Tutorial 1

2015-11-02 (document version: 2015-11-03)

1 Problems

Problem 1. (a) Show that for *any* Lorentz transformation $\Lambda^{\mu}{}_{\nu}$ it always hold,

$$\det(\Lambda^{\mu}{}_{\nu}) = \pm 1 \quad \text{and} \quad \left|\Lambda^{0}{}_{0}\right| \geq 1. \tag{1.1}$$

(b) (Schwartz, 2.4) Is the transformation $Y : (t, x, y, z) \to (t, x, -y, z)$ a Lorentz transformation? If so, why is it not considered with P and T as a discrete Lorentz transformation? If not, why not? Supplement: What about $Z : (t, x, y, z) \to (-t, -x, y, z)$?

Problem 2. (a) Show that for any $X_{\alpha\beta}$,

$$\frac{\partial}{\partial X_{\mu\nu}} X_{[\alpha\beta]} = \delta^{\mu}_{[\alpha}\delta^{\nu}_{\beta]}. \tag{1.2}$$

(b) Show that,

$$\frac{\partial}{\partial(\partial_\mu A_\nu)} \left(\frac{1}{4} F_{\rho\sigma}F^{\rho\sigma}\right) = F^{\mu\nu}, \tag{1.3}$$

where,

$$F_{\mu\nu} := \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{1.4}$$

Problem 3. (Schwartz 2.6) Lorentz invariance.

(a) Show that

$$\int_{-\infty}^{\infty} \mathrm{d}k^0\, \delta\left(k^2 - m^2\right)\, \theta(k^0) = \frac{1}{2\omega_k}, \tag{1.5}$$

where $\theta(x)$is the unit step function and $\omega_k := \sqrt{\mathbf{k}^2 + m^2}$.

(b) Show that the integration measure d^4k is Lorentz invariant.

(c) Finally, show that

$$\int \frac{\mathrm{d}^3k}{2\omega_k}, \tag{1.6}$$

is Lorentz invariant.

Purpose

The first problem addresses: (i) Vector and dual spaces. Metric. SR spacetime, Minkowski metric. Vectors/covectors, index notation, raising/lowering indices. (ii) Isometries of SR, translations, rotations, boosts, parity/time reflections. Lorentz transformation, definition, decomposition P-T-B-R, orthochronous+proper=restricted (continuous), space and time inversions (discrete transformations). Lorentz invariance.

The second problem displays index gymnastics. It also demonstrates the importance of raising/lowering indices, e.g. when taking derivatives of $F_{\rho\sigma}F^{\rho\sigma} = F_{\rho_1\sigma_1}g^{\rho_1\rho_2}g^{\sigma_1\sigma_2}F_{\rho_2\sigma_2}$. Also some questions to answer: Is $F_{\rho\sigma}F^{\rho\sigma}$ Lorentz invariant? What terms in the Lagrangian density can we have?

The third problem addresses: (i) The delta function and its properties, e.g. $\delta(g(x))$. Usage of the last property in (P3.a) and cross-sections (with references to M&S). Background for (P3.b). It addresses also the Lorentz invariance, the volume element, a Jacobian of the transformation (and tensor densities. If we have time: Why it's wrong $(\delta(x))^3$ and what does mean $\delta^3(\vec{x})$ vs $\delta^{(3)}(x)$.

2 Recap: SR and Lorentz Invariance

In the following, the Einstein summation convention is used, which says that a sum is understood over any repeated index. Also, symmetrization and anti-symmetrization are often denoted by parentheses and brackets around the indices concerned. For instance, for any tensor $T_{\alpha\beta}$,

$$T_{(\alpha\beta)} = \frac{1}{2}\left(T_{\alpha\beta} + T_{\beta\alpha}\right), \quad T_{[\alpha\beta]} = \frac{1}{2}\left(T_{\alpha\beta} - T_{\beta\alpha}\right). \tag{2.1}$$

The arena of SR is the Minkowski spacetime (usually denoted as $\mathbb{R}^{1,3}$, or sometimes M^4) which is a pseudo-Euclidean space with the Minkowski metric η having the signature,

$$\eta_{\mu\nu} = \mathrm{diag}(+1, -1, -1, -1), \tag{2.2}$$

(There is also an alternative signature $(-, +, +, +)$; the most literature in GR uses the latter while particle physicists tend to use the former.)

The inverse of the metric is,

$$(\eta^{-1})^{\mu\sigma}\eta_{\sigma\nu} = \delta^{\mu}_{\nu}, \tag{2.3}$$

for which we use the same symbol (as there is no confusion),

$$\eta^{\mu\nu} := (\eta^{-1})^{\mu\nu}, \tag{2.4}$$

The inverse of the Minkowski metric has the same components,

$$\eta_{\mu\nu} = \mathrm{diag}(+1, -1, -1, -1). \tag{2.5}$$

We distinguish two vector spaces, one containing ordinary vectors, identified by the components having the indices up, e.g.,

$$x^{\mu} = (x^0, x^1, x^2, x^3) = (x^0, x^i) = (t, \mathbf{x}), \tag{2.6}$$
$$p^{\mu} = (p^0, p^1, p^2, p^3) = (p^0, p^i) = (E, \mathbf{p}), \tag{2.7}$$

and the corresponding dual vector space with covectors (i.e., linear functionals) that we identify by the components with the indices down, e.g.,

$$\partial_{\mu} = \frac{\partial}{\partial x^{\mu}} = (\partial_0, \partial_1, \partial_2, \partial_3) = (\partial_0, \partial_i) = (\partial_t, \boldsymbol{\nabla})$$
$$p_{\mu} = (p_0, p_1, p_2, p_3) = (p_0, p_i) = (E, -\mathbf{p}).$$

We use the metric to raise/lower indices, e.g.,

$$p_{\mu} = \eta_{\mu\nu}p^{\nu}, \qquad p^{\mu} = \eta^{\mu\nu}p_{\nu}. \tag{2.8}$$

The Lorentz transformations,

$$x^\mu \to x'^\mu = \Lambda^\mu{}_\nu x^\nu, \tag{2.9}$$

is the homogeneous isometry of Minkowski spacetime which leave the square of the length of a vector invariant. The matrix Λ representing the transformation is a constant matrix; x^μ and x'^μ are the coordinates of the same event in two different inertial frames. The matrix Λ must satisfy the condition,

$$\eta = \Lambda^{\mathtt{T}}\eta\Lambda, \tag{2.10}$$

implying,

$$\det\Lambda = \pm 1. \tag{2.11}$$

Moreover, we necessarily have (see Problem 1),

$$\left|\Lambda^0{}_0\right| \geq 1. \tag{2.12}$$

The Lorentz transformation can be decomposed into the symmetric part (boost) and the orthogonal part (rotations) which can be continuously parametrized, plus two discrete transformations: time inversion and space reflections. The Lorentz transformation with $\det\Lambda = 1$ are called proper and with $\Lambda^0{}_0 \geq 1 > 0$ are called orthochronous. The orthochronous and proper are also known as the reduced Lorentz transformations.

Under the Lorentz transformation,

$$x^\mu \to x'^\mu = \Lambda^\mu{}_\nu x^\nu, \qquad x' = \Lambda x, \quad x = \Lambda^{-1}x', \tag{2.13}$$

the scalar field $\phi(x)$ transform as,

$$\phi(x) \to \phi'(x') = \phi(x) = \phi(\Lambda^{-1}x'), \tag{2.14}$$

the vector field $V^\mu(x)$ transform as,

$$V^\mu(x) \to V'^\mu(x') = \Lambda^\mu{}_\nu V^\nu(x) = \Lambda^\mu{}_\nu V^\nu(\Lambda^{-1}x'), \tag{2.15}$$

the covector field $A_\mu(x)$ transform as,

$$A_\mu(x) \to A'_\mu(x') = (\Lambda^{-1,\mathtt{T}})_\mu{}^\nu A_\nu(x) = (\Lambda^{-1,\mathtt{T}})_\mu{}^\nu A_\nu(\Lambda^{-1}x') \tag{2.16}$$

and the Maxwell tensor field $F_{\mu\nu}(x)$ transforms accordingly,

$$F_{\mu\nu}(x) \to F'_{\mu\nu}(x') = (\Lambda^{-1,\mathtt{T}})_\mu{}^\rho F_{\rho\sigma}(x)(\Lambda^{-1})^\sigma{}_\nu \tag{2.17}$$

$$= (\Lambda^{-1,\mathtt{T}})_\mu{}^\rho F_{\rho\sigma}(\Lambda^{-1}x')(\Lambda^{-1})^\sigma{}_\nu. \tag{2.18}$$

3 Recap: Fourier transform, Notation;[1] Dirac delta function

The Fourier transform has many forms, where in physics, we mostly use the one written in terms of angular frequency ω,

$$\hat{f}(\omega) = \int_{\mathbb{R}^n} \mathrm{d}^n x\, f(x)\, \mathrm{e}^{-\mathrm{i}\omega\cdot x}. \tag{3.1}$$

[1]This Fourier transform was not covered during Tutorial 1 (it will be covered later).

Under this convention, the inverse transform becomes,

$$f(x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \mathrm{d}^n\omega\, \hat{f}(\omega) \mathrm{e}^{\mathrm{i}\omega\cdot x}. \tag{3.2}$$

When the Fourier transform is defined this way, it is no longer a unitary transformation on $L^2(\mathbb{R}^n)$. There is also less symmetry between the formulae for the Fourier transform and its inverse. Another convention is to split the factor of $(2\pi)^n$ evenly between the Fourier transform and its inverse, which leads to definitions (convention used in Sakurai, e.g.),

$$\hat{f}(\omega) = \frac{1}{(2\pi)^{n/2}} \int \mathrm{d}^n x\, f(x)\, \mathrm{e}^{-\mathrm{i}\omega\cdot x}, \quad f(x) = \frac{1}{(2\pi)^{n/2}} \int \mathrm{d}^n\omega\, \hat{f}(\omega) \mathrm{e}^{\mathrm{i}\omega\cdot x}. \tag{3.3}$$

Fourier transform in braket notation. Sakurai & Napolitano - *Modern Quantum Mechanics*

$$\begin{gathered}
\langle \mathbf{x}' | \alpha \rangle = \psi_\alpha(\mathbf{x}'), \qquad \langle \mathbf{p}' | \alpha \rangle = \phi_\alpha(\mathbf{p}') \\
\int \mathrm{d}^3\mathbf{x}' \, |\mathbf{x}'\rangle \langle \mathbf{x}'| = \mathbb{1}, \quad \langle \mathbf{x}' | \mathbf{x}'' \rangle = \delta^3(\mathbf{x}' - \mathbf{x}'') \\
\int \mathrm{d}^3\mathbf{p}' \, |\mathbf{p}'\rangle \langle \mathbf{p}'| = \mathbb{1}, \quad \langle \mathbf{p}' | \mathbf{p}'' \rangle = \delta^3(\mathbf{p}' - \mathbf{p}'') \\
\langle \mathbf{x}' | \mathbf{p}' \rangle = \frac{1}{(2\pi\hbar)^{3/2}} \cdot \exp\left(\mathrm{i}\mathbf{p}' \cdot \mathbf{x}'/\hbar\right) \\
\langle \mathbf{p}' | \mathbf{x}' \rangle = \langle \mathbf{x}' | \mathbf{p}' \rangle^* = \frac{1}{(2\pi\hbar)^{3/2}} \cdot \exp\left(-\mathrm{i}\mathbf{p}' \cdot \mathbf{x}'/\hbar\right) \\
\text{cf. discrete: } \sum_n |n\rangle \langle n| = \mathbb{1}, \quad \langle n|m\rangle = \delta_{nm}, \quad \langle n|\alpha\rangle = u_n
\end{gathered}$$

Now, compare the plane wave solutions $\phi(x)$ for the real Klein-Gordon-Fock equation,

$$(\Box + m^2)\phi(x) = (\partial_\mu \partial^\mu + m^2)\phi(x) = 0, \quad \partial_\mu\partial^\mu = \partial_t^2 - \boldsymbol{\nabla}^2, \tag{3.4}$$

derived as the equations of motion from the Lagrangian density,

$$\mathcal{L} = \frac{1}{2}\left(\partial_\mu\phi(x)\right)\left(\partial^\mu\phi(x)\right) - \frac{1}{2}m^2\phi^2(x). \tag{3.5}$$

- Mandl and Shaw, Eq. (3.7)
  (N.b. $\phi(x)$ is an operator , as well as a, $a^\dagger$!)

$$\phi(x) = \phi^+(x) + \phi^-(x) = \sum_{\mathbf{k}} \left(\frac{\hbar c^2}{2V\omega_{\mathbf{k}}}\right)^{1/2} \left(a(\mathbf{k})\mathrm{e}^{-\mathrm{i}kx} + a^\dagger(\mathbf{k})\mathrm{e}^{\mathrm{i}kx}\right) \tag{3.6}$$

$$\phi^+(x) = \sum_{\mathbf{k}} \left(\frac{\hbar c^2}{2V\omega_{\mathbf{k}}}\right)^{1/2} a(\mathbf{k})\mathrm{e}^{-\mathrm{i}kx} \tag{3.7}$$

$$\phi^+(x) = \sum_{\mathbf{k}} \left(\frac{\hbar c^2}{2V\omega_{\mathbf{k}}}\right)^{1/2} a^\dagger(\mathbf{k})\mathrm{e}^{\mathrm{i}kx} \tag{3.8}$$

- Schwartz, Eq. (2.78)

$$\phi(x) = \int \frac{\mathrm{d}^3p}{(2\pi)^3} \frac{1}{\sqrt{2\omega_{\mathbf{p}}}} \left(a_{\mathbf{p}} \mathrm{e}^{-\mathrm{i}px} + a_{\mathbf{p}}^\dagger \mathrm{e}^{\mathrm{i}px} \right) \tag{3.9}$$

- Peskin and Schroeder Eq. (2.47)

$$\phi(x) = \int \frac{\mathrm{d}^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_{\mathbf{p}}}} \left(a_{\mathbf{p}} \mathrm{e}^{-\mathrm{i}p\cdot x} + a_{\mathbf{p}}^\dagger \mathrm{e}^{\mathrm{i}p\cdot x} \right) |_{p^0 = E_{\mathbf{p}}} \tag{3.10}$$

Dirac Delta Function, properties. Jackson - *Classical Electrodynamics*, 3rd ed
The Dirac delta function, or δ function, is a generalized function, or distribution, on the real number line that is zero everywhere except at zero, with an integral of one over the entire real line. It has many representations, one the most common is given in (7°) below.

$$1^\circ\text{(definition)} \int \mathrm{d}x\, \delta(x)\, f(x) = f(0), \quad \text{(i.e. } \langle \delta, f \rangle = f(0))$$
$$2^\circ\ \delta(-x) = \delta(x), \quad 3^\circ\ \delta(ax) = \frac{\delta(x)}{|a|}, \quad 4^\circ \int \mathrm{d}x\, f(x)\delta(x-a) = f(a),$$
$$5^\circ\ \delta(g(x)) = \sum_{x_i \in g^{-1}(\{0\})} \frac{\delta(x - x_i)}{\left|\frac{\partial g(x)}{\partial x}\right|_{x = x_i}}, \quad x_i \in g^{-1}(\{0\}) = \{\, x \,|\, g(x) = 0 \,\},$$
$$6^\circ\ \delta^3(\mathbf{x}) = \delta(x)\delta(y)\delta(z), \quad 7^\circ\ \delta(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{d}k\, \mathrm{e}^{\mathrm{i}xk}$$

Questions: Is $(\delta(\mathbf{x}))^3$ valid? What does mean $\delta^3(\mathbf{x})$? What about $\delta^{(3)}(\mathbf{x})$? What is dimension of $\delta^3(\mathbf{x})$, if $\mathbf{x}$ is in meters? (Hint: for the point electrical charge distribution we have: $\rho(\mathbf{x}) = \sum_i Q_i \delta^3(\mathbf{x} - \mathbf{x}_i)$.)

One of the most useful properties of the delta function is the composition with a function (Property 5° above),

$$\delta(g(x)) = \sum_{\text{zeros } f(x_0)=0} \frac{\delta(x - x_0)}{g'(x)|_{x=x_0}}, \tag{3.11}$$

Later on in the course, this expression will be used, e.g., in deriving expressions for the differential cross-sections. See Mandl and Shaw Eqs. (8.11) and (8.14) used to derive (8.15). Compare the following expressions (touched upon in Problem 3),

- Schwartz, Problem (2.6) // cf. Mandl and Shaw Eq. (8.11)

$$\int \mathrm{d}k^0 \delta(k^2 - m^2)\theta(k^0) = \frac{1}{2\omega_{\mathbf{k}}}, \qquad \omega_{\mathbf{k}} = \sqrt{\mathbf{k}^2 + m^2} > 0, \tag{3.12}$$

- Peskin & Schroeder Eq. (2.40)

$$\int \frac{\mathrm{d}^3p}{(2\pi)^3} \frac{1}{E_{\mathbf{p}}} = \int \frac{\mathrm{d}^4p}{(2\pi)^4} 2\pi\delta(p^2 - m^2)\big|_{p^0>0}, \qquad E_{\mathbf{p}} = \sqrt{\mathbf{p}^2 + m^2} > 0, \tag{3.13}$$

- Srednicki, Eq. (3.16)

$$\int \mathrm{d}k^0 \delta(k^2 + m^2)\theta(k^0) = \frac{1}{2\omega}, \qquad \omega = \sqrt{\mathbf{k}^2 + m^2} > 0, \tag{3.14}$$

- Reg Fourier transform and Lorentz invariant form, Schwartz, Eq. (2.72) // cf. Sakurai

$$\langle \mathbf{p}|\mathbf{k}\rangle = 2\omega_{\mathbf{k}}(2\pi)^3\delta^3(\mathbf{p}-\mathbf{k}), \tag{3.15}$$

$$\mathbb{1} = \int \frac{d^3p}{(2\pi)^3}\frac{1}{2\omega_{\mathbf{p}}}|\mathbf{p}\rangle\langle\mathbf{p}| \tag{3.16}$$

At the end, a few questions for clarifying the delta function notation:

1. Is $(\delta(\mathbf{x}))^3$ valid?
2. What does mean $\delta^3(\mathbf{x})$?
3. What about $\delta^{(3)}(x)$?

Answers:

1. Multiplying distributions is invalid (in the sense we define them and use in physics, that is Schwartz theory of distributions. In fact, Laurent Schwartz proved this. Distributions may be multiplied by infinitely differentiable functions, but it is not possible to define a product of general distributions that extends the usual pointwise product of functions and has the same algebraic properties. There are recent attempts to reconcile this.) Thus, if you derive any expression containing powers of a distribution, your derivation somewhere went wrong!

2. Initially we have defined $\langle\delta,\varphi\rangle = \varphi(0)$ by acting in the space of test functions $\varphi(x)$ of one dimensional x. Since $\varphi(\mathbf{x}) = \varphi(x,y,z)$, where

$$|\mathbf{x}\rangle = |x\rangle\otimes|y\rangle\otimes|z\rangle\,, \tag{3.17}$$

$$|\varphi\rangle_{\mathbf{x}} = |\varphi\rangle_x\otimes|\varphi\rangle_y\otimes|\varphi\rangle_z \tag{3.18}$$

with, e.g., $|\varphi\rangle_y$ denotes $\varphi(x,y,z)$ when considered only as a function of y with x,z kept constant; consequently, we can define,

$$\delta^3(\mathbf{x}) = \left[\langle\delta|_x\otimes\langle\delta|_y\otimes\langle\delta|_z\right]\left[|\varphi\rangle_x\otimes|\varphi\rangle_y\otimes|\varphi\rangle_z\right] \tag{3.19}$$

$$= \langle\delta,\varphi\rangle_x\langle\delta,\varphi\rangle_y\langle\delta,\varphi\rangle_z = \delta(x)\delta(y)\delta(z). \tag{3.20}$$

In such case if dimension of $[x] = \mathsf{L}$ then $[\delta^3(\mathbf{x})] = \mathsf{L}^{-3}$.

3. $\delta^{(n)}(x)$ denotes a derivative. The derivative of a distribution is a well definable property (which can be *observed* by using integration by parts),

$$\langle\delta',\varphi\rangle = -\langle\delta,\varphi'\rangle\,. \tag{3.21}$$

Note that the Heaviside step function can be defined $\langle\theta,\varphi'\rangle = -\varphi(0)$, that is $\delta(x) = \theta'(x)$.

4 Solved Problems

Problem 1. (a) Show that for *any* Lorentz transformation $\Lambda^\mu{}_\nu$ it always hold,

$$\det(\Lambda^\mu{}_\nu) = \pm 1 \quad \text{and} \quad \left|\Lambda^0{}_0\right| \geq 1. \tag{4.1}$$

(b) (Schwartz, 2.4) Is the transformation $Y : (t, x, y, z) \to (t, x, -y, z)$ a Lorentz transformation? If so, why is it not considered with P and T as a discrete Lorentz transformation? If not, why not? Supplement: What about $Z : (t, x, y, z) \to (-t, -x, y, z)$?

Proposed solution:
(a) We start from the defining property of the Lorentz transformation,

$$\eta_{\mu\nu} = \Lambda_\mu{}^\rho \eta_{\rho\sigma} \Lambda^\sigma{}_\nu, \qquad \text{in matrix notation: } \eta = \Lambda^\mathrm{T} \eta \Lambda. \tag{4.2}$$

From this,

$$\det \eta = \det \Lambda^2 \det \eta \qquad \Longrightarrow \qquad \det \Lambda = \pm 1. \tag{4.3}$$

Next, from $\eta_{\mu\nu} = \Lambda_\mu{}^\rho \eta_{\rho\sigma} \Lambda^\sigma{}_\nu$, set $\mu = \nu = 0$, and concentrating on the component $\eta_{00} = 1$:

$$1 = \eta_{00} = \Lambda_0{}^\rho \eta_{\rho\sigma} \Lambda^\sigma{}_0 = \Lambda_0{}^0 \Lambda^0{}_0 - \Lambda_0{}^i \delta_{ij} \Lambda^j{}_0. \tag{4.4}$$

From this

$$\Lambda_0{}^0 \Lambda^0{}_0 = 1 + \Lambda_0{}^i \delta_{ij} \Lambda^j{}_0 \geq 1, \tag{4.5}$$

since δ_{ij} is positive definite (the term $\Lambda_0{}^i \delta_{ij} \Lambda^j{}_0$ is the norm of the vector $\Lambda_0{}^i$ squared: $|\Lambda_0{}^i|^2$). Note that any Lorentz transformation can be decomposed into a possible time reversal $t = \pm 1$, a boost B and any rotation R,

$$\begin{aligned}
\Lambda_0{}^0 \Lambda^0{}_0 &= 1 + \Lambda_0{}^i \delta_{ij} \Lambda^j{}_0, && (4.6)\\
\Lambda_0{}^0 \Lambda^0{}_0 = t^2 \left(B_0{}^0\right)^2 &= 1 + t^2 B_0{}^i R_i{}^k \delta_{kl} R^l{}_j B^j{}_0, && (4.7)\\
\left(\Lambda_0{}^0\right)^2 &= 1 + t^2 B_0{}^i \delta_{ij} B^j{}_0 = 1 + t^2 \left|B_0{}^i\right|^2 \geq 1. && (4.8)
\end{aligned}$$

(b) From $Y : (t, x, y, z) \to (t, x, -y, z)$, we first reconstruct the transformation in matrix notation,

$$Y = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & -1 & \\ & & & 1 \end{pmatrix}. \tag{4.9}$$

From here we can easily verify that it holds $\eta = Y^\mathrm{T} \eta Y$, hence Y is a Lorentz transformation. Since $Y_0{}^0 = 1$, the transformation is orthochronous and with no boost (observing $\Lambda^i{}_0 = 0$ for all i). On the other hand, from $\det Y = -1$, we conclude that the rotation is improper. It is a reflection with respect to yz plane. It is not considered as a seperate transformation, as it can be decomposed into space reflection P and rotation R,

$$Y = PR = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & 1 & \\ & & & -1 \end{pmatrix}, \tag{4.10}$$

$$R = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & 1 & \\ & & & -1 \end{pmatrix} = \begin{pmatrix} 1 & & & \\ & \cos\pi & 0 & \sin\pi \\ & 0 & 1 & 0 \\ & -\sin\pi & 0 & \cos\pi \end{pmatrix}. \tag{4.11}$$

Problem 2. (a) Show that for any $X_{\alpha\beta}$,

$$\frac{\partial}{\partial X_{\mu\nu}} X_{[\alpha\beta]} = \delta^{\mu}_{[\alpha}\delta^{\nu}_{\beta]}. \tag{4.12}$$

(b) Show that,

$$\frac{\partial}{\partial(\partial_\mu A_\nu)}\left(\frac{1}{4}F_{\rho\sigma}F^{\rho\sigma}\right) = F^{\mu\nu}, \tag{4.13}$$

where,

$$F_{\mu\nu} := \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{4.14}$$

Proposed solution:
(a) First we note,

$$\frac{\partial}{\partial X_{\mu\nu}} X_{\alpha\beta} = \delta^{\mu}_{\alpha}\delta^{\nu}_{\beta}, \tag{4.15}$$

hence,

$$\frac{\partial}{\partial X_{\mu\nu}} X_{[\alpha\beta]} = \frac{1}{2}\left(\frac{\partial}{\partial X_{\mu\nu}}X_{\alpha\beta} - \frac{\partial}{\partial X_{\mu\nu}}X_{\beta\alpha}\right) = \frac{1}{2}\left(\delta^{\mu}_{\alpha}\delta^{\nu}_{\beta} - \delta^{\mu}_{\beta}\delta^{\nu}_{\alpha}\right) = \delta^{\mu}_{[\alpha}\delta^{\nu}_{\beta]}. \tag{4.16}$$

(b) Observe that we cannot directly differentiate the expression,

$$F_{\rho\sigma}F^{\rho\sigma} = \left(\partial_\rho A_\sigma - \partial_\sigma A_\rho\right)\left(\partial^\rho A^\sigma - \partial^\sigma A^\rho\right), \tag{4.17}$$

with respect to $\partial_\rho A_\sigma$. We must first lower the indices of $F^{\rho\sigma}$.

$$\begin{aligned}
\frac{\partial}{\partial(\partial_\mu A_\nu)}\left(\frac{1}{4}F_{\rho\sigma}F^{\rho\sigma}\right) &= \frac{\partial}{\partial(\partial_\mu A_\nu)}\left(\frac{1}{4}\left(\partial_{\rho_1}A_{\sigma_1} - \partial_{\sigma_1}A_{\rho_1}\right)\left(\partial^{\rho_1}A^{\sigma_1} - \partial^{\sigma_1}A^{\rho_1}\right)\right) && (4.18)\\
&= \frac{1}{4}\frac{\partial}{\partial(\partial_\mu A_\nu)}\left(\left(\partial_{\rho_1}A_{\sigma_1} - \partial_{\sigma_1}A_{\rho_1}\right)\eta^{\rho_1\rho_2}\eta^{\sigma_1\sigma_2}\left(\partial_{\rho_2}A_{\sigma_2} - \partial_{\sigma_2}A_{\rho_2}\right)\right) && (4.19)\\
&= \frac{1}{4}\left(\delta^{\mu}_{\rho_1}\delta^{\nu}_{\sigma_1} - \delta^{\mu}_{\sigma_1}\delta^{\nu}_{\rho_1}\right)\eta^{\rho_1\rho_2}\eta^{\sigma_1\sigma_2}\left(\partial_{\rho_2}A_{\sigma_2} - \partial_{\sigma_2}A_{\rho_2}\right) + && (4.20)\\
&\quad + \frac{1}{4}\left(\partial_{\rho_1}A_{\sigma_1} - \partial_{\sigma_1}A_{\rho_1}\right)\eta^{\rho_1\rho_2}\eta^{\sigma_1\sigma_2}\left(\delta^{\mu}_{\rho_2}\delta^{\nu}_{\sigma_2} - \delta^{\mu}_{\sigma_2}\delta^{\nu}_{\rho_2}\right) && (4.21)\\
&= \frac{1}{2}\left(\delta^{\mu}_{\rho_1}\delta^{\nu}_{\sigma_1} - \delta^{\mu}_{\sigma_1}\delta^{\nu}_{\rho_1}\right)\eta^{\rho_1\rho_2}\eta^{\sigma_1\sigma_2}\left(\partial_{\rho_2}A_{\sigma_2} - \partial_{\sigma_2}A_{\rho_2}\right) && (4.22)\\
&= \frac{1}{2}\left(\delta^{\mu}_{\rho_1}\delta^{\nu}_{\sigma_1} - \delta^{\mu}_{\sigma_1}\delta^{\nu}_{\rho_1}\right)\left(\partial^{\rho_1}A^{\sigma_1} - \partial^{\sigma_1}A^{\rho_1}\right) && (4.23)\\
&= \frac{1}{2}\left(\left(\partial^\mu A^\nu - \partial^\nu A^\mu\right) - \left(\partial^\nu A^\mu - \partial^\mu A^\nu\right)\right) && (4.24)\\
&= \partial^\mu A^\nu - \partial^\nu A^\mu = F^{\mu\nu}. && (4.25)
\end{aligned}$$

This could be solved a bit faster using the result from (a), identifying $X_{\rho\sigma} = \partial_\rho A_\sigma$, we get,

$$\frac{\partial}{\partial(\partial_\mu A_\nu)}\left(\frac{1}{4}F_{\rho\sigma}F^{\rho\sigma}\right) = \frac{\partial}{\partial(\partial_\mu A_\nu)}\left(\frac{1}{4}\times 2\partial_{[\rho}A_{\sigma]}\,2\partial^{[\rho}A^{\sigma]}\right) = \frac{\partial}{\partial(\partial_\mu A_\nu)}\left(\partial_{[\rho}A_{\sigma]}\partial^{[\rho}A^{\sigma]}\right) \tag{4.26}$$

$$= 2\delta^\mu_{[\rho}\delta^\nu_{\sigma]}\partial^{[\rho}A^{\sigma]} = 2\delta^\mu_\rho\delta^\nu_\sigma\partial^{[\rho}A^{\sigma]} = 2\partial^{[\mu}A^{\nu]} = F^{\mu\nu}. \tag{4.27}$$

Problem 3. (Schwartz 2.6) Lorentz invariance.

(a) Show that

$$\int_{-\infty}^{\infty} \mathrm{d}k^0\,\delta\left(k^2 - m^2\right)\,\theta(k^0) = \frac{1}{2\omega_k}, \tag{4.28}$$

where $\theta(x)$is the unit step function and $\omega_k := \sqrt{\mathbf{k}^2 + m^2}$.

(b) Show that the integration measure d^4k is Lorentz invariant.

(c) Finally, show that

$$\int \frac{\mathrm{d}^3k}{2\omega_k}, \tag{4.29}$$

is Lorentz invariant.

Proposed solution:

(a) We shall use the property of the delta function in case of change variables,

$$\int \mathrm{d}x\,\delta(g(x))\,f(x) = \sum_{\xi\in g^{-1}(\{0\})}\frac{f(\xi)}{\left|\frac{\partial g(x)}{\partial x}\right|_{x=\xi}}, \quad \delta(g(x)) = \sum_{\xi\in g^{-1}(\{0\})}\frac{\delta(x-\xi)}{\left|\frac{\partial g(x)}{\partial x}\right|_{x=\xi}}, \tag{4.30}$$

where $\xi \in g^{-1}(\{0\})$ denotes denotes all zeros of $g(x)$, that is, all ξ satisfying the equation $g(\xi) = 0$. In this case we have $x = k^0$, and

$$g(k^0) = k^2 - m^2 \tag{4.31}$$

$$= k^\mu k_\mu - m^2 = k^0k_0 + k^ik_i - m^2 = (k^0)^2 - \mathbf{k}^2 - m^2 = (k^0)^2 - \omega_0^2 \tag{4.32}$$

$$= (k^0 - \omega_0)(k^0 + \omega_0) \tag{4.33}$$

where

$$\omega_0 := +\sqrt{\mathbf{k}^2 + m^2}. \tag{4.34}$$

Then, the equation $g(k^0) = 0$ gives two solutions,

$$k^0 \in \{-\omega_0, \omega_0\}. \tag{4.35}$$

On the other hand, we have the derivative of g wrt k^0,

$$\partial_{k^0} g(k^0) = \frac{\partial}{\partial k^0}\left[(k^0 - \omega_0)(k^0 + \omega_0)\right] = 2k^0, \tag{4.36}$$

so the expression becomes,

$$I = \int_{-\infty}^{\infty} \mathrm{d}k^0\,\delta\underbrace{\left(k^2 - m^2\right)}_{g(k^0)}\,\underbrace{\theta(k^0)}_{f(k^0)} \tag{4.37}$$

$$= \sum_{k^0\in g^{-1}(\{0\})}\frac{f(k^0)}{\left|\frac{\partial g(k^0)}{\partial k^0}\right|} \tag{4.38}$$

$$= \sum_{k^0 \in \{-\omega_0, \omega_0\}} \frac{\theta(k^0)}{2k^0} = \frac{1}{2\omega_0}. \tag{4.39}$$

(b) Using the change of variables, for the integration measure d^4k we have,

$$\mathrm{d}^4k' = \mathrm{d}k'^0 \mathrm{d}k'^1 \mathrm{d}k'^2 \mathrm{d}k'^3 = J\, \mathrm{d}k^0 \mathrm{d}k^1 \mathrm{d}k^2 \mathrm{d}k^3$$

where,

$$J = \det\left(\frac{\partial(x'^\mu)}{\partial(x^\nu)}\right) = \det(\Lambda^\mu{}_\nu), \tag{4.40}$$

is the Jacobian determinant of the transformation. In the case of restricted Lorentz transformations, $J = 1$, hence the integration measure is Lorentz invariant.

(c) Now simply,

$$\int \frac{\mathrm{d}^3k}{2\omega_k} = \int \mathrm{d}^3k \int_{-\infty}^{\infty} \mathrm{d}k^0\, \delta\left(k^2 - m^2\right)\, \theta(k^0) \tag{4.41}$$

$$= \int \mathrm{d}^4k\, \delta\left(k^2 - m^2\right)\, \theta(k^0) \tag{4.42}$$

$$= \int \mathrm{d}^4k\, \delta\left(k^\mu k_\mu - m^2\right)\, \theta(k^0). \tag{4.43}$$

Since d^4k and $k^\mu k_\mu - m^2$ are Lorentz invariant, and since k^0 never changes the sign under restricted Lorentz transformations, the whole expression is Lorentz invariant.

A Appendix: Earlier years notes

A.1 Special Relativity, Essentials

Further down in the course, we will extensively be working in a covariant formulation of QFT, that is, a formulation that is the same in all inertial frames. For this purpose, we will use the notation from special relativity (SR), and frequently use the 4-vector notation.

Of the special interest are the two kind of objects (in the introductory lectures, we had $x^0 = ct$, here $c = 1$), *vectors*, like the 4-position,

$$x^\alpha \to \left\{x^0, x^1, x^2, x^3\right\} = \{t, x, y, z\} \ ,$$

and *covectors*, like a 'differential' (a one-form) given by the components,

$$\partial_\alpha \to \{\partial_0, \partial_1, \partial_2, \partial_3\} = \left\{\frac{\partial}{\partial x^0}, \frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \frac{\partial}{\partial x^3}\right\} = \left\{\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right\} .$$

By definition, we have

$$\partial_\beta x^\alpha = x^\alpha{}_{,\beta} = \frac{\partial}{\partial x^\beta} x^\alpha = \delta_\beta{}^\alpha = \delta^\alpha{}_\beta = \delta^\alpha_\beta \, . \tag{A.1}$$

These two objects are of special interest since they transform in different ways under a general coordinate transformation, and especially Lorentz transformations. The x^α transform under general transform as $x^\alpha \to x^{\bar\alpha} = x^{\bar\alpha}\left(\{x^\alpha\}\right)$. For the increment this means

$$\mathrm{d}x^\alpha \ \to \ \mathrm{d}x^{\bar\alpha} = \frac{\partial x^{\bar\alpha}}{\partial x^\beta} \mathrm{d}x^\beta = M^{\bar\alpha}{}_\beta \mathrm{d}x^\beta \, , \tag{A.2}$$

and

$$\partial_\alpha \ \to \ \partial_{\bar\alpha} = \frac{\partial x^\beta}{\partial x^{\bar\alpha}} \partial_\beta = \left(M^{-1,\mathrm{T}}\right)_{\bar\alpha}{}^{\beta} \partial_\beta \, . \tag{A.3}$$

For now M is any transformation matrix $x^{\bar\alpha}{}_{,\beta} = {\partial x^{\bar\alpha}}/{\partial x^\beta} =: M^{\bar\alpha}{}_\beta$, where $x^{\bar\alpha}$ are the new coordinates and x^α are the old ones. (Sidenote: We also have $x^\alpha{}_{,\bar\beta} = {\partial x^\alpha}/{\partial x^{\bar\beta}} =: M^\alpha{}_{\bar\beta}$, which is the inverse of $x^{\bar\alpha}{}_{,\beta}$. The coordinate transformation is singular at the point if its Jacobian vanishes there.)

Notice that $\mathrm{d}x^\alpha$ and ∂_α transform in different ways. The first one (A.2) is called *contravariant* and the second one (A.3) is called *covariant*.

Sidenote: Are we talking here about the transformations of the coordinate basis $\{\partial_\alpha\}$ for the vectors and the coordinate basis $\{\mathrm{d}x^\alpha\}$ for the covectors, respectively? – that is,

$$\tilde\omega^\alpha = \tilde{\mathrm{d}}x^\alpha \ \to \ \tilde\omega^{\bar\alpha} = \tilde{\mathrm{d}}x^{\bar\alpha} = \frac{\partial x^{\bar\alpha}}{\partial x^\beta} \tilde{\mathrm{d}}x^\beta = M^{\bar\alpha}{}_\beta \tilde{\mathrm{d}}x^\beta \, ,$$

$$\vec e_\alpha = \vec\partial_\alpha \ \to \ \vec e_{\bar\alpha} = \vec\partial_{\bar\alpha} = \frac{\partial x^\beta}{\partial x^{\bar\alpha}} \vec\partial_\beta = \left(M^{-1,\mathrm{T}}\right)_{\bar\alpha}{}^{\beta} \vec\partial_\beta \, ,$$

In such case, there might be a confusion since $\mathrm{d}x^\alpha$ are basis one-forms having the covariant components $(\mathrm{d}x^\alpha)_\beta$, and similarly ∂_α are basis vectors having the contravariant components $(\partial_\alpha)^\beta$, transforming as:

$$(\tilde\omega^\alpha)_\beta = (\tilde{\mathrm{d}}x^\alpha)_\beta \ \to \ (\tilde\omega^\alpha)_{\bar\beta} = (\tilde{\mathrm{d}}x^\alpha)_{\bar\beta} = \frac{\partial x^\beta}{\partial x^{\bar\alpha}} (\tilde{\mathrm{d}}x^\alpha)_\beta = \left(M^{-1,\mathrm{T}}\right)_{\bar\alpha}{}^{\beta} (\tilde{\mathrm{d}}x^\alpha)_\beta \, ,$$

$$(\vec{e}_\alpha)^\beta = (\vec{\partial}_\alpha)^\beta \;\rightarrow\; (\vec{e}_\alpha)^{\bar{\beta}} = (\vec{\partial}_\alpha)^{\bar{\beta}} = \frac{\partial x^{\bar{\alpha}}}{\partial x^\beta}(\vec{\partial}_\alpha)^\beta = M^{\bar{\alpha}}{}_\beta (\vec{\partial}_\alpha)^\beta \,.$$

(end of sidenote)

Here we make use of the Einstein summation convention: any repeated index (one up, one down) is summed over:

$$A_\beta B^\beta = \sum_{\beta=0}^{3} A_\beta B^\beta = A_0 B^0 + A_1 B^1 + A_2 B^2 + A_3 B^3$$

The operation $A_\alpha B^\alpha$ is called a *contraction*. As long as A_α transforms as (A.3) and B^α transforms as (A.2) then $A_\beta B^\beta$ is always invariant.

This since (cf. Eq. (3.10) in Schutz)

$$A_{\bar{\alpha}} B^{\bar{\alpha}} = \left(\left(M^{-1,\mathrm{T}}\right)_{\bar{\alpha}}^{\;\;\mu} A_\mu\right)\left(M^{\bar{\alpha}}{}_\nu B^\nu\right) = \overbrace{\left(\left(M^{-1}\right)^{\mu}_{\;\;\bar{\alpha}} M^{\bar{\alpha}}{}_\nu\right)}^{\delta^\mu_\nu}(A_\mu B^\nu) = A_\mu B^\mu = A_\alpha B^\alpha \,.$$

Here we used that $M^\beta{}_\alpha = \left(M^{\mathrm{T}}\right)_\alpha^{\;\;\beta}$ is a transpose.[2]

Sidenote: Organizing components of tensors like matrices makes life complicated. The matrix transpose works properly only for (1,1)-type tensors (linear transformations). In general case, a transpose of a tensor is defined as any permutation of the indices, which is related to the ordering sequence of the vector (and dual vector) spaces in the tensor product, e.g. in case of $\mathbf{T} = T^{\alpha\beta}{}_\gamma{}^\delta\, \vec{e}_\alpha \otimes \vec{e}_\beta \otimes \tilde{\omega}^\gamma \otimes \vec{e}_\delta$, that is $\mathbf{T}(\cdot,\cdot,\cdot,\cdot) = T^{\alpha\beta}{}_\gamma{}^\delta\, \vec{e}_\alpha(\cdot) \otimes \vec{e}_\beta(\cdot) \otimes \tilde{\omega}^\gamma(\cdot) \otimes \vec{e}_\delta(\cdot)$, $\mathbf{T} : V^\star \times V^\star \times V \times V^\star \to \mathbb{R}$, $\mathbf{T} \in V \otimes V \otimes V^\star \otimes V$, we might have transposed 2nd and 3rd slots: $\mathbf{T}' = T^\alpha{}_\gamma{}^{\beta\delta}\, \vec{e}_\alpha \otimes \tilde{\omega}^\gamma \otimes \vec{e}_\beta \otimes \vec{e}_\delta$.

Note also that, in abstract notation, the contraction may be done on any two indices (slots) between two tensors that connect a vector space and its dual (not only as in the last one and the first one, like in the matrix product). In such case, when expressing tensors with the components, the only requirement is that two indices are one up and the other down, in whatever order. (end of sidenote)

The index in a *contraction* is called a *dummy index* and can always we renamed. The constructions $A_\alpha B_\alpha$ or $A^\alpha B^\alpha$ are ill defined and do not have any meaning, and thus forbidden. If $A = B$ we may use the shorthand notation $A_\alpha A^\alpha = A^2$. If $A \neq B$, and it does not warrant confusion, we may write $AB = A \cdot B := \mathbf{g}(A, B) = A_\alpha B^\alpha$.

What makes the Lorentz transformations special is that there is a relation between A_μ and A^μ in SR. The connection between then is given through the metric $g^{\mu\nu}$.

The road to raise of lower an index is through the metric $g^{\mu\nu}$ and its inverse $g_{\mu\nu}$. Using the matrix an upper index can be turned to a lower one, and vice versa

$$A_\alpha = g_{\alpha\beta} A^\beta \,,$$

and

$$B^\beta = g^{\beta\mu} B_\mu \,.$$

[2] The same thing written in matrix notation would have been: $B' = MB$ and $A' = M^{-1,\mathrm{T}} A$ such that $A'_\alpha B'^\alpha = A'^{\mathrm{T}} B' = \left(M^{-1,\mathrm{T}} A\right)^{\mathrm{T}} MB = A^T M^{-1} M B = AB = A_\alpha B^\alpha$.

It works on tensors to

$$T_{\alpha\beta} = g_{\beta\mu} T_{\alpha}{}^{\mu} \,,$$

and

$$T^{\eta\mu} = g^{\eta\alpha} T_{\alpha}{}^{\mu} \,.$$

The metric inverse $\mathbf{g}^{-1}$ can by definition be written as $\left(\mathbf{g^{-1}}\right)_{\mu\nu} = g^{\mu\nu}$ such that

$$g_{\mu\nu} g^{\nu\eta} = g_{\mu}{}^{\eta} = \delta_{\mu}{}^{\eta}.$$

It is this special metric that gives the Lorenz transformations their structure. Because $A_\alpha = g_{\alpha\beta} A^\beta$, if we want $A_{\bar{\alpha}} B^{\bar{\alpha}} = A_\alpha B^\alpha$ we must have (note that, if used, the metric should be also transformed)

$$\begin{aligned}
A_{\bar{\beta}} B^{\bar{\beta}} &= A^{\bar{\alpha}} g_{\bar{\alpha}\bar{\beta}} B^{\bar{\beta}} = \left(\Lambda^{\bar{\alpha}}{}_{\alpha} A^{\alpha}\right) \left(\Lambda_{\bar{\alpha}}{}^{\mu} g_{\mu\nu} \Lambda^{\nu}{}_{\bar{\beta}}\right) \left(\Lambda^{\bar{\beta}}{}_{\beta} B^{\beta}\right) \\
&= \left(\Lambda_{\bar{\alpha}}{}^{\mu} \Lambda^{\bar{\alpha}}{}_{\alpha}\right) \left(\Lambda^{\nu}{}_{\bar{\beta}} \Lambda^{\bar{\beta}}{}_{\beta}\right) \left(A^{\alpha} g_{\mu\nu} B^{\beta}\right) \\
&= \left(\delta^{\mu}_{\alpha}\right) \left(\delta^{\nu}_{\beta}\right) \left(A^{\alpha} g_{\mu\nu} B^{\beta}\right) \\
&= A^{\alpha} g_{\alpha\beta} B^{\beta} = A_{\beta} B^{\beta}.
\end{aligned}$$

This gives the extra condition $\left(\Lambda^{T}\right)_{\bar{\alpha}}^{\;\alpha} g_{\alpha\beta} \Lambda^{\beta}{}_{\bar{\beta}} = g_{\bar{\alpha}\bar{\beta}}$ that defines the Lorenz transformation.

The N-dimensional Euclidean case, where $g_{\beta\alpha} = 1$, the transforms are the $\mathrm{O}(N)$ rotations. The requirement on Λ can be rewritten as $\left(\Lambda^{\mathrm{T}}\right)_{\mu}^{\;\beta} \Lambda_{\beta}{}^{\gamma} = \delta_{\mu}{}^{\gamma}$ or $\Lambda^{\beta}{}_{\mu} \Lambda_{\beta}{}^{\gamma} = \delta_{\mu}{}^{\gamma}$ or $\Lambda^{\delta}{}_{\mu} = \left(\Lambda^{-1}\right)_{\mu}^{\;\delta}$. In SR the metric is the simple Minkowski metric (of flat spacetime)

$$(g_{\mu\nu}) = (\eta_{\mu\nu}) = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix}$$

Because of this simple structure of the Minkowski metric has the special property; $(\eta)_{\mu\nu} = \eta_{\mu\nu} = \eta^{\mu\nu} = (\eta^{-1})^{\mu\nu}$ (only component-wise, not as a tensor equation!). We can destinguish between Greek index $\alpha, \beta, \ldots \in \{0,1,2,3\}$ numbering the 4-vectors, and Latin indices $i, j, k, \dots \in \{1,2,3\}$ numbering the spatial components. Because of this, the raising and lowering of an index is associated with a minus sign $(-)$ if the indexes are Latin, and a plus sign $(+)$ is the index is 0 (n.b. signature). Specifically for x^α and ∂_α we have,

$$\begin{aligned}
x^{\alpha} &= (x^0, x^1, x^2, x^3)^{\mathrm{T}} = (t, \mathbf{x})^{\mathrm{T}} \\
x_{\alpha} &= (x_0, x_1, x_2, x_3) = (x^0, -x^1, -x^2, -x^3) = (t, -\mathbf{x})
\end{aligned}$$

and

$$\begin{aligned}
\partial_{\alpha} &= (\partial_0, \partial_1, \partial_2, \partial_3) = (\partial_t, \nabla) \\
\partial^{\alpha} &= (\partial^0, \partial^1, \partial^2, \partial^3)^{\mathrm{T}} = (\partial_0, -\partial_1, -\partial_2, -\partial_3)^{\mathrm{T}} = (\partial_t, -\nabla)^{\mathrm{T}}
\end{aligned}$$

From this we can construct two important **operators** $s^2 = x_\alpha x^\alpha$ and $\Box = \partial_\alpha \partial^\alpha$. These can explicitly be written as

$$s^2 = x^2 = x_\alpha x^\alpha = (t, -\mathbf{x}) \cdot (t, \mathbf{x})^{\mathrm{T}} = t^2 - |\mathbf{x}|^2$$

and

$$\Box = \partial^2 = \partial_\alpha \partial^\alpha = (\partial_t, \nabla) \cdot (\partial_t, -\nabla)^{\mathrm{T}} = \partial_t^2 - \nabla^2$$

The first is the invariant distance in SR. The second is the wave operator in field theory.

A.2 The Four-Momentum

A vector of extra importance is the 4-momentum p^α with components

$$\begin{aligned} p^\alpha &\rightarrow (p^0, p^1, p^2, p^3)^{\mathrm{T}} = \begin{pmatrix} E \\ \mathbf{p} \end{pmatrix} \\ p_\alpha &\rightarrow (p_0, p_1, p_2, p_3) = (E, -\mathbf{p}) \end{aligned}$$

because of the relativistic dispersion relation $E^2 = m^2c^4 + |\mathbf{p}|^2 c^2$ the norm of p is p

$$p^2 = E^2 - |\mathbf{p}|^2 = m^2$$

This relation always holds for particles that are measurable, also called *on the mass shell* or simply *on shell*.

For virtual particles this is usually not true, but we will return to this issue later in the course. Finally to tie things in with quantum mechanics we look why $\hat{p}^\alpha = \mathrm{i}\hbar\,\partial^\alpha$. From QM you remember that the time dependence of a plane wave propagating with momentum $\mathbf{p}$ and energy E is

$$\psi_{\mathbf{p}}(\mathbf{x}, t) = \mathrm{e}^{\mathrm{i}(\mathbf{p}\cdot\mathbf{x} - Et)\hbar} = \mathrm{e}^{-\mathrm{i}p_\alpha x^\alpha/\hbar} = \mathrm{e}^{-\mathrm{i}px/\hbar}.$$

Now acting with $\hat{p}^\alpha$ on $\psi_{\mathbf{p}}(\mathbf{x}, t)$ we get

$$\hat{p}^\alpha \psi_{\mathbf{p}}(\mathbf{x}, t) = (\mathrm{i}\hbar\,\partial^\alpha)\,\mathrm{e}^{-\mathrm{i}p_\beta x^\beta/\hbar} = \left(\partial^\alpha p_\beta x^\beta\right)\mathrm{e}^{-\mathrm{i}p_\alpha x^\alpha/\hbar}.$$

cf. Sakurai p. 50;

$\langle \mathbf{x}'|\alpha\rangle = \psi_\alpha(\mathbf{x}')$,$\langle \mathbf{p}'|\alpha\rangle = \phi_\alpha(\mathbf{p}')$, or with $\mathbf{k} := \mathbf{p}/\hbar$, $\langle \mathbf{k}'|\alpha\rangle = \phi_\alpha(\mathbf{k}')$,

$\langle \mathbf{x}'|\mathbf{p}'\rangle = \dfrac{1}{(2\pi\hbar)^{3/2}} \cdot \exp(\mathrm{i}\mathbf{k}' \cdot \mathbf{x}')$, $\langle \mathbf{k}'|\mathbf{x}'\rangle = \langle \mathbf{x}'|\mathbf{k}'\rangle^* = \dfrac{1}{(2\pi)^{3/2}} \cdot \exp(-\mathrm{i}\mathbf{k}' \cdot \mathbf{x}')$,

$\langle x'|\,\hat{p}\,|\alpha\rangle = -\mathrm{i}\hbar \dfrac{\partial}{\partial x'} \langle x'|\alpha\rangle$, $\langle \mathbf{x}'|\,\hat{\mathbf{p}}\,|\alpha\rangle = -\mathrm{i}\hbar\nabla\langle \mathbf{x}'|\alpha\rangle$.

The derivative is the eigenvalue of $\hat{p}^\alpha$ an is easily computed using (A.1) as

$$\partial^\alpha p_\beta x^\beta = \partial^\alpha p^\beta x_\beta = p^\beta \delta^\alpha{}_\beta = p^\alpha.$$

Thus $\hat{p}^\alpha \psi_{\mathbf{p}}(\mathbf{x}, t) = p^\alpha \psi(\mathbf{x}, t)$. It is also clear that $\psi_{\mathbf{p}}(\mathbf{x}, t)$ solves the equation $\hbar^2 \Box \psi = -m^2 \psi$ since

$$\begin{aligned} \hbar^2 \Box \psi_{\mathbf{p}} &= \hbar^2 \partial_\alpha \partial^\alpha \psi_{\mathbf{p}} = -\hat{p}_\alpha \hat{p}^\alpha \psi_{\mathbf{p}} = -p_\alpha p^\alpha \psi_{\mathbf{p}} \\ &= -m^2 \psi_{\mathbf{p}}. \end{aligned}$$