E COVARIANT ELECTRODYNAMICS

E.1 *Covariant formulation of electrodynamics*

Relativity provides the tools to formulate the Maxwell-equations very compactly, elegantly, and in a Lorentz-covariant way. For this purpose, one needs to construct a differential operator ∂_μ for derivatives with respect to the coordinates, which themselves form a Lorentz-vector x^μ.

⚠ *Sometimes, ∂^μ is used, defined as $\partial^\mu = \eta^{\mu\nu}\partial_\nu$, but please avoid notations like $\partial^\mu = \partial/\partial x_\mu$.*

$$\partial_\mu = \frac{\partial}{\partial x^\mu} = (\partial_{ct}, +\partial_i) \tag{E.329}$$

For consistency, the divergence $\partial_\mu x^\mu$ needs to be equal to the dimensionality

$$\partial_\mu x^\mu = \frac{\partial x^\mu}{\partial x^\mu} = \partial_{ct}(ct) + \partial_i x^i = 4 \tag{E.330}$$

which comes out naturally. With this differential form ∂_μ, the d'Alembert-operator is given as a Lorentz-square,

$$\Box = \eta^{\mu\nu}\partial_\mu\partial_\nu = \partial_{ct}^2 - \gamma^{ij}\partial_i\partial_j = \partial_{ct}^2 - \Delta, \tag{E.331}$$

and is in fact a Lorentz-scalar, as shown by the orthogonality relation of the Lorentz-transforms,

$$\Box = \eta^{\mu\nu}\partial_\mu\partial_\nu \to \underbrace{\eta^{\mu\nu}\Lambda_\mu{}^\alpha\Lambda_\nu{}^\beta}_{=\eta^{\alpha\beta}}\partial_\alpha\partial_\beta = \eta^{\alpha\beta}\partial_\alpha\partial_\beta = \Box, \tag{E.332}$$

reflecting the fact that wave propagation according to $\Box$ takes place at the velocity c in every frame, which was the defining principle of the Lorentz transforms. The transformation property $\partial_\mu \to \Lambda_\mu{}^\alpha\partial_\alpha$ generalises the transformation $\partial_i \to \mathrm{R}_i{}^j\partial_j$ to the full Lorentz group. In the same way as Δ is invariant under rotations, $\Box$ becomes invariant under combined rotations and Lorentz transforms.

With the operator ∂_μ it is straightforward to formulate the continuity equation for the charge density:

⚠ *$\jmath^\mu$ contains the electric charge density ρ and the current density $\jmath^i$ as a vector.*

$$\jmath^\mu = \begin{pmatrix} \rho c \\ \jmath^i \end{pmatrix} \quad \text{with} \quad \partial_\mu \jmath^\mu = \partial_{ct}(\rho c) + \partial_i \jmath^i = 0 \tag{E.333}$$

where it is interesting to see, that $\jmath^t = \rho c$ has the same units as $\jmath^i$, reflecting the consistency of the units in ∂_{ct} and ∂_i, with the additional benefit that a charge at rest in a given frame has a nonzero t-component $\jmath^t = c\rho$, as it moves with the velocity c along the ct-axis!

As a Lorentz-vector, the 4-current density transforms according to

$$\jmath^\mu \to \Lambda^\mu{}_\alpha \jmath^\alpha \tag{E.334}$$

and necessarily inversely to ∂_μ, such that $\partial_\mu \jmath^\mu$ is indeed a Lorentz-scalar and has the same value in all Lorentz-frames: The derivative transforms according to $\partial_\mu \to \Lambda_\mu{}^\alpha\partial_\alpha$ and the vectorial $\jmath^\mu$ inversely, $\jmath^\mu \to \Lambda^\mu{}_\alpha \jmath^\alpha$, such that

$$\partial_\mu J^\mu \to \Lambda_\mu{}^\alpha \Lambda^\mu{}_\beta \partial_\alpha J^\beta = \delta^\alpha_\beta \partial_\alpha J^\beta = \partial_\alpha J^\alpha \tag{E.335}$$

with $\Lambda_\mu{}^\alpha \Lambda^\mu{}_\beta = \delta^\alpha_\beta$, as the two Lorentz-transforms are inverse to each other.

This differential formulation with its clear Lorentz-invariance has a giant advantage over an integral formulation within a given frame: Earlier, we would have written

$$\frac{d}{dt} \int_V dV\, \rho = -\int_{\partial V} dS_i\, J^i. \tag{E.336}$$

Observed from a different Lorentz frame, the integration volume V is relativistically contracted by a Lorentz-factor γ, while the charge density ρ is larger by the same factor, as the charge is squeezed into a seemingly smaller volume. The two effects compensate each other, after all, it is the same charge within V. The surface ∂V of the volume is smaller by γ, too, for this to be true one can easily imagine a cuboid which is contracted by γ along the direction of motion. But for the same reasons as for the charge density, the current density J is changed by the inverse factor. Lastly, there is relativistic time dilation appearing in d/dt as well as in the current density J^i, again compensating each other: One sees all charge carries changing position at a slower rate due to their dilated proper time, leading to smaller fluxes J^i and smaller rates of change of ρ.

E.2 *Maxwell's equations*

E.2.1 Inhomogeneous Maxwell equations

The inhomogeneous Maxwell-equations are first of all a divergence $\partial_i D^i = 4\pi\rho$ and a rotation $\epsilon^{ijk}\partial_j H_k = +\partial_{ct} D^i + 4\pi/c\, J^i$. But with the help of the dual tensor $H^{ij} = \epsilon^{ijk} H_k$ the first term of Ampère's law becomes a divergence as well, $\epsilon^{ijk}\partial_j H_k = \partial_j H^{ij}$. This motivates to package the two equations into a single divergence-like tensorial relation,

$$\partial_\mu G^{\mu\nu} = \frac{4\pi}{c} J^\nu, \quad \text{in components} \quad G^{\mu\nu} = \begin{pmatrix} 0 & +D^x & +D^y & +D^z \\ -D^x & 0 & +H^z & -H^y \\ -D^y & -H^z & 0 & +H^x \\ -D^z & +H^y & -H^x & 0 \end{pmatrix} \tag{E.337}$$

with the antisymmetric field tensor $G^{\mu\nu}$. When inspecting the coordinates separately, one obtains $\partial_\mu G^{\mu t} = \partial_i D^i = 4\pi/c\, J^t = 4\pi\rho$ and $\partial_\mu G^{\mu i} = -\partial_{ct} D^i + \epsilon^{ijk}\partial_j H_k = 4\pi/c\, J^i$.

⚠ $G^{\mu\nu}$ *contains the fields* D^i *and* H_i *(effectively as* $H^{ij} = \epsilon^{ijk} H_k$*) in matter.*

One of the first conclusion we drew from the Maxwell-equations was that the field respected charge conservation, which becomes very apparent in this formalism:

$$\partial_\mu G^{\mu\nu} = \frac{4\pi}{c} J^\nu \to \partial_\nu \partial_\mu G^{\mu\nu} = \frac{4\pi}{c} \partial_\nu J^\nu = 0 \tag{E.338}$$

implying that the continuity equation $\partial_\nu J^\nu = 0$ is valid because of the contraction of the symmetric operator $\partial_\nu \partial_\mu$ with an antisymmetric tensor $G^{\mu\nu}$. With 6 free entries as an antisymmetric tensor, $G^{\mu\nu}$ can accommodate 3 components of the electric field D^i and 3 components of the magnetic field H_i.

⚠ *It follows from the antisymmetry of* $G^{\mu\nu}$ *that in* $n+1$ *dimensions, there would be* n *components for* D^i *but* $n(n-1)/2$ *components for* H_i.

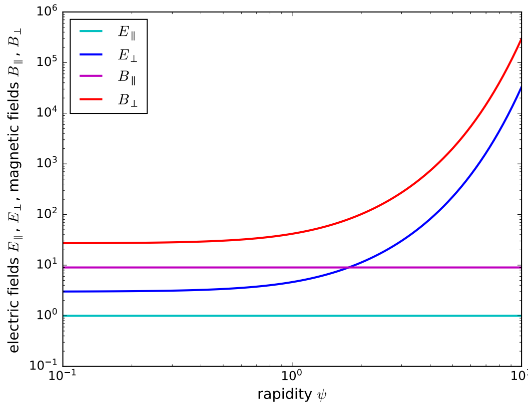

Figure 24: Electric and magnetic field components under Lorentz boosting $\tilde{F}^{\alpha\beta} \rightarrow \Lambda^{\alpha}{}_{\mu}\Lambda^{\beta}{}_{\nu}\tilde{F}^{\mu\nu}$ as a function of rapidity ψ.

E.2.2 Homogeneous Maxwell equations

Writing the two homogeneous Maxwell-equations as divergences requires a similar construction: For that purpose, one defines the dual field tensor $\tilde{F}^{\mu\nu}$ with a suitable arrangement of the fields E_i and B^i: The rotation appearing in the induction law is recast into a divergence $\epsilon^{ijk}\partial_j E_k = \partial_j \epsilon^{ijk} E_k = \partial_j E^{ij}$ with the dual $E^{ij} = \epsilon^{ijk} E_k$. Combining the electric field components in a similar alternating fashion with the magnetic field components leads to,

⚠ $\tilde{F}^{\mu\nu}$ *contains the fields* B^i *and* E_i *(effectively as* $E^{ij} = \epsilon^{ijk}E_k$*) in vacuum.*

$$\partial_\mu \tilde{F}^{\mu\nu} = 0, \quad \text{in components} \quad \tilde{F}^{\mu\nu} = \begin{pmatrix} 0 & -B_x & -B_y & -B_z \\ +B_x & 0 & +E_z & -E_y \\ +B_y & -E_z & 0 & +E_x \\ +B_z & +E_y & -E_x & 0 \end{pmatrix}. \tag{E.339}$$

With this definition of the dual field tensor, one can write analogously $\partial_\mu \tilde{F}^{\mu t} = \partial_i B^i = 0$ (the overall minus-sign does not matter) and $\partial_\mu \tilde{F}^{\mu i} = \partial_{ct} B^i + \epsilon^{ijk}\partial_j E_k = 0$. Electromagnetic duality in vacuum now amounts simply to interchanging $G^{\mu\nu}$ and $\tilde{F}^{\mu\nu}$, because $\partial_\mu G^{\mu\nu} = \partial_\mu \tilde{F}^{\mu\nu} = 0$ as soon as $\jmath^\nu = 0$.

⚠ *Again, antisymmetry of* $\tilde{F}^{\mu\nu}$ *requires that in* $n+1$ *dimensions, there would be* n *components for* B^i *but* $n(n-1)/2$ *components for* E_i*.*

Both field tensors transform under boosting according to $\tilde{F}^{\mu\nu} \rightarrow \Lambda^{\mu}{}_{\alpha}\Lambda^{\nu}{}_{\beta}\tilde{F}^{\alpha\beta}$ and $G^{\mu\nu} \rightarrow \Lambda^{\mu}{}_{\alpha}\Lambda^{\nu}{}_{\beta}G^{\alpha\beta}$, which has a strong effect $\propto \gamma^2$ (in fact, two Λs needed because the tensors have two indices) on the fields, as illustrated in Fig. 24.

E.3 *Relativistic potentials and gauging*

The next step would be to package the potentials Φ and A_i into a 4-potential, according to

$$A_\mu = (\Phi, -A_i), \tag{E.340}$$

which allows to write the Lorenz gauge-condition in a very compact way as a divergence:

$$\eta^{\mu\nu}\partial_\mu A_\nu = \partial_{ct}\Phi + \gamma^{ij}\partial_i A_j = 0, \tag{E.341}$$

⚠ A_μ contains the electric potential Φ and the magnetic potential A_i as a linear form.

where the minus signs from the spatial part of the metric $\eta^{\mu\nu}$ and of the spatial part of A_μ cancel each other. Defining the potential A_μ as in eqn. (E.340) allows to write wave equation in Lorenz-gauge in a very compact form,

$$\Box A_\mu = \frac{4\pi}{c}\eta_{\mu\nu}\jmath^\nu, \tag{E.342}$$

which at the same time explains the minus-sign in the spatial part of A_μ as well as the cancellation of the additional factor of c in $\jmath^t = \rho c$.

Linking the potential A_μ to the Faraday tensor $F_{\mu\nu}$ is possible by writing

$$\partial_\mu A_\nu - \partial_\nu A_\mu = F_{\mu\nu}, \tag{E.343}$$

because then the the electric field components would be given as $F_{it} = \partial_i A_t - \partial_{ct}A_i = -\partial_i\Phi - \partial_{ct}A_i = E_i$ as well as $F_{ij} = \partial_i A_j - \partial_j A_i$ with mutually different indices (ijk). It is interesting to see, how the requirement of antisymmetry reduces the number of free field components from initially 16 in $\partial_\mu A_\nu$ to 6, corresponding to 3 components of the electric and 3 components of the magnetic field. Weirdly enough, it's a bit of a coincidence that in 3 + 1 dimensions there are as many components of the electric and of the magnetic field, allowing to write B^i as a vector:

$$B^i = \epsilon^{ijk}F_{jk} = \epsilon^{ijk}\left(\partial_j A_k - \partial_k A_j\right) \tag{E.344}$$

albeit with a small caveat: Under parity transform $\mathcal{P}$, B^i does not change its sign, because both ∂_i and A_j change their signs. In contrast, E_i does change its sign, because in $\partial_i\Phi$ only ∂_i changes its sign, and in $\partial_{ct}A_i$ only A_i! Consequently, one calls E_i a polar vector and B^i an axial vector.

Applying gauge transformations would change the potentials, $A_\mu \to A_\mu + \partial_\mu\chi$, but leaves the Faraday tensor $F_{\mu\nu}$ invariant, as

$$F_{\mu\nu} \to \partial_\mu\left(A_\nu + \partial_\nu\chi\right) - \partial_\nu\left(A_\mu + \partial_\mu\chi\right) = \partial_\mu A_\nu - \partial_\nu A_\mu + \underbrace{\partial_\mu\partial_\nu\chi - \partial_\nu\partial_\mu\chi}_{=0} = F_{\mu\nu} \tag{E.345}$$

as partial derivatives interchange. The same result applies to the tensor $G^{\mu\nu}$ as it originates from $F_{\mu\nu}$ through a linear transform. It is well possible to derive $\tilde{F}^{\mu\nu}$ from the potential directly, through

$$\tilde{F}^{\mu\nu} = \frac{1}{2}\epsilon^{\mu\nu\alpha\beta}F_{\alpha\beta} = \frac{1}{2}\epsilon^{\mu\nu\alpha\beta}\left(\partial_\alpha A_\beta - \partial_\beta A_\alpha\right) = \epsilon^{\mu\nu\alpha\beta}\partial_\alpha A_\beta, \tag{E.346}$$

using an antisymmetry-argument in the second step. Gauge transforms on the potential imply

$$\epsilon^{\mu\nu\alpha\beta}\partial_\alpha A_\beta \rightarrow \epsilon^{\mu\nu\alpha\beta}\partial_\alpha\left(A_\beta + \partial_\beta\chi\right) = \epsilon^{\mu\nu\alpha\beta}\partial_\alpha A_\beta + \epsilon^{\mu\nu\alpha\beta}\partial_\alpha\partial_\beta\chi = \epsilon^{\mu\nu\alpha\beta}\partial_\alpha A_\beta = \tilde{F}^{\mu\nu} \tag{E.347}$$

with the contraction of the symmetric $\partial_\alpha\partial_\beta$ with the antisymmetric $\epsilon^{\mu\nu\alpha\beta}$ vanishes. In consequence, not only $F_{\mu\nu}$ but also $\tilde{F}^{\mu\nu}$ is gauge-invariant, and by extension $\tilde{G}_{\mu\nu}$.

An interesting manipulation shows a derivative relation for $F_{\mu\nu}$ as it originates from the potential. Composing a cyclic permutation of indices in $\partial_\lambda F_{\mu\nu}$ yields

$$\partial_\lambda F_{\mu\nu} + \partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} = \partial_\lambda(\partial_\mu A_\nu - \partial_\nu A_\mu) + \partial_\mu(\partial_\nu A_\lambda - \partial_\lambda A_\nu) + \partial_\nu(\partial_\lambda A_\mu - \partial_\mu A_\lambda) = 0 \tag{E.348}$$

with a pairwise cancellation of the terms. This derivative relation is called the Bianchi-identity and is in fact equivalent to the field equation $\partial_\mu\tilde{F}^{\mu\nu} = 0$ for the dual tensor $\tilde{F}^{\mu\nu}$,

$$\partial_\mu\tilde{F}^{\mu\nu} = \frac{\partial_\mu}{2}\epsilon^{\mu\nu\alpha\beta}F_{\alpha\beta} = \frac{\partial_\mu}{2}\epsilon^{\mu\nu\alpha\beta}(\partial_\alpha A_\beta - \partial_\beta A_\alpha) = \partial_\mu\epsilon^{\mu\nu\alpha\beta}\partial_\alpha A_\beta = \epsilon^{\mu\nu\alpha\beta}\partial_\mu\partial_\alpha A_\beta = 0, \tag{E.349}$$

with the well-used argument that a contraction between a symmetric and an antisymmetric index pair, here (α, μ), has to vanish. One sees immediately, that working with a potential is enabled by the condition $\partial_\mu\tilde{F}^{\mu\nu} = 0$ instead of $\tilde{F}^{\mu\nu}$ being sourced by a magnetic charge density $\imath^\nu$, in the spirit of

$$\partial_\mu\tilde{F}^{\mu\nu} = -\frac{4\pi}{c}\imath^\nu, \tag{E.350}$$

with an associated conservation law $\partial_\nu\imath^\nu = 0$. Only then can we make the argument that a potential A_μ invalidates a nonzero divergence of $\tilde{F}^{\mu\nu}$.

The field tensor $G^{\mu\nu}$ containing D^i and H_i can be related to the field tensor $F_{\mu\nu}$ containing E_i and B^i by means of a generalised constitutive relation,

$$G^{\alpha\beta} = X^{\alpha\beta\mu\nu}F_{\mu\nu} \quad\leftrightarrow\quad F_{\alpha\beta} = X_{\alpha\beta\mu\nu}G^{\mu\nu} \tag{E.351}$$

with the orthogonality relation

$$X^{\alpha\beta\mu\nu}X_{\mu\nu\gamma\delta} = \delta^\alpha_\gamma\,\delta^\beta_\delta, \quad\text{implying}\quad G^{\alpha\beta} = X^{\alpha\beta\mu\nu}X_{\mu\nu\gamma\delta}\,G^{\gamma\delta} = \delta^\alpha_\gamma\delta^\beta_\delta\,G^{\gamma\delta} = G^{\alpha\beta} \tag{E.352}$$

The tensor $X^{\alpha\beta\mu\nu}$ is antisymmetric in each index pair (α, β), (μ, ν) and maps an antisymmetric linear form $F_{\mu\nu}$ to an antisymmetric vectorial tensor $G^{\mu\nu}$. Tensors of that type can be written as being proportional to proper antisymmetrisations of the metric,

$$X^{\alpha\beta\mu\nu} = \frac{\eta^{\alpha\mu}\eta^{\beta\nu} - \eta^{\alpha\nu}\eta^{\beta\mu}}{2}, \tag{E.353}$$

allowing us to convert the divergence $\partial_\mu G^{\mu\nu} = 4\pi/c\ \jmath^\nu$ into a wave equation for the potentials,

$$\partial_\mu G^{\mu\nu} = \partial_\mu X^{\mu\nu\alpha\beta} F_{\alpha\beta} = \partial_\mu X^{\mu\nu\alpha\beta}(\partial_\alpha A_\beta - \partial_\beta A_\alpha) = 2\partial_\mu X^{\mu\nu\alpha\beta}\partial_\alpha A_\beta =$$
$$\left(\eta^{\alpha\mu}\eta^{\beta\nu} - \eta^{\alpha\nu}\eta^{\beta\mu}\right)\partial_\mu\partial_\alpha A_\beta = \underbrace{\eta^{\alpha\mu}\partial_\mu\partial_\alpha}_{=\Box}\eta^{\beta\nu}A_\beta - \eta^{\alpha\nu}\partial_\alpha\underbrace{\eta^{\beta\mu}\partial_\mu A_\beta}_{=0} = \Box\eta^{\beta\nu}A_\beta = \frac{4\pi}{c}\jmath^\nu. \tag{E.354}$$

In summary, under the assumption of Lorenz-gauge, the wave equation

$$\Box A_\beta = \frac{4\pi}{c}\eta_{\beta\nu}\jmath^\nu \tag{E.355}$$

relates potential and source, where we have already discussed solutions in terms of Liénard-Wichert retarded potentials. Effectively, with the time-component of the source being $c\rho$, and the overall coupling constant being $4\pi/c$, one can combine both potentials into a single linear form and all sources into a single vector.

E.4 *Dual field tensors and the Bianchi-identity*

The duality transformation interchanges the positions of the electric and magnetic field components when transitioning from $F_{\mu\nu}$ to $\tilde{F}^{\mu\nu}$ and vice versa:

$$\tilde{F}^{\alpha\beta} = -\frac{1}{2}\epsilon^{\alpha\beta\mu\nu}F_{\mu\nu} \quad \leftrightarrow \quad F_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\alpha\beta}\tilde{F}^{\alpha\beta} \tag{E.356}$$

making $F_{\mu\nu}$ autodual

$$\tilde{\tilde{F}}_{\mu\nu} = -\frac{1}{4}\epsilon_{\mu\nu\alpha\beta}\epsilon^{\alpha\beta\gamma\delta}F_{\gamma\delta} = \delta^{\gamma\delta}_{\mu\nu}F_{\gamma\delta} = \frac{1}{2}\left(\delta^\gamma_\mu\delta^\delta_\nu - \delta^\gamma_\nu\delta^\delta_\mu\right)F_{\gamma\delta} = \frac{1}{2}\left(F_{\mu\nu} - F_{\nu\mu}\right) = F_{\mu\nu}, \tag{E.357}$$

where analogous formulas apply to $\tilde{F}^{\mu\nu}$. For the contraction between the two Levi-Civita symbols we have used the relation

$$\epsilon^{i_1\ldots i_q k_1\ldots k_p}\epsilon_{k_1\ldots k_p j_1\ldots j_q} = -p!q!\delta^{i_1\ldots i_q}_{j_1\ldots j_q}, \tag{E.358}$$

valid for Minkowksi-spaces, with the dimension $n = p + q$ and the overlap p between the indices to be contracted. Specifically, we need $p = 2 = q$ in $n = 4$. $\delta^{i_1\ldots i_q}_{j_1\ldots j_q}$ refers to the generalised Kronecker symbol. In complete analogy, there is a dual $\tilde{G}_{\mu\nu}$ of the field tensor $G^{\mu\nu}$,

$$\tilde{G}_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\alpha\beta}G^{\alpha\beta} \quad \leftrightarrow \quad G^{\alpha\beta} = -\frac{1}{2}\epsilon^{\alpha\beta\mu\nu}\tilde{G}_{\mu\nu}. \tag{E.359}$$

To make things more concrete, one can follow through how the duality transform reorganises the tensors $\tilde{F}^{\mu\nu}$ and $G^{\mu\nu}$ isolated from the homogeneous and inhomogeneous Maxwell-equations. First of all, $\epsilon_{\alpha\beta\mu\nu}\tilde{F}^{\mu\nu}$ maps the antisymmetric (μ,ν) index pair to an object $F_{\alpha\beta}$, which is likewise antisymmetric, this time in (α,β). For a non-vanishing contribution, all indices in the Levi-Civita-symbol need to be different, which implies that there is no linear combination being formed, but simply a remapping of all components: For instance, choosing $(\alpha, \beta) = (t, x)$ for $F_{\alpha\beta}$ can only acquire a combination from $\tilde{F}^{\mu\nu}$ for $(\mu, \nu) = (y, z)$ or (z, y). But $\tilde{F}^{y,z} = -\tilde{F}^{z,y}$ due to the antisymmetry of the field tensor, therefore the two are equal, and are added twice, which in turn is remedied by the prefactor of 1/2.

Specifically F_{tx} will be set equal to $\tilde{F}^{yz} = E_x$, and F_{xy} will become $\tilde{F}^{tz} = -B_z$: We observe, how the first row and the first column of $F_{\alpha\beta}$ will accommodate the electric field components which had been stored in the interior of the tensor $\tilde{F}^{\mu\nu}$, while the first row and first column of $\tilde{F}^{\mu\nu}$ get scattered into the interior of the tensor $F_{\alpha\beta}$: Effectively, the magnetic and electric field components get interchanged up to a sign, leading to:

$$F_{\mu\nu} = \begin{pmatrix} 0 & +E_x & +E_y & +E_z \\ -E_x & 0 & -B_z & +B_y \\ -E_y & +B_z & 0 & -B_x \\ -E_z & -B_y & +B_x & 0 \end{pmatrix}. \tag{E.360}$$

The same rearrangement takes place in the duality transform of the tensor $G^{\alpha\beta}$:

$$\tilde{G}_{\mu\nu} = \begin{pmatrix} 0 & -H^x & -H^y & -H^z \\ +H^x & 0 & -D^z & +D^y \\ +H^y & +D^z & 0 & -D^x \\ +H^z & -D^y & +D^x & 0 \end{pmatrix} \tag{E.361}$$

with the replacement of D^i and H_i, again with a sign change: This sign change is very important, as it recovers the idea of duality of electromagnetism in vacuum, where under the replacement of electric and magnetic fields the Maxwell equations do not change.

The duality transform respects the antisymmetry of $\tilde{F}^{\mu\nu}$ and $F_{\mu\nu}$, which is important because it links charge conservation to gauge invariance of the potentials: Nature has chosen to have $\imath^\mu = 0$ and $\partial_\mu \imath^\mu = 0$ which has important implications, as we can now differentiate between the inhomogeneous and homogeneous Maxwell equations, which read:

$$\partial_\mu G^{\mu\nu} = \frac{4\pi}{c} \jmath^\nu \quad \text{and} \quad \partial_\mu \tilde{F}^{\mu\nu} = 0 \tag{E.362}$$

With $F_{\mu\nu}$ following from a potential A_μ in an antisymmetrised, gauge-invariant way,

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{E.363}$$

the homogeneous Maxwell equation is automatically fulfilled, as

$$\partial_\mu \tilde{F}^{\mu\nu} = \frac{1}{2} \partial_\mu \epsilon^{\mu\nu\alpha\beta} F_{\alpha\beta} = \epsilon^{\mu\nu\alpha\beta} \partial_\mu \partial_\alpha A_\beta = 0 \tag{E.364}$$

through the contraction of the antisymmetric Levi-Civita symbol over the symmetric index pair (α, μ).

The equivalence of the Bianchi-identity

$$\partial_\mu F_{\alpha\beta} + \partial_\beta F_{\mu\alpha} + \partial_\alpha F_{\beta\mu} = 0 \tag{E.365}$$

and the divergence-like field equation for the dual tensor $\tilde{F}^{\mu\nu}$

$$\partial_\mu \tilde{F}^{\mu\nu} = 0 \tag{E.366}$$

can be shown as follows:

$$\partial_\mu \tilde{F}^{\mu\nu} = -\frac{1}{2}\epsilon^{\mu\nu\alpha\beta}\partial_\mu F_{\alpha\beta} = +\frac{1}{2}\epsilon^{\nu\mu\alpha\beta}\partial_\mu F_{\alpha\beta} \tag{E.367}$$

by substituting the definition of the duality transform and by interchanging $\mu \leftrightarrow \nu$ in the last step, which brings in a minus-sign because of the antisymmetry of ϵ. In fact, any cyclic permutation of the indices does not change anything, so that one can write

$$\ldots = \frac{1}{6}\left[\epsilon^{\nu\mu\alpha\beta} + \epsilon^{\nu\alpha\beta\mu} + \epsilon^{\nu\beta\mu\alpha}\right]\partial_\mu F_{\alpha\beta} = \frac{1}{6}\epsilon^{\nu\mu\alpha\beta}\Big(\underbrace{\partial_\mu F_{\alpha\beta} + \partial_\beta F_{\mu\alpha} + \partial_\alpha F_{\beta\mu}}_{=0}\Big) = 0 \tag{E.368}$$

making $\partial_\mu \tilde{F}^{\mu\nu} = 0$ equivalent with eqn. (E.365), after renaming the indices in the second and third term.

The Bianchi-identity is particularly interesting because it provides a propagation mechanism for electromagnetic waves: Acting on eqn. (E.365) with the derivative $\eta^{\mu\nu}\partial_\nu$ gives

$$\eta^{\mu\nu}\partial_\nu\left(\partial_\mu F_{\alpha\beta} + \partial_\beta F_{\mu\alpha} + \partial_\alpha F_{\beta\mu}\right) = \underbrace{\eta^{\mu\nu}\partial_\nu\partial_\mu}_{=\Box} F_{\alpha\beta} + \partial_\beta \underbrace{\eta^{\mu\nu}\partial_\nu F_{\mu\alpha}}_{=0} - \partial_\alpha \underbrace{\eta^{\mu\nu}\partial_\nu F_{\mu\beta}}_{=0} = 0, \tag{E.369}$$

and substituting the field equation for vacuum twice has us arrive at a wave equation for the fields,

$$\Box F_{\alpha\beta} = 0. \tag{E.370}$$

It can be solved with a wave ansatz $F_{\alpha\beta} \propto \exp(\pm i k_\mu x^\mu)$, leading to the null-condition

$$\eta^{\mu\nu}k_\mu k_\nu = 0 \quad \text{equivalent with} \quad \left(\frac{\omega}{c}\right)^2 - \gamma^{ij}k_i k_j = 0 \rightarrow \omega = \pm ck \tag{E.371}$$

such that group velocity $d\omega/dk$ and phase velocity ω/k are both c, and dispersion can not occur.

The wave equation for a non-vacuum situation looks a bit weird: Substituting the sources $\jmath^\alpha$ and $\jmath^\beta$ gives

$$\Box F_{\alpha\beta} = \frac{4\pi}{c}\left(\partial_\alpha \eta_{\beta\mu}\jmath^\mu - \partial_\beta \eta_{\alpha\mu}\jmath^\mu\right), \tag{E.372}$$

where it is interesting to see that the antisymmetry in the index pair (α, β) appears consistently in the sources on the right side. The same result could have been derived from the potentials, too, as $\Box A_\mu = 4\pi/c\ \eta_{\mu\nu}\jmath^\nu$ in e.g. Lorenz-gauge becomes

$$\Box F_{\alpha\beta} = \Box\left(\partial_\alpha A_\beta - \partial_\beta A_\alpha\right) = \partial_\alpha \Box A_\beta - \partial_\beta \Box A_\alpha = \frac{4\pi}{c}\left(\partial_\alpha \eta_{\beta\mu}\jmath^\mu - \partial_\beta \eta_{\alpha\mu}\jmath^\mu\right). \tag{E.373}$$

Actually, eqn. (E.373) is able to explain an interesting fact: Naively, one would think that it is not entirely clear how the six components of $G^{\mu\nu}$ are sourced by the four components of $\jmath^\mu$, and only going through the potential A_μ resolves the issue: There is, in particular in Lorenz-gauge (just for illustration, any gauging term $\partial_\mu\chi$ would drop from the expression), a one-to-one relation linking A_μ to $\jmath^\mu$ in $\Box A_\mu = 4\pi/c\;\eta_{\mu\nu}\jmath^\nu$, and the definition of $F_{\mu\nu}$ as $\partial_\mu A_\nu - \partial_\nu A_\mu$ then generates six mutually independent field components, to be related linearly to the six free components of $G^{\mu\nu}$ through the constitutive relation.

On the other hand, eqn. (E.373) may be interpreted in a way that it is not the current density $\jmath^\alpha$ that sources $F_{\alpha\beta}$, but rather its antisymmetric derivative $\partial_\alpha\eta_{\beta\mu}\jmath^\mu - \partial_\beta\eta_{\alpha\mu}\jmath^\mu$. Its six components determine each individually and independently the six components of $F_{\alpha\beta}$, even in a physical and gauge independent way.

A summary of the two field tensors and their duals, along with all four possible quadratic Lorentz-invariants (three of which are distinct, and reduce to two in vaccum) is given by this diagram:

$$
\begin{tikzcd}
A_\nu \arrow[dr, "\epsilon^{\alpha\beta\mu\nu}\partial_\mu"] & & \tilde{F}^{\alpha\beta}\tilde{G}_{\alpha\beta} = E_iD^i - H_iB^i & & \\
& \tilde{F}^{\alpha\beta} \arrow[ur, dashed] \arrow[dl, dashed] \arrow[dr, "\partial_\alpha"] & & \tilde{G}_{\gamma\delta} \arrow[ll, "\tilde{X}^{\alpha\beta\gamma\delta}"'] \arrow[ul, dashed] \arrow[dr, dashed] \arrow[dd, "-\epsilon^{\mu\nu\gamma\delta}/2"'] & \\
\tilde{F}^{\alpha\beta}F_{\alpha\beta} \propto E_iB^i & & 0 & & \tilde{G}_{\mu\nu}G^{\mu\nu} \propto H_iD^i \\
& F_{\gamma\delta} \arrow[uu, "-\epsilon^{\alpha\beta\gamma\delta}/2"] \arrow[ur, "\partial_\alpha F_{\gamma\delta}+\partial_\gamma F_{\delta\alpha}+\partial_\delta F_{\alpha\gamma}"'] \arrow[ul, dashed] \arrow[rr, "X^{\mu\nu\gamma\delta}"'] \arrow[dr, dashed] & & G^{\mu\nu} \arrow[ur, dashed] \arrow[dl, dashed] \arrow[dr, "\partial_\mu"] & \\
A_\delta \arrow[ur, "\partial_\gamma A_\delta - \partial_\delta A_\gamma"] & & F_{\mu\nu}G^{\mu\nu} = E_iD^i - H_iB^i & & \jmath^\nu
\end{tikzcd}
\tag{E.374}
$$

E.5 *Covariant electrodynamics*

Summarising the results from the previous chapters shows that there is a clear conceptual picture defining Maxwell-electrodynamics:

	vector	*form*
matter	$G^{\mu\nu}$	$\tilde{G}_{\mu\nu}$
vacuum	$\tilde{F}^{\mu\nu}$	$F_{\mu\nu}$

- The 4-potential A_μ and the 4-current $\jmath^\mu$ are a Lorentz-linear form and a Lorentz-vector, respectively.
- The inhomogeneous Maxwell-equation take on the form $\partial_\mu G^{\mu\nu} = 4\pi/c\;\jmath^\nu$ and the homogeneous Maxwell-equations are written as $\partial_\mu\tilde{F}^{\mu\nu} = 0$, as there are no magnetic charges.
- Equivalent to the homogeneous Maxwell equation is the Bianchi-identity, $\partial_\lambda F_{\mu\nu} + \partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} = 0$, which is automatically fulfilled if $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$.

- Charge is conserved and the inhomogeneous Maxwell-equation $\partial_\mu j^\mu = 0$ respects it through the antisymmetry of $G^{\mu\nu}$.

- Gauging with a gauge function χ implies the transformation $A_\mu \rightarrow A_\mu + \partial_\mu\chi$, leaving the Faraday tensor $F_{\mu\nu}$ invariant through its antisymmetry.

- Under the Lorenz-gauge condition $\eta^{\mu\nu}\partial_\mu A_\nu = 0$ one obtains a typical wave equation $\Box A_\mu = 4\pi/c\ \eta_{\mu\nu} j^\nu$ from the inhomogeneous Maxwell-equation, with Lorentz-invariant propagation speed c.

- The geometry is defined by the metric tensor $\eta_{\mu\nu}$ which is relevant for the vacuum fields in $F_{\mu\nu}$. The constitutive relation $X^{\alpha\beta\mu\nu}$ links $G^{\mu\nu}$ to $F_{\mu\nu}$ and falls back onto the metric in vacuum.

It is amazing to see how clearly gauge-transforms and Lorentz-transforms are incorporated into the formalism, and how the mathematical structure of the Maxwell-equations results from the antisymmetry of the field tensor, as well as its gauge-independence. It's worthwhile to contemplate, how the Lorenz-gauge condition $\eta^{\mu\nu}\partial_\mu A_\nu = 0$ is at the same time a Lorentz-invariant: As a Lorentz-scalar it has the same value, zero in this case, in all frames. The electromagnetic field, too, possesses Lorentz-invariants, which are necessarily quadratic or of higher order in the fields, as all contractions ${F^\mu}_\mu = \eta^{\mu\nu}F_{\mu\nu} = 0$, ${\tilde{F}^\mu}_\mu = \eta_{\mu\nu}\tilde{F}^{\mu\nu} = 0$, ${\tilde{G}^\mu}_\mu = \eta^{\mu\nu}\tilde{G}_{\mu\nu} = 0$ and lastly ${G^\mu}_\mu = \eta_{\mu\nu}G^{\mu\nu} = 0$ vanish because of the antisymmetry of $F_{\mu\nu}$, $G^{\mu\nu}$ and their respective duals.

Quadratic invariants are first of all

$$F_{\mu\nu}G^{\mu\nu} = \tilde{F}^{\mu\nu}\tilde{G}_{\mu\nu} = E_i D^i - H_i B^i, \tag{E.375}$$

which is a properly scalar quantity which is in addition parity-positive: The product of two parity-even magnetic fields is parity-even and the product of two parity-odd electric fields is likewise parity-even. Mixed contractions involving a single dual,

$$\tilde{F}^{\mu\nu}F_{\mu\nu} = 4E_i B^i \quad \text{and} \quad \tilde{G}_{\mu\nu}G^{\mu\nu} = 4H_i D^i \tag{E.376}$$

are parity negative, as products of a parity-even magnetic field and a parity-odd electric field.

⚠ *The Maxwell-field has a single, scalar quadratic invariant, $F_{\mu\nu}G^{\mu\nu}$; there are two pseudoscalar quadratic invariants, $\tilde{F}^{\mu\nu}F_{\mu\nu}$ and $\tilde{G}_{\mu\nu}G^{\mu\nu}$, where the last two coincide in vacuum.*

In particular the first invariant does not reflect an energy density $T^{tt} \propto E_i D^i + H_i B^i$, which should depend on the choice of frame and can not be invariant. Its numerical value is actually zero for all vacuum solutions, as can be quickly verified by considering a plane wave: The electric and magnetic energy densities are equal at every point and instant, $E_i D^i = H_i B^i$, making sure that $F_{\mu\nu}G^{\mu\nu} = 0$. Furthermore, the electric and magnetic fields are orthogonal to each other, such that $E_i B^i = 0$ and $H_i D^i = 0$.

The invariant discussed above are contractions between the vectorial tensors $G^{\mu\nu}$ and $\tilde{F}^{\mu\nu}$ on one side and the linear forms $F_{\mu\nu}$ and $\tilde{G}_{\mu\nu}$ on the other. In a vacuum situation, all vectorial quantities are trivially related to their linear forms through the Minkowski-metric, so it is possible to construct 4 more invariants

$$\eta^{\alpha\mu}\eta^{\beta\nu}F_{\alpha\beta}F_{\mu\nu} = \eta^{\alpha\mu}\eta^{\beta\nu}\tilde{G}_{\alpha\beta}\tilde{G}_{\mu\nu} = \eta_{\alpha\mu}\eta_{\beta\nu}G^{\alpha\beta}G^{\mu\nu} = \eta_{\alpha\mu}\eta_{\beta\nu}\tilde{F}^{\alpha\beta}\tilde{F}^{\mu\nu} \propto \gamma^{ij}E_iE_j - \gamma_{ij}B^iB^j. \tag{E.377}$$

E.6 *Lagrange-density for the dynamics of fields*

To our knowledge, all fundamental physical theories can be derived from variational principles, and electrodynamics is no exception. At the basis of all variational principles is the notion that the action is invariant under a certain relativity principle, in our case Lorentz-relativity, which leads to a covariant equation of motion, where all quantities are consistently behaving under changes in the frame: This was already the case for Galilean dynamics, as a rotationally invariant Lagrange-function $\mathcal{L}(x^i, \dot{x}^i) = \gamma_{ij}\dot{x}^i\dot{x}^j/2 - \Phi(x^i)$ with the Euclidean, rotationally invariant scalar product $\gamma_{ij}\dot{x}^i\dot{x}^j$ gave rise to a equation of motion $\ddot{x}^i = -\gamma^{ij}\partial_j\Phi$ relating two vectors to each other. From this point of view one would hope to arrive at a Lorentz-covariant equation of motion from a Lorentz-invariant Lagrange function. As the Euler-Lagrange-equation usually reduces the powers by one in the derivative process, one would like to begin with quadratic Lorentz-invariants in order to arrive at a linear field equation which respects the superposition principle. Then, if the Lagrange-function does not depend explicitly on the coordinates x^μ, i.e. if x^μ is a cyclic variable, one has reasons to expect that the theory is conserving energy and momentum. And lastly, charge conservation should result from gauge-invariance as the symmetry principle.

⚠ *invariance/covariance principle: covariant field equations from invariant Lagrange functions*

⚠ *Quadratic Lagrange functions lead to linear field equations: superposition principle*

⚠ *coordinates as cyclic variables imply energy-momentum conservation*

⚠ *gauge symmetry is related to charge conservation*

E.6.1 Scalar field on a Euclidean background

Let's illustrate how variational principles work with a simpler example than the full Maxwell-theory. Electrostatics is fully characterised by a potential Φ which is linked to the source ρ by means of the Poisson-equation $\Delta\Phi = -4\pi\rho$, in other words: We're looking for a variational principle for a scalar field φ on a Euclidean background, that is coupled to a source and does not have any dynamics on its own. Writing the action S as an integral over a Lagrange-density $\mathcal{L}$ would give

$$S = \int_V d^3x\, \mathcal{L}(\varphi, \partial_i\varphi) \tag{E.378}$$

and Hamilton's principle $\delta S = 0$ then suggests the variation

$$\delta S = \delta\int_V d^3x\, \mathcal{L} = \int_V d^3x \left(\frac{\partial\mathcal{L}}{\partial\varphi}\delta\varphi + \frac{\partial\mathcal{L}}{\partial\partial_i\varphi}\delta\partial_i\varphi\right) \tag{E.379}$$

Interchanging the partial derivative and the variation, $\delta\partial_i\varphi = \partial_i\delta\varphi$, allows an integration by parts. One can isolate the Euler-Lagrange-equation for a scalar field φ

$$\delta S = \int_V d^3x \left(\frac{\partial\mathcal{L}}{\partial\varphi} - \partial_i\frac{\partial\mathcal{L}}{\partial\partial_i\varphi}\right)\delta\varphi = 0 \quad\rightarrow\quad \partial_i\frac{\partial\mathcal{L}}{\partial\partial_i\varphi} = \frac{\partial\mathcal{L}}{\partial\varphi} \tag{E.380}$$

because the variation $\delta\varphi$ is zero by construction on the boundary ∂V,

$$\int_V dV\,\partial_i\left(\frac{\partial\mathcal{L}}{\partial\partial_i\varphi}\delta\varphi\right) = \int_{\partial V} dS_i\left(\frac{\partial\mathcal{L}}{\partial\partial_i\varphi}\delta\varphi\right) = 0 \quad\text{as}\quad \frac{\partial\mathcal{L}}{\partial\partial_i\varphi}\delta\varphi\Big|_{\partial V} = 0. \tag{E.381}$$

The Poisson-equation as a second order partial differential equation should result from an action that contains squares of first derivatives of the potential φ, for instance

$$\mathcal{L}(\varphi, \partial_i\varphi) = \frac{\gamma^{ab}}{2}\partial_a\varphi\partial_b\varphi - 4\pi\rho\varphi. \tag{E.382}$$

Concerning the invariance-covariance principle, we note that the first term is as a scalar product, invariant under rotations. Substitution into the Euler-Lagrange equation gives

$$\frac{\partial\mathcal{L}}{\partial\varphi} = -4\pi\rho \tag{E.383}$$

as well as (please always rename the indices when you're doing this)

$$\frac{\partial\mathcal{L}}{\partial\partial_i\varphi} = \frac{\gamma^{ab}}{2}\left(\frac{\partial\partial_a\varphi}{\partial\partial_i\varphi}\partial_b\varphi + \partial_a\varphi\frac{\partial\partial_b\varphi}{\partial\partial_i\varphi}\right) =$$
$$\frac{\gamma^{ab}}{2}\left(\delta^i_a\partial_b\varphi + \partial_a\varphi\delta^i_b\right) = \frac{1}{2}\left(\gamma^{ib}\partial_b\varphi + \gamma^{ai}\partial_a\varphi\right) = \gamma^{ib}\partial_b\varphi \tag{E.384}$$

such that one arrives precisely at the Poisson-equation

$$\partial_i\frac{\partial\mathcal{L}}{\partial\partial_i\varphi} = \partial_i\gamma^{ib}\partial_b\varphi = \Delta\varphi = \frac{\partial\mathcal{L}}{\partial\varphi} = -4\pi\rho. \tag{E.385}$$

where the Laplace-operator Δ is scalar and does not change under rotations.

E.6.2 Scalar field on a Lorentz background

Repeating the entire derivation for a relativistic field theory with the Lagrange density

$$\mathcal{L}(\varphi, \partial_\mu\varphi) = \frac{\eta^{\mu\nu}}{2}\partial_\mu\varphi\partial_\nu\varphi + 4\pi\rho\varphi \tag{E.386}$$

leads with the Euler-Lagrange equation

$$\partial_\alpha\frac{\partial\mathcal{L}}{\partial\partial_\alpha\varphi} = \frac{\partial\mathcal{L}}{\partial\varphi} \tag{E.387}$$

for varying the action

$$S = \int_V \mathrm{d}^4x\, \mathcal{L}(\varphi, \partial_\mu\varphi) \tag{E.388}$$

that results as an integral over the spacetime volume $\mathrm{d}^4x = c\mathrm{d}t\mathrm{d}^3x$. Carrying out the variation $\delta S = 0$ implies the wave equation

$$\Box\varphi = 4\pi\rho \quad \text{with} \quad \Box = \eta^{\mu\nu}\partial_\mu\partial_\nu. \tag{E.389}$$

as a generalisation to the Poisson-equation.

E.6.3 Maxwell field on a Lorentz background

The Maxwell-equations expressed in terms of the potential A_μ are likewise second order differential equations, where the action should contain squares of first derivatives of the potential. The new aspect now is that the potential has (4) internal degrees of freedom and is not scalar as in the previous two examples. The squares of the first derviatives of A_μ should be Lorentz-invariants, and we will only utilise the parity-positive one for the time being.

Driven by analogy, one would write for a vacuum situation

please keep in mind that the Lagrange-density is invariant under affine transforms, $\mathcal{L} \rightarrow \alpha\mathcal{L}+\beta$, therefore only the ratio of prefactors matters.

$$S = \int_V d^4x\, \mathcal{L}(A_\mu, \partial_\mu A_\nu) = \int_V d^4x \quad \underbrace{\eta^{\alpha\mu}\eta^{\beta\nu}F_{\alpha\beta}F_{\mu\nu}}_{\text{square of first derivatives}} \quad + \quad \underbrace{\frac{16\pi}{c}A_\mu \jmath^\mu}_{\text{coupling to the source}} \tag{E.390}$$

Please keep in mind that it is only through broken duality and the non-existence of magnetic charges that the potentials A_μ exist such which ultimately enables a Lagrangian description as in eqn. (E.390). A suitable Euler-Lagrange equation would result from variation δS of the action S with respect to δA, which becomes

$$\delta S = \delta \int_V d^4x\, \mathcal{L} = \int_V d^4x \left(\frac{\partial \mathcal{L}}{\partial A_\gamma}\delta A_\gamma + \frac{\partial \mathcal{L}}{\partial \partial_\gamma A_\delta}\delta \partial_\gamma A_\delta\right) =$$
$$\int_V d^4x \left(\frac{\partial \mathcal{L}}{\partial A_\delta} - \partial_\gamma \frac{\partial \mathcal{L}}{\partial \partial_\gamma A_\delta}\right)\delta A_\delta = 0 \tag{E.391}$$

where as always we wrote $\delta\partial_\gamma A_\delta = \partial_\gamma \delta A_\delta$ for the integration by parts, finally allowing the extraction of the Euler-Lagrange equation by means of Hamilton's principle $\delta S = 0$:

$$\partial_\gamma \frac{\partial \mathcal{L}}{\partial \partial_\gamma A_\delta} = \frac{\partial \mathcal{L}}{\partial A_\delta}, \tag{E.392}$$

again keeping the variation δA_δ fixed on the boundary,

$$\int_V dV\, \partial_\gamma \left(\frac{\partial \mathcal{L}}{\partial \partial_\gamma A_\delta}\delta A_\delta\right) = \int_{\partial V} dS_\gamma \left(\frac{\partial \mathcal{L}}{\partial \partial_\gamma A_\delta}\delta A_\delta\right) = 0 \quad \text{as} \quad \frac{\partial \mathcal{L}}{\partial \partial_\gamma A_\delta}\delta A_\delta\Big|_{\partial V} = 0. \tag{E.393}$$

Substitution of the Lagrange-density $\mathcal{L}$ is rather straightforward for the ∂A_δ-derivative,

$$\frac{\partial \mathcal{L}}{\partial A_\delta} = \frac{16\pi}{c}\frac{\partial A_\mu}{\partial A_\delta}\jmath^\mu = \frac{16\pi}{c}\delta^\delta_\mu \jmath^\mu = \frac{16\pi}{c}\jmath^\delta \tag{E.394}$$

but involves handling many indices for the derivatives with respect to $\partial_\gamma A_\delta$.

Instead, one can rewrite the derivative as

$$\frac{\partial}{\partial\partial_\gamma A_\delta} = \frac{\partial F_{\sigma\tau}}{\partial\partial_\gamma A_\delta}\frac{\partial}{\partial F_{\sigma\tau}} = \frac{\partial(\partial_\sigma A_\tau - \partial_\tau A_\sigma)}{\partial\partial_\gamma A_\delta}\frac{\partial}{\partial F_{\sigma\tau}} = \left(\frac{\partial\partial_\sigma A_\tau}{\partial\partial_\gamma A_\delta} - \frac{\partial\partial_\tau A_\sigma}{\partial\partial_\gamma A_\delta}\right)\frac{\partial}{\partial F_{\sigma\tau}} =$$
$$\left(\delta_\sigma^\gamma\delta_\tau^\delta - \delta_\tau^\gamma\delta_\sigma^\delta\right)\frac{\partial}{\partial F_{\sigma\tau}} = \frac{\partial}{\partial F_{\gamma\delta}} - \frac{\partial}{\partial F_{\delta\gamma}} = 2\frac{\partial}{\partial F_{\gamma\delta}}. \quad \text{(E.395)}$$

In both cases, the elementary derivatives give either 0 or 1 according to

$$\frac{\partial\partial_\mu A_\nu}{\partial\partial_\gamma A_\delta} = \delta_\mu^\gamma\delta_\nu^\delta \quad \text{as well as} \quad \frac{\partial A_\mu}{\partial A_\gamma} = \delta_\mu^\gamma, \quad \text{(E.396)}$$

because the field components and their derivatives into the different coordinate directions are all independent. The derivatives $\partial F_{\alpha\beta}/\partial F_{\mu\nu}$ of the field tensor with respect to itself are slightly more involved, because of the antisymmetry of both $F_{\alpha\beta}$ and $F_{\mu\nu}$. The necessary (anti-)symmetrisation reads

$$\frac{\partial F_{\alpha\beta}}{\partial F_{\mu\nu}} = \frac{1}{4}\left(\delta_\alpha^\mu\delta_\beta^\nu - \delta_\alpha^\nu\delta_\beta^\mu - \delta_\beta^\mu\delta_\alpha^\nu + \delta_\beta^\nu\delta_\alpha^\mu\right) = \frac{1}{2}\left(\delta_\alpha^\mu\delta_\beta^\nu - \delta_\alpha^\nu\delta_\beta^\mu\right) \quad \text{(E.397)}$$

with a simplification due to the pairwise identity of terms.

Then, application of the differentiations to the kinetic term required by the Euler-Lagrange equation yields:

$$\frac{\partial\mathcal{L}}{\partial\partial_\gamma A_\delta} = 2\frac{\partial}{\partial F_{\gamma\delta}}\eta^{\alpha\mu}\eta^{\beta\nu}F_{\alpha\beta}F_{\mu\nu} = 2\eta^{\alpha\mu}\eta^{\beta\nu}\left(\frac{\partial F_{\alpha\beta}}{\partial F_{\gamma\delta}}F_{\mu\nu} + F_{\alpha\beta}\frac{\partial F_{\mu\nu}}{\partial F_{\gamma\delta}}\right) =$$
$$\eta^{\alpha\mu}\eta^{\beta\nu}\left(\left(\delta_\alpha^\gamma\delta_\beta^\delta - \delta_\alpha^\delta\delta_\beta^\gamma\right)F_{\mu\nu} + F_{\alpha\beta}\left(\delta_\mu^\gamma\delta_\nu^\delta - \delta_\mu^\delta\delta_\nu^\gamma\right)\right) = 4\eta^{\gamma\mu}\eta^{\delta\nu}F_{\mu\nu}. \quad \text{(E.398)}$$

Collection of all results suggests as the field equation the relation

$$\partial_\gamma\frac{\partial\mathcal{L}}{\partial_\gamma A_\delta} = 4\partial_\gamma\eta^{\gamma\mu}\eta^{\delta\nu}F_{\mu\nu} = \frac{\partial\mathcal{L}}{\partial A_\delta} = \frac{16\pi}{c}\jmath^\delta \quad \rightarrow \quad \eta^{\gamma\mu}\partial_\gamma F_{\mu\nu} = \frac{4\pi}{c}\eta_{\delta\nu}\jmath^\delta \quad \text{(E.399)}$$

which one immediately recognises as the inhomogeneous Maxwell-equation in vacuum, with the divergence of the field tensor being equated to the source. The invariance of the Lagrangian description and the covariance of the field equation is summarised by this diagram,

$$\begin{array}{ccccc} S = \int_V d^4x & \eta^{\alpha\mu}\eta^{\beta\nu}F_{\alpha\beta}F_{\mu\nu} & + & \frac{16\pi}{c}A_\mu\jmath^\mu & \\ & \Big\downarrow{\scriptstyle \delta S=0} & & \Big\downarrow{\scriptstyle \delta S=0} & \\ & \eta^{\alpha\mu}\eta^{\beta\nu}\partial_\mu F_{\alpha\beta} & - & \frac{4\pi}{c}\jmath^\nu & = 0, \end{array} \quad \text{(E.400)}$$

and substitution of the expression for $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ finally leads to

$$\eta^{\gamma\mu}\partial_\gamma F_{\mu\nu} = \eta^{\gamma\mu}\partial_\gamma\left(\partial_\mu A_\nu - \partial_\nu A_\mu\right) = \underbrace{\eta^{\gamma\mu}\partial_\gamma\partial_\mu A_\nu}_{=\Box A_\nu} - \partial_\nu \underbrace{\eta^{\gamma\mu}\partial_\gamma A_\mu}_{=0} = \frac{4\pi}{c}\eta_{\delta\nu}\jmath^\delta, \qquad \text{(E.401)}$$

which clearly demonstrates a covariant wave equation with the potential A_ν as a linear form related to the source $\eta_{\delta\nu}\jmath^\delta$, a vector converted into a linear form, with the assumption of Lorenz-gauge $\eta^{\gamma\mu}\partial_\gamma A_\mu = 0$ for making the second term disappear.

Formal application of the variation to the action integral would be an expression

$$\delta S = \delta\int_V d^4x\ \eta^{\alpha\mu}\eta^{\beta\nu}F_{\alpha\beta}F_{\mu\nu} = 2\int_V d^4x\ \eta^{\alpha\mu}\eta^{\beta\nu}F_{\alpha\beta}\ \delta F_{\mu\nu} = 0 \qquad \text{(E.402)}$$

where one can interpret the requirement of Hamilton's principle, namely $\delta S = 0$, as an orthogonality condition between $F_{\alpha\beta}$ and its variation $\delta F_{\alpha\beta}$, as a modern embodiment of the principle of virtual work.

It might be an interesting endeavour to understand how exactly the structure $\eta^{\alpha\mu}\eta^{\beta\nu}F_{\alpha\beta}F_{\mu\nu}$ in the kinetic term of the Lagrange density is to be interpreted, beyond the fact that it is a quadratic Lorentz-invariant. With the antisymmetry of $F_{\mu\nu} = -F_{\nu\mu}$ one can write

$$S = \int_V d^4x\ \eta^{\alpha\mu}\eta^{\beta\nu}F_{\alpha\beta}F_{\mu\nu} = \int_V d^4x\ \eta^{\alpha\mu}\eta^{\beta\nu}\frac{1}{2}\left(F_{\alpha\beta}F_{\mu\nu} - F_{\alpha\beta}F_{\nu\mu}\right) \qquad \text{(E.403)}$$

which becomes, after renaming the indices $\mu \leftrightarrow \nu$ in the second term,

$$S = \frac{1}{2}\int_V d^4x\ \eta^{\alpha\mu}\eta^{\beta\nu}F_{\alpha\beta}F_{\mu\nu} - \eta^{\alpha\nu}\eta^{\beta\mu}F_{\alpha\beta}F_{\mu\nu} = \int_V d^4x\ \frac{\eta^{\alpha\mu}\eta^{\beta\nu} - \eta^{\alpha\nu}\eta^{\beta\mu}}{2}F_{\alpha\beta}F_{\mu\nu} \qquad \text{(E.404)}$$

which can be written as

$$S = \int_V d^4x\ X^{\alpha\beta\mu\nu}F_{\alpha\beta}F_{\mu\nu} \quad \text{with a measure} \quad X^{\alpha\beta\mu\nu} = \frac{\eta^{\alpha\mu}\eta^{\beta\nu} - \eta^{\alpha\nu}\eta^{\beta\mu}}{2} \qquad \text{(E.405)}$$

as tensor with two antisymmetric index pairs (α,μ) and (β,ν). Perhaps the index structure reminds you of the Grassmann-relation $\gamma_{il}\epsilon^{ijk}\epsilon^{lmn} = \gamma^{jm}\gamma^{kn} - \gamma^{jn}\gamma^{km}$ of a square of a vector product, which quantifies the area spanned by two vectors: In some sense, the same happens in the Lagrange density, which is an abstract measure of the area between ∂_μ and A_ν, induced by the metric $\eta^{\mu\nu}$.

E.6.4 Maxwell field in matter

For the behaviour of the Maxwell field in matter a suitable starting point could be the action

$$S = \int_V d^4x\ F_{\mu\nu}G^{\mu\nu} + \frac{16\pi}{c}A_\mu\jmath^\mu \qquad \text{(E.406)}$$

where the Lorentz invariant in matter constitutes the kinetic term. Expressed in terms of the fields it reads $F_{\mu\nu}G^{\mu\nu} = E_iD^i - H_iB^i$. On possible pathway to carry out the variation and to perform the derivatives with respect to A_γ and $\partial_\gamma A_\delta$ is provided by the constitutive relation,

$$G^{\alpha\beta} = X^{\alpha\beta\mu\nu} F_{\mu\nu}, \tag{E.407}$$

that relates the fields D^i and H_i contained in $G^{\mu\nu}$ to the vacuum fields E_i and B^i in $F_{\mu\nu}$. After all, only $F_{\mu\nu}$ follows from the derivation of the potential A_μ and is accessible to variation. As both tensors are antisymmetric, $X^{\alpha\beta\mu\nu}$ has to be antisymmetric in each index pair, $X^{\alpha\beta\mu\nu} = -X^{\alpha\beta\nu\mu} = -X^{\beta\alpha\mu\nu} = X^{\beta\alpha\nu\mu}$. Then, the action integral reads

$$S = \int_V d^4x\, X^{\alpha\beta\mu\nu} F_{\alpha\beta} F_{\mu\nu} + \frac{16\pi}{c} A_\mu \jmath^\mu \tag{E.408}$$

Variation proceeds as in the previous case, as

$$\frac{\partial \mathcal{L}}{\partial \partial_\gamma A_\delta} = 2\frac{\partial}{\partial F_{\gamma\delta}} X^{\alpha\beta\mu\nu} F_{\alpha\beta} F_{\mu\nu} = 2X^{\alpha\beta\mu\nu}\left(\frac{\partial F_{\alpha\beta}}{\partial F_{\gamma\delta}} F_{\mu\nu} + F_{\alpha\beta}\frac{\partial F_{\mu\nu}}{\partial F_{\gamma\delta}}\right) =$$
$$X^{\alpha\beta\mu\nu}\left(\left(\delta^\gamma_\alpha\delta^\delta_\beta - \delta^\delta_\alpha\delta^\gamma_\beta\right) F_{\mu\nu} + F_{\alpha\beta}\left(\delta^\gamma_\mu\delta^\delta_\nu - \delta^\delta_\mu\delta^\gamma_\nu\right)\right) = 4X^{\gamma\delta\mu\nu} F_{\mu\nu} = 4G^{\gamma\delta}. \tag{E.409}$$

Combined with the previous result on the derivative with respect to A_δ, the Euler-Lagrange equation yields:

$$\partial_\gamma \frac{\partial \mathcal{L}}{\partial_\gamma A_\delta} = 4\partial_\gamma X^{\gamma\delta\mu\nu} F_{\mu\nu} = 4\partial_\gamma G^{\gamma\delta} = \frac{\partial \mathcal{L}}{\partial A_\delta} = \frac{16\pi}{c}\jmath^\delta \quad\rightarrow\quad \partial_\gamma G^{\gamma\delta} = \frac{4\pi}{c}\jmath^\delta, \tag{E.410}$$

which is in fact the Maxwell field equation in matter. While the Lagrange density eqn. (E.406) is the source of the field equation and links ultimately of the fields D^i and H_i to the sources, the dynamics of the dual field tensor $\tilde{F}^{\mu\nu}$ with E_i and B^i is already fixed by the Bianchi-identity.

E.7 *Optics*

It is fair to say that the covariant constitutive relation falls back in isotropic media on the antisymmetrised metric,

$$X^{\alpha\beta\mu\nu} = \frac{\eta^{\alpha\mu}\eta^{\beta\nu} - \eta^{\alpha\nu}\eta^{\beta\mu}}{2} \tag{E.411}$$

possibly with $(\epsilon\mu)$ as a prefactor in isotropic media in the spatial part of the metric. In fact, in isotropic media one gets for the effective metric

⚠ Please keep in mind that $D^i = \epsilon^{ij}E_j \propto \epsilon\gamma^{ij}E_j$ and $H_i = \mu_{ij}B^j \propto \mu\gamma_{ij}B^j$, so the mapping from the vacuum fields E_i, B^i in $F_{\mu\nu}$ to D^i, H_i in $G^{\mu\nu}$ picks up a factor of $\epsilon\mu = n^2$ in the spatial components.

$$\eta_{\mu\nu} = \begin{pmatrix} 1 & & & \\ & -n^2 & & \\ & & -n^2 & \\ & & & -n^2 \end{pmatrix} \quad\leftrightarrow\quad \eta^{\mu\nu} = \begin{pmatrix} 1 & & & \\ & -n^{-2} & & \\ & & -n^{-2} & \\ & & & -n^{-2} \end{pmatrix} \tag{E.412}$$

with the refractive index $n = \sqrt{\epsilon\mu}$.

In this particular case, a plane-wave ansatz $\exp(\pm \mathrm{i} k_\alpha x^\alpha)$ yields a modified null-condition

$$\eta^{\mu\nu} k_\mu k_\nu = 0 = \left(\frac{\omega}{c}\right)^2 - \frac{k^2}{n^2} \quad \rightarrow \quad \omega = \pm \frac{ck}{n} \tag{E.413}$$

Consequently, the velocities are diminished by the refractive index n,

$$\upsilon_{\mathrm{gr}} = \frac{\mathrm{d}\omega}{\mathrm{d}k} = \frac{c}{n} = \frac{\omega}{k} = \upsilon_{\mathrm{ph}} \tag{E.414}$$

and the light cone becomes narrower by the factor n. As constitutive tensor $X^{\alpha\beta\mu\nu}$ is composed of the two contributions, namely the permissivity tensor ϵ^{ij} and the permeability tensor μ^{ij}, on the spatial components are affected: This effectively means that in a medium, the wave length $\lambda = 2\pi/k$ is affected by the refractive index and not the angular frequency ω.

The notion, that wave length changes in a medium according to $\lambda \rightarrow n\lambda$ with the refractive index n, paving the way for Fermat's principle for refraction: The optical path length is effectively increased by the same factor of n. The *spatial* distance between two point A and B is given by

$$s = \int_{\mathrm{A}}^{\mathrm{B}} \mathrm{d}s \quad \rightarrow \quad \int_{\mathrm{A}}^{\mathrm{B}} \mathrm{d}s\, n = \int_{\mathrm{A}}^{\mathrm{B}} \mathrm{d}\lambda \sqrt{\gamma_{ij} \frac{\mathrm{d}x^i}{\mathrm{d}\lambda} \frac{\mathrm{d}x^j}{\mathrm{d}\lambda}}\, n(x^i) \tag{E.415}$$

and is extremised according to $\delta s = 0$ to yield the actual light path, technically through application of the Euler-Lagrange equation albeit for a rather unusual form of the Lagrange-function

$$\mathcal{L} = \sqrt{\gamma_{ij} \frac{\mathrm{d}x^i}{\mathrm{d}\lambda} \frac{\mathrm{d}x^j}{\mathrm{d}\lambda}}\, n(x^i) \tag{E.416}$$

with no additive separation in a kinetic and potential part. Instead, in applying the Euler-Lagrange equation (abbreviating $\dot{x}^i = \mathrm{d}x^i/\mathrm{d}\lambda$)

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} \frac{\partial \mathcal{L}}{\partial \dot{x}^a} = \frac{\partial \mathcal{L}}{\partial x^a} \tag{E.417}$$

one needs to be careful because after the $\partial \dot{x}^i$-differentiation, $\mathcal{L}$ still depends on x^i, which yields additional terms involving $\dot{x}^i$ in the $\mathrm{d}\lambda$-differentiation, in particular the gradient of the refractive index $\mathrm{d}n/\mathrm{d}\lambda = \partial_a n\, \dot{x}^a$. The first two derivatives are

$$\frac{\partial \mathcal{L}}{\partial x^a} = \sqrt{\gamma_{ij} \dot{x}^i \dot{x}^j}\, \partial_a n, \quad \text{followed by} \quad \frac{\partial \mathcal{L}}{\partial \dot{x}^a} = \frac{n \gamma_{ai} \dot{x}^i}{\sqrt{\gamma_{mn} \dot{x}^m \dot{x}^n}}, \tag{E.418}$$

but increase dramatically in their complexity in the $\mathrm{d}\lambda$-differentiaton. Ultimately, these equations lead to the concept of Lagrangian optics and can only be solved sensibly either through numerical methods or in approximations. While we commonly assumed homogeneous media, the formalism is still applicable in the limit of geometric optics where the scale on which n changes is large compared to the scale on which the fields vary, i.e. the wave length λ.

While it is clear that the metric in an anisotropic medium can show different light propagation speeds along the three coordinate directions, the constitutive tensor $X^{\alpha\beta\mu\nu}$: The wave equation in the most general case reads

$$\partial_\alpha G^{\alpha\beta} = X^{\alpha\beta\mu\nu}\partial_\alpha F_{\mu\nu} = 2X^{\alpha\beta\mu\nu}\partial_\alpha\partial_\mu A_\nu = 0, \tag{E.419}$$

which suggest for an ansatz $A_\mu \propto A_\mu^{(0)}\exp(\pm ik_\gamma x^\gamma)$, with an amplitude $A_\mu^{(0)}$ that contains information about polarisation. Then, the null-condition reads

$$X^{\alpha\beta\mu\nu}A_\nu^{(0)}\, k_\alpha k_\mu = 0 \tag{E.420}$$

and is effectively a polarisation-dependent dispersion relation, with differences in propagation speeds even into the same direction for different polarisations: This phenomenon is known as birefringence, and can be observed in e.g. calcite crystals.

Please note how the null-condition requires a summation over the pair (α, μ) and not (μ, ν) *which would be trivially zero.*

E.8 *Gauge-invariance and charge conservation*

Gauge-invariance of the term $\eta^{\alpha\mu}\eta^{\beta\nu}F_{\alpha\beta}F_{\mu\nu}$ is clearly given, as $F_{\mu\nu}$ does not change under gauge-transformation anyways. But it is interesting to see how gauge-invariance is recovered in the entire Lagrange-formalism. In fact, with $A_\mu \rightarrow A_\mu + \partial_\mu\chi$ one obtains

$$S = \int_V d^4x\, \mathcal{L} \rightarrow S + \frac{16\pi}{c}\int_V d^4x\, \partial_\mu\chi\, \jmath^\mu = S + \frac{16\pi}{c}\int_V d^4x\, \left[\partial_\mu(\chi\jmath^\mu) - \chi\partial_\mu\jmath^\mu\right] \tag{E.421}$$

where $\partial_\mu\jmath^\mu = 0$ due to continuity of the charge density. The first term can be converted into a surface integral with the Gauß-theorem,

$$S \rightarrow S + \frac{16\pi}{c}\int_{\partial V} dS_\mu\, (\chi\jmath^\mu) = S \tag{E.422}$$

i.e. one recovers gauge invariance when assuming a localised charge distribution: moving the integration surface ∂V out leads to $\chi\jmath^\mu$ vanishing faster than ∂V increases, and consequently, the integral approaches zero. Hence, the action is gauge invariant if charge is conserved. To show the opposite is impossible for our current understanding of charge as a source of the electromagnetic field and requires a more detailed model for the charge-carrying matter in the form of a quantum theory.

Please note that there are different concepts at play to have terms vanish in S (locality of the charge distribution) and in δS (fixed variation on boundary).

E.9 *Conservation of energy and momentum*

E.9.1 Scalar field on a Lorentz background

The Lagrange-density of the electromagnetic field does not depend explicitly on the coordinates x^μ, meaning that it is truly universal: The way in which the field is coupled to its charges and the internal dynamics is the same everywhere and at every time. As a consequence of the translation invariance along the ct- and x^i-coordinates, energy and momentum are conserved, which we should derive first for a scalar field φ. There, the Lagrange-density is given by $\mathcal{L}(\varphi, \partial_\mu\varphi)$ but *not* by $\mathcal{L}(\phi, \partial_\mu\varphi, x^\mu)$. The Euler-Lagrange equations follow from the variation of the action S

$$S = \int_V d^4x\, \mathcal{L}(\varphi, \partial_\mu\varphi) \quad \rightarrow \quad \delta S = \int_V d^4x \left(\frac{\partial\mathcal{L}}{\partial\varphi} - \partial_\mu \frac{\partial\mathcal{L}}{\partial\partial_\mu\varphi}\right)\delta\varphi = 0 \tag{E.423}$$

such that Hamilton's principle $\delta S = 0$ implies

$$\partial_\mu \frac{\partial\mathcal{L}}{\partial\partial_\mu\varphi} = \frac{\partial\mathcal{L}}{\partial\varphi} \tag{E.424}$$

If the Lagrange density $\mathcal{L}$ depends only on the fields themselves and not on the position, meaning the functional principle of the field theory as defined by $\mathcal{L}$ is the same everywhere and at very time, there is only one way in which the Lagrange density can change is moving through spacetime to a new point where the fields and their derivatives are different: The fields themselves need to change. This implies that under an infinitesimal shift in the coordinates into the direction ϵ^μ,

$$x^\mu \rightarrow x^\mu + \epsilon^\mu, \tag{E.425}$$

one expects a variation of the field $\delta\varphi$ to be

$$\delta\varphi = \varphi(x^\mu + \epsilon^\mu) - \varphi(x^\mu) = \epsilon^\alpha \partial_\alpha \varphi \tag{E.426}$$

and the corresponding variation of the Lagrange density would become

$$\delta\mathcal{L} = \epsilon^\alpha \partial_\alpha \mathcal{L} \tag{E.427}$$

On the other hand, the variation of the Lagrange density is given by

$$\delta\mathcal{L} = \frac{\partial\mathcal{L}}{\partial\varphi}\delta\varphi + \frac{\partial\mathcal{L}}{\partial\partial_\mu\varphi}\delta\partial_\mu\varphi = \left(\frac{\partial\mathcal{L}}{\partial\varphi}\delta\varphi - \partial_\mu \frac{\partial\mathcal{L}}{\partial\partial_\mu\varphi}\right)\delta\varphi + \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial\partial_\mu\varphi}\delta\varphi\right) \tag{E.428}$$

using the Leibnitz-rule. As the physical fields fulfil the Euler-Lagrange equation in the first term, only the second term remains, implying

$$\delta\mathcal{L} = \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial\partial_\mu\varphi}\delta\varphi\right) \tag{E.429}$$

Assembling the final expression from the variation $\delta\mathcal{L}$ in eqn. (E.429) with the expression eqn. (E.427) and the variation $\delta\varphi$ in eqn. (E.426) leads to

$$\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial\partial_\mu\varphi}\delta\varphi\right) - \epsilon^\alpha \partial_\alpha \mathcal{L} = 0 \tag{E.430}$$

such that, using $\partial_\alpha = \delta^\mu_\alpha \partial_\mu$,

$$\epsilon^\alpha\, \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial\partial_\mu\varphi}\partial_\alpha\varphi - \delta^\mu_\alpha \mathcal{L}\right) = 0 \tag{E.431}$$

implying that there is a covariant divergence which vanishes,

$$\partial_\mu T_\alpha{}^\mu = 0 \tag{E.432}$$

with the energy-momentum tensor $T_\alpha{}^\mu$

$$T_\alpha{}^\mu = \frac{\partial\mathcal{L}}{\partial\partial_\mu\varphi}\partial_\alpha\varphi - \delta_\alpha^\mu\mathcal{L}. \tag{E.433}$$

Effectively, this suggests a multidimensional Legendre-transform with the canonical field momentum π^μ

$$\pi^\mu = \frac{\partial\mathcal{L}}{\partial\partial_\mu\varphi} \quad \text{such that} \quad T_\alpha{}^\mu = \pi^\mu\,\partial_\alpha\varphi - \delta_\alpha^\mu\mathcal{L}(\varphi,\pi^\mu) \tag{E.434}$$

where the structural similarity to the relation $\mathcal{H} = p_i\dot{x}^i - \mathcal{L}$ from classical mechanics is quite apparent.

If the Lagrange density had an additional dependence on the coordinates x^μ, it's variation (E.428) when transitioning form x^μ to $x^\mu + \epsilon^\mu$ would not only be caused by the different field amplitudes and their derivatives, but there would be a new term Q_α,

$$\delta\mathcal{L} = \epsilon^\alpha\partial_\alpha\mathcal{L}(\text{field variation}) + \epsilon^\alpha Q_\alpha(\text{explicit coordinate dependence}) \tag{E.435}$$

where this new term is effectively a source term to the otherwise vanishing continuity equation,

$$\partial_\mu T_\alpha{}^\mu = Q_\alpha. \tag{E.436}$$

The identification of $T_\alpha{}^\mu$ with the energy-momentum tensor becomes sensible for the case of a standard Lagrange-density for a scalar field φ,

$$\mathcal{L}(\varphi,\partial_\mu\varphi) = \frac{\eta^{\mu\nu}}{2}\partial_\mu\varphi\partial_\nu\varphi - V(\varphi) \tag{E.437}$$

with a self-interaction potential $V(\varphi)$ that would contain e.g. a coupling to sources. Variation by substitution into the Euler-Lagrange equation yields directly the Klein-Gordon equation

$$\Box\varphi = -\frac{\partial V}{\partial\varphi} \quad \text{because} \quad \pi^\mu = \frac{\partial\mathcal{L}}{\partial_\mu\varphi} = \eta^{\mu\nu}\partial_\nu\varphi \quad \text{and} \quad \frac{\partial\mathcal{L}}{\partial\varphi} = -\frac{\partial V}{\partial\varphi} \tag{E.438}$$

with the next differentiation generating $\Box\varphi = \partial_\mu\pi^\mu = \eta^{\mu\nu}\partial_\mu\partial_\nu\varphi$. Then, the tensor $T^\mu{}_\alpha$ becomes

$$T_\alpha{}^\mu = \pi^\mu\partial_\alpha\varphi - \delta_\alpha^\mu\mathcal{L}(\varphi,\pi^\mu) = \eta^{\mu\nu}\partial_\nu\varphi\partial_\alpha\varphi - \delta_\alpha^\mu\frac{\eta^{\gamma\delta}}{2}\partial_\gamma\varphi\partial_\delta\varphi + \delta_\alpha^\mu V(\varphi) \tag{E.439}$$

with the sign change in front of $V(\varphi)$ which is typical for the Legendre transform.

E.9.2 Maxwell field on a Lorentz background

There is a very important detail in the derivation of the energy-momentum tensor of the electromagnetic field, which otherwise proceeds exactly as in the case of the scalar field φ: When shifting the potential to compute δA_δ one should not use the derivative $\partial_\alpha A_\delta$ for forming $\delta A_\delta = \epsilon^\alpha \partial_\alpha A_\delta$ because it is not gauge-invariant. Rather, the variation should be given by the antisymmetrised form,

$$\delta A_\delta = \epsilon^\alpha \partial_\alpha A_\delta \to \epsilon^\alpha \left(\partial_\alpha A_\delta - \partial_\delta A_\alpha\right) = \epsilon^\alpha F_{\alpha\delta} \tag{E.440}$$

as the Faraday tensor $F_{\alpha\delta}$ is the gauge-invariant derivative of A_δ . The variation in the Lagrange-density becomes formally

⚠ *watch out for gauge-independence in the derivative*

$$\delta\mathcal{L} = \epsilon^\alpha \partial_\alpha \mathcal{L} \tag{E.441}$$

but expressed in terms of the fields, by virtue of the Leibnitz-rule,

$$\delta\mathcal{L} = \frac{\partial\mathcal{L}}{\partial A_\delta}\delta A_\delta + \frac{\partial\mathcal{L}}{\partial\partial_\gamma A_\delta}\delta\partial_\gamma A_\delta = \left(\frac{\partial\mathcal{L}}{\partial A_\delta} - \frac{\partial\mathcal{L}}{\partial\partial_\gamma A_\delta}\right)\delta A_\delta + \partial_\gamma\left(\frac{\partial\mathcal{L}}{\partial\partial_\gamma A_\delta}\delta A_\delta\right), \tag{E.442}$$

where the first bracket disappears as it fulfils the Euler-Lagrange equation, that appears after the usual replacement $\delta\partial_\gamma A_\delta = \partial_\gamma \delta A_\delta$. The divergence in the second term can be reformulated as

$$\partial_\gamma\left(\frac{\partial\mathcal{L}}{\partial\partial_\gamma A_\delta}\delta A_\delta\right) = \epsilon^\alpha\partial_\gamma\left(\frac{\partial\mathcal{L}}{\partial\partial_\gamma A_\delta}F_{\alpha\delta}\right) = \delta\mathcal{L} = \epsilon^\alpha\partial_\alpha\mathcal{L} = \epsilon^\alpha\delta^\gamma_\alpha\partial_\gamma\mathcal{L} = \epsilon^\alpha\partial_\gamma\delta^\gamma_\alpha\mathcal{L} \tag{E.443}$$

so that the combination of the second and the sixth term suggest, as the shift ϵ^α was arbitrary:

$$\partial_\gamma\left(\frac{\partial\mathcal{L}}{\partial\partial_\gamma A_\delta}F_{\alpha\delta} - \delta^\gamma_\alpha\mathcal{L}\right) = 0, \tag{E.444}$$

i.e. a conservation law for the energy momentum tensor,

$$\partial_\gamma {T_\alpha}^\gamma = 0, \quad \text{with} \quad {T_\alpha}^\gamma = \frac{\partial\mathcal{L}}{\partial\partial_\gamma A_\delta}F_{\alpha\delta} - \delta^\gamma_\alpha\mathcal{L}. \tag{E.445}$$

The energy-momentum tensor ${T_\mu}^\nu$ is the relativistic generalisation of the Maxwell-tensor ${T_i}^j$, which makes up the spatial part of it. In vacuum, it is symmetric, ${T_\mu}^\nu = {T_\nu}^\mu$ and traceless, ${T_\mu}^\mu = \eta^{\mu\nu}T_{\mu\nu} = 0$: The physical meaning of this is not straightforward to understand, but essentially corresponds to the fact that there is no mass associated with the photons, i.e with excitations of the electromagnetic field. The components of ${T_\mu}^\nu$ contain the energy density, ${T_t}^t = E_i D^i - H_i B^i = w_{\mathrm{el}} + w_{\mathrm{mag}}$ and the Poynting-vector, $4\pi/c\, P^i = {T_t}^i$. In particular, the formulation of the Poynting-law would become $\partial_\mu {T_t}^\mu = \partial_t(w_{\mathrm{el}} + w_{\mathrm{mag}}) + \partial_i P^i = 0$.

Perhaps it's a weird and funny thought that 👁 Kirchhoff's 🖋 mesh and knot rules for electric circuits are essentially reflections of the coordinate-independence of the Lagrange-function $\mathcal{L}$ giving rise to energy conservation, and of the gauge invariance of $\mathcal{L}$ compatible with charge conservation. And as a last remark in this context I would

		$\mathcal{C}$	$\mathcal{PT}$	$\mathcal{CPT}$
derivative	∂_μ	+	−	−
electric 4-current	$\jmath^\mu$	−	+	−
magnetic 4-current	$\imath^\mu$	−	−	+
Faraday tensor	$F_{\mu\nu}$	−	+	−
field tensor	$\tilde{F}^{\mu\nu}$	−	−	+

Table 2: Summary of the behaviour of all fields and sources in extended electrodynamics with electric and magnetic sources.

like to add that the construction with the infinitesimal shift of the Lagrange-density is in some sense a trick: Actually, one would like to construct a gradient $\partial\mathcal{L}$ of $\mathcal{L}$ which is caused by the fact that the fields and their derivatives have gradients. But one usually works with the convention that partial derivatives of functionals only apply to their explicit dependence on the coordinates, not their "indirect" position-dependence through the fields (and their derivatives). With this convention, $\partial_\mu\mathcal{L}$ would be zero, even though of course $\mathcal{L}$ changes as a function of position, because the fields do change. On a larger scale, the derivation of a conserved energy-momentum tensor from the Lagrange-density or the action is an example of a Lie-derivative.

E.10 *Maxwell's equations under discrete symmetries, revisited*

The behaviour of the Maxwell-equations under the three discrete symmetries charge conjugation $\mathcal{C}$, parity inversion $\mathcal{P}$ and time reversal $\mathcal{P}$ was already the subject of Sect. A.7, but can be extended to deal with covariant objects like $G^{\mu\nu}$, $\tilde{F}^{\mu\nu}$ or ∂_μ in a straightforward way. As before, we will treat the general case with electric charges $\jmath^\mu$ as well as magnetic charges $\imath^\mu$:

$$\partial_\mu G^{\mu\nu} = +\frac{4\pi}{c}\jmath^\nu \quad \text{and} \quad \partial_\mu \tilde{F}^{\mu\nu} = -\frac{4\pi}{c}\imath^\nu \tag{E.446}$$

In both cases the antisymmetry of the field tensors $G^{\mu\nu}$ and $\tilde{F}^{\mu\nu}$ makes sure that the currents are conserved, i.e. $\partial_\nu \jmath^\nu = 0$ and $\partial_\nu \imath^\nu = 0$.

∂_μ, combining spatial and temporal derivatives, transforms sensibly only under the combined $\mathcal{PT}$-operation: Clearly, $\mathcal{PT}x^\mu = -x^\mu$ and in consequence, $\mathcal{PT}\partial_\mu = -\partial_\mu$. The electric 4-current $\jmath^\mu$ transforms under $\mathcal{PT}$ like a velocity, $\mathcal{PT}\jmath^\mu = \jmath^\mu$, and under $\mathcal{C}$ as $\mathcal{C}\jmath^\mu = -\jmath^\mu$, and therefore $\mathcal{CPT}\jmath^\mu = -\jmath^\mu$ under the full $\mathcal{CPT}$ transform. Magnetic charges, however are pseudoscalar such that $\mathcal{PT}\imath^\mu = -\imath^\mu$, but in fact the additional minus sign does not matter when considering the continuities $\partial_\mu \jmath^\mu = 0$ and $\partial_\mu \imath^\mu = 0$.

Please note that one can only invoke arguments that relate $G^{\mu\nu}$ to the potential A_μ if there are no magnetic charges and duality is broken. It will be sufficient to consider the Faraday tensor $F_{\mu\nu}$ as its properties are identical to $G^{\mu\nu}$ because the two are related in a linear way by a mere prefactor. If there are only electric charges, $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ suggests that $\mathcal{PT}F_{\mu\nu} = -F_{\mu\nu}$, using the wave equation $\Box A_\mu = 4\pi/c\ \eta_{\mu\nu}\jmath^\nu$, such that A_μ inherits its properties from $\jmath^\mu$, in summary $\mathcal{PT}A_\mu = +A_\mu$. This is consistent with the field equation $\partial_\mu G^{\mu\nu} = 4\pi/c\ \jmath^\mu$, as the minus signs brought in by ∂_μ and $G^{\mu\nu}$ cancel. Similarly, $\mathcal{PT}\tilde{F}^{\mu\nu} = +\tilde{F}^{\mu\nu}$ to reflect the plus-sign in $\mathcal{PT}\imath^\mu = +\imath^\mu$.

E.11 *Links to particle physics*

E.11.1 Axions and pseudoscalar particles

There is a second quadratic field invariant, $F_{\mu\nu}\tilde{F}^{\mu\nu} \propto E_i B^i$, which is pseudo-scalar: despite being "just" a number, it changes its sign under application of parity-transforms $\mathcal{P}$ and time reversal $\mathcal{T}$. This is the reason why we disregarded this particular term, despite being quadratic, as a contender for the Lagrange density $\mathcal{L}$ for electrodynamics. But multiplying with a field θ which itself is pseudoscalar, would amend this problem:

$$\mathcal{L} = \frac{\eta^{\alpha\mu}\eta^{\beta\nu}}{4} F_{\alpha\beta}F_{\mu\nu} + \alpha\theta F_{\mu\nu}\tilde{F}^{\mu\nu} + \frac{4\pi}{c} A_\mu \jmath^\mu + \frac{\eta^{\mu\nu}}{2}\partial_\mu\theta\,\partial_\nu\theta - V(\theta) \tag{E.447}$$

with a coupling strength α. This axion field θ needs its own dynamics and interacts with itself through the potential $V(\theta)$. Looking at the Taylor-expansion of $V(\theta)$ one can only admit even powers

$$V(\theta) = \sum_{n=0} \frac{\alpha_{2n}}{(2n)!}\theta^{2n} \tag{E.448}$$

as only those are invariant under parity transform: Essentially, this is a very strong restriction on the form of the potential for self-interaction of the axion field: it is necessarily an even function. Please note that a mass term of the type

$$V(\theta) = \frac{m^2}{2}\theta^2 \tag{E.449}$$

would be naturally contained in the interaction potential $V(\theta)$ even in the restriction to parity positive terms, by setting $\alpha_2 = m^2$ for $n = 1$.

Variation of the Lagrange-density with respect to A_μ yields an extension to the Maxwell field-equation, and the variation with respect to θ a corresponding equation of motion for θ, which is coupled to $F_{\mu\nu}$, i.e. a modified field equation

$$\eta^{\alpha\mu}\eta^{\nu\beta}\partial_\alpha F_{\mu\beta} = \frac{4\pi}{c}\jmath^\nu + \alpha\partial_\mu\left(\theta\tilde{F}^{\mu\nu}\right) = \frac{4\pi}{c}\jmath^\nu + \alpha\partial_\mu\theta\cdot\tilde{F}^{\mu\nu} \tag{E.450}$$

because $\partial_\mu\tilde{F}^{\mu\nu} = 0$, if duality is properly broken, and alongside a dynamical equation for θ,

$$\Box\theta = \alpha F_{\mu\nu}\tilde{F}^{\mu\nu} - \frac{dV}{d\theta} \tag{E.451}$$

Therefore, θ obeys a wave equation that is coupled to the electromagnetic field and driven by the potential gradient $-dV/d\theta$. Experiments with axions are always great and fascinating, for instance light through wall-type experiments. There, one tries to take a very strong photon source, such as a laser beam, and convert the photons by means of the $\theta F_{\mu\nu}\tilde{F}^{\mu\nu}$-term to axions. Clearly, $F_{\mu\nu}\tilde{F}^{\mu\nu}$ is zero for a plane electromagnetic wave, so one provides an additional magnetic field to make the scalar product between the laser's electric field E_i and the external magnetic field B^i nonzero, enabling the conversion. The experimental setup continues then to block the laser beam by a wall and invert the conversion behind the wall, hopefully recoving photons from the axion field by supplying again a strong magnetic field.

E.11.2 Massive fields, Proca-terms and the Higgs-mechanism

For a scalar field φ it is rather straightforward to make it massive. In fact, it suffices to add a term that is quadratic in the field amplitude φ to the Lagrange-density. Then,

$$\mathcal{L}(\varphi, \partial_\mu\varphi) = \frac{\eta^{\alpha\beta}}{2}\partial_\alpha\varphi\partial_\beta\varphi - \frac{m^2}{2}\varphi^2. \tag{E.452}$$

Substitution into the Euler-Lagrange equation gives the equation of motion, which now reads

$$\Box\varphi = m^2\varphi \tag{E.453}$$

and a plane-wave ansatz of the type $\varphi \propto \exp(\pm \mathrm{i}k_\alpha x^\alpha)$ would yield as a dispersion relation

$$\eta^{\mu\nu}k_\mu k_\mu = \left(\frac{\omega}{c}\right)^2 - \gamma^{ij}k_i k_j = m^2 \quad \text{such that} \quad \omega = \pm c\sqrt{k^2 + m^2} \tag{E.454}$$

The wave number k_μ has clearly a time-like normalisation, $\eta^{\mu\nu}k_\mu k_\nu = m^2 > 0$, such that the propagation takes place inside the future light cone, as expected from a massive object. In addition, the group- and phase velocities are

$$\upsilon_{\mathrm{gr}} = \frac{\mathrm{d}\omega}{\mathrm{d}k} = c\frac{k}{\sqrt{k^2 + m^2}} < c \quad \text{and} \quad \upsilon_{\mathrm{ph}} = \frac{\omega}{k} = c\frac{\sqrt{k^2 + m^2}}{k} > c \tag{E.455}$$

because $\sqrt{k^2 + m^2} > k$, but their geometric mean is exactly

$$\upsilon_{\mathrm{ph}} \times \upsilon_{\mathrm{gr}} = c^2 \tag{E.456}$$

i.e. the phase velocity is superluminal, but the group velocity which is associated to the propagation speed of wave packets representing massive particles, remains subluminal. This is nicely illustrated by Fig. 25, where both velocities reach the same limiting value of c for $k \to \infty$, i.e. for $k \gg m$, as the mass becomes less and less relevant in that limit.

Motivated by this example one could think of a modified Lagrange density for the Maxwell field of the form

$$\mathcal{L} = \frac{\eta^{\alpha\mu}\eta^{\beta\nu}}{4}\mathrm{F}_{\alpha\beta}\mathrm{F}_{\mu\nu} + \frac{m^2}{2}\eta^{\alpha\mu}\mathrm{A}_\alpha\mathrm{A}_\mu \tag{E.457}$$

with a so-called Proca-term $\eta^{\alpha\mu}\mathrm{A}_\alpha\mathrm{A}_\mu$. Variation with the corresponding Euler-Lagrange equation would yield a seemingly sensible result, as

$$\eta^{\alpha\mu}\partial_\alpha\mathrm{F}_{\mu\nu} = \Box\mathrm{A}_\nu = m^2\mathrm{A}_\nu \tag{E.458}$$

in Lorenz-gauge, where the same plane-wave ansatz $\exp(\pm\mathrm{i}k_\alpha x^\alpha)$ would give a time-like normalisation $\eta^{\alpha\mu}k_\alpha k_\mu = m^2 > 0$ that corresponds to subluminal motion inside the light cone. But there is a fundamental problem already present in the Lagrange density: It is not gauge-invariant,

$$\eta^{\alpha\mu}\mathrm{A}_\alpha\mathrm{A}_\mu \to \eta^{\alpha\mu}\mathrm{A}_\alpha\mathrm{A}_\mu + 2\eta^{\alpha\mu}\partial_\alpha\chi\,\mathrm{A}_\mu + \eta^{\alpha\mu}\partial_\alpha\chi\,\partial_\mu\chi, \tag{E.459}$$

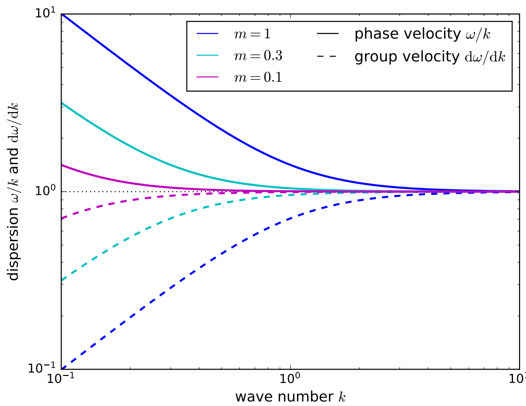

Figure 25: Dispersion relation, i.e. group and phase velocity as a function of wave number, for different masses.

so the choice of a suitable gauge is not possible. In fact, the problem of generating masses dynamically in a gauge-invariant way is solved only by the Higgs-mechanism for field theories and misses yet a complete solution for massive gravity. Electrodynamics as a theory without masses is backed up by stringent experimental upper bounds on the photon mass.

One should add, though, that Lorenz-gauge is still a very sensible choice for cases with a non-zero Proca-mass. Clearly, constructing the action S from the Lagrange-density eqn. (E.457) includes the additional terms

$$S = \int_V \mathrm{d}^4x \left(2\eta^{\alpha\mu}\partial_\alpha\chi A_\mu + \eta^{\alpha\mu}\partial_\alpha\chi\partial_\mu\chi\right) = -\int_V \mathrm{d}^4x \left(2\chi\,\underbrace{\eta^{\alpha\mu}\partial_\alpha A_\mu}_{=0} + \chi\,\underbrace{\eta^{\alpha\mu}\partial_\alpha\partial_\mu\chi}_{=\Box\chi}\right) \quad \text{(E.460)}$$

after integration by parts: In fact, Lorenz-gauge then makes the first term disappear and forces the gauge field χ to obey a wave-equation $\Box\chi = 0$.

E.11.3 Modifications of the Coulomb-potential

Scalar fields φ, even in the case of linear field equations, show an interesting phenomenology on large scales: Starting from the most general Lagrange-density $\mathcal{L}(\varphi, \partial_\mu\varphi)$ including all terms up to φ^2 would ensure a linear field equation, as in the variation process the powers get reduced by one:

$$\mathcal{L}(\varphi, \partial_\mu\varphi) = \frac{\gamma^{\mu\nu}}{2}\partial_\mu\varphi\,\partial_\nu\varphi - \frac{m^2}{2}\varphi^2 - 4\pi\rho\varphi + \lambda\varphi, \quad \text{(E.461)}$$

leading by variation to the field equation

$$(\Box + m^2)\varphi = 4\pi\rho + \lambda \quad \text{(E.462)}$$

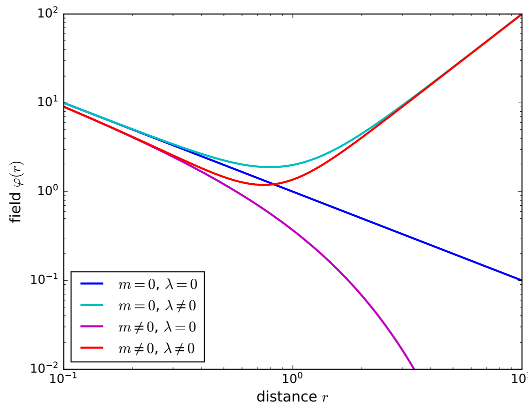

Figure 26: Field amplitude $\varphi(r)$ for the most general linear scalar field theory.

where one admits a source term ρ and an inhomogeneity λ, which would be present even in a charge free space and which would, in a gravitational theory, correspond to the gravitational constant. Focusing on a static, spherically symmetric situation for a point charge one recovers from the field equation

$$(\Delta - m^2)\varphi = -4\pi\rho - \lambda \quad \text{with} \quad \Delta\varphi = \frac{1}{r^2}\partial_r\left(r^2\partial_r\varphi\right), \tag{E.463}$$

depending on the choice of the two parameters, the classical Coulomb-potential

$$\varphi(r) = \frac{1}{r} \tag{E.464}$$

for $m = 0 = \lambda$. Admitting a nonzero mass leads to the Yukawa-potential

$$\varphi(r) = \frac{\exp(-mr)}{r} \tag{E.465}$$

for $m \neq 0 = \lambda$, where the field amplitude φ is suppressed at large r. The full theory implies

$$\varphi(r) = \frac{\exp(-mr)}{r} + \lambda r^2 \tag{E.466}$$

for $m \neq 0 \neq \lambda$, with modifications large scales, while λ alone gives rise to

$$\varphi(r) = \frac{1}{r} + \lambda r^2 \tag{E.467}$$

for $m = 0 \neq \lambda$, which would, up to a sign, be the gravitational potential of a point mass in the classical limit including a cosmological constant. Common to the results are the definition of two additional length scales $1/m$ and $1/\sqrt{\lambda}$, which modify the otherwise scale-free Coulomb-solution. Fig. 26 summarises these modifications.

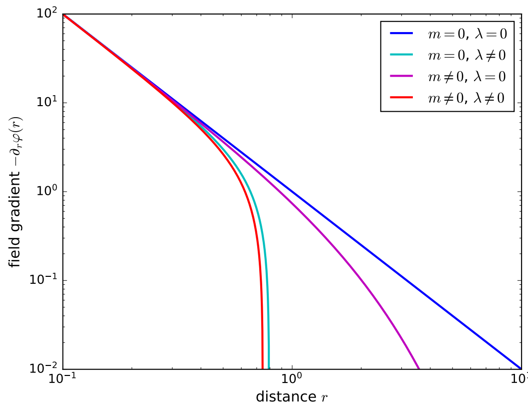

Figure 27: Field gradients $-\mathrm{d}\varphi/\mathrm{d}r$ *for the most general linear scalar field theory.*

The gradient $-\partial\varphi/\partial r$ would, if φ is interpreted as a potential, accelerate a test charge. The acceleration as a function of r is shown in Fig. 27, illustrating how on small scales $r \ll 1/m$ and $r \ll 1/\sqrt{\lambda}$, the unaffected Coulomb-potential is recovered, while there are modifications on large scales $r \gg 1/m$ and $r \gg 1/\sqrt{\lambda}$. It should be noted that the generalised inhomogeneity λ is not admissible in a non-scalar theory like electrodynamics, as a term linear in the 4-potential λA_μ is clearly non-scalar.

E.12 *Conformal invariance of the Maxwell-theory*

Apart from Lorentz- and gauge-invariance, and the spacetime shift symmetries of the Lagrange-density of Maxwell-electrodynamics there is, at least for vacuum-solutions, a weird scale-symmetry. Applying a rescaling of the spacetime coordinates

$$x^\alpha \rightarrow \lambda x^\alpha \quad \text{and consequently,} \quad \partial_\alpha \rightarrow \frac{1}{\lambda}\partial_\alpha. \tag{E.468}$$

The fields obey homogeneous wave equations in vacuum,

$$\Box F_{\mu\nu} = 0 \quad \text{and} \quad \Box G^{\mu\nu} = 0, \tag{E.469}$$

where in fact the λ^{-2} factor generated in $\Box \rightarrow \Box/\lambda^2$ drops out because of the vanishing right hand side of the two equations. This is an example of conformal symmetry. It is broken because the charge density ρ changes under the scaling $\propto \lambda^{-3}$ instead of $\propto \lambda{-}2$ as the differential operators.

E.13 *Gauge-invariance as a geometric concept*

The relationship between the fields and the derivatives in a relativistic notation are summarised by this diagram: The potential A_ν has an antisymmetric derivative $\tilde{F}^{\alpha\beta} = \epsilon^{\alpha\beta\mu\nu}\partial_\mu A_\nu$, and this dual is divergence-free in fulfilment of the Bianchi-identity: $\partial_\alpha\tilde{F}^{\alpha\beta} = \epsilon^{\alpha\beta\mu\nu}\partial_\alpha\partial_\mu A_\nu$, with an exchange symmetry in the index pair (α,μ) which makes the expression disappear in conjunction with the antisymmetry of the Levi-

Civita symbol: This is quite important because of two reasons: Not only does one recover the homogeneous Maxwell-equations, but it is clear that the potential A_μ is incompatible with a hypothetical nonzero magnetic source $\imath^\beta$.

Converting $\tilde{F}^{\alpha\beta}$ into $F_{\gamma\delta}$ and bringing in the constitutive relation yields the field tensor $G^{\gamma\delta}$. The divergence $\partial_\gamma G^{\gamma\delta}$ is the source $\jmath^\delta$, as an expression of the field equation. And finally, charge conservation in the sense of $\partial_\delta \jmath^\delta = 0$ is ensured by $\partial_\gamma\partial_\delta G^{\gamma\delta}$, again with a contraction of a symmetric with an antisymmetric tensor.

The gauge function χ changes A_μ, but leaves $F_{\mu\nu}$ invariant: This is accomplished by the derivative $\epsilon^{\alpha\beta\nu\mu}\partial_\nu\partial_\mu\chi = 0$, as the two derivatives interchange, $\partial_\mu\partial_\nu = \partial_\nu\partial_\mu$, but $\epsilon^{\alpha\beta\nu\mu}$ is antisymmetric in the index pair (μ, ν).

$$\begin{tikzcd}
 & & F_{\gamma\delta}, G^{\gamma\delta} \arrow[r, "\partial_\gamma"] & \jmath^\delta & \text{Maxwell} \\
 & A_\nu \arrow[r, "\epsilon^{\alpha\beta\mu\nu}\partial_\mu"] & \tilde{F}^{\alpha\beta} \arrow[u, dashed, "\epsilon_{\alpha\beta\gamma\delta}"] \arrow[r, "\partial_\alpha"] & 0 & \text{Bianchi} \\
\chi \arrow[r, "\partial_\nu"] & \partial_\nu\chi \arrow[u, squiggly, "+"] \arrow[r, "\epsilon^{\alpha\beta\mu\nu}\partial_\mu"] & 0 \arrow[u, squiggly, "+"] & & \text{gauging}
\end{tikzcd} \tag{E.470}$$

Finding a gauge function χ for a given gauge condition, usually a derivative property of the potential like a particular value for $\eta^{\mu\nu}\partial_\mu A_\nu$ as in the Lorenz gauge requires the solution of a wave equation: Substitution $A_\mu \rightarrow A_\mu + \partial_\mu\chi$ into the gauge condition leads to $\eta^{\mu\nu}\partial_\mu\partial_\nu\chi = -b$ with $b = \eta^{\mu\nu}\partial_\mu A_\nu$. Wave-equations of this type are readily solvable by means of the retarded Green-functions.

$$\begin{tikzcd}
 & A_\mu \arrow[r, "\eta^{\mu\nu}\partial_\nu"] \arrow[d, squiggly, "-"'] & b \arrow[d, squiggly, "-"'] & \text{condition} \\
\chi \arrow[r, "\partial_\mu"] & \partial_\mu\chi \arrow[r, "\eta^{\mu\nu}\partial_\nu"] & \Box\chi & \text{gauging}
\end{tikzcd} \tag{E.471}$$

The same diagram with identical arguments can be more concisely expressed in the language of differential forms: Starting from the 4-potential A_μ as a one-form A, application of the exterior derivative d leads to the two-form F, corresponding to the field tensor $F_{\mu\nu}$. The co-differential δ, which can be expressed as $\star d \star$ with the Hodge-star operator $\star$, leads to the source $\jmath$, again a one-form. The Hodge-dual of the field two-form F would be $\star F$, whose co-differential $\delta\star F = \star d\star\star F = \star dF = \star ddA = 0$, recovering the Bianchi-identity. The gauge field χ has an exterior derivative $d\chi$, which can be added to the one-form A without changing the observable fields contained in the two-form F, as $dA \rightarrow d(A + d\chi) = dA + dd\chi = dA$. On the other hand, $F = dA$ is only sensible if $\delta\star F = 0$ physically, i.e. that the magnetic charges are non-existent: The existence of a potential A requires broken duality.

$$
\begin{tikzcd}
 & & \star\mathrm{F} \arrow[r, "\delta"] & 0 & \text{Bianchi} \\
 & \mathrm{A} \arrow[r, "\mathrm{d}"] & \mathrm{F},\mathrm{G} \arrow[u, dashed, "\star"] \arrow[r, "\delta"] & \jmath & \text{Maxwell} \\
\chi \arrow[r, "\mathrm{d}"] & \mathrm{d}\chi \arrow[u, squiggly, "+"] \arrow[r, "\mathrm{d}"] & 0 \arrow[u, squiggly, "+"] & & \text{gauging}
\end{tikzcd}
\tag{E.472}
$$

The construction of a (scalar) gauge function χ for ensuring e.g. Lorenz-gauge $\delta\mathrm{A} = 0$ implies that $\mathrm{d}\chi$, now a one-form, is added to A and leads to $\delta\mathrm{d}\chi = -b$, with a source $b = \delta\mathrm{A}$ after substitution into the gauge condition. $\delta\mathrm{d}\chi$, however is the Laplace-de Rham-operator, which for our case of a Lorentzian metric background is the d'Alembert-operator $\Box$, up to a symmetrisation.

$$
\begin{tikzcd}
 & \mathrm{A} \arrow[r, "\delta"] \arrow[d, squiggly, "-"'] & b \arrow[d, squiggly, "-"'] & \text{condition} \\
\chi \arrow[r, "\mathrm{d}"] & \mathrm{d}\chi \arrow[r, "\delta"] & \Box\chi & \text{gauging}
\end{tikzcd}
\tag{E.473}
$$

E.14 *Motion of particles through spacetime*

E.14.1 Fermat's or Hamilton's principle?

The relativistic expression for the arc length s through spacetime, as mapped out by proper time, can be amended by a second term, $q\mathrm{A}_\mu\mathrm{d}x^\mu$ which should incorporate the accelerating effects of electric and magnetic fields on a test particle with charge q:

$$
s = \int_{\mathrm{A}}^{\mathrm{B}} \mathrm{d}\tau\, mc\sqrt{\eta_{\mu\nu}u^\mu u^\nu} + q\mathrm{A}_\mu\mathrm{d}x^\mu \quad \rightarrow \quad \mathcal{L}(x^\mu, u^\mu) = mc\sqrt{\eta_{\mu\nu}u^\mu u^\nu} + q\mathrm{A}_\mu u^\mu, \tag{E.474}
$$

where in isolating the Lagrange function one rewrites $\mathrm{d}x^\mu = u^\mu\mathrm{d}\tau$, from the definition $u^\mu = \mathrm{d}x^\mu/\mathrm{d}\tau$. Variation of the arc-length, now a function of both u^μ and x^μ (through the coordinate dependence of A_μ) is achieved with the Euler-Lagrange equation,

⚠ *The term $\mathrm{A}_\mu u^\mu$ emphasises how natural velocity-dependent forces in relativity are!*

$$
\frac{\mathrm{d}}{\mathrm{d}\tau}\frac{\partial\mathcal{L}}{\partial u^\alpha} = \frac{\partial\mathcal{L}}{\partial x^\alpha}. \tag{E.475}
$$

The expression (E.474) for the relativistic arc length is remarkable, as it combines the metric distance in the first term with a second distance measure $\mathrm{A}_\mu\mathrm{d}x^\mu$ mediated by the vector potential, called Finsler geometry.

An intuitive (but gauge-dependent) picture might be, that different paths through spacetime have the particle change its proper time according to the magnitude and direction of A_μ relative to its velocity u^μ, like a tail- or headwind that changes travel time. The necessary derivatives are

$$\frac{\partial \mathcal{L}}{\partial x^\alpha} = q u^\mu \partial_\alpha A_\mu \tag{E.476}$$

and

$$\frac{\partial \mathcal{L}}{\partial u^\alpha} = m\eta_{\mu\alpha} u^\mu + qA_\alpha \quad \rightarrow \quad \frac{\mathrm{d}}{\mathrm{d}\tau}\frac{\partial \mathcal{L}}{\partial u^\alpha} = m\eta_{\mu\alpha}\frac{\mathrm{d}u^\mu}{\mathrm{d}\tau} + q u^\mu \partial_\mu A_\alpha \tag{E.477}$$

where the last term appears in the time derivative through the coordinate dependence of A_α:

$$\frac{\mathrm{d}A_\alpha}{\mathrm{d}\tau} = \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\frac{\partial A_\alpha}{\partial x^\mu} = u^\mu\, \partial_\mu A_\alpha. \tag{E.478}$$

Collecting all results yields

$$m\eta_{\mu\alpha}\frac{\mathrm{d}u^\mu}{\mathrm{d}\tau} = q\left(\partial_\alpha A_\mu - \partial_\mu A_\alpha\right) u^\mu = qF_{\alpha\mu} u^\mu \tag{E.479}$$

by identifying the Faraday-tensor in the last step: Finally, we recover the Lorentz equation of motion, and the appearance of $F_{\mu\nu}$ makes sure that the acceleration does not depend on gauge. Multiplying both sides of the equation with u^α leads to an interesting result:

$$\eta_{\mu\alpha} u^\alpha \frac{\mathrm{d}u^\mu}{\mathrm{d}\tau} = \frac{m}{2}\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\eta_{\mu\alpha} u^\alpha u^\mu\right) = qF_{\alpha\mu} u^\alpha u^\mu = 0, \tag{E.480}$$

where the last term is necessarily zero as the contraction between the symmetric tensor $u^\alpha u^\mu$ and the antisymmetric $F_{\alpha\mu}$. This safeguards the norm $\eta_{\mu\alpha} u^\alpha u^\mu = c^2$ from any changes, and keeps the particle from being accelerated to superluminal velocities outside the light cone.

While the equation of motion is perfectly gauge-invariant (and Lorentz-covariant), the gauge-invariance of the Lagrange-function requires additional arguments: Performing a gauge transform $A_\mu \rightarrow A_\mu + \partial_\mu \chi$ with a gauge function χ changes the relativistic arc length according to

$$s = \int_A^B \mathrm{d}\tau \left(mc\sqrt{\eta_{\mu\nu} u^\mu u^\nu} + qA_\mu u^\mu\right) \quad \rightarrow \quad s + q\int_A^B \mathrm{d}\tau\, \partial_\mu \chi\, u^\mu. \tag{E.481}$$

This new term can be rewritten, by falling back onto the form how it was introduced,

$$\int_A^B \mathrm{d}\tau\, \partial_\mu \chi\, u^\mu = \int_A^B \mathrm{d}x^\mu \partial_\mu \chi = \int_A^B \mathrm{d}\chi = (\chi_B - \chi_A), \tag{E.482}$$

using $\mathrm{d}\chi = \partial_\mu \chi \mathrm{d}x^\mu$. In summary, there is a constant, additive term that becomes irrelevant for the variation for obtaining the trajectory.

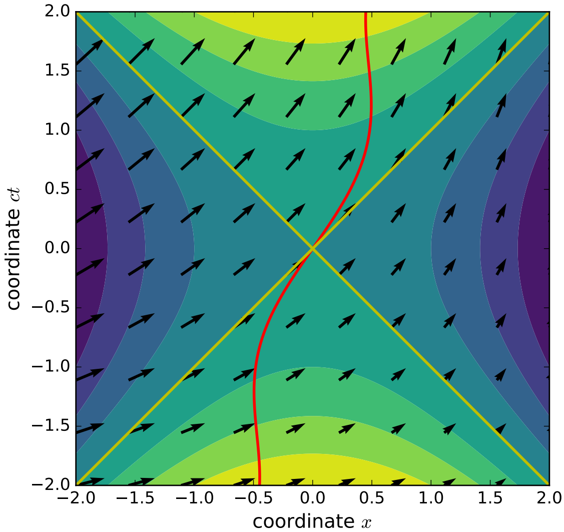

Figure 28: Trajectory through spacetime, with the metric contribution $\eta_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu$ to the line element $\mathrm{d}s^2$ in the background shading and the Finsler contribution $A_\mu\mathrm{d}x^\mu$ generated by the potential A_μ as arrows.

An impression on the contributions to the line element $\mathrm{d}s^2$ given by the metric $\eta_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu$ and the Finsler-term $A_\mu\mathrm{d}x^\mu$ is given in Fig. 28.

E.14.2 Relativistic horizons

We can probe the limits of special relativity by looking at accelerated, non-inertial motion through spacetime. Starting from the coordinates x^μ we already defined the 4-velocity u^μ,

$$u^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} = \frac{\mathrm{d}}{\mathrm{d}\tau}\begin{pmatrix} ct \\ x \end{pmatrix} = \frac{\mathrm{d}t}{\mathrm{d}\tau}\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} ct \\ x \end{pmatrix} = \gamma\begin{pmatrix} c \\ \upsilon \end{pmatrix} \tag{E.483}$$

with $\upsilon = \dot{x}$ and $\gamma = \mathrm{d}t/\mathrm{d}\tau$. Repeating this argument one computes the 4-acceleration a^μ as

$$a^\mu = \frac{\mathrm{d}u^\mu}{\mathrm{d}\tau} = \frac{\mathrm{d}}{\mathrm{d}\tau}\gamma\begin{pmatrix} c \\ \upsilon \end{pmatrix} = \frac{\mathrm{d}t}{\mathrm{d}\tau}\frac{\mathrm{d}}{\mathrm{d}t}\gamma\begin{pmatrix} c \\ \upsilon \end{pmatrix} = \frac{\upsilon a}{c^2}\gamma^4\begin{pmatrix} c \\ \upsilon \end{pmatrix} + \gamma^2\begin{pmatrix} 0 \\ a \end{pmatrix} \tag{E.484}$$

with $a = \dot{\upsilon} = \ddot{x}$, and the derivative $\mathrm{d}\gamma/\mathrm{d}t = \gamma^3\upsilon a/c^2$. This system of equations can be integrated numerically for e.g. an assumed constant acceleration a, giving a parametric solution $(ct(\tau), x(\tau))$. The resulting trajectories in $x^\mu(\tau)$ are shown in Fig. 29, where the accelerated trajectory evades light signals that are emitted at $x = 0$ later than $ct \geq 4$, which is impossible for inertial motion. Effectively, evading light signals means that there is a relativistic horizon between the emitter of light signals and the accelerated particle.

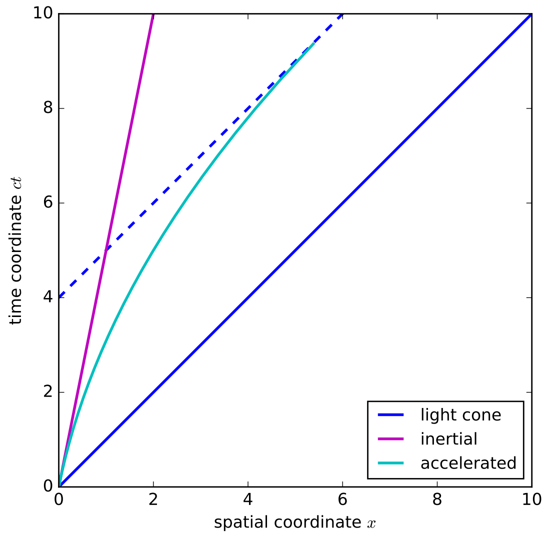

Figure 29: Paths through spacetime at constant velocity, and in comparison a path with constant acceleration, with the emergence of a relativistic horizon.

The 4-acceleration a^μ is always perpendicular to the 4-velocity u^μ,

$$\eta_{\mu\nu} u^\mu a^\nu = 0, \tag{E.485}$$

as a direct computation with the above expression shows. This has in fact dramatic consequences, as

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\eta_{\mu\nu} u^\mu u^\nu\right) = \eta_{\mu\nu}\left(\frac{\mathrm{d}u^\mu}{\mathrm{d}\tau} u^\nu + u^\mu \frac{\mathrm{d}u^\nu}{\mathrm{d}\tau}\right) = 2\eta_{\mu\nu} u^\mu a^\nu = 0, \tag{E.486}$$

implying that the (timelike) norm $\eta_{\mu\nu} u^\mu u^\nu = c^2 > 0$ of u^μ is conserved. At this point it is worth mentioning that many texts attribute the impossibility of accelerating a massive object past c to the relativistic mass increase, which is really superfluous as a concept as it is completely covered by the geometric, kinematical structure of spacetime. Proper acceleration is defined in terms of proper time τ, which is dilated relative to the coordinate time t by the Lorentz-factor γ. A faster-moving system reacts to an accelerating force as if it had more inertia and therefore a higher mass, but it is really the conversion between proper time and coordinate time that brings in the Lorentz-factor, and one does not need to invoke a new relativistic effect on mass, and surely the number of atoms inside an object would be unchanged under Lorentz transforms!

"The concept of "relativistic mass" is subject to misunderstanding. That's why we don't use it. First, it applies the name mass – belonging to the magnitude of a 4-vector – to a very different concept, the time component of a 4-vector. Second, it makes increase of energy of an object with velocity or momentum appear to be connected with some change in internal structure of the object. In reality, the increase of energy with velocity originates not in the object but in the geometric properties of spacetime itself.", E. F. Taylor and J. A. Wheeler, Spacetime Physics

E.14.3 Tachyons and tardyons

Tachyons are hypothetical, superluminally moving particles with 4-velocities u^μ outside the light cone, $\eta_{\mu\nu} u^\mu u^\nu = -c^2 < 0$. On the other side, tardyons are

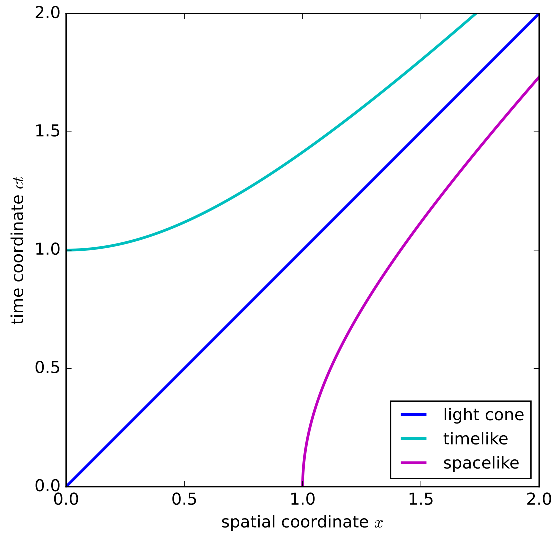

Figure 30: Curves of constant Minkowski-norm $\eta_{\mu\nu}x^\mu x^\nu = \pm 1$, or equivalently, curves traced out by the endpoint of a timelike and spacelike unit vector under Lorentz-transforms.

conventional, massive particles with subluminal velocities inside the light cone, $\eta_{\mu\nu}u^\mu u^\nu = +c^2 > 0$. Naturally, these norms are conserved under Lorentz-transforms, as illustrated by Fig. 30, where the hyperbolic curves traced out by the unit vectors along the x- and ct-axes never leave their associated timelike or spacelike quadrants. For a timelike vector this would be,

$$\begin{pmatrix} ct \\ x \end{pmatrix} = \begin{pmatrix} \cosh\psi & \sinh\psi \\ \sinh\psi & \cosh\psi \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cosh\psi \\ \sinh\psi \end{pmatrix} \tag{E.487}$$

and for a spacelike vector correspondingly,

$$\begin{pmatrix} ct \\ x \end{pmatrix} = \begin{pmatrix} \cosh\psi & \sinh\psi \\ \sinh\psi & \cosh\psi \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \sinh\psi \\ \cosh\psi \end{pmatrix}. \tag{E.488}$$

For a particle moving on a spacelike trajectory one would write down a line element

$$c^2 d\tau^2 = \eta_{\mu\nu} dx^\mu dx^\nu = c^2 dt^2 - \gamma_{ij} dx_i dx_j = \left(c^2 - \gamma_{ij} \upsilon^i \upsilon^j\right) dt^2 \tag{E.489}$$

with $\upsilon^i = dx^i/dt$. Negative norms would then imply that $\gamma_{ij}\upsilon^i\upsilon^j > c^2$, and hence that the magnitude of υ exceeds c. The velocity u^μ for such a particle would necessarily have the same negative norm, as one writes $u^\mu = dx^\mu/d\tau$, and because $c^2 d\tau^2 = \eta_{\mu\nu}u^\mu u^\nu d\tau^2$ has to have the same overall sign.

The relativistic dispersion relation $\mathcal{H}^2 = (cp)^2 + (mc^2)^2$ suggests the definition of a relativistic 4-momentum p_μ (as a linear form), whose norm is positive for tardyonic and negative for tachyonic particles, according to the location of the corresponding velocities in the respective quadrants in a spacetime diagram,

$$p_\mu = \begin{pmatrix} \mathcal{H}, & cp_i \end{pmatrix} \quad \text{with} \quad \eta^{\mu\nu} p_\mu p_\nu = \mathcal{H}^2 - c^2\gamma^{ij} p_i p_j = \mathcal{H}^2 - (cp)^2 = \pm(mc^2)^2, \tag{E.490}$$

resulting in a funny shape of the dispersion relation,

$$\mathcal{H}(p) = \sqrt{(cp)^2 \pm (mc^2)^2}, \tag{E.491}$$

for the negative sign: This is in fact consistent with their superluminality, as p^2 is bounded from below by $(mc)^2$: Tachyons need to be faster than the speed of light, and if they brake down to approach the speed of light from above, they can only reach mc. In a weird sense, this is analogous to the non-vanishing energy associated with the rest mass for normal, tardyonic particles: While for them the energy is nonzero even for vanishing momenta, tachyons have a minimal momentum even at zero energies. To some degree of overinterpretation, tachyons have a minimal momentum mc whereas the tardyons have a minimal energy mc^2. Reexpressing the tachyonic dispersion relation in terms of wave number and angular frequency would be

$$\omega = \pm c\sqrt{k^2 - m^2} \tag{E.492}$$

Group and phase velocities for tachyons come out as

$$\upsilon_{\mathrm{gr}} = \frac{\mathrm{d}\omega}{\mathrm{d}k} = \frac{ck}{\sqrt{k^2 - m^2}} > c \quad \text{and} \quad \upsilon_{\mathrm{ph}} = \frac{\omega}{k} = \frac{c\sqrt{k^2 - m^2}}{k} < c, \tag{E.493}$$

exactly inverted compared to massive particles: The group velocity, associated with particle propagation, is always superluminal because $\sqrt{k^2 - m^2} < k$, and the phase velocity subluminal. Their geometric average, though, comes out as

$$\upsilon_{\mathrm{gr}} \times \upsilon_{\mathrm{ph}} = c^2. \tag{E.494}$$

Of course one should keep in mind that outside the light cone there is no causal ordering due to the relativity of simultaneity, so it would be problematic to have tachyons influence the causal world inside the light cone. To conclude, there is no place for tachyons in a Galilean world: In the formal limit of $c \to \infty$, the future light cone opens up: The timelike region increases while the spacelike region decreases, until all of spacetime reaches an absolute causal ordering according to the universal, Galilean time. And of course, every velocity is subluminal as $c \to \infty$.