (4.4.5). Even though the Lorenz gauge Maxwell equations are fully 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 (4.4.4) are in fact a collection of free simple harmonic oscillators.

Now, how does one solve eq. (4.4.5)? Explain why the inverse of $(\mathrm{d}/\mathrm{d}t)^2 + k^2$ is

$$G_{\text{SHO}}(t-t',k) = -\int_{\mathbb{R}}\frac{\mathrm{d}\omega}{2\pi}\frac{e^{-i\omega(t-t')}}{\omega^2-k^2}. \tag{4.4.6}$$

That is, verify that this equation satisfies

$$\left(\frac{\mathrm{d}^2}{\mathrm{d}t^2}+k^2\right)G_{\text{SHO}}(t-t',k) = \left(\frac{\mathrm{d}^2}{\mathrm{d}t'^2}+k^2\right)G_{\text{SHO}}(t-t',k) = \delta(t-t'). \tag{4.4.7}$$

[5]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. (4.2.24), but propagates signals on the past light cone: $t = t'-R$.

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

If one tries to integrate ω over the real line in eq. (4.4.6), one runs into trouble – explain the issue. In other words, eq. (4.4.6) is actually ambiguous as it stands.

Now evaluate the Green's function G^+_{SHO} in eq. (4.4.6) using the contour running just slightly above the real line, i.e., $\omega \in (-\infty + i0^+, +\infty + i0^+)$. You should find

$$G^+_{\text{SHO}}(t-t',k) = \Theta(t-t')\frac{\sin\left(k(t-t')\right)}{k}. \tag{4.4.8}$$

Here, Θ is the step function

$$\begin{aligned}\Theta(x) &= 1, \qquad \text{if } x > 0, &(4.4.9)\\ &= 0, \qquad \text{if } x < 0. &(4.4.10)\end{aligned}$$

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

$$\widetilde{A}_\mu(t,\vec{k}) = \int_{-\infty}^{t} \mathrm{d}t' G^+_{\text{SHO}}(t-t',k)\widetilde{J}_\mu(t',\vec{k}). \tag{4.4.11}$$

By performing an inverse-Fourier transform, namely

$$A_\mu(x) = \int_{\mathbb{R}^{3,1}} \mathrm{d}^4x' G^+_4(x-x')J_{\mu'}(x'), \tag{4.4.12}$$

arrive at the expression in eq. (4.4.1) □

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

$$\epsilon^\pm{}_\mu(k)\exp(-ik\cdot x) = \epsilon^\pm{}_\mu(k)\exp(-ik_jx^j)\exp(-i\omega t), \qquad \omega \equiv |\vec{k}|, \tag{4.4.13}$$

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. (4.4.13) is a scalar, in the $\{x'\}$ Lorentz frame

$$-ik_\alpha x^\alpha = -ik_\alpha\Lambda^\alpha{}_\mu x'^\mu = -i(k_\alpha\Lambda^\alpha{}_0)t' - i(k_\alpha\Lambda^\alpha{}_i)x'^i. \tag{4.4.14}$$

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

$$\omega' = k_\alpha\Lambda^\alpha{}_0 = \omega\left(\Lambda^0{}_0 + \widehat{k}_i\Lambda^i{}_0\right) \tag{4.4.15}$$

$$\widehat{k}_i \equiv k_i/|\vec{k}| = k_i/\omega. \tag{4.4.16}$$

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:

$$u^\mu\partial_\mu S = \partial_0(k_\alpha x^\alpha) = \omega. \tag{4.4.17}$$

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

$$\begin{aligned}u'^\alpha\partial_{\alpha'}S = \partial_{t'}S &= \Lambda^\alpha{}_0\partial_\alpha(k\cdot x)\\ &= \Lambda^\alpha{}_0 k_\alpha = \omega\left(\Lambda^0{}_0 + \widehat{k}_i\Lambda^i{}_0\right). &(4.4.18)\end{aligned}$$

Problem 4.27. 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 $\widehat{n}_i$. $\square$

4.5 Symmetry and Conservation Laws for Free Photons

Free Photons We turn to the free photon, satisfying the free Maxwell equation

$$\partial_\mu F^{\mu\nu} = 0 \qquad \text{with} \qquad F_{\mu\nu} \equiv \partial_{[\mu} A_{\nu]}, \tag{4.5.1}$$
$$0 = \partial_{[\alpha} F_{\beta\gamma]}; \tag{4.5.2}$$

which in turn arises from the Lagrangian

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu}. \tag{4.5.3}$$

Here, there is a subtlety. While $F_{\mu\nu}$ is a Lorentz rank-2 tensor; the transformation properties of A_μ depends on the choice of gauge. If a non-covariant gauge is chosen, there is no reason to claim A_μ is a rank-1 Lorentz tensor. To this end, for technical convenience, we shall choose the Lorentz covariant Lorenz gauge

$$\partial^\mu A_\mu = 0. \tag{4.5.4}$$

Translations Under spacetime translations, the vector potential then transforms as follows.

$$A_\mu(x) \to A_\mu(x+a) = A_\mu(x) + a^\nu \partial_\nu A_\mu(x) \tag{4.5.5}$$
$$\partial_\mu A_\nu(x) \to \partial_\mu A_\nu(x+a) = \partial_\mu A_\nu(x) + a^\sigma \partial_\sigma \partial_\mu A_\nu(x). \tag{4.5.6}$$

The Lagrangian density transforms as

$$\mathcal{L}(x) \to \mathcal{L}(x) + a^\sigma \partial_\sigma \mathcal{L}(x) \tag{4.5.7}$$
$$\to \mathcal{L} + \partial_\mu (a^\sigma \delta^\mu_\sigma \mathcal{L}). \tag{4.5.8}$$

Whereas, if we had expanded it via its fields,

$$\mathcal{L} \to \mathcal{L} + \frac{\partial \mathcal{L}}{\partial \partial_\lambda A_\tau} a^\sigma \partial_\sigma \partial_\lambda A_\tau(x) \tag{4.5.9}$$
$$= \mathcal{L} + \partial_\lambda \left(\frac{\partial \mathcal{L}}{\partial \partial_\lambda A_\tau} a^\sigma \partial_\sigma A_\tau(x) \right) - \partial_\lambda \left(\frac{\partial \mathcal{L}}{\partial \partial_\lambda A_\tau} \right) a^\sigma \partial_\sigma A_\tau(x). \tag{4.5.10}$$

The Noether current evaluated on the EoM $\partial_\lambda(\partial \mathcal{L}/\partial(\partial_\lambda A_\tau)) = 0$ is

$$J^\mu{}_\nu = \frac{\partial \mathcal{L}}{\partial \partial_\mu A_\tau} \partial_\nu A_\tau(x) - \delta^\mu{}_\nu \mathcal{L} \tag{4.5.11}$$
$$= -F^{\mu\tau} \partial_\nu A_\tau + \frac{1}{4} \delta^\mu{}_\nu F_{\lambda\tau} F^{\lambda\tau}. \tag{4.5.12}$$

Notice this is not symmetric nor gauge-invariant![7] Let's backtrack a little.

$$a^\sigma \partial_\sigma \mathcal{L} = \delta \mathcal{L} = \frac{\partial \mathcal{L}}{\partial(\partial_\lambda A_\tau)} a^\sigma \partial_\sigma \partial_\lambda A_\tau \tag{4.5.13}$$

$$= -a^\sigma F^{\lambda\tau} \partial_\sigma \partial_\lambda A_\tau = -\frac{1}{2} a^\sigma F^{\lambda\tau} \partial_\sigma F_{\lambda\tau}. \tag{4.5.14}$$

Let us use the Bianchi identity $\partial_{[\sigma} F_{\lambda\tau]} = 0$ to deduce

$$\partial_\sigma F_{\lambda\tau} = \partial_\lambda F_{\sigma\tau} + \partial_\tau F_{\lambda\sigma} = -\partial_{[\lambda} F_{\tau]\sigma}. \tag{4.5.15}$$

Inserting this back into the first order variation of the Lagrangian,

$$a^\sigma \partial_\mu (\delta^\mu_\sigma \mathcal{L}) = \frac{1}{2} a^\sigma F^{\lambda\tau} \partial_{[\lambda} F_{\tau]\sigma} \tag{4.5.16}$$

$$= a^\sigma F^{\lambda\tau} \partial_\lambda F_{\tau\sigma} = -a^\sigma F^{\lambda\tau} \partial_\lambda F_{\sigma\tau} \tag{4.5.17}$$

$$= -a^\nu \partial_\mu \left(F^{\mu\sigma} F_{\nu\sigma} \right). \tag{4.5.18}$$

Our gauge invariant and symmetric Noether current is therefore

$$J^{\mu\nu} = -F^{\mu\sigma} F^\nu{}_\sigma + \frac{1}{4} \eta^{\mu\nu} F^{\sigma\rho} F_{\sigma\rho} \equiv T^{\mu\nu} \tag{4.5.19}$$

$$0 = \partial_\mu J^{\mu\nu}. \tag{4.5.20}$$

Lorentz transformations We now turn to the Noether current of Lorentz transformations. For *fixed* (α, β), we have

$$\partial_{\mu'} A_{\nu'}(x') = \frac{\partial x^\sigma}{\partial x'^\mu} \frac{\partial x^\rho}{\partial x'^\nu} \partial_\sigma A_\rho (x = \Lambda x') \tag{4.5.21}$$

$$\begin{aligned} &= \partial_\mu A_\nu (x \to x') + i\omega \left(\widehat{J}^{\alpha\beta} \right)^\sigma{}_\mu \partial_\sigma A_\nu (x \to x') + i\omega \left(\widehat{J}^{\alpha\beta} \right)^\sigma{}_\nu \partial_\mu A_\sigma (x \to x') \\ &\quad - i\omega \left(\widehat{J}^{\alpha\beta} \right)^\sigma{}_\rho x'^\rho \partial_{\sigma'} \partial_\mu A_\nu (x \to x'). \end{aligned} \tag{4.5.22}$$

The Lagrangian, being a Lorentz scalar, transforms as

$$\mathcal{L}(x = \Lambda x') = \mathcal{L}(x \to x') - i\omega \left(\widehat{J}^{\alpha\beta} \right)^\sigma{}_\rho x'^\rho \partial_{\sigma'} \mathcal{L}(x \to x'). \tag{4.5.23}$$

This means, under $x \to \Lambda x$, we have

$$x^{[\beta} \partial^{\alpha]} \mathcal{L} = \frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} \left(\delta^{[\alpha}_\mu \partial^{\beta]} A_\nu + \delta^{[\alpha}_\nu \partial_\mu A^{\beta]} + x^{[\beta} \partial^{\alpha]} \partial_\mu A_\nu \right) \tag{4.5.24}$$

$$= -F^{\mu\nu} \left(\delta^{[\alpha}_\mu \partial^{\beta]} A_\nu + \delta^{[\alpha}_\nu \partial_\mu A^{\beta]} + x^{[\beta} \partial^{\alpha]} \partial_\mu A_\nu \right). \tag{4.5.25}$$

[7] Several sources, including the *An Introduction to Quantum Field Theory* textbook by Peskin and Schroeder and the *Classical Field Theory* book by Davison Soper, stops here. The latter goes on to assert the electromagnetic Noether current generated by spacetime translations is not gauge-invariant nor symmetric because of 'internal spin'. It is true that the Noether current is not unique; a $\partial_\sigma K^{\mu\sigma\nu}$ may always be added to $J^{\mu\nu}$, while still maintaining $\partial_\mu J^{\mu\nu} = 0$, if K is anti-symmetric in the first two indices: $K^{\mu\sigma\nu} = -K^{\sigma\mu\nu}$. However, we will soon witness, if the index anti-symmetries and Bianchi identities are proprerly employed, a symmetric and gauge invariant free Maxwell stress tensor may be obtained.

Using the explicit expression $F^{\mu\nu} = \partial^{[\mu} A^{\nu]}$, and via a direct calculation, we may verify the first two terms on the last line vanish.

$$\partial_\mu \left(x^{[\beta} \eta^{\alpha]\mu} \mathcal{L} \right) = -F^{\mu\nu} x^{[\beta} \partial^{\alpha]} \partial_\mu A_\nu = -\frac{1}{2} F^{\mu\nu} x^{[\beta} \partial^{\alpha]} F_{\mu\nu} \tag{4.5.26}$$

$$= \frac{1}{2} F^{\mu\nu} x^{[\beta} \partial_{[\mu} F_{\nu]}{}^{\alpha]} = -F^{\mu\nu} x^{[\beta} \partial_\mu F^{\alpha]}{}_\nu. \tag{4.5.27}$$

Bianchi was employed in the second line. Recalling the free photon EoM,

$$\partial_\mu \left(x^{[\beta} \eta^{\alpha]\mu} \mathcal{L} \right) = -\partial_\mu \left(F^{\mu\nu} x^{[\beta} F^{\alpha]}{}_\nu \right) + F^{\mu\nu} \delta^{[\beta}_\mu F^{\alpha]}{}_\nu \tag{4.5.28}$$

$$= \partial_\mu \left(-x^{[\beta} F^{\alpha]}{}_\nu F^{\mu\nu} \right) + F^{[\beta}{}_\mu F^{\alpha]}{}_\nu \eta^{\mu\nu}. \tag{4.5.29}$$

The second term on the final line is zero, because we are anti-symmetrizing the indices of a symmetric tensor. Comparison with the stress tensor of the free electromagnetic field, we obtain the angular momentum Noether tensor

$$M^{\mu\alpha\beta} = x^{[\alpha} T^{\beta]\mu} = x^{[\alpha} \left(-F^{\beta]\sigma} F^\mu{}_\sigma + \frac{1}{4} \eta^{\beta]\mu} F_{\sigma\rho} F^{\sigma\rho} \right) \tag{4.5.30}$$

$$0 = \partial_\mu M^{\mu\alpha\beta}. \tag{4.5.31}$$

Problem 4.28. Components of EM Stress Tensor Starting from the electromagnetic stress tensor in eq. (4.5.19), verify that

$$T^{00}[F] = \frac{\vec{E}^2 + \vec{B}^2}{2} \qquad T^{0i}[F] = T^{i0}[F] = \vec{E} \times \vec{B} \tag{4.5.32}$$

$$\text{and} \qquad T^{ij}[F] = \frac{\delta^{ij}}{2} \left(\vec{E}^2 + \vec{B}^2 \right) - E^i E^j - B^i B^j. \tag{4.5.33}$$

□

4.6 Electromagnetic fields of a point charge

The electromagnetic current of a point charge q is

$$J^\mu(x) = q \int_{\mathbb{R}} \mathrm{d}\tau \frac{\mathrm{d}z^\mu}{\mathrm{d}\tau} \delta^{(4)}(x - z). \tag{4.6.1}$$

We may compute

$$F_{\mu\nu} \tag{4.6.2}$$

$$= q \int_{x'} \Theta[t - t'] \frac{\partial_{[\mu} \delta[\sigma]}{4\pi} \int \mathrm{d}\tau \frac{\mathrm{d}z_{\nu]}}{\mathrm{d}\tau} \delta^{(4)}[x' - z] \tag{4.6.3}$$

$$= q \int \mathrm{d}\tau \frac{1}{\dot{z} \cdot (z - x)} \frac{\mathrm{d}}{\mathrm{d}\tau} \frac{\delta_+[\sigma]}{4\pi} (x - z)_{[\mu} \frac{\mathrm{d}z_{\nu]}}{\mathrm{d}\tau}$$

$$= \frac{q}{4\pi} \int \mathrm{d}\tau \delta_+[\sigma] \frac{\mathrm{d}}{\mathrm{d}\tau} \left(\frac{(x - z)_{[\mu}}{\dot{z} \cdot (x - z)} \frac{\mathrm{d}z_{\nu]}}{\mathrm{d}\tau} \right) \tag{4.6.4}$$

$$
\begin{aligned}
&= \frac{q}{4\pi}\int \mathrm{d}\tau \frac{\delta_+[\tau-\tau_{\rm ret}]}{|\mathrm{d}\sigma/\mathrm{d}\tau|}\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{(x-z)_{[\mu}}{\dot{z}\cdot(x-z)}\frac{\mathrm{d}z_{\nu]}}{\mathrm{d}\tau}\right) \\
&= \frac{q}{4\pi|\dot{z}\cdot(z-x)|}\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{(x-z)_{[\mu}}{\dot{z}\cdot(x-z)}\frac{\mathrm{d}z_{\nu]}}{\mathrm{d}\tau}\right)_{\tau=\tau_{\rm ret}} \\
&= \frac{q}{4\pi|\mathrm{d}\sigma/\mathrm{d}t'|}\frac{\mathrm{d}}{\mathrm{d}t'}\left(\frac{(x-z)_{[\mu}}{(1,\vec{v})\cdot(x-z)}\frac{\mathrm{d}z_{\nu]}}{\mathrm{d}t'}\right)_{t'=t_{\rm ret}} = \frac{q}{4\pi|(1,\vec{v})\cdot(z-x)|}\frac{\mathrm{d}}{\mathrm{d}t'}\left(\frac{(x-z)_{[\mu}}{(1,\vec{v})\cdot(x-z)}\frac{\mathrm{d}z_{\nu]}}{\mathrm{d}t'}\right)_{t'=t_{\rm ret}} \\
&= \frac{q}{4\pi|\vec{x}-\vec{z}|(1-\vec{v}\cdot\widehat{R})}\left(\frac{(x-z)_{[\mu}}{(1,\vec{v})\cdot(x-z)}\frac{\mathrm{d}^2z_{\nu]}}{\mathrm{d}t'^2} - \frac{(x-z)_{[\mu}}{((1,\vec{v})\cdot(x-z))^2}\frac{\mathrm{d}z_{\nu]}}{\mathrm{d}t'}\left((0,\vec{a})\cdot(x-z)-(1,\vec{v})^2\right)\right)_{t'=t_{\rm ret}} \\
&= \frac{q}{4\pi R(1-\vec{v}\cdot\widehat{R})}\left(\frac{(R,-\vec{R})_{[\mu}(0,-\vec{a})_{\nu]}}{R-\vec{v}\cdot\vec{R}} + \frac{(R,-\vec{R})_{[\mu}(1,-\vec{v})_{\nu]}}{(R-\vec{v}\cdot\vec{R})^2}\left(1-\vec{v}^2+\vec{a}\cdot\vec{R}\right)\right)_{t'=t_{\rm ret}} \\
&= \frac{q}{4\pi}\left(\frac{(1,-\widehat{R})_{[\mu}(0,-\vec{a})_{\nu]}}{R(1-\vec{v}\cdot\widehat{R})^2} + \frac{(1,-\widehat{R})_{[\mu}(1,-\vec{v})_{\nu]}}{(1-\vec{v}\cdot\widehat{R})^3}\left(\frac{1-\vec{v}^2}{R^2}+\frac{\vec{a}\cdot\widehat{R}}{R}\right)\right)_{t'=t_{\rm ret}} && (4.6.5)
\end{aligned}
$$

$$
F^{\mu\nu} = \frac{q}{4\pi}\left(\frac{(1,\widehat{R})^{[\mu}(0,\vec{a})^{\nu]}}{R(1-\vec{v}\cdot\widehat{R})^2} + \frac{(1,\widehat{R})^{[\mu}(1,\vec{v})^{\nu]}}{(1-\vec{v}\cdot\widehat{R})^3}\left(\frac{1-\vec{v}^2}{R^2}+\frac{\vec{a}\cdot\widehat{R}}{R}\right)\right)_{t'=t_{\rm ret}} \qquad (4.6.6)
$$

The electric fields are

$$
\begin{aligned}
F^{i0} = E^i &= -\frac{q}{4\pi}\left(\frac{a^i}{R(1-\vec{v}\cdot\widehat{R})^2} + \frac{v^i-\widehat{R}^i}{(1-\vec{v}\cdot\widehat{R})^3}\left(\frac{1-\vec{v}^2}{R^2}+\frac{\vec{a}\cdot\widehat{R}}{R}\right)\right)_{t'=t_{\rm ret}} && (4.6.7) \\
&= -\frac{q}{4\pi R}\left(\frac{a^i}{(1-\vec{v}\cdot\widehat{R})^2} + \frac{(v^i-\widehat{R}^i)\left(\vec{a}\cdot\widehat{R}\right)}{(1-\vec{v}\cdot\widehat{R})^3}\right) - \frac{q(1-\vec{v}^2)}{4\pi R^2}\frac{v^i-\widehat{R}^i}{(1-\vec{v}\cdot\widehat{R})^3}\Bigg|_{t'=t_{\rm ret}} && (4.6.8)
\end{aligned}
$$

The magnetic fields are

$$
F^{ij} = \frac{q}{4\pi}\left(\frac{\widehat{R}^{[i}a^{j]}}{R(1-\vec{v}\cdot\widehat{R})^2} + \frac{\widehat{R}^{[i}v^{j]}}{(1-\vec{v}\cdot\widehat{R})^3}\left(\frac{1-\vec{v}^2}{R^2}+\frac{\vec{a}\cdot\widehat{R}}{R}\right)\right)_{t'=t_{\rm ret}} = -\widehat{R}^{[i}E^{j]}. \qquad (4.6.9)
$$

Here, we have defined $\vec{v} \equiv \mathrm{d}\vec{z}/\mathrm{d}t$, $\vec{a} \equiv \mathrm{d}^2\vec{z}/\mathrm{d}t^2$, $\vec{R} \equiv \vec{x}-\vec{z}$, $R \equiv |\vec{x}-\vec{z}|$ and $\widehat{R} \equiv \vec{R}/R$.

Problem 4.29. Vector Form Verify that

$$
\begin{aligned}
\vec{E} &= \frac{q}{4\pi R}\frac{\widehat{R}\times\left((\widehat{R}-\vec{v})\times\vec{a}\right)}{(1-\vec{v}\cdot\widehat{R})^3} + \frac{q(1-\vec{v}^2)}{4\pi R^2}\frac{\widehat{R}-\vec{v}}{(1-\vec{v}\cdot\widehat{R})^3}\Bigg|_{t'=t_{\rm ret}} && (4.6.10) \\
\vec{B} &= \widehat{R}\times\vec{E}. && (4.6.11)
\end{aligned}
$$

In particular, note that the $1/R$ piece of the electric field can be expressed as a double cross product; and the electric and magnetic fields are mutually perpendicular. □

5 Last update: May 18, 2025

References