1 Special Theory of Relativity

Here we use the convention $\eta_{\mu\nu} = (-1, 1, 1, 1)$, but the other convention $\eta_{\mu\nu} = (1, -1, -1, -1)$ is also popular in literature.

Maxwell's equations says that the E&M phenomena are described by the speed c of light, and all the known waves always propagate through a medium. *Aether* was proposed to be such medium for light propagation, and the light propagates at c in *Aether*. However, according to the Galilean transformation, the speed of light will be different, if the laboratory or the source is moving. The Michelson-Morley experiment in 1895 was devised to detect the Earth motion relative to *Aether* by measuring the change in the speed of light.

In 1895, Lorentz, Fitzgerald, Poincaré explained the null results of the Michelson-Morley experiment by adopting two *ad hoc* hypotheses: contraction of rigid bodies and time dilation relative to *Aether*, and Lorentz provided a concise formula (or the *Lorentz transformation*) that captures it in a mathematical form. In fact, all the essence of special relativity is contained in the Lorentz transformation. Einstein in 1905 re-evaluated the idea of spacetime and re-derived the Lorentz transformation based on two principles, providing its correct physical interpretations and highlighting the limitation of classical mechanics in terms of simultaneity.

Einstein's special relativity is based on two principles:

1 laws of physics are the same in any inertial frames (same from old days),

2 the speed of light is the same to all observers in any inertial frames.

1.1 Lorentz Transformation and Physical Interpretations

A general coordinate transformation from x^μ to $\tilde{x}^\mu(x^\nu)$ can be expanded for an infinitesimal change:

$$\tilde{x}^\mu = x^\mu + \xi^\mu \,, \tag{1.1}$$

and one can integrate to obtain a finite coordinate transformation, where $\mu, \nu, \cdots$ represent the spacetime indices, while $i, j, k, \cdots$ represent the spatial indices. In classical Newtonian mechanics, Newton's law is invariant (or valid) only in the inertial frames, and the inertial frames are related in terms of Galilean transformation:

$$\tilde{x}^i = R^i{}_j x^j + v^i t + c^i \,, \qquad \tilde{t} = t + c^0 \,, \tag{1.2}$$

where R is a rotation matrix, and v^i, c^i, c^0 are constant. In short, the inertial frames are related to each other by a simple spatial and temporal translation c^i, c^0, a rotation R, or a relative motion with constant speed v^i. Note that in Galilean transformation there is no upper limit to the relative speed.

However, it is clear that the Galilean transformation does not preserve the speed of light, and hence the Maxwell's equations appear different in different inertial frames. The Lorentz transformation preserves the speed of light and the metric $\eta_{\mu\nu}$ in all inertial frames:

$$\tilde{x}^\mu = \Lambda^\mu{}_\nu x^\nu + c^\mu \,, \tag{1.3}$$

where the constant shift c^μ is a space-time translation. The Lorentz transformation is made of rotation and (Lorentz) boost, and the Poincaré transformation includes translation in addition to the Lorentz transformation. Given the full generality in Eq. (1.1), we consider only very specific transformations (or Lorentz transformations) in Minkowski spacetime. Furthermore, since the spacetime is intertwined, it is clear that we have to consider four (spacetime) vectors for physical quantities, rather than ordinary three (spatial) vectors.

1.1.1 Transformation Laws

The Lorentz transformation can be derived by finding a transformation that leave the Minkowski metric $\eta_{\mu\nu}$ and the spacetime interval ds^2 invariant:

$$\tilde{\eta}_{\rho\sigma} = \eta_{\mu\nu}\Lambda^\mu{}_\rho\Lambda^\nu{}_\sigma = \eta_{\rho\sigma} \,, \qquad ds^2 := \eta_{\mu\nu}dx^\mu dx^\nu = \tilde{\eta}_{\mu\nu}d\tilde{x}^\mu d\tilde{x}^\nu \,. \tag{1.4}$$

For a null path $ds = 0$, the speed of propagation is therefore always unity in all inertial frames.

The transformation law is determined by the (symmetric) Lorentz matrix $\mathbf{\Lambda}$ (its components are denoted as $\Lambda^\mu{}_\nu$). The matrix $\mathbf{\Lambda}$ describes the transformation of a four vector in the rest frame to a four vector in the frame moving with the velocity β^i relative to the rest frame:

$$\Lambda^0{}_0 = \gamma \; , \qquad \Lambda^0{}_i = \Lambda^i{}_0 = -\gamma\beta^i \; , \qquad \Lambda^i{}_j = \delta^i_j + \frac{\gamma^2}{1+\gamma}\beta^i\beta_j = \delta^i_j + (\gamma - 1)\frac{\beta^i\beta_j}{\beta^2} \; , \tag{1.5}$$

where we used the relation

$$\gamma := \left(1 - \beta^i\beta_i\right)^{-1/2} \; , \qquad \gamma^2\beta^2 = \gamma^2 - 1 \; . \tag{1.6}$$

Please see the lecture note for the derivation of the Lorentz transformation. The components of four vectors represent those measured in each rest frame, and the Lorentz transformation of the four vectors provides the relation between two frames. It is symmetric, in the sense that when the velocity vector β^i is replaced with $-\beta^i$, the same formula can be used for the inverse transformation.

For example, consider the relative velocity to be aligned along the z-direction, and the matrix becomes

$$\Lambda^\mu{}_\nu = \begin{pmatrix} \gamma & 0 & 0 & -\gamma\beta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\gamma\beta & 0 & 0 & \gamma \end{pmatrix} \; , \qquad \beta^i := \beta\, \delta^i_z \; , \tag{1.7}$$

and two coordinates of an event are related as[1]

$$\tilde{t} = \gamma(t - \beta z) \; , \qquad \tilde{z} = \gamma(z - \beta t) \; . \tag{1.9}$$

While in Newtonian mechanics two simultaneous events can be specified only in terms of its spatial separation, the simultaneity is now a relative concept, and the events are now characterized in terms of its spacetime interval:

$$s^2 := -\tilde{t}^2 + \tilde{z}^2 = -t^2 + z^2 \; , \tag{1.10}$$

invariant under Lorentz transformations. The Lorentz transformation can also be derived by starting with the invariance of the spacetime interval.

Now consider three inertial frames with two relative velocities β_{12} and β_{23}. The relation between one and three frames should be described by the Lorentz transformation, which leads to

$$\beta_{13} = \frac{\beta_{12} + \beta_{23}}{1 + \beta_{12}\beta_{23}} \; . \tag{1.11}$$

Note that this relation to the lowest order is just an addition of two velocities. Furthermore, the Lorentz transformation allows only three possibilities, strictly separated by the speed of light $\beta = 1$: sub-luminal, luminal or super-luminal in all inertial frames.

• ***Manipulations of the Lorentz transformation***.— The Minkowski metric is invariant under the Lorentz transformation (by construction):

$$\eta_{\mu\nu} = \eta_{\alpha\beta}\Lambda^\alpha{}_\mu\Lambda^\beta{}_\nu \; , \qquad \mathbf{\Lambda} \equiv \Lambda^\alpha{}_\beta \; , \qquad \boldsymbol{\eta} \equiv \eta_{\mu\nu} \qquad \mapsto \qquad \boldsymbol{\eta} = \mathbf{\Lambda}^t\boldsymbol{\eta}\mathbf{\Lambda} \; , \tag{1.12}$$

where the Lorentz matrix is symmetric. Using the transformation of the inverse Minkowski metric and the identity matrix

$$\eta^{\alpha\beta} = \Lambda^\alpha{}_\mu\Lambda^\beta{}_\nu\eta^{\mu\nu} \; , \qquad \delta^\alpha_\gamma \equiv \eta_{\gamma\beta}\eta^{\alpha\beta} = \Lambda^\alpha{}_\mu\left(\eta_{\gamma\beta}\Lambda^\beta{}_\nu\eta^{\nu\mu}\right) \equiv \Lambda^\alpha{}_\mu\left(\Lambda^{-1}\right)^\mu{}_\gamma \; , \tag{1.13}$$

[1] Since the un-tilde coordinate is moving with β in the opposite direction relative to the tilde coordinate, we can readily derive the inverse transformation relation:

$$t = \gamma(\tilde{t} + \beta\tilde{z}) \; , \qquad z = \gamma(\tilde{z} + \beta\tilde{t}) \; . \tag{1.8}$$

we derive the expression for the inverse transformation

$$\mathbf{\Lambda}^{-1} \equiv (\Lambda^{-1})^\mu{}_\nu = \boldsymbol{\eta}\mathbf{\Lambda}\boldsymbol{\eta}^{-1} . \tag{1.14}$$

In Minkowski space, the indices are raised and lowered by $\boldsymbol{\eta}$:

$$\boldsymbol{\eta}^{-1} := \eta^{\mu\nu} = \eta_{\mu\nu} = \boldsymbol{\eta} . \tag{1.15}$$

Furthermore, we can also define two variants of the Lorentz matrix, satisfying

$$\Lambda_{\sigma\mu} := \eta_{\sigma\nu}\Lambda^\nu{}_\mu , \qquad \Lambda_\sigma{}^\rho := \Lambda_{\sigma\mu}\eta^{\mu\rho} , \qquad \Lambda_\sigma{}^\rho\Lambda^\sigma{}_\nu \equiv \delta^\rho{}_\nu , \tag{1.16}$$

to find that

$$\therefore \quad (\Lambda^{-1})^\mu{}_\nu \equiv \Lambda_\nu{}^\mu . \tag{1.17}$$

Note that $\Lambda_{\mu\nu}$ is not symmetric, while $\Lambda^\mu{}_\nu$ is symmetric; the Lorentz matrix is not orthogonal or unitary:

$$\mathbf{\Lambda} = \mathbf{\Lambda}^t , \qquad \text{orthogonal} \ : \ \mathbf{\Lambda}^{-1} \neq \mathbf{\Lambda} , \qquad \text{unitary} \ : \ \mathbf{\Lambda}^{-1} \neq \mathbf{\Lambda}^\dagger . \tag{1.18}$$

Unitary matrices preserve the amplitude of complex vectors, and real unitary matrices are orthogonal.

Physical phenomena are described by tensors (vectors and scalars), and the relation between two frames with relative velocity is described by the Lorentz transformation $\mathbf{\Lambda}$:

$$\tilde p^\mu = \Lambda^\mu{}_\nu p^\nu . \tag{1.19}$$

A co-vector p_μ can be obtained by lowering the index with $\eta_{\mu\nu}$, and hence it transforms with the inverse Lorentz matrix:

$$p_\mu := \eta_{\mu\nu}p^\nu , \qquad \tilde p_\mu = \eta_{\mu\nu}\tilde p^\nu = \eta_{\mu\nu}\Lambda^\nu{}_\rho(\eta^{\rho\sigma}p_\sigma) = \left(\Lambda^{-1}\right)^\sigma{}_\mu p_\sigma = \Lambda_\mu{}^\sigma p_\sigma . \tag{1.20}$$

The notation convention is constructed to facilitate the manipulation. Note that only the four vectors and tensors transform with the Lorentz matrix; one *cannot* apply the Lorentz transformation to three spatial vectors.

1.1.2 Length Contraction and Time-Dilation

• ***Length contraction***.— Consider an observer at rest in a tilde coordinate, but moving relative to an un-tilde coordinate, and consider a small stick with length L_p, described by the observer as located in coordinates $\tilde z_1$ and $\tilde z_2$ with $L_p := \tilde z_2 - \tilde z_1$. The proper length L_p is the length measured by the observer in the rest frame with the stick. Another observer in a un-tilde coordinate would then see the stick as located in z_1 and z_2, with these coordinate points transforming according to Eq. (1.9) as

$$\tilde z_1 = \gamma(z_1 - \beta t_1) , \qquad \tilde z_2 = \gamma(z_2 - \beta t_2) . \tag{1.21}$$

At some time t, this observer measures the length L of the stick by measuring two end points z_1 and z_2 at the same time ($t = t_1 = t_2$):

$$\tilde L_p = \tilde z_2 - \tilde z_1 = \gamma(z_2 - z_1) = \gamma L , \qquad L = \frac{1}{\gamma}\, L_p \leq L_p . \tag{1.22}$$

The length of the stick is smaller than that in the rest frame. This is *not an illusory effect*, but a real effect due to different simultaneity.

• ***Time dilation***.— Consider instead a clock at rest in a tilde coordinate, and emits signals at $\tilde t_1$ and $\tilde t_2$ to indicate that the clock is ticking:

$$t_1 = \gamma(\tilde t_1 + \beta\tilde z_0) , \qquad t_2 = \gamma(\tilde t_2 + \beta\tilde z_0) . \tag{1.23}$$

The proper time interval $dt_p := \tilde t_2 - \tilde t_1$ is the time interval measured by the observer in the rest frame with the clock. The observer in a un-tilde coordinate would then see the clock tick slower as

$$dt = t_2 - t_1 = \gamma\, dt_p \geq dt_p . \tag{1.24}$$

Again, the time dilation is *not an illusory effect*. For example, muons decay with half-life $\tau \approx 2.2\ \mu$ sec. When accelerated to $\gamma \gg 1$ in particle accelerators, the change in their half-life can be measured. Also, muons are produced at the atmosphere from a collision with high-energy cosmic rays. With 2.2 μ sec, muons can only travel 0.6 km, even at the speed of light, but those muons created in the atmosphere hit the Earth surface all the time, due to time dilation.

1.1.3 Transformation of Velocities and Accelerations

Consider a particle moving with some velocity $v^i = dx^i/dt$, and consider two observers with relative velocity β, measuring the particle velocity in their own rest frames. Using Eq. (1.9), we can readily derive

$$\tilde{v}_z = \frac{\widetilde{dz}}{\widetilde{dt}} = \frac{v_z - \beta}{1 - v_z\beta} \,, \qquad \tilde{v}_x = \frac{\widetilde{dx}}{\widetilde{dt}} = \frac{v_x}{\gamma(1 - v_z\beta)} \,, \tag{1.25}$$

where the relation for v_y is similar. The velocity v_z along the direction of the moving observer is to the leading order just a Galilean transformation. The perpendicular direction is also affected due to the change in dt.

Instead of three velocity $v^i = dx^i/dt$, we consider a four velocity in terms of proper time

$$u^\mu := \frac{dx^\mu}{d\tau} = \left(\gamma_u,\ \gamma_u v^i\right) \,, \qquad -1 = \eta_{\mu\nu}u^\mu u^\nu = -\gamma_u^2 + \gamma_u^2 v^i v^i \,, \tag{1.26}$$

where γ for u^μ is one for the moving object with v^i. We can show that the four velocity transforms as a four-vector under the Lorentz transformation.

$$\tilde{u}^\mu = \Lambda^\mu{}_\nu u^\nu \,, \tag{1.27}$$

which implies

$$\tilde{\gamma}_u = \Lambda_0 \gamma_u \,, \qquad \tilde{v}^i = \Lambda_0^{-1}\left(\Lambda^i{}_0 + \Lambda^i{}_j v^j\right) \,, \tag{1.28}$$

where we defined

$$\Lambda_0 := \Lambda^0{}_0 + \Lambda^0_i v^i = \gamma\left(1 - \boldsymbol{\beta}\cdot\boldsymbol{v}\right) \,. \tag{1.29}$$

To put it in a vector form, we obtain the general expression

$$\tilde{\boldsymbol{v}} = \left(\frac{\boldsymbol{v}\cdot\hat{\boldsymbol{\beta}} - \beta}{1 - \boldsymbol{\beta}\cdot\boldsymbol{v}}\right)\hat{\boldsymbol{\beta}} + \frac{\boldsymbol{v} - (\boldsymbol{v}\cdot\hat{\boldsymbol{\beta}})\hat{\boldsymbol{\beta}}}{\gamma(1 - \boldsymbol{\beta}\cdot\boldsymbol{v})} \,, \qquad 1 - \tilde{v}^2 = \frac{(1 - v^2)^2}{(1 - \boldsymbol{\beta}\cdot\boldsymbol{v})^2} \,. \tag{1.30}$$

In the same way, we consider an acceleration $a^i = dv^i/dt$ of a particle, measured by two observers with relative velocity β. In the same way we derived the transformation of the velocity, we first compute

$$\widetilde{dv}_z = \frac{1}{\gamma^2}\frac{dv_z}{(1 - v_z\beta)^2} \,, \qquad \widetilde{dv}_x = \frac{dv_x}{\gamma(1 - v_z\beta)} + \frac{v_x\beta\, dv_z}{\gamma(1 - v_z\beta)^2} \,, \tag{1.31}$$

and then obtain the transformation of three-acceleration

$$\tilde{a}_z = \frac{a_z}{\gamma^3(1 - v_z\beta)^3} \,, \qquad \tilde{a}_x = \frac{a_x}{\gamma^2(1 - v_z\beta)^2} + \frac{v_x\beta a_z}{\gamma^2(1 - v_z\beta)^3} \,. \tag{1.32}$$

Similarly, the relative velocity between two observers affects not only the parallel component a_z, but also the perpendicular component a_x, a_y. One can also compute the four acceleration: $a^\mu := du^\mu/d\tau$.

In the Newtonian dynamics (or Galilean transformation), the position and the velocity are relative, while the time and the acceleration are absolute (or *identical*) in all inertial frames. In special relativity, the time (or simultaneity) is now relative, but the acceleration is still absolute in all inertial frames, while its value is *not* invariant for all inertial observers, i.e., non-vanishing accelerations are observed by all inertial observers, which is *not true* in general relativity.

1.1.4 Uniform Acceleration

Consider two observers with relative velocity (in fact infinitely many observers) and a moving object that is *uniformly* accelerating. Mind that the amount of acceleration depends on observers. As a special case, consider an observer moving with β, which is identical to the velocity of the moving object at a given moment (note β of the observer is constant, while v of the object is changing). At this instance, the object is at rest in the observer rest frame, and the observer measures the acceleration:

$$\beta = v \,, \qquad \therefore\ \tilde{a} = \gamma^3 a \,, \tag{1.33}$$

and this acceleration $\tilde{a}$ is constant. This uniform acceleration would correspond to a situation, in which a rocket is constantly accelerating in its rest frame by ejecting the same amount of rocket propellants. Note that $v_\perp = a_\perp = 0$ in this special case.

At a moment later, the rocket moves ahead of the observer, so we have to consider another observer, who is initially moving faster and ahead of the object. The object will soon catch up this observer, when their velocities are identical and this observer will again measure the acceleration. Note that with $\tilde{a}$ constant, the acceleration a in an un-tilde coordinate continuously decreases as the speed of the object increases, and the object will never reach the speed of light. Note that the acceleration is not bounded, while the speed is. The acceleration is a dimension of inverse length, such that when the acceleration is stronger (or the length of interest is larger than the scale of acceleration), (general) relativistic effects become more important (like curvature). For an earth gravity, the length scale of the acceleration is a light year. In comparison, a typical acceleration in nucleus is $a \sim v^2/r \sim 10^{28}$ earth gravity. Gravity is very weak on Earth!

Now we derive the trajectory $x(t)$ of the uniformly accelerating object. First, consider the identities

$$\frac{d\gamma}{dt} = \frac{d}{dt}\left(\frac{1}{\sqrt{1-\beta^2(t)}}\right) = \beta\gamma^3\frac{d\beta}{dt} , \qquad v = \beta , \tag{1.34}$$

hence we derive

$$\frac{d(\gamma\beta)}{dt} = \beta\frac{d\gamma}{dt} + \gamma\frac{d\beta}{dt} = \gamma\frac{d\beta}{dt}\left(\beta^2\gamma^2+1\right) = \gamma^3\frac{d\beta}{dt} = \tilde{a} . \tag{1.35}$$

Given the constant acceleration $\tilde{a}$ along the trajectory, we can integrate the above equation to obtain the solution for the velocity $\beta(t)$ as

$$\beta = \frac{\tilde{a}\,t}{\gamma(\beta)} , \qquad \therefore\ \beta(t) = \frac{\tilde{a}\,t}{\sqrt{1+\tilde{a}^2\,t^2}} \le 1 , \tag{1.36}$$

and one more integration of the solution yields

$$z(t) = \frac{\sqrt{1+\tilde{a}^2t^2}}{\tilde{a}} , \tag{1.37}$$

where the integral constant is set zero as a boundary condition. The trajectory of the uniformly accelerating object is then

$$z^2 - t^2 = \frac{1}{\tilde{a}^2} := L^2 . \tag{1.38}$$

The uniformly accelerating object follows a hyperbolic trajectory i.e., it starts at $z = \infty$ in the past $t = -\infty$, moving almost at the speed of light toward $z = 0$, but with constant acceleration $\tilde{a}$ toward positive direction (or slowing down); it reaches the closest point $z = L$ at $t = 0$ and the velocity at the moment is zero; it then turns around and moves to $z = \infty$ again with increasing velocity. The trajectories are bounded by two light cones that pass through $z = t = 0$, such that anything in the upper quadrant cannot be seen for those objects, acting as a horizon. Similarly, no accelerating observer cannot send signals to the lower quadrant. This trajectory of a uniformly accelerating object is used to describe Rindler coordinates.

1.1.5 Null Vectors: Photon Wavevectors

The photon propagation is described by a null vector k^μ, and it is fully characterized in terms of its frequency ν, direction, and polarization:

$$k^\mu = \omega\left(1, n^i\right) , \qquad 0 = k_\mu k^\mu , \qquad \omega := 2\pi\nu , \tag{1.39}$$

where n^i is a unit vector toward the propagation direction. Consider a light source in a tilde frame, moving with relative velocity with respect to an observer in an un-tilde frame. The Lorentz transformation $\tilde{k}^\mu = \Lambda^\mu{}_\nu k^\nu$ yields the physical relation of what the observer measures to the physical quantities in the source rest frame:

$$\tilde{\omega} = \Lambda_\omega\omega , \qquad \tilde{n}^i = \Lambda_\omega^{-1}\left(\Lambda^i{}_0 + \Lambda^i{}_j n^j\right) , \tag{1.40}$$

where we defined (which is the same as Λ_0)

$$\Lambda_\omega := \Lambda^0{}_0 + \Lambda^0{}_i n^i = \gamma\left(1 - \boldsymbol{\beta}\cdot\boldsymbol{n}\right) . \tag{1.41}$$

The ratio of angular frequencies is the relativistic Doppler effect:

$$\frac{\tilde{\omega}}{\omega} = \frac{\tilde{\nu}}{\nu} = \frac{\lambda}{\tilde{\lambda}} = \Lambda_\omega . \tag{1.42}$$

For the propagation direction, we can express the transformation relation in a vector form (again the same as $\tilde{\boldsymbol{v}}$)

$$\tilde{\boldsymbol{n}} = \left(\frac{\boldsymbol{n}\cdot\hat{\boldsymbol{\beta}} - \beta}{1 - \boldsymbol{n}\cdot\boldsymbol{\beta}}\right)\hat{\boldsymbol{\beta}} + \frac{\boldsymbol{n} - (\boldsymbol{n}\cdot\hat{\boldsymbol{\beta}})\hat{\boldsymbol{\beta}}}{\gamma(1 - \boldsymbol{n}\cdot\boldsymbol{\beta})} \,. \tag{1.43}$$

As an example, consider a source moving approaching the observer along z-direction with photons emitted with an angle θ. The source is at rest in a *tilde* coordinate, and for convenience the emission direction is $\tilde{\boldsymbol{n}} = (\sin\tilde{\theta}, 0, \cos\tilde{\theta})$ confined in x-z plane ($\tilde{\phi} \equiv 0$). The observer is at rest in an *un-tilde* coordinate. Using the formulas above, we derive

$$\Lambda_\omega = \gamma(1 - \beta\cos\theta) \,, \qquad \cos\tilde{\theta} = \frac{\cos\theta - \beta}{1 - \beta\cos\theta} \,, \qquad \sin\tilde{\theta} = \frac{\sin\theta}{\gamma(1 - \beta\cos\theta)} \,, \tag{1.44}$$

or in terms of observables in un-tilde coordinate

$$\omega = \tilde{\omega}\gamma(1 + \beta\cos\tilde{\theta}) \,, \qquad \cos\theta = \frac{\cos\tilde{\theta} + \beta}{1 + \beta\cos\tilde{\theta}} \,, \qquad \sin\theta = \frac{\sin\tilde{\theta}}{\gamma(1 + \beta\cos\tilde{\theta})} \,. \tag{1.45}$$

Given that in a un-tilde coordinate the observer is at rest and the source is moving with $\boldsymbol{\beta}$ (in this case z-direction), the source can be approaching the observer or moving away from the observer, depending on its location relative to the observer.

If the source is approaching the observer and photons are emitted at $\tilde{\theta} = 0$ parallel to the moving direction, the ratio of the observed frequency ω to the one emitted in the source rest frame is blueshifted

$$\frac{\omega}{\tilde{\omega}} = \gamma(1 + \beta) = \sqrt{\frac{1 + \beta}{1 - \beta}} \approx 1 + \beta + \mathcal{O}(\beta^2) \,, \qquad \theta = 0 \,, \tag{1.46}$$

which reduces to the classical Doppler effect in the non-relativistic limit. For photons emitted at $\tilde{\theta} = \pi/2$ perpendicular to the moving direction, we derive

$$\frac{\omega}{\tilde{\omega}} = \gamma \approx 1 + \frac{1}{2}\beta^2 + \mathcal{O}(\beta^3) \,, \qquad \cos\theta = \beta \,, \qquad \sin\theta = \frac{1}{\gamma} \,. \tag{1.47}$$

In case that the source is moving relativistically $\gamma \gg 1$, the observed propagation angle becomes super narrow

$$\sin\theta \approx \theta \approx \frac{1}{\gamma} \ll 1 \,, \tag{1.48}$$

which is known as the *relativistic beaming effect*. Beyond this angle $\theta \ll 1$ the moving object is hardly visible, while the frequency is highly boosted within the angle (or along the motion).

If the source is moving away from the observer, the photons emitted along $\tilde{\theta} = \pi$ can be received by the observer, and their frequencies are redshifted:

$$\frac{\omega}{\tilde{\omega}} = \gamma(1 - \beta) = \sqrt{\frac{1 - \beta}{1 + \beta}} \approx 1 - \beta + \mathcal{O}(\beta^2) \,, \qquad \theta = \pi \,. \tag{1.49}$$

Now consider a source moving perpendicular to the light propagation direction

$$\frac{\omega}{\tilde{\omega}} = \gamma = \frac{1}{\sqrt{1 - \beta^2}} \approx 1 + \mathcal{O}(\beta^2) \,, \tag{1.50}$$

where it is noted that the formula for Λ_ω depends only on $\boldsymbol{\beta}\cdot\tilde{\boldsymbol{n}}$. Mind that the photons are received by the observer only in the case the source is approaching the observer. The observed frequency is again blueshifted, and this purely relativistic effect is called the transverse Doppler effect.

Further details in astrophysical applications of the relativistic effects can be found in Rybicki and Lightman (1979).

• ***Perpendicular components***.— Now we define a projection tensor $P^{ij} = \delta^{ij} - n^i n^j$ with n^i and consider how the perpendicular components of a vector A^μ orthogonal to a null vector k^μ transform. Given the constraint $0 = A_\mu k^\mu$, we can write

$$A^\mu =: \left(-A_\parallel, A^i_\perp + n^i A_\parallel\right) , \qquad A^i_\perp = P^{ij}A^j , \qquad A^i_\perp n^i = 0 , \tag{1.51}$$

and in the tilde coordinate the expression takes the same form. From the transformation law

$$\eta_{\mu\nu} = \eta_{\rho\sigma}\Lambda^\rho{}_\mu\Lambda^\sigma{}_\nu = -\Lambda^0{}_\mu\Lambda^0{}_\nu + \Lambda^i{}_\mu\Lambda^i{}_\nu , \tag{1.52}$$

we can derive useful relations

$$\Lambda^i{}_0\Lambda^i{}_0 = -1 + \left(\Lambda^0{}_0\right)^2 = -1 + \gamma^2\beta^2 , \qquad \Lambda^j{}_0\Lambda^j{}_i = \Lambda^0{}_0\Lambda^0{}_i = -\gamma^2\beta_i , \tag{1.53}$$

$$\Lambda^k{}_i\Lambda^k{}_j = \delta_{ij} + \Lambda^0{}_i\Lambda^0{}_j = \delta_{ij} + \gamma^2\beta_i\beta_j . \tag{1.54}$$

Under the transformation, the perpendicular components transform as

$$\tilde{A}^i_\perp = (\Lambda_\perp)^i{}_j A^j_\perp , \tag{1.55}$$

with

$$(\Lambda_\perp)^i{}_j := \Lambda^i{}_j - \frac{\Lambda^i{}_0 - \Lambda^i{}_k n^k}{\Lambda^0{}_0 - \Lambda^0{}_l n^l}\,\Lambda^0{}_j = \Lambda^i{}_j - \Lambda_\parallel^{-1}\left(\Lambda^i{}_0 - \Lambda^i{}_k n^k\right)\Lambda^0{}_j , \tag{1.56}$$

and we can show that

$$\tilde{X}^i_\perp\tilde{Y}^i_\perp = (\Lambda_\perp)^i{}_k X^k_\perp(\Lambda_\perp)^i{}_l Y^l_\perp = X^i_\perp Y^i_\perp , \tag{1.57}$$

where we need to replace all the summed Λ's, e.g., $\Lambda^i{}_\mu\Lambda^i{}_\nu$ by using the above relations and use the orthogonality $0 = X^i_\perp n_i = Y^i_\perp n_i$. The derivation is relatively straightforward. The inner product of $X_\perp$ is invariant under the transformation.

1.2 Some Paradoxes in Relativity

1.2.1 Twin Paradox

A twin sets a journey to a remote star by accelerating the rocket and maintaining a constant speed for most of the journey and decelerating just before arrival. Once arrives at the star, the twin reverses the engine and comes back to the other twin at Earth, finding that the person at Earth has already died old.

It is called the paradox, because one can switch the argument by saying things are relative, such that one at Earth is moving relative to the one on the rocket. However, the outcome is not symmetric. This is the case due to the acceleration, no matter how short the period of acceleration can be made.

Check the diagram for the simultaneity in § 3.9 D'Inverno (1992) or p.88 in C. Will 2005.

1.2.2 Causality & Tachyons

In case one can move faster than the speed of light, how is the causality violated? Particles moving along a spacelike geodesic are called *Tachyons*. Their existence violates the causality. But note that even though one can move faster than the speed of light, nothing strange happens in its rest frame. However, to other time-like observers, the sequence of events is not preserved.

For example, consider Alice at A and Bob at B at rest. They send a signal via Tachyon, i.e., faster than the speed of light, so that it is almost simultaneous transition from Alice to Bob (same time coordinate). Now consider Charlie and Dave moving relative to Alice and Bob at a constant speed. In their rest frame of Charlie and Dave, they can also communicate via Tachyon, which travels almost simultaneously in their rest frame. Note that their simultaneous surface is tilted relative to Alice and Bob's, and the simultaneous transition between Charlie and Dave is a transition backward in time in Alice and Bob's. In the same way, Alice and Bob's simultaneous transition is a transition backward in time in Charlie and Dave's.

According to special relativity, tachyons with $v > 1$ have imaginary energy, which is not defined, so that one consider instead an imaginary mass with real energy. As it gains more energy, the speed is reduced asymptotically to the speed of light, while at its low energy the speed approaches to infinity. It was shown that tachyons are spin zero, but follow Fermi-Dirac statistics (hence many issues). Arnold Sommerfeld was the first to conceive tachyons, but its name was given by someone else later.

1.2.3 Barn-Pole Paradox & Bell's Spaceship Paradox

• ***Barn-Pole paradox***.— Consider a barn (or a garage), whose size is 10m long on side, and a 20m-long pole. The pole cannot fit in the barn. However, a fast runner runs with the pole at $\gamma = 2$, such that from the barn's rest frame, the pole is length contracted to 10m to be fit in the barn, while from the runner's rest frame, the barn is length contracted to 5m too small to be fit. This paradox can be made more interesting, by closing the doors in the barn, when the pole fits the barn.

• ***Bell's spaceship paradox***.— Originally by Dewan & Beran, and later modified by Bell: In the inertial frame S, there are two spaceships with separation d at rest, but they are connected by a string. They both accelerate equally at the same time to some speed in S, hence they maintain the same speed and the same separation at each moment with respect to the frame S. Therefore, the string should remain unaffected, which is the ordinary point of view. However, due to length contraction in special relativity, all lengths contract equally, i.e., two spaceships are smaller, and the string is smaller. That is the paradox. In this lab frame, the string is torn apart due to length contraction. In the rest frame of first spaceship with the final speed, two spaceships did not accelerate at the same time. The first one accelerates first, and then the second one later. Hence the distance between two spaceships increased, breaking apart the string. In this frame, the string is broken by external forces.

1.3 Relativistic Mechanics

1.3.1 Relativistic Mass and Energy

As the length and time measurements depend on observers, the mass measurements should also depend on observers. Consider two identical particles of mass m_0 at rest. One particle is moving with β and of relativistic mass $m(\beta)$, and the other particle is at rest. After collision, two particles are stuck, moving at u with mass $M(u)$. Assuming the conservation of the relativistic mass and the momentum, we demand the relativistic mass terms satisfy

$$m(\beta) + m_0 = M(u) \,, \qquad\qquad m(\beta)\beta + 0 = M(u)u \,. \tag{1.58}$$

Removing $M(u)$, we obtain

$$m(\beta) = m_0 \left(\frac{u}{\beta - u}\right) \,. \tag{1.59}$$

Now consider a center of mass frame, where the combined particles are at rest after collision. This frame has the relative velocity u, compared to the lab frame. The first particle in the center of mass frame should also move with velocity u, because the second particle that is at rest in the lab frame moves with velocity u in the opposite direction. Using Eq. (1.11), we derive

$$\beta = \frac{2u}{1+u^2} \,, \qquad\qquad \therefore \quad u = \frac{1}{\beta}\left(1 - \sqrt{1-\beta^2}\right) = \frac{\gamma - 1}{\beta\gamma} \,, \tag{1.60}$$

and the relativistic mass is

$$m(\beta) = \gamma m_0 \,. \tag{1.61}$$

Observing that

$$E := m(\beta)c^2 \approx m_0c^2 + \frac{1}{2}m_0(\beta c)^2 + \mathcal{O}(\beta^4) \,, \tag{1.62}$$

the conservation of the relativistic mass includes the classical energy conservation, and the total energy of the particle is $E = \gamma m_0 c^2$. The relativistic mass and momentum conservation is recast as the conservation of four momentum.

The photon is always moving at the speed of light. For any energy E for photons, the rest mass is zero:

$$m_0 = \frac{E}{\gamma c^2} = 0 \,. \tag{1.63}$$

1.3.2 Particle Dynamics

We generalize the Newtonian dynamics in special relativity. The energy-momentum four vector p^μ and four velocity u^μ are defined as

$$p^\mu := m\frac{dx^\mu}{d\tau} = mu^\mu \,, \qquad\qquad u^\mu := \frac{dx^\mu}{d\tau} \,, \tag{1.64}$$

where m hereafter represents the rest mass (without subscript 0) and $d\tau$ is the proper time interval for the particle:

$$d\tau = \sqrt{dt^2 - dx^2} = \frac{1}{\gamma}\, dt \,, \qquad \beta^i = \frac{dx^i}{dt} \,. \tag{1.65}$$

The particle moves on a time-like path, fixing the normalization:

$$-d\tau^2 = \eta_{\mu\nu} dx^\mu dx^\nu \,, \qquad -1 \equiv \eta_{\mu\nu} u^\mu u^\nu \,, \qquad u^0 \equiv \gamma \,, \tag{1.66}$$

which yields

$$\frac{p^i}{E} = \beta^i \,, \qquad p^0 = \gamma\, m = E \,, \qquad -m^2 = \eta_{\mu\nu} p^\mu p^\nu \,. \tag{1.67}$$

Hence, the four-momentum conservation in relativity combines the energy and the momentum conservation in Newtonian mechanics.

The Newton's force-law should then be

$$m \frac{d^2 x^\mu}{d\tau^2} = f^\mu \,, \tag{1.68}$$

which becomes in the rest frame

$$m \frac{d^2 \xi^\mu}{dt^2} = f^\mu_{\text{rest}} \,, \qquad f^0_{\text{rest}} = 0 \,, \qquad f^i_{\text{rest}} \,, \tag{1.69}$$

with $\vec{f}_{\text{rest}}$ being the ordinary forces (three vector) in Newtonian dynamics, where we use ξ^μ to represent a coordinate where the particle is at rest. Given a force in the rest frame such as E&M force, the four force f^μ can then be obtained by Lorentz transformation:

$$f^\mu = \Lambda^\mu{}_\nu f^\nu_{\text{rest}} \,, \qquad f^i = f^i_{\text{rest}} + (\gamma - 1)\beta^i \,\frac{\beta \cdot f_{\text{rest}}}{\beta^2} \,, \qquad f^0 = \gamma\, \beta \cdot f_{\text{rest}} = \beta^i f^i \,. \tag{1.70}$$

1.3.3 Energy-Momentum Tensor

Consider a system of point particles, and we define the energy-momentum tensor as the usual macroscopic way as the energy density $T^{00} \propto \rho$, the energy flux $T^{0i} \propto p^i$ (or momentum density), and the pressure $T^{ij} \propto p^i p^j$. Note that the energy density and pressure depend on the relative motion between the observer and the fluid. For a macroscopic fluid or perfect fluid, the energy-momentum tensor in the rest frame (tilde coordinate) is

$$\tilde{T}^{00} = \hat{\rho} \,, \qquad \tilde{T}^{0i} = 0 \,, \qquad \tilde{T}^{ij} = \hat{p}\, \delta^{ij} \,, \tag{1.71}$$

where we put hat to indicate that those quantities are measured in the rest frame of the fluid. In a lab frame (or untilde coordinate), the fluid is moving with relative velocity v^i, and the Lorentz transform to a general (untilde) coordinate is

$$\Lambda^0{}_0 = \gamma \,, \qquad \Lambda^i{}_0 = \gamma v^i \,, \qquad \Lambda^i{}_j = \delta^i_j + v^i v^j \,\frac{\gamma - 1}{v^2} \,, \tag{1.72}$$

(mind the sign for v^i). This yields the energy-momentum tensor

$$T^{\mu\nu} = \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma \tilde{T}^{\rho\sigma} = \begin{pmatrix} \gamma^2(\hat{\rho} + \hat{p}v^2) & \gamma^2(\hat{\rho} + \hat{p})v^i \\ \gamma^2(\hat{\rho} + \hat{p})v^i & \hat{p}\,\delta_{ij} + (\hat{\rho} + \hat{p})\gamma^2 v^i v^j \end{pmatrix} = \hat{\rho} u^\mu u^\nu + \hat{p}\mathcal{H}^{\mu\nu} \,, \tag{1.73}$$

where the projection tensor is

$$\mathcal{H}_{\mu\nu} := \eta_{\mu\nu} + u_\mu u_\nu \,, \qquad \mathcal{H}_{\mu\nu} u^\nu = 0 \,. \tag{1.74}$$

Given $T^{\mu\nu}$, we can readily derive

$$\begin{aligned} &T_{\mu\nu} = \eta_{\mu\rho}\eta_{\nu\sigma} T^{\rho\sigma} = \eta_{\mu\mu}\eta_{\nu\nu} T^{\mu\nu} \,, \qquad T_{00} = T^{00} \,, \qquad T_{0i} = -T^{0i} \,, \qquad T_{ij} = T^{ij} \,, \\ &T^\mu{}_\nu = T^{\mu\rho}\eta_{\rho\nu} = T^{\mu\nu}\eta_{\nu\nu} \,, \qquad T^0{}_0 = -T^{00} \,, \qquad T^0{}_i = T^{0i} = -T^i{}_0 \,, \qquad T^i{}_j = T^{ij} \,, \end{aligned} \tag{1.75}$$

where no summation in the second equality over μ, ν. Mind that $T^\mu{}_\nu$ is not symmetric over indices.

Using the energy-momentum tensor, we can define the angular momentum tensor $J^{\mu\nu}$ and the spin four-vector S_μ:

$$M^{\rho\mu\nu} := x^\mu T^{\nu\rho} - x^\nu T^{\mu\rho} \,, \qquad J^{\mu\nu} := \int d^3\mathbf{x}\, M^{0\mu\nu} = -J^{\nu\mu} \,, \qquad S_\mu := \frac{1}{2}\, \varepsilon_{\mu\nu\rho\sigma}\, J^{\nu\rho} u^\sigma \,, \tag{1.76}$$

and the conservation equation guarantees that the angular momentum tensor is constant in time:

$$\partial_\rho M^{\rho\mu\nu} = \partial_\nu T^{\mu\nu} = 0 \,. \tag{1.77}$$

The spin vector is the intrinsic angular momentum that does not vanish in the rest frame:

$$\text{rest} - \text{frame}: \quad S_1 = J^{23} \,, \qquad S_2 = J^{31} \,, \qquad S_3 = J^{12} \,, \qquad S_0 = 0 \,. \tag{1.78}$$

1.4 Electromagnetism

Here we provide a concise review of E&M in view of special relativity.

1.4.1 E&M Basics

The physical laws governing the electromagnetism are

$$\text{Coulomb} \;:\; F_1 =: k_1 \frac{qq'}{r^2} \,, \qquad E := k_1 \frac{q}{r^2} \,, \tag{1.79}$$

$$\text{Ampere} \;:\; \frac{dF_2}{dl} =: 2k_2 \frac{II'}{r} \,, \qquad I := \frac{dq}{dt} \,, \qquad \frac{k_1}{k_2} = c^2 \,, \tag{1.80}$$

$$\text{Faraday} \;:\; \nabla \times E + k_3 \frac{\partial B}{\partial t} = 0 \,, \qquad B := \frac{2k_2}{k_3} \frac{I}{r} \,, \tag{1.81}$$

where three dimensionful proportionality constants k_1, k_2, k_3 are left unspecified, q, q' are electric charges, and I, I' are electric currents ($I = dq/dt$). The standard units are

$$\text{cgs units}: \quad k_1 = 1 \,, \qquad k_2 = \frac{1}{c^2} \,, \qquad k_3 = \frac{1}{c} \,, \qquad q = 4.803 \times 10^{-10}\ \text{esu} \,, \tag{1.82}$$

$$\text{MKS units}: \quad k_1 = \frac{1}{4\pi\epsilon_0} = 10^{-7} c^2 \,, \qquad k_2 = \frac{\mu_0}{4\pi} = 10^{-7} \,, \qquad k_3 = 1 \,, \qquad q = 1\ \text{coulomb} \,,$$

where cgs units are sometimes called Gaussian cgs units. The Maxwell's equations are

$$\nabla \cdot \mathbf{E} = 4\pi k_1 \rho \,, \qquad \nabla \cdot \mathbf{B} = 0 \,, \qquad \nabla \times \mathbf{B} - \frac{1}{k_3 c^2} \frac{\partial \mathbf{E}}{\partial t} = \frac{4\pi k_2}{k_3} \mathbf{J} \,, \qquad \nabla \times \mathbf{E} + k_3 \frac{\partial \mathbf{B}}{\partial t} = 0 \,. \tag{1.83}$$

In this note, we use the cgs units, and the Maxwell's equations in the cgs units become dynamical equations with source

$$\nabla \cdot \mathbf{E} = 4\pi\rho \,, \qquad \nabla \times \mathbf{B} - \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} = \frac{4\pi}{c} \mathbf{J} \,, \tag{1.84}$$

and the source-free equations

$$\nabla \cdot \mathbf{B} = 0 \,, \qquad \nabla \times \mathbf{E} + \frac{1}{c}\frac{\partial \mathbf{B}}{\partial t} = 0 \,, \tag{1.85}$$

in conjunction to the continuity equation and the Lorentz force

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0 \,, \qquad \mathbf{F} = q\left(\mathbf{E} + \frac{\mathbf{v}}{c} \times \mathbf{B}\right) \,. \tag{1.86}$$

The macroscopic field variables are

$$\mathbf{D} = \mathbf{E} + 4\pi\mathbf{P} \equiv \epsilon\mathbf{E} \,, \qquad \mathbf{H} = \mathbf{B} - 4\pi\mathbf{M} \,. \tag{1.87}$$

1.4.2 Field Strength Tensor

Here we work with $(-,+,+,+)$, but compare to $(+,-,-,-)$ convention with tilde. We consider a four current and a four vector:

$$J^\mu := (c\rho, \mathbf{J}) \equiv \tilde{J}^\mu \,, \qquad A^\mu := (\phi, \mathbf{A}) \equiv \tilde{A}^\mu \,, \qquad A_\mu = (-\phi, \mathbf{A}) \equiv -\tilde{A}_\mu \,, \tag{1.88}$$

and they are related as

$$0 = \partial_\mu J^\mu = \frac{\partial\rho}{\partial t} + \nabla\cdot\mathbf{J} \,, \qquad \mathbf{E} =: -\frac{1}{c}\frac{\partial\mathbf{A}}{\partial t} - \nabla\phi \,, \qquad \mathbf{B} =: \nabla\times\mathbf{A} \,, \qquad Q := \int d^3x\, J^0 \,. \tag{1.89}$$

The field strength tensor is then defined as

$$F_{\mu\nu} := 2A_{[\nu,\mu]} = \partial_\mu A_\nu - \partial_\nu A_\mu \,, \qquad \tilde{F}_{\mu\nu} := 2\tilde{A}_{[\nu,\mu]} \equiv -F_{\mu\nu} \,, \qquad F^{\mu\nu} = -\tilde{F}^{\mu\nu} \,. \tag{1.90}$$

With $\eta_{\mu\nu}$, mind that

$$F^{0i} = -F_{0i} = \tilde{F}_{0i} = -\tilde{F}^{0i} \,, \qquad F^{ij} = F_{ij} = -\tilde{F}_{ij} = -\tilde{F}^{ij} \,. \tag{1.91}$$

Using the above relations of A^μ to the electric and the magnetic fields, we derive

$$\mathbf{E} = -F_{0i} = \tilde{F}_{0i} \,, \qquad \mathbf{B} = \frac{1}{2}\epsilon_{ijk}F^{jk} = -\frac{1}{2}\epsilon_{ijk}\tilde{F}^{jk} \,, \tag{1.92}$$

where for $\mathbf{E}$ and $\mathbf{B}$ we do not distinguish their components with upper index or lower index, as it is defined in the Euclidean space, and note that the Levi-Civita symbols in 3D satisfy

$$\epsilon_{ijk} = \tilde{\epsilon}_{ijk} \,, \qquad \epsilon^{ijk} = \eta^{ii}\eta^{jj}\eta^{kk}\epsilon_{ijk} = -\tilde{\epsilon}^{ijk} \,, \qquad 1 = \epsilon_{123} = \epsilon^{123} = \tilde{\epsilon}_{123} \,, \qquad \tilde{\epsilon}^{123} = -1 \,. \tag{1.93}$$

In the matrix notation, we derive

$$F^{\mu\nu} = \begin{pmatrix} 0 & E_x/c & E_y/c & E_z/c \\ -E_x/c & 0 & B_z & -B_y \\ -E_y/c & -B_z & 0 & B_x \\ -E_z/c & B_y & -B_x & 0 \end{pmatrix} = -\tilde{F}^{\mu\nu} \,, \qquad F_{\mu\nu} = \begin{pmatrix} 0 & -E_x/c & -E_y/c & -E_z/c \\ E_x/c & 0 & B_z & -B_y \\ E_y/c & -B_z & 0 & B_x \\ E_z/c & B_y & -B_x & 0 \end{pmatrix} = -\tilde{F}_{\mu\nu} \,. \tag{1.94}$$

It is convenient to define the dual tensor

$$F^*_{\mu\nu} := \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}F^{\rho\sigma} = \frac{1}{2}\epsilon_{\rho\sigma\mu\nu}F^{\rho\sigma} = \begin{pmatrix} 0 & -B_x/c & -B_y/c & -B_z/c \\ B_x/c & 0 & -E_z & E_y \\ B_y/c & E_z & 0 & -E_x \\ B_z/c & -E_y & E_x & 0 \end{pmatrix} = -\tilde{F}^*_{\mu\nu} \,, \qquad \tilde{F}^*_{\mu\nu} := \frac{1}{2}\tilde{\epsilon}_{\mu\nu\rho\sigma}\tilde{F}^{\rho\sigma} \,, \tag{1.95}$$

where the totally anti-symmetric Levi-Civita symbol in 4D is

$$\epsilon^{0123} := 1 \,, \qquad \epsilon_{0123} = -1 \,, \qquad \epsilon_{\mu\nu\rho\sigma} = \tilde{\epsilon}_{\mu\nu\rho\sigma} \,, \qquad \epsilon^{\mu\nu\rho\sigma} = -\epsilon_{\mu\nu\rho\sigma} \,, \qquad \tilde{\epsilon}^{\mu\nu\rho\sigma} = -\tilde{\epsilon}_{\mu\nu\rho\sigma} \,. \tag{1.96}$$

The field strength tensor satisfies (Lorentz invariant)

$$F_{\mu\nu}F^{\mu\nu} = -F^*_{\mu\nu}F^{*\mu\nu} = 2\left(B^2 - \frac{E^2}{c^2}\right) \,, \qquad \det F = \frac{1}{c^2}\left(\mathbf{E}\cdot\mathbf{B}\right)^2 \,, \qquad F^*_{\mu\nu}F^{\mu\nu} = -\frac{4}{c^2}\left(\mathbf{E}\cdot\mathbf{B}\right) \,. \tag{1.97}$$

1.4.3 Maxwell's Equations and E&M Action

The Maxwell's equations with the source terms in Eq. (1.84) become a field equation

$$\partial_\nu F^{\mu\nu} = 4\pi J^\mu \,, \qquad \text{or} \qquad \Box A_\mu - \partial_\mu\partial_\nu A^\nu = -4\pi J_\mu \,, \tag{1.98}$$

and the remaining source-free equations (1.85) become the Bianchi identity

$$0 = \epsilon^{\rho\sigma\mu\nu}\partial_\sigma F_{\mu\nu} \ , \tag{1.99}$$

where 0- and i-components become

$$0 = \partial_x F_{yz} + \partial_y F_{zx} + \partial_z F_{xy} \ , \qquad\qquad 0 = \epsilon^{ijk}\partial_j F_{k0} - \frac{1}{2}\epsilon^{ijk}\partial_0 F_{jk} \ . \tag{1.100}$$

Note that the Bianchi identity can be re-arranged to be the source-free Maxwell's equation in terms of the dual tensor $\tilde{F}^{\mu\nu}$:

$$0 = \epsilon^{\rho\sigma\mu\nu}\partial_\sigma F_{\mu\nu} = \partial_{[\mu}F_{\nu\lambda]} = \partial_\mu \tilde{F}^{\mu\nu} \ , \tag{1.101}$$

where

$$[abc] := \frac{1}{3!}\left(abc - acb + bca - bac + cab - cba\right) \ . \tag{1.102}$$

• ***Action***.— The Lagrangian for the E&M is and the Noether current (under spacetime translation invariance) is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + 4\pi J^\mu A_\mu \ , \qquad\qquad t_{\mu\nu} = \eta_{\mu\sigma}\left(-\frac{\partial\mathcal{L}}{\partial(\partial_\sigma A_\rho)}\right)\partial_\nu A_\rho + \eta_{\mu\nu}\mathcal{L} = F_{\mu\rho}\partial_\nu A^\rho - \frac{1}{4}\eta_{\mu\nu}F_{\rho\sigma}F^{\rho\sigma} \ . \tag{1.103}$$

However, the Noether current is anti-symmetric, nor gauge-invariant, such that the correct energy-momentum tensor is[2]

$$T_{\mu\nu} = t_{\mu\nu} - \frac{i}{2}\left(-2iF_{\mu\rho}\partial^\rho A_\nu\right) = F_{\mu\rho}F_\nu{}^\rho - \frac{1}{4}\eta_{\mu\nu}F_{\rho\sigma}F^{\rho\sigma} = \begin{pmatrix} \frac{1}{2}\left(E^2+B^2\right) & -\left(\mathbf{E}\times\mathbf{B}\right)_i \\ -\left(\mathbf{E}\times\mathbf{B}\right)_i & \sigma_{ij} \end{pmatrix} \ , \tag{1.105}$$

where T_{00} is the energy density, T_{0i} is the Poynting vector, the spatial components σ_{ij} are the momentum density, and we used

$$\frac{1}{4}\eta_{\mu\nu}F_{\rho\sigma}F^{\rho\sigma} = \frac{1}{2}\eta_{\mu\nu}\left(B^2 - E^2\right) \ , \qquad\qquad \sigma_{ij} \equiv \frac{1}{2}\delta_{ij}\left(E^2 + B^2\right) - E_iE_j - B_iB_j \ . \tag{1.106}$$

Note that the energy-momentum tensor of E&M is not conserved

$$\partial_\nu T^{\mu\nu}_{\text{EM}} = -F^\mu{}_\nu J^\nu \ , \tag{1.107}$$

but in the presence of E&M (or external force other than gravity) the energy-momentum tensor of (charged) particles is subject to

$$\partial_\nu T^{\mu\nu} = F^\mu{}_\nu J^\nu \ , \tag{1.108}$$

so that the total energy-momentum tensor is conserved.

• ***Gauge choice***.— Finally, in terms of the field strength tensor, the Lorentz force becomes

$$\mathbf{F} = q\left(\mathbf{E} + \frac{\mathbf{v}}{c}\times\mathbf{B}\right) \qquad \longmapsto \qquad F^\mu = qF^{\mu\nu}u_\nu \ . \tag{1.109}$$

The electric and the magnetic fields are invariant under the gauge transformation of the four vector

$$\phi \quad \mapsto \quad \phi - \frac{\partial\psi}{\partial t} \ , \qquad\qquad \mathbf{A} \quad \mapsto \quad \mathbf{A} + \nabla\psi \ . \tag{1.110}$$

[2]The correct energy-momentum tensor is obtained by the Belinfante-Rosenfeld procedure:

$$T_{\mu\nu} = t_{\mu\nu} - \frac{i}{2}\partial^\lambda\left(S_{\mu\nu\lambda} + S_{\lambda\mu\nu} - S_{\nu\lambda\mu}\right) \ , \tag{1.104}$$

where $S_{\mu\nu\lambda}$ is the contribution of the intrinsic angular momentum (see EnM.2016.pdf or Maggiore).

Popular choices of gauge condition are as follows:

$$\text{Lorenz} \quad : \quad 0 = \partial_\mu A^\mu = \frac{\partial \phi}{\partial t} + \nabla \cdot \mathbf{A} \qquad \rightarrow \qquad \nabla^2 \phi - \partial_t^2 \phi = \partial^\mu \partial_\mu \phi = \Box \phi = -4\pi\rho \,, \tag{1.111}$$

$$\text{Coulomb} \quad : \quad 0 = \nabla \cdot \mathbf{A} \;\; \rightarrow \;\; \nabla^2 \phi = -4\pi\rho \,, \tag{1.112}$$

where Lorenz and Lorentz are different.

• ***Particle in E&M Field***.— The Lagrangian for a charged particle is

$$L = \frac{1}{2} m \dot{\boldsymbol{x}}^2 - q \left(\phi - \frac{\dot{\boldsymbol{x}}}{c} \cdot \mathbf{A} \right) \,, \tag{1.113}$$

and a gauge transformation for ϕ and $\boldsymbol{A}$ yields a shift in the action at the boundary:

$$S = \int dt \, L \quad \mapsto \quad S + q \left(\psi_f - \psi_i \right) \,, \tag{1.114}$$

while the equation of motion remains invariant. The canonical momentum and the Hamiltonian are then

$$\boldsymbol{p} = m \dot{\boldsymbol{x}} + \frac{q}{c} \mathbf{A} \,, \qquad\qquad H = \frac{1}{2m} \left(\boldsymbol{p} - \frac{q}{c} \mathbf{A} \right)^2 + q\phi \,. \tag{1.115}$$

Mind that the canonical momentum is often different from the physical mechanical momentum, and observe that the Hamiltonian is just the mechanical kinetic energy plus electrostatic energy. The potential changes under a gauge transformation, but the equation of motion is invariant.