Lecture Notes on Electromagnetism and Gauge Invariance

Nick Kaiser

September 3, 2020

Contents

1 Maxwell's equations

The physical content of Maxwell's equations is illustrated in figure 1. In the 'rationalised' SI system they are

$$\nabla \cdot \mathbf{E} = \rho/\epsilon_0 \qquad \nabla \cdot \mathbf{B} = 0$$
$$\nabla \times \mathbf{E} = -\partial \mathbf{B}/\partial t \qquad \nabla \times \mathbf{B} = \mu_0 \mathbf{j} + c^{-2}\partial \mathbf{E}/\partial t \tag{1}$$

where

- $\mathbf{E}$ and $\mathbf{B}$ are the electric and magnetic fields
- ϵ_0 is the *permittivity of free space*
  - defined in terms of the electrostatic force between two charges
- μ_0 is the *permeability of free space*
  - defined in terms of the magnetostatic force between two current element (Biot & Savart)
- $c = 1/\sqrt{\epsilon_0 \mu_0}$
- ρ is the charge density
- $\mathbf{j}$ is the current density

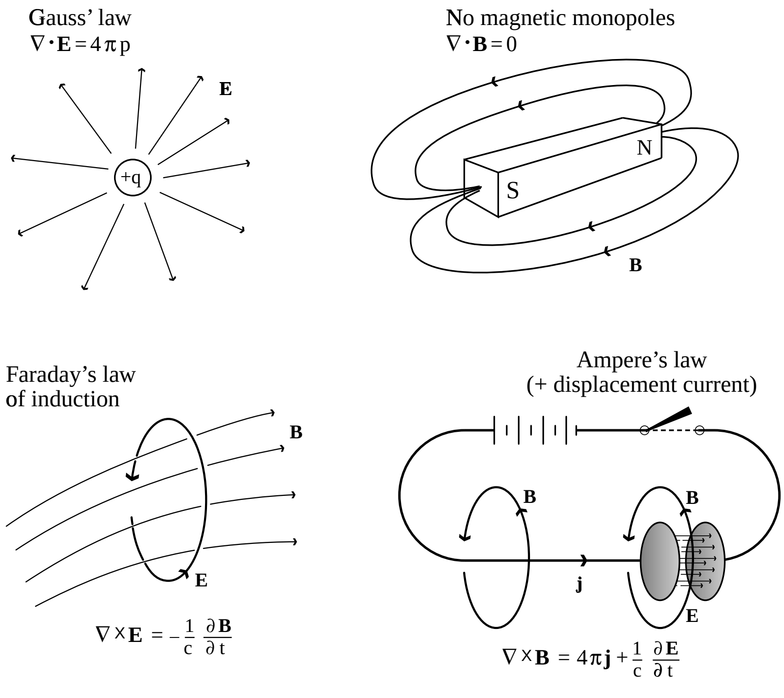

Figure 1: Maxwell's equations. These are in 'Gaussian units'.

Special relativity is 'built in' to Maxwell's electromagnetism. Most fundamentally from the fact that these equations permit wave-like solutions that propagate with speed c – determined by the empirical constants of electro- and magneto-statics – that is apparently independent of any frame of reference.

2 Gauge Invariance in electromagnetism and the gauge principle

2.1 The electromagnetic 4-potential

- Maxwell's equation $\nabla \cdot \mathbf{B} = 0$ implies
  - $\boxed{\mathbf{B} = \nabla \times \mathbf{A}}$
  - with $\mathbf{A}$ the *vector potential*,

- while $\boldsymbol{\nabla} \times \mathbf{E} + \partial_t \mathbf{B} = 0$ (together with the above) implies that $\mathbf{E} + \partial_t \mathbf{A}$ is the gradient of some scalar
  - $\boxed{\mathbf{E} = -\boldsymbol{\nabla}\varphi - \partial_t \mathbf{A}}$
  - where φ is called the *scalar potential*
- the scalar and vector (or magnetic) potentials are the time and space components of a 4-vector
  - $\boxed{\vec{A} \longrightarrow A^\mu = (\phi, \mathbf{A})}$

2.2 Invariance of electromagnetic fields under a gauge transformation

The potentials φ and $\mathbf{A}$ are not unique as $\mathbf{E}$ and $\mathbf{B}$ are invariant under a *gauge transformation* of the 4-potential $A_\mu \equiv (-\varphi, \mathbf{A}) \to A_\mu + \xi_{,\mu}$ for arbitrary $\xi(\vec{x})$.[1]

2.3 The gauge principle; quantum mechanics and charge conservation

The *gauge-principle* is often described as originating from quantum mechanics and as having to do with the essential non-observability of the phase θ of the wave-function $\psi = |\psi|e^{i\theta}$ for a particle. The argument typically goes like this:

- Schrödinger's equation $i\hbar\partial_t\psi = -\hbar^2\nabla^2\psi/2m + V\psi$, and observables like the probability density $\rho = \psi\psi^\star$ and its current $\mathbf{j} = \frac{i\hbar}{2m}(\psi\boldsymbol{\nabla}\psi^\star - \psi^\star\boldsymbol{\nabla}\psi)$, which obey the law of conservation of total probability $\partial_t\rho + \boldsymbol{\nabla}\cdot\mathbf{j} = 0$, are invariant under a *global gauge transformation* $\psi \to \psi' = \psi e^{i\delta\theta}$ where the phase shift $\delta\theta$ is independent of position
- if, however, we make a *local* gauge transformation $\psi \to \psi' = \psi e^{i\delta\theta(\vec{x})}$ the current is not invariant ($\mathbf{j} \to \mathbf{j}' = \mathbf{j} + \frac{\hbar}{m}\rho\boldsymbol{\nabla}\theta$) and ψ' is no longer a solution of Schrödinger's equation (hereafter the S-equation)
- But if we modify the S-equation by changing the spatio-temporal partial derivatives according to $\hbar\partial_\mu \to \hbar\partial_\mu - iqA_\mu$ then
  - the S-equation *is* invariant under a position dependent change of the phase of the wave-function $\delta\theta(\vec{x}) = (q/\hbar)\xi(\vec{x})$, or $\psi \to \psi' = \psi e^{i(q/\hbar)\xi(\vec{x})}$, provided that we make a simultaneous gauge transformation $A_\mu \to A'_\mu = A_\mu + \xi_{,\mu}$
  - this follows from the fact that $(\hbar\partial_\mu - iqA'_\mu)\psi' = e^{i(q/\hbar)\xi}(\hbar\partial_\mu + iq\xi_{,\mu} - iqA'_\mu)\psi = e^{i(q/\hbar)\xi}(\hbar\partial_\mu - iqA_\mu)\psi$. Thus the phase factor $e^{i(q/\hbar)\xi}$ 'passes through' the differential operator, with the result that applying the combined transformations to the S-equation it is just multiplied by $e^{i(q/\hbar)\xi}$.
  - this works also with the relativistic S-equation (obtained from $E^2 = p^2c^2 + m^2c^4$ or $\vec{p}\cdot\vec{p} = -m^2c^4$ with the replacement of the Hamiltonian $H = E + q\varphi$ and the canonical momentum $\mathbf{P} = \mathbf{p} + q\mathbf{A}$ by their operator equivalents $i\hbar\partial_t$ and $-i\hbar\boldsymbol{\nabla}$ respectively).

Thus, it is argued, promoting the invariance of Schrödinger's equation with respect to a global change of the phase of the wave-function to invariance with respect to a local one requires the existence of a 'gauge field' $\vec{A}$ introduced into the S-equation in the prescribed manner, and which must have the gauge invariance properties of Maxwell's electromagnetism. Demanding invariance with respect to a local phase transformation somehow conjures electromagnetism into existence.

Charge conservation, which is built into the form of Maxwell's equations at a rather fundamental level, is in this view the conservation law arising via Noether's theorem from the symmetry of the quantum mechanical S-equation under gauge transformations.

[1] In these notes we assume the units of length and time are such that the speed of light $c = 1$ and we use the convention that the flat-spacetime metric (used for raising and lowering indices on 4-vectors and tensors) is $\eta_{\mu\nu} = \mathrm{diag}\{-1, 1, 1, 1\}$. A spacetime 4-displacement $\vec{x}$ has 'contravariant' components $\vec{x} \to x^\mu = (t, x, y, z) = (t, \mathbf{x})$ and 'covariant' components $\vec{x} \to x_\mu = \eta_{\mu\nu}x^\nu = (-t, x, y, z) = (-t, \mathbf{x})$, displaying the Einstein summation convention, while the gradient operator $\vec{\partial}$ has covariant components $\vec{\partial} \to \partial_\mu = (\partial_t, \partial_x, \partial_y, \partial_z) = (\partial_t, \boldsymbol{\nabla})$.

2.4 Charge conservation in classical field theory

A different approach is to notice that if we have two classical fields a and b – perhaps long-wavelength displacements on some kind of crystalline lattice – obeying $\ddot{a} - \gamma^2\nabla^2 a + \mu^2 a = 0$ and similarly for b (and assuming identical speed and frequency parameters γ and μ) then it follows, on multiplying the first by b and the second by a and then subtracting, while using $a\nabla^2 b = \boldsymbol{\nabla}\cdot(a\boldsymbol{\nabla}b) - \boldsymbol{\nabla}a\cdot\boldsymbol{\nabla}b$ and $a\ddot{b} = \partial_t(a\dot{b}) - \dot{a}\dot{b}$, that

$$\partial_t(\dot{a}b - \dot{b}a) + \gamma^2\boldsymbol{\nabla}\cdot(a\boldsymbol{\nabla}b - b\boldsymbol{\nabla}a) = 0. \tag{2}$$

This is evidently a conservation law of the form $\partial_t\rho + \boldsymbol{\nabla}\cdot\mathbf{j} = 0$ with conserved 'charge' density $\rho = \dot{a}b - \dot{b}a = a\partial^t b - b\partial^t a$ and current density $\mathbf{j} = \gamma^2(a\boldsymbol{\nabla}b - b\boldsymbol{\nabla}a)$.

Thus, at a purely classical level, a two-component field obeying equations of this kind, and which we can conveniently express as a single complex valued field $\phi \equiv a + ib$ – in terms of which $\rho = \phi\partial^t\phi^\star - \phi^\star\partial^t\phi$ and $\mathbf{j} = \gamma^2(\phi\boldsymbol{\nabla}\phi^\star - \phi^\star\boldsymbol{\nabla}\phi)$ – has a conserved charge-like quantity (in addition to the energy and the 3 components of the wave-momentum).

With $\gamma = c$ the above equations are the *Klein-Gordon equation* for a classical relativistic scalar field $\partial_\mu\partial^\mu\phi - \mu^2\phi = 0$ (for which the quantum excitations are particles with Compton frequency μ). This is for a free field, which is rather sterile. The question is whether one can introduce a coupling to an EM potential so that e.g. a charged 'matter beam' (or a wave-packet), would be deflected in the manner expected for a beam of charged particles (or a single particle), and in such a way that there is still a conserved charge. The answer is yes: if one introduces the coupling by modifying the spatio-temporal partial derivatives as above. There is still an exactly conserved charge, and matter beams and wave-packets have their momenta and energies modified in the appropriate manner.

From this perspective, there is nothing essentially quantum mechanical about gauge invariance. The reason these are usually associated is that it was only with the invention of QM that physicists came back to the idea that charged matter should be thought of as a wave (having dropped the idea that "cathode rays" were waves after J.J. Thomson showed that they carry electric charge).

Discussions of gauge invariance in terms of the quantum mechanical S-equation are somewhat anachronistic. In quantum field theory one starts with a classical field equation which one then proceeds to 'second quantise'. Gauge invariance, if present in the theory, does not emerge in the quantisation; it was there already in the classical equation.

3 Relativistic particle electrodynamics

3.1 Hamiltonian dynamics of a charged particle

3.1.1 The Lagrangian and the action

Let's guess that the relativistic Lagrangian for a particle of mass m and charge q moving with velocity $\dot{\mathbf{x}}$ in an electromagnetic potential A_μ is

$$L(\mathbf{x}, \dot{\mathbf{x}}, t) = -m/\gamma - q\varphi + q\dot{\mathbf{x}}\cdot\mathbf{A} \tag{3}$$

where $\gamma \equiv 1/\sqrt{1 - |\dot{\mathbf{x}}|^2}$. We may also write this as

$$\boxed{L(\mathbf{x}, \dot{\mathbf{x}}, t) = -m/\gamma + q\dot{x}^\mu A_\mu.} \tag{4}$$

At first sight this might seem like a bad choice as it is evidently not gauge invariant, nor is it Lorentz invariant. On the positive side:

- it generates an *action* $S = \int dtL$ with differential $dS = -md\tau + A_\mu dx^\mu$ which *is* a Lorentz scalar (here the proper time interval is $d\tau = dt/\gamma$)
- and, as we shall see, it gives the empirical Lorentz force law $d\mathbf{p}/dt = q(\mathbf{E} + \dot{\mathbf{x}}\times\mathbf{B})$ and work equation $dE/dt = \mathbf{E}\cdot\dot{\mathbf{x}}$

3.1.2 The canonical and mechanical momenta

The *canonical* or 'generalised' momentum is defined, in general, by

$$\boxed{\mathbf{P} \equiv \partial L / \partial \dot{\mathbf{x}}} \tag{5}$$

which here is given, from (3), by

$$\boxed{\mathbf{P} = \mathbf{p} + q\mathbf{A}} \tag{6}$$

where

$$\boxed{\mathbf{p} \equiv \gamma m \dot{\mathbf{x}}} \tag{7}$$

is the usual relativistic 3-momentum. We will call this the *mechanical momentum*.

3.1.3 The Euler-Lagrange equation

The *Euler-Lagrange equation* obtained from the requirement that the particle path extremises the action is, in general,

$$\boxed{d\mathbf{P}/dt = \partial L / \partial \mathbf{x}} \tag{8}$$

and here, from (4), the *generalised force* $\partial L / \partial \mathbf{x}$ has components

$$\boxed{dP_i/dt = q\dot{x}^\mu A_{\mu,i}} \tag{9}$$

so a particle in a *spatially uniform* potential (i.e. one for which $A_{\mu,i} = 0$) has constant $\mathbf{P}$.

3.1.4 $d\mathbf{p}/dt$ and the Faraday tensor

From (6), the rate of change of the mechanical momentum is $d\mathbf{p}/dt = d\mathbf{P}/dt - qd\mathbf{A}/dt$. The *convective derivative* of $\mathbf{A}$ here is $d\mathbf{A}/dt = \partial_t \mathbf{A} + (\dot{\mathbf{x}} \cdot \boldsymbol{\nabla})\mathbf{A}$ or, in components, $dA_i/dt = A_{i,t} + \dot{x}_j A_{i,j} = \dot{x}^\mu A_{i,\mu}$.

Using (9) for dP_i/dt gives

$$\boxed{dp_i/dt = d(P_i - qdA_i)/dt = q\dot{x}^\mu F_{\mu i}} \tag{10}$$

where

$$\boxed{F_{\mu\nu} \equiv A_{[\mu,\nu]} \equiv A_{\mu,\nu} - A_{\nu,\mu}} \tag{11}$$

is the definition of the *Faraday tensor* contains the $\mathbf{E}$ and $\mathbf{B}$ fields.

3.1.5 The Hamiltonian and the energy-momentum relation

The *Hamiltonian* is defined, in general, to be

$$\boxed{H \equiv \dot{\mathbf{x}} \cdot \mathbf{P} - L(\mathbf{x}, \dot{\mathbf{x}}, t)} \tag{12}$$

which here, from (3) and (6), is $H = \dot{\mathbf{x}} \cdot (\mathbf{p} + q\mathbf{A}) + m/\gamma + q\varphi - q\dot{x} \cdot \mathbf{A} = \gamma m(|\dot{\mathbf{x}}|^2 + 1/\gamma^2) + q\varphi$ or

$$H = \gamma m + q\varphi. \tag{13}$$

Despite its appearance, H is formally, and in general, only a function of $\mathbf{x}$, $\mathbf{P}$ and t since the differential of (12) is

$$dH = \dot{\mathbf{x}} \cdot d\mathbf{P} - \dot{\mathbf{P}} \cdot d\mathbf{x} - (\partial L / \partial t) dt \tag{14}$$

the terms $\mathbf{P} d\dot{\mathbf{x}} - (\partial L / \partial \dot{\mathbf{x}}) d\dot{\mathbf{x}}$ having cancelled by virtue of the definition (5) of $\mathbf{P}$.

To express H explicitly in terms of $\mathbf{x}$ and $\mathbf{P}$ we can note that the definition $\mathbf{p} \equiv \gamma m \dot{\mathbf{x}}$ implies $|\mathbf{p}|^2 + m^2 = m^2(\gamma^2 |\dot{\mathbf{x}}|^2 + 1) = \gamma^2 m^2$, or, defining $E \equiv \gamma m$, which is the mechanical (i.e. rest mass plus kinetic) energy,

$$\boxed{E^2 = |\mathbf{p}|^2 + m^2} \tag{15}$$

which is the familiar relativistic energy momentum relation.

But according to (13) $E = \gamma m = H - q\varphi$, so, along with $\mathbf{p} = \mathbf{P} - q\mathbf{A}$ from (6), the energy-momentum relation, in terms of H and $\mathbf{P}$, is

$$(H - q\varphi)^2 = |\mathbf{P} - q\mathbf{A}|^2 + m^2. \tag{16}$$

and consequently

$$\boxed{H(\mathbf{x}, \mathbf{P}, t) = \sqrt{m^2 + |\mathbf{P} - q\mathbf{A}|^2} + q\varphi} \tag{17}$$

with the dependence of H on $\mathbf{x}$ and t coming *via* $\mathbf{A}(\mathbf{x}, t)$ and $\varphi(\mathbf{x}, t)$.

3.1.6 Hamilton's equations and dH/dt

Inspection of the coefficients of $d\mathbf{P}$ and $d\mathbf{x}$ in (14) provide us, in general, with *Hamilton's equations*

$$\dot{\mathbf{x}} = \boxed{\frac{\partial H}{\partial \mathbf{P}}} \quad \text{and} \quad \boxed{\dot{\mathbf{P}} = -\frac{\partial H}{\partial \mathbf{x}}} \tag{18}$$

which are effectively equivalent to the Euler-Lagrange equation.

Equation (14) for dH is valid for arbitrary $d\mathbf{x}$ and $d\mathbf{P}$. If these differentials are those for an actual particle obeying the equation of motion the first two terms $\dot{\mathbf{x}} \cdot d\mathbf{P} - \dot{\mathbf{P}} \cdot d\mathbf{x}$ in (14) cancel, and we have the equation for the evolution of the Hamiltonian which, in general, is

$$\boxed{dH/dt = -\partial L/\partial t} \tag{19}$$

and here is, from (4),

$$\boxed{dH/dt = q\dot{x}^\mu A_{\mu,t}} \tag{20}$$

so in a *static* potential (i.e. one for which $A_{\mu,t} = 0$) the Hamiltonian is constant.

3.1.7 Equations of motion for the canonical and mechanical 4-momenta

The equations of motion and that for the evolution of the energy are succinctly expressed in 4-notation. Defining the 4-velocity $u^\mu = (\gamma, \gamma\dot{\mathbf{x}})$ and the *canonical 4-momentum* $P^\mu \equiv (H, \mathbf{P})$ we have, for the 4-vector $dP^\mu/d\tau = \gamma dP^\mu/dt$:

$$\boxed{dP^\mu/d\tau = qu^\nu A_\nu{}^{,\mu}} \tag{21}$$

while, for the *mechanical 4-momentum* $p^\mu \equiv mu^\mu = (\gamma m, \gamma m\dot{\mathbf{x}})$,

$$\boxed{dp^\mu/d\tau = qu^\nu F_\nu{}^\mu.} \tag{22}$$

Note that if we 'dot' this with p_μ we have $p_\mu dp^\mu/d\tau = qp_\mu u^\nu F_\nu{}^\mu = qmu^\mu u^\nu F_{\nu\mu}$ which is the trace of the product of a symmetric tensor ($qmu^\mu u^\nu$) and an anti-symmetric one ($F_{\nu\mu}$) and so vanishes, so $d\vec{p}/dt$ is orthogonal to $\vec{p}$, something we could have inferred directly from the fact that, according to (15), the norm of the 4-momentum is $\vec{p} \cdot \vec{p} = p_\mu p^\mu = -m^2$ which is constant.

3.1.8 Gauge invariance of particle electrodynamics

The Lagrangian (4) is not gauge invariant (it changes if $A_\mu \rightarrow A_\mu + \xi_{,\mu}$). Neither is the action $S[x(t)] = \int dt L(x, \dot{x}, t)$ nor are the canonical momentum $\mathbf{P}$ and energy H (their equations of motion (21) not being gauge invariant).

But the Faraday tensor $F_{\mu\nu}$ *is* gauge invariant since $F'_{\mu\nu} = A'_{[\mu,\nu]} = A_{[\mu,\nu]} + \xi_{,\mu\nu} - \xi_{,\nu\mu} = A_{[\mu,\nu]} = F_{\mu\nu}$. And so the equation of motion for the 'mechanical' 4-momentum $dp^\mu/d\tau = qu^\nu F_\nu{}^\mu$ is gauge invariant and consequently the 4-momenta p^μ and the paths $x^\mu(\tau)$, which are obtained by integrating the $dp^\mu/d\tau = md^2x^\mu/d\tau^2 = qF_\nu{}^\mu dx^\nu/d\tau$, are also gauge invariant.

This can be seen directly from the definition of the particle paths as those which extremise the action. For a variation of a path $\mathbf{x}(t) \rightarrow \mathbf{x}(t) + \delta\mathbf{x}(t)$ the variation of the action is

$$\delta S = \int_{t_1}^{t_2} dt \left(\delta\mathbf{x} \cdot \frac{\partial L}{\partial \mathbf{x}} + \delta\dot{\mathbf{x}} \cdot \frac{\partial L}{\partial \dot{\mathbf{x}}} \right) = \left[\delta\mathbf{x} \cdot \frac{\partial L}{\partial \dot{\mathbf{x}}} \right]_{t_1}^{t_2} + \int_{t_1}^{t_2} dt \delta\mathbf{x} \cdot \left(\frac{\partial L}{\partial \mathbf{x}} - \frac{d}{dt} \frac{\partial L}{\partial \dot{\mathbf{x}}} \right) \tag{23}$$

where we have integrated by parts to eliminate $\delta\dot{\mathbf{x}}$.

In the integral here we have, from (4), the generalised force $\partial L/\partial x_i = q\dot{x}^\mu A_{\mu,i}$ and the generalised momentum $\partial L/\partial \dot{x}_i = \gamma m\dot{x}_i + qA_i$ whose time derivative is $d(\partial L/\partial \dot{x}_i)/dt = d(\gamma m\dot{x}_i)/dt + q\dot{x}^\mu A_{i,\mu}$. Consequently

$$\delta S = \left[\delta\mathbf{x} \cdot \frac{\partial L}{\partial \dot{\mathbf{x}}} \right]_{t_1}^{t_2} - \int_{t_1}^{t_2} dt \delta x_i \left[\frac{d}{dt} \frac{m\dot{x}_i}{\sqrt{1 - |\dot{\mathbf{x}}|^2}} - q(F_{ti} + \dot{x}_j F_{ji}) \right] \tag{24}$$

where the integrand – whose vanishing for extremal paths implies the Euler-Lagrange equation – is manifestly gauge invariant. The 'boundary term' $[\delta\mathbf{x} \cdot (\partial L/\partial \dot{\mathbf{x}})]$, on the other hand, is gauge-dependent.

3.1.9 The Hamilton-Jacobi equation

Requiring that δS vanish for two paths that begin and end at the same end points – so that the boundary term above vanishes – gives the equations of motion. On the other hand, if we consider different paths obeying the equations of motion – so the integral term vanishes – and assume that these have the same starting point but have different end points, we have $\delta S = \delta \mathbf{x} \cdot \partial L/\partial \dot{\mathbf{x}} = \delta \mathbf{x} \cdot \mathbf{P}$, so $\mathbf{P}$ is evidently the rate at which S changes with position (at fixed t_2):

$$\boxed{\mathbf{P} = \partial S/\partial \mathbf{x}} \tag{25}$$

And, since $dS = Ldt = (\partial S/\partial t)dt + (\partial S/\partial \mathbf{x}) \cdot d\mathbf{x} = (\partial S/\partial t + \mathbf{P} \cdot \dot{\mathbf{x}})dt$, it follows that $\partial S/\partial t = L - \mathbf{P} \cdot \dot{\mathbf{x}} = -H$, so

$$\boxed{H = -\partial S/\partial t} \tag{26}$$

the two above expressions being the space and time components of the covariant expression

$$\boxed{P^\mu = \partial^\mu S} \tag{27}$$

Using (25) and (26) to replace $\mathbf{P}$ and H in the energy-momentum relation (16) gives us the *Hamilton-Jacobi equation*:

$$\boxed{(\partial S/\partial t - q\varphi)^2 = |\partial S/\partial \mathbf{x} - q\mathbf{A}|^2 + m^2} \tag{28}$$

3.1.10 The wave-function and S-equation from Hamilton, Jacobi, Dirac and Feynman

We saw that for a collection of particles that start from the same position with a range of initial momenta their action $S = \int dtL$ has partial derivatives with respect to final times and positions given by $H = -\partial S/\partial t$ and $\mathbf{P} = \partial S/\partial \mathbf{x}$. It follows that the action of one of these particles, in the vicinity of some space-time point $(t_0, \mathbf{x}_0)$, which we take, temporarily, to be the origin of our coordinate system, is $S = S_0 - Ht + \mathbf{P} \cdot \mathbf{x}$.

According to the formulation of QM of Dirac and Feynman, the quantum mechanical amplitude to be at $(t, \mathbf{x})$, or, in other words, the *wave-function* $\psi(t, \mathbf{x})$, is proportional to $e^{iS/\hbar}$.

So here the wave-function is $\psi \propto e^{i(-Ht + \mathbf{P}\cdot\mathbf{x})/\hbar)}$, leading one to $i\hbar\partial_t\psi = H\psi$ and $-i\hbar\boldsymbol{\nabla}\psi = \mathbf{P}\psi$ and therefore the identification of H with the operator $i\hbar\partial_t$ (or, equivalently, of $P^t = H$ with $-i\hbar\partial^t$) and of $\mathbf{P}$ with the operator $-i\hbar\boldsymbol{\nabla}$.

In general, there may be more than one classical path leading to $(t_0, \mathbf{x}_0)$, in which case there will be interference in the total wave-function (this being, according to Feynman, the sum over all paths). Putting that aside, and using $P^\mu = (H, \mathbf{P}) \to -i\hbar\partial^\mu$ in the energy-momentum relation $p^\mu p_\mu = (P^\mu - qA^\mu)(P_\mu - qA_\mu) = -m^2$ gives the relativistic Schrödinger-equation alluded to above,

$$(\hbar\partial^\mu - iqA^\mu)(\hbar\partial_\mu - iqA_\mu)\psi = m^2\psi \tag{29}$$

which is gauge invariant in the sense that it has the property that, if $\psi(\vec{x})$ is a solution for some $A_\nu(\vec{x})$, then $\psi'(\vec{x}) = \psi(\vec{x})e^{i(q/\hbar)\xi(\vec{x})}$ is a solution for a different, but, to the extent that the phase of the wave function not be observable, physically equivalent, potential $A'_\nu(\vec{x}) = A_\nu(\vec{x}) + \xi_{,\nu}(\vec{x})$. This can be seen most readily, as already remarked, by noticing that $(\hbar\partial_\mu - iqA'_\mu)\psi' = e^{i(q/\hbar)\xi}(\hbar\partial_\mu - iqA_\mu)\psi$, so the phase factor $e^{i(q/\hbar)\xi}$ in ψ' 'passes through' the operator $\hbar\partial_\mu - iqA'_\mu$, converting it, in the process, back to $\hbar\partial_\mu - iqA_\mu$, with the result that, under a gauge transformation, (29) is just multiplied by the phase factor and remains valid.

3.2 The components of the Faraday tensor

The scalar and vector potentials were defined in terms of $\mathbf{E} = -\boldsymbol{\nabla}\varphi - \partial_t\mathbf{A}$ and $\mathbf{B} = \boldsymbol{\nabla} \times \mathbf{A}$. From the former we have $E_i = A_{t,i} - A_{i,t} = F_{ti}$ which, together with $F_{tt} = 0$, determines the top row and left column of $F_{\mu\nu}$ (we adopt the convention that the first index labels the rows and the second the columns). From the second we have $\mathbf{B} = (A_{[z,y]}, A_{[x,z]}, A_{[y,x]})$ which determines the spatial 3 by 3 sub-matrix of the Faraday tensor:

$$F_{\mu\nu} \equiv A_{[\mu,\nu]} = \begin{bmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -B_z & B_y \\ -E_y & B_z & 0 & -B_x \\ -E_z & -B_y & B_x & 0 \end{bmatrix}. \tag{30}$$

3.3 The Lorentz force law and the work equation

One can readily verify from (30) that the spatial components of (22) $dp_i/dt = q\dot{x}^\mu A_{[\mu,i,]}$ are the *Lorentz force law*:

$$\boxed{d\mathbf{p}/dt = q(\mathbf{E} + \dot{\mathbf{x}} \times \mathbf{B})} \tag{31}$$

while the time component $dp^t/dt = q\dot{x}^\nu F_\nu{}^t = -q\dot{x}^\nu F_{\nu t}$ implies the *work equation*

$$\boxed{dE/dt = q\dot{\mathbf{x}} \cdot \mathbf{E}} \tag{32}$$

where, as before, $E \equiv \gamma m = H - q\varphi$ is the mechanical (i.e. rest-mass plus kinetic) energy, and which states that it is only the $\mathbf{E}$ field that does work on a charged particle, the magnetic force being perpendicular to the particle's trajectory.

3.4 Maxwell's equations in terms of the Faraday tensor

Maxwell's equations relate $\mathbf{E}$, $\mathbf{B}$ and the charge and current density. How are they expressed in terms of the Faraday tensor?

One can readily verify, from the definition $F_{\mu\nu} \equiv A_{[\mu,\nu]}$, the 'cyclic' identity

$$\boxed{F_{\mu\nu,\gamma} + F_{\gamma\mu,\nu} + F_{\nu\gamma,\mu} = 0} \tag{33}$$

The only non-trivial combinations are when all indices are different, of which there are four, depending on whether the missing index is time or one of the 3 spatial components. The former yields $\boldsymbol{\nabla} \cdot \mathbf{B} = 0$ while the latter give the 3 equations $\boldsymbol{\nabla} \times \mathbf{E} + \partial_t \mathbf{B} = 0$, so (33) usefully encodes the 'homogeneous' pair of Maxwell's equations.

The 'inhomogeneous' (or 'sourced') pair of Maxwell's equations

$$\nabla \cdot \mathbf{E} = \rho \quad \text{and} \quad \boldsymbol{\nabla} \times \mathbf{B} - c^{-2} \partial_t \mathbf{E} = \mu_0 \mathbf{j}, \tag{34}$$

where ρ is the electric charge density and $\mathbf{j}$ is the electric current density, are encoded in

$$\boxed{F^{\mu\nu}{}_{,\mu} = j^\nu} \tag{35}$$

where the 4-current density is $j^\nu \equiv (\rho, \mathbf{j})$.

So Maxwell's equations, of which there are 8 in total, are, in terms of $F_{\mu\nu}$, given by the two 4-vector equations (33) and (35).

3.5 Conservation of electric charge

The inhomogeneous Maxwell's equations tell us that the divergence of the current is $j^\nu{}_{,\nu} = F^{\mu\nu}{}_{,\mu\nu}$. But this is invariant if we interchange the dummy indices $\mu \leftrightarrow \nu$, i.e. $F^{\mu\nu}{}_{,\mu\nu} = F^{\nu\mu}{}_{,\nu\mu}$. And $F^{\mu\nu}$ is anti-symmetric under $\mu \leftrightarrow \nu$ so applying this to the right hand side gives $F^{\mu\nu}{}_{,\mu\nu} = -F^{\mu\nu}{}_{,\nu\mu}$. But the partial derivatives in the last term commute so we have $F^{\mu\nu}{}_{,\mu\nu} = -F^{\mu\nu}{}_{,\mu\nu}$, implying that $F^{\mu\nu}{}_{,\mu\nu} = 0$ and therefore that

$$\boxed{j^\nu{}_{,\nu} = 0} \tag{36}$$

In 3+1 form this equation says that the rate of change of the charge density ρ is minus the divergence of the current $\mathbf{j}$, so the form of Maxwell's equations *require*, in some sense, that electric charge be conserved.

This is remarkable. Maxwell's equations, along with the Lorentz force and work equations, encapsulate empirical facts (the laws of Gauss, Ampère and Faraday and the dipole nature of magnets), based on measurements of forces. Another empirical fact of electromagnetism – one that pre-dates Maxwell – is that charge does not seem to spontaneously appear or disappear, and this, together with the constancy of charge to mass ratio for elementary particles, encouraged the idea that charge is carried by particles and is an intrinsic and fixed attribute. The above result shows that this does not have to be added to the equations as an additional assumption; it is already 'built in'. In a universe where charge were not precisely conserved, the equations of EM would have to be different from Maxwell's.

Writing (36) as $\partial_t\rho + \boldsymbol{\nabla}\cdot\mathbf{j} = 0$ and defining $Q \equiv \int d^3x\rho$ we have

$$\frac{dQ}{dt} = \frac{d}{dt}\int d^3x\rho = \int d^3x\partial_t\rho = -\int d^3x\boldsymbol{\nabla}\cdot\mathbf{j} = -\int dxdydz(\partial_x j_x + \partial_y j_y + \partial_z j_z) \tag{37}$$

but $\int_{x_1}^{x_2} dx\partial_x j_x = [j_x]_{x_1}^{x_2}$ and similarly for the equivalent integrals involving $\partial_y j_y$ and $\partial_z j_z$, so the right hand side vanishes if we let the limits of integration tend to $\pm\infty$ and we have that the total charge $Q \equiv \int d^3x\rho$ is conserved. If we have a bounded charge distribution, so we can draw some surface around it where $\mathbf{j} = 0$, then the charge within the surface is constant.

3.6 The relation between the 4-current density and particle density.

There are various useful ways to express the 4-current for a collection of particles, as described in detail in appendix (A). One is in terms of the spatial density of particles, which we can write as a sum of 3-dimensional Dirac δ-functions: $n(\mathbf{x}, t) = \sum_P \delta^{(3)}(\mathbf{x} - \mathbf{x}_P(t))$. In terms of these, the 4-current is

$$j^\nu(\mathbf{x}, t) = \sum_P q_P \dot{x}_P^\nu(t)\delta^{(3)}(\mathbf{x} - \mathbf{x}_P(t)). \tag{38}$$

Another is in terms of the space-time density of particles: a set of filaments $n(\vec{x}) = \sum_P \int d\tau\delta^{(4)}(\vec{x} - \vec{x}_P(\tau))$, where $\vec{x}_P(\tau)$ is the world-line, parameterised by proper time, and in relation to which

$$j^\nu(\vec{x}) = \sum_P q_P \int d\tau u_P^\nu(\tau)_P\delta^{(4)}(\vec{x} - \vec{x}_P(\tau)). \tag{39}$$

A third is in terms of the density in 6-dimensional phase space $f(\mathbf{x}, \mathbf{p}, t) = \sum_P \delta^{(3)}(\mathbf{x} - \mathbf{x}_P(t))\delta^{(3)}(\mathbf{p} - \mathbf{p}_P(t))$ for which, for particles of identical charge,

$$j^\nu(\vec{x}) = q\int d^3p\,\dot{x}^\nu f(\mathbf{x}, \mathbf{p}, t), \tag{40}$$

with the total current being, in general, the sum of this over species of particles with different charges.

All the above are equivalent. In the appendix (A) we show how the continuity equation (36) is implicit in each of the above definitions.

3.7 Transformation of the 4-current under a Lorentz boost

The charge current-density j^ν and particle current-density n^ν are both 4-vectors (they transform like x^ν under boosts), the number of particles in a spatial volume δV is $\delta N = n^t\delta V$, for example, being a scalar even though both n^t and δV are frame dependent. This is relatively easily understood; if we have a cubical volume of side L in the 'lab-frame' then, in the frame of particles moving in the x-direction relative to the lab, the x-separation between two events that occur at $x = \pm L/2$ – i.e. on opposite the ends of the cube – and at the same time – in the lab-frame – will have particle-frame x-separation enhanced by a factor γ and the particle-frame volume is γL^3. The number of particles in the box in the particle frame will be (γL^3) times the proper number density. The number of particles being Lorentz invariant it must be that the number density of particles with this velocity will be enhanced in the lab-frame by a factor γ relative to the proper number density, just like the enhancement of the time component of the particles' 4-momentum, and the mean volume per particle δV is therefore decreased by a factor $1/\gamma$ as compared to the proper value. The is consistent with the statement that the particle current density $\vec{n}$ is a 4-vector, having a proper – i.e. particle-frame – value $\vec{n} \to n^\nu = (n_{\rm proper}, 0, 0, 0)$, so its value in the lab-frame has time component $n^t = \gamma n_{\rm proper}$, and so $n^t\delta V$ is frame independent.

It is very common to refer to the 4-current-density as the '4-current'. This is a little dangerous, as one might get confused with $q\dot{x}^\mu$ that appears in the Lagrangian (4), or with δJ^ν in (121), whose spatial components are an electrical current (i.e. electrical charge times velocity), but which do not transform as 4-vectors.

3.8 Liouville's theorem for charged particles

In terms of the phase space density $f(\mathbf{x}, \mathbf{P})$, conservation of particles implies $\partial_t f + \boldsymbol{\nabla}_{\mathbf{x}} \cdot (\dot{\mathbf{x}} f) + \boldsymbol{\nabla}_{\mathbf{P}} \cdot (\dot{\mathbf{P}} f) = 0$. I.e. the rate of change of density at some point in phase-space is the 6-divergence of the particle 6-current $(\dot{\mathbf{x}} f, \dot{\mathbf{P}} f)$. Equivalently, the convective derivative of f is

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \dot{\mathbf{x}} \cdot \boldsymbol{\nabla}_{\mathbf{x}} f + \dot{\mathbf{P}} \cdot \boldsymbol{\nabla}_{\mathbf{P}} f = -f[\boldsymbol{\nabla}_{\mathbf{x}} \cdot \dot{\mathbf{x}} + \boldsymbol{\nabla}_{\mathbf{P}} \cdot \dot{\mathbf{P}}]. \tag{41}$$

But Hamilton's equations ($\dot{\mathbf{P}} = -\boldsymbol{\nabla}_{\mathbf{x}} H$ and $\dot{\mathbf{x}} = \boldsymbol{\nabla}_{\mathbf{P}} H$) say the right hand side is $-f \times (\boldsymbol{\nabla}_{\mathbf{x}} \cdot \boldsymbol{\nabla}_{\mathbf{P}} - \boldsymbol{\nabla}_{\mathbf{P}} \cdot \boldsymbol{\nabla}_{\mathbf{x}}) H$ which, because partial derivatives commute, vanishes and we have $df/dt = 0$.

In terms of the phase space density $f(\mathbf{x}, \mathbf{p})$ we have $\partial_t f + \boldsymbol{\nabla}_{\mathbf{x}} \cdot (\dot{\mathbf{x}} f) + \boldsymbol{\nabla}_{\mathbf{p}} \cdot (\dot{\mathbf{p}} f) = 0$ and

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \dot{\mathbf{x}} \cdot \boldsymbol{\nabla}_{\mathbf{x}} f + \dot{\mathbf{p}} \cdot \boldsymbol{\nabla}_{\mathbf{p}} f = -f[\boldsymbol{\nabla}_{\mathbf{x}} \cdot \dot{\mathbf{x}} + \boldsymbol{\nabla}_{\mathbf{p}} \cdot \dot{\mathbf{p}}] \tag{42}$$

We cannot now simply appeal to Hamilton's equation to show it, but this also vanishes: On the right hand side, $\boldsymbol{\nabla}_{\mathbf{x}} \cdot \dot{\mathbf{x}}$ is, by definition, the rate of change of $\dot{\mathbf{x}}$ with respect to position $\mathbf{x}$ at fixed $\mathbf{p}$. But $\mathbf{p} = m\dot{\mathbf{x}}/\sqrt{1 - |\dot{\mathbf{x}}|^2}$ implies $\dot{\mathbf{x}} = \mathbf{p}/\sqrt{m^2 + |\mathbf{p}|^2}$ which is solely a function of $\mathbf{p}$, so the rate of change of $\dot{\mathbf{x}}$ at fixed $\mathbf{p}$ is the same as rate of change of $\dot{\mathbf{x}}$ at fixed $\dot{\mathbf{x}}$, i.e. zero, so $\boldsymbol{\nabla}_{\mathbf{x}} \cdot \dot{\mathbf{x}} = 0$. The equation of motion is $\dot{p}_i = qF_{ti}(\mathbf{x}) + q\dot{x}_j F_{ji}(\mathbf{x})$ so the second term in brackets on the right hand side of (42) (being, by definition, the rate of change of $\dot{\mathbf{p}}$ with respect to $\mathbf{p}$ at fixed $\mathbf{x}$), is $\boldsymbol{\nabla}_{\mathbf{p}} \cdot \dot{\mathbf{p}} = qF_{ji} \partial \dot{x}_j / \partial p_i$. But from $\dot{\mathbf{x}} = \mathbf{p}/\sqrt{m^2 + |\mathbf{p}|^2}$ we have $\partial \dot{x}_j / \partial p_i = (\gamma m)^{-1} (\delta_{ij} - \dot{x}_i \dot{x}_j)$ which is symmetric. But F_{ji} is anti-symmetric, so (using an exactly analogous argument to that used above, flipping dummy indices etc.) we have $\boldsymbol{\nabla}_{\mathbf{p}} \cdot \dot{\mathbf{p}} = 0$.

Thus, even though the Lorentz force on a particle is velocity, and hence momentum, dependent, so, in general, $\partial \dot{p}_i / \partial p_j \neq 0$, the trace of this, which is the momentum-space divergence $\boldsymbol{\nabla}_{\mathbf{p}} \cdot \dot{\mathbf{p}} = 0$. Therefore both terms on the right of (42) vanish, and we have, as a consequence, *Liouville's theorem:*

$$\boxed{df/dt = 0} \tag{43}$$

i.e. the phase space density along any particle trajectory for charged particles moving under the influence of an external electromagnetic field $F_{\mu\nu}(\mathbf{x}, t)$ (but ignoring their collisions with one another) is constant. Nothing here is gauge dependent.

3.9 The stress-energy tensor for charged particles

As well as electric charge, particles can transport other attributes. In particular they transport their 4-momentum. That means that just as there is a current-density 4-vector j^ν there is a current for the 4-momentum. But as the quantity being transported is itself a 4-vector, its current is a rank two tensor $T^{\mu\nu}$, being defined as the rate at which the ν^{th} component of 4-momentum is being transported along the μ^{th} coordinate axis. Another distinction with charge is that the 4-momentum is not, in general, constant along the particle trajectory.

3.9.1 Stress-energy in terms of the 3, 4 or 6 dimensional particle densities

This motivates one to define the 4-momentum current for particles, also known as the stress-energy tensor, simply by replacing the charge q in the formulae for the 4-current-density j^μ by the particles' 4-momenta $\vec{p} = m\vec{u}$. As with the charge 4-current density, this can be expressed in terms of the 3, 4 or 6-dimensional particle densities:

$$\begin{aligned} T^{\mu\nu} &= m \sum_P \dot{x}_P^\mu(t) u_P(t) \delta^{(3)}(\mathbf{x} - \mathbf{x}_P(t)) \\ &= m \sum_P \int d\tau u_P^\mu(\tau) u_P^\nu(\tau) \delta^{(4)}(\vec{x} - \vec{x}_P(t)) \\ &= \int d^3p f(\mathbf{x}, \mathbf{p}, t) \dot{x}^\mu(\mathbf{p}) p^\nu(\mathbf{p}) = \int \frac{d^3p}{p^t} f(\mathbf{x}, \mathbf{p}, t) p^\mu p^\nu \end{aligned} \tag{44}$$

3.9.2 Continuity of energy and momentum

If the Faraday tensor vanishes, the 4 components of the 4-momentum of each particle are independent of time, so the four 4-current densities $T^{\mu t}$, $T^{\mu x}$, $T^{\mu y}$ and $T^{\mu z}$, each analogous to the charge 4-current density j^μ, are all conserved, or

$$T^{\mu\nu}{}_{,\mu} = 0. \tag{45}$$

Each of the $\nu = t, x, y, z$ components of this equation are saying that the rate of change of density $T^{t\nu}$ of ν-momentum in a volume is the (integral over the volume of the) divergence of the ν-momentum flux density $T^{i\nu}$.

If $F_{\mu\nu} \neq 0$, the EM field will be transferring 4-momentum to the particles at a rate, per particle, given by $\dot{p}^\nu = F_\mu{}^\nu \dot{x}^\mu_P$, so there will be an additional rate of change in the amount of ν-momentum in a volume δV, over and above that being convected in or out, given by the sum over the particles of $\sum\limits_{P \in \delta V} \dot{p}^\nu = \sum\limits_{P \in \delta V} \dot{x}^\mu F_\mu{}^\nu$. So there will be an addition rate of change in the *density* of ν-momentum equal to this divided by the volume: $F_\mu{}^\nu \delta V^{-1} \sum\limits_{P \in \delta V} \dot{x}^\mu = F_\mu{}^\nu j^\mu$ and we therefore have

$$\boxed{T^{\mu\nu}{}_{,\mu} = j^\mu F_\mu{}^\nu} \tag{46}$$

This may also be derived, more arduously, directly from the various expressions for $T^{\mu\nu}$ in (44) as shown in appendix (B).

If $F_{\mu\nu} = 0$, the four continuity equations $T^{\mu\nu}{}_{,\mu} = 0$ imply four globally conserved quantities $p^\nu_{\text{tot}} = \int d^3x T^{t\nu}$ each analogous to the globally conserved charge Q. While if $F_{\mu\nu} \neq 0$ there will be a transfer of energy and momentum to the particles given, per unit volume, in (46) by $j^\mu F_\mu{}^\nu$, the time and space components of which are the rate-of-work density and the Lorentz force density respectively.

We could have also computed the *canonical stress-energy tensor* $T_{\text{c}}^{\mu\nu}$ defined as the flux density of canonical 4-momentum P^ν. This is gauge dependent and has a divergence $T_{\text{c}}^{\mu\nu}{}_{,\mu} = j^\mu A_\mu{}^{,\nu}$ with a source which is gauge dependent also. Both sides of (46), on the other hand, are gauge invariant.

3.10 The stress-energy tensor for EM radiation

3.10.1 Energy density of the electromagnetic field

Consider the work done pulling two capacitor plates – with separation along the x-axis – apart. Assuming, for simplicity, a separation much less than their size of the plates, the field between the plates will be $\mathbf{E} = \hat{\mathbf{x}} E_x$ with, by Gauss's law, $E_x = \Sigma$, the charge density. The force is $F = A E_x \Sigma / 2$ with A the area of the plates – the factor 1/2 coming from the fact that the field ramps from zero to the inter-plate value passing through the plate so the mean field is $E_x/2$ – and so the work done in increasing the separation by δx is $\delta W = AF\delta x = A\delta x E_x^2/2 = \delta V E_x^2/2$. Equating this to the change in the volume times the energy density $\mathcal{E}$ of the field gives $\mathcal{E} = |\mathbf{E}|^2/2$.

It is crucial to recognise that we are here attributing the energy entirely to the field. In this process there was no change in the mechanical energy of the plates as they started and ended at rest. If we release the plates, they will gain mechanical energy at the expense of the field and field plus mechanical energy will be conserved. The *canonical* energy H of the particles in the plates, on the other hand, contains, in addition to the mechanical energy, contributions from the term $q\varphi$, so canonical particle energy plus field energy (as defined above) would not be conserved. It is somewhat arbitrary how we assign the energy; whether we say it resides in the field or whether it resides in the particles.

Similar considerations can be applied to the magnetic field. Consider a long solenoid of length L with radius r and with N turns. Ampère's law says the field is $B = NJ/L$, where J is the current. If we ramp the current up to increase the field there will be an induced EMF (according to Faraday's law) such that $E \times 2\pi r = AdB/dt$ where $A = \pi r^2$ is the area. The power required to drive the current J against the induced electric field is $dW/dt = EJ \times 2\pi r N = AL \times BdB/dt$. It follows that the work done is $dW = VdB^2/2$, with $V = AL$ the volume, so, attributing this to the the energy density of the magnetic field gives $\mathcal{E} = |\mathbf{B}|^2/2$. The sum of the electric and magnetic field densities is therefore $\mathcal{E} = (|\mathbf{E}|^2 + |\mathbf{B}|^2)/2$.

3.10.2 Poynting's theorem

Dotting Faraday's law $\partial_t \mathbf{B} = -\boldsymbol{\nabla} \times \mathbf{E}$ with $\mathbf{B}$ and dotting the Ampère-Maxwell equation $\partial_t \mathbf{E} = \boldsymbol{\nabla} \times \mathbf{B} - \mathbf{j} \cdot \mathbf{E}$ with $\mathbf{E}$ and adding gives $\partial_t (|\mathbf{E}|^2 + |\mathbf{B}|^2)/2 = \mathbf{E} \cdot (\boldsymbol{\nabla} \times \mathbf{B}) - \mathbf{B} \cdot (\boldsymbol{\nabla} \times \mathbf{E}) - \mathbf{j} \cdot \mathbf{E}$. Using the identity

$\nabla \cdot (\mathbf{a} \times \mathbf{b}) = \mathbf{b} \cdot (\nabla \times \mathbf{a}) - \mathbf{a} \cdot (\nabla \times \mathbf{b})$ gives *Poynting's theorem:*

$$\boxed{\partial_t \mathcal{E} + \nabla \cdot \mathbf{S} = -\mathbf{j} \cdot \mathbf{E}} \tag{47}$$

where the *Poynting vector* (or *Poynting flux*) is defined as

$$\boxed{\mathbf{S} \equiv \mathbf{E} \times \mathbf{B}} \tag{48}$$

The right hand side of (47) is minus the rate at which the particles are gaining energy from the field. So Poynting's theorem expresses conservation of total energy in which we identify the Poynting flux with the energy flux density.

3.10.3 Maxwell's electromagnetic stress tensor

Considering again a capacitor, now static with its plates kept apart by being pulled by springs. It is evident that since the flux of momentum must be continuous then, as there is a flux of momentum in the stretched springs, there must be a flux associated with the **E**-field between the plates also. If we let the separation of the plates $\mathbf{d}$ (and therefore also the field $\mathbf{E}$) be parallel to $\hat{\mathbf{x}}$ and their area be A, then the momentum flux is $-AE_x^2/2$. This is negative since the springs (and the field also) are in tension, so the spring at negative $\mathbf{x}$ is delivering negative p_x to the plate at $-\mathbf{d}/2$; this is transferred within the plate to the field (the plates being static so there being no change of their momentum) and this is delivered to the plate at $+\mathbf{d}/2$, which it, in turn, transfers to the spring at positive $\mathbf{x}$. The rate at which x-momentum is being transported along the x-axis per unit area (i.e. the component of the stress (or momentum flux density) tensor T_{xx}) is, in this situation, $-E_x^2/2$.

If we consider a charged sphere, then, at the North pole, there is a field in the z-direction, so there is a (negative) flux density of z-momentum. But that flux density falls off as the square of the field strength, or as $1/r^4$. This may seem strange as this does not look like momentum is being conserved since the flux density times an element of area dA subtending a certain solid angle $d\Omega$ at the centre of the sphere is falling off like $1/r^2$. What this is telling us is that there is, in fact, not just a flux of momentum along the field lines, there must be momentum flowing in the direction perpendicular to the field. That there must be such a transverse momentum flux for a magnetic field is apparent in a solenoid where there is an outward $\mathbf{j} \times \mathbf{B}$ force acting on the coils, so if there is continuity of momentum flux there must be a positive flux density of momentum within the solenoid in the directions transverse to the field.

To elucidate this, consider a skinny conical plug-shaped cylinder of height δh sitting atop the North pole. From the foregoing, more z-momentum is flowing out of this volume through its bottom than is flowing in through its top. If that were all the momentum transport, the amount of z momentum in the volume would be changing at a rate $\dot{p}_z \equiv dp_z/dt = \delta(\pi r_p^2(r) \times E_z^2(r)/2)$. The plug radius is $r_p(r) = r_p \times r/r_s$ and the field is $E_z(r) = E_z \times r_s^2/r^2$ (where r_s is the radius of the sphere and where r_p and E_p without arguments represent the values at the surface of the sphere) so $\dot{p}_z = \pi r_s^2 r_p^2 E_z^2 \delta(1/2r^2)$. Assuming $\delta h \ll r_s$ this is $\dot{p}_z = -2\pi r_s^{-1} r_p^2 \delta h E_z^2$. For momentum to be conserved, there must evidently be a negative flux of z-momentum flowing out the sides of the cylinder (i.e. z-momentum flowing in).

The area of the sides of the plug is $2\pi r_p \delta h$ so in modulus the flux density of z-momentum inward at the wall must be rate of change of momentum divided by area, or $|\dot{p}_z|/(2\pi r_p \delta h) = (r_p/r_s)E_z^2$. This must point towards the polar axis, or (assuming positive E_z) anti-parallel to the vector $\mathbf{E}_\perp \equiv (E_x, E_y, 0) = (x_p, y_p, 0) \times E_z/r_s$. It follows that the three components of the flux density of z-momentum are $T_{iz} = -E_z(E_x, E_y, E_z/2)$. There is nothing special about the z-axis, of course, so the components of the flux density of x-momentum is, in general, $T_{ix} = -E_x(E_x/2, E_y, E_z)$ while that of y-momentum is $T_{iy} = -E_y(E_x, E_y/2, E_z)$. Combining these gives, for the flux density of j-momentum along the i^{th} coordinate axis for an arbitrary electric field

$$T_{ij} = -E_i E_j + \tfrac{1}{2}\delta_{ij}|\mathbf{E}|^2. \tag{49}$$

It is interesting to take the divergence of (49). Assuming $\partial_t \mathbf{A} = 0$ (so the magnetic field, if any, is static) we have $E_i = -\varphi_{,i}$, so the divergence of the j^{th} component of (49) is $T_{ij,i} = -\varphi_{,ii}\varphi_{,j} - \varphi_{,i}\varphi_{,ji} + \delta_{ij}\varphi_{,k}\varphi_{,ki}$. The last two terms cancel, so we have $T_{ij,i} = -\varphi_{,ii}\varphi_{,j} = E_j \nabla^2 \varphi$. According to Poisson's equation, this vanishes if there are no charges present, and the momentum flux density is then divergence-free. Otherwise, if integrated over a region containing charges, this is equal to the rate at which those charges are losing j-momentum. Invoking the divergence theorem, this says that the inward flux of field j-momentum across a surface (being minus the integral of the divergence of the j-momentum flux density over the enclosed volume) is equal to the rate at which charges in the volume are gaining j-momentum as a result of the presence of the E-field.

Identical considerations apply to magnetic fields. One might, for instance, consider the stresses in the monopole-like field around one end of a long skinny solenoid. Such considerations led Maxwell to his electromagnetic stress tensor $\boldsymbol{\sigma}$, which is, aside from a minus sign, the sum of (49) and a identical expression for the magnetic field with $\mathbf{E}$ replaced by $\mathbf{B}$ so the combined stress tensor is

$$\boxed{T_{ij} = -\sigma_{ij} = -E_i E_j - B_i B_j + \tfrac{1}{2}\delta_{ij}(|\mathbf{E}|^2 + |\mathbf{B}|^2)} \tag{50}$$

Returning to the case of a constant E-field along the x-axis, the stress tensor is $T_{ij} = \frac{1}{2} E_x^2 \mathrm{diag}(-1, 1, 1)$ so there is tension along the field line and pressure in the perpendicular direction (i.e. positive flux of y-momentum in the y-direction and similarly for z). This can be understood qualitatively in terms of the work needed (or released) if we change the field configuration. The tension along the field lines follows from the fact that pulling the capacitor plates apart requires energy whereas if we were to increase the area of the plates for a fixed amount of charge the field strength times the area would be constant so the total field energy (being the field energy density times the area) would go down and this would release energy. As discussed above, such transverse stress for a magnetic field can be inferred from the outward force on the coils of a solenoid.

3.10.4 The momentum density of the electromagnetic field

Poynting's theorem suggests that the left (i.e. energy) column of the stress energy tensor for radiation has components $T_\mathrm{r}^{\mu t} = (\mathcal{E}, \mathbf{S})$ since the theorem then says that $T_\mathrm{r}^{\mu t}{}_{,\mu} = -\mathbf{j} \cdot \mathbf{E} = -j^\mu F_\mu{}^t = -T_\mathrm{m}^{\mu t}{}_{,\mu}$ where $T_\mathrm{m}^{\mu t}$ is the energy column of the mechanical stress energy tensor. So the sum of the 4-divergences of the energy columns for the particles and the field vanishes and consequently the total (mechanical plus field) energy is conserved.

Maxwell's stress tensor $-\boldsymbol{\sigma}$ provides the spatial parts of $T_\mathrm{r}^{\mu\nu}$; the momentum flux density. All that remains to be determined is the spatial part of the top row (i.e. the momentum density). One might guess that, like $T_\mathrm{m}^{\mu\nu}$, the radiation stress tensor $T_\mathrm{r}^{\mu\nu}$ should be symmetric, so $T_\mathrm{r}^{t\mu} = (\mathcal{E}, \mathbf{S})$ and hence the momentum density be simply the Poynting (energy) flux density $\mathbf{S}$.

That is correct. With $T_\mathrm{r}^{tj} = S_j$ and $T_\mathrm{r}^{ij} = -\sigma_{ij}$, the 4-divergence of the j^th column turns out to be $T_\mathrm{r}^{\mu j}{}_{,\mu} = -(\rho\mathbf{E} + \mathbf{j} \times \mathbf{B})_j$, which is minus the Lorentz force density. And this is $-j^\mu F_\mu{}^j$ which is $-T_\mathrm{m}^{\mu j}{}_{,\mu}$. So total j^th component of spatial momentum is conserved. To prove this, consider first the time-time component $T_\mathrm{r}^{tj}{}_{,t} = \partial_t S_j$. This is

$$\partial_t \mathbf{S} = \partial_t(\mathbf{E} \times \mathbf{B}) = (\partial_t \mathbf{E} \times \mathbf{B}) + (\mathbf{E} \times \partial_t \mathbf{B}) = -\mathbf{j} \times \mathbf{B} - \mathbf{E} \times (\boldsymbol{\nabla} \times \mathbf{E}) - \mathbf{B} \times (\boldsymbol{\nabla} \times \mathbf{B}) \tag{51}$$

where we have used Maxwell's equations. Using the identity $\mathbf{a} \times (\boldsymbol{\nabla} \times \mathbf{a}) = \frac{1}{2}\boldsymbol{\nabla}|\mathbf{a}|^2 - (\mathbf{a} \cdot \boldsymbol{\nabla})\mathbf{a}$ we find, in component form,

$$T_\mathrm{r}^{tj}{}_{,t} = \partial_t S_j = -(\mathbf{j} \times \mathbf{B})_j - E_i E_{i,j} + E_i E_{j,i} - B_i B_{i,j} + B_i B_{j,i}. \tag{52}$$

The rest of the 4-divergence is the 3-divergence of the spatial part. This is

$$T_\mathrm{r}^{ij}{}_{,i} = -\sigma_{ij,i} = -E_j E_{i,i} + E_i E_{i,j} - E_i E_{j,i} - B_j B_{i,i} + B_i B_{i,j} - B_i B_{j,i}. \tag{53}$$

adding (52) and (53), four of the terms in common cancel, while $B_{i,i} = \boldsymbol{\nabla} \cdot \mathbf{B} = 0$ and $E_{i,i} = \boldsymbol{\nabla} \cdot \mathbf{E} = \rho$, so we have, finally, for $T_\mathrm{r}^{\mu j}{}_{,\mu} = T_\mathrm{r}^{tj}{}_{,t} + T_\mathrm{r}^{ij}{}_{,i}$:

$$T_\mathrm{r}^{\mu j}{}_{,\mu} = -(\rho\mathbf{E} + \mathbf{j} \times \mathbf{B})_j = -j^\mu F_\mu{}^j. \tag{54}$$

The above equation is the analogue for the 3-momentum of Poynting's theorem for energy conservation.

We have thus determined that the stress-energy tensor for the radiation is

$$\boxed{T_\mathrm{r}^{\mu\nu} = \begin{bmatrix} \frac{1}{2}(|\mathbf{E}|^2 + |\mathbf{B}|^2) & \mathbf{S} \\ \mathbf{S} & -\boldsymbol{\sigma} \end{bmatrix}} \tag{55}$$

the left column coming from Poynting's theorem and consisting of the field energy density $\mathcal{E} = (|\mathbf{E}|^2 + |\mathbf{B}|^2)/2$ and the energy flux density $\mathbf{S}$. The upper right is the 3-momentum density (also equal to $\mathbf{S}$) and the bottom right is the flux density of 3-momentum: (minus) the Maxwell stress. We will return to this presently. Next, however, we will see how this emerges from a Lagrangian field theory approach.

3.10.5 The Lagrangian for electromagnetism in the presence of charges

The Lagrangian (4) is that of a free particle $L = -m/\gamma$ plus an interaction term $L_{\rm int} = q\dot{x}^\mu A_\mu$. Generalising to a collection of particles, or a continuous distribution of charge, we have $L_{\rm int} = \sum q\dot{x}^\mu A_\mu = \int d^3x A_\mu q \int d^3p f(\mathbf{x},\mathbf{p})\dot{x}^\mu = \int d^3x A_\mu j^\mu$. So $L_{\rm int} = \int d^3x \mathcal{L}_{\rm int}$ with interaction Hamiltonian density $\mathcal{L}_{\rm int} = j^\mu A_\mu$. Adding this to the free-field Lagrangian density for radiation $-F^{\mu\nu}F_{\mu\nu}/4$ we have the Lagrangian density for the electromagnetic field in the presence of a current $j^\mu(\vec{x})$

$$\mathcal{L}(A_\alpha, A_{\alpha,\beta}, x^\gamma) = -\tfrac{1}{4}F^{\mu\nu}F_{\mu\nu} + j^\mu A_\mu. \tag{56}$$

As justification for this, we may note that from this and the definition $F_{\mu\nu} \equiv A_{\mu,\nu} - A_{\nu,\mu}$ we find $\partial\mathcal{L}/\partial A_{\alpha,\beta} = F^{\beta\alpha}$ while $\partial\mathcal{L}/\partial A_\alpha = j^\alpha$ so the Euler-Lagrange equations

$$\frac{\partial}{\partial x^\beta}\left(\frac{\partial\mathcal{L}}{\partial A_{\alpha,\beta}}\right) = \frac{\partial\mathcal{L}}{\partial A_\alpha} \tag{57}$$

become simply ${F^{\beta\alpha}}_{,\beta} = j^\alpha$ which are the inhomogeneous Maxwell's equations.

3.10.6 The canonical stress-energy tensor for the radiation

To obtain a continuity equation for energy and momentum of the radiation à la Noether we take the partial derivative of $\mathcal{L}(x^\gamma) = \mathcal{L}(A^\alpha(x^\gamma), {A^\alpha}_{,\beta}(x^\gamma), x^\gamma)$ with respect to x^ν. Using the chain rule gives

$$\frac{\partial\mathcal{L}(\vec{x})}{\partial x^\nu} = \frac{\partial\mathcal{L}}{\partial A_\alpha}A_{\alpha,\nu} + \frac{\partial\mathcal{L}}{\partial A_{\alpha,\mu}}A_{\alpha,\mu\nu} + \frac{\partial\mathcal{L}}{\partial x^\nu} \tag{58}$$

where, just to be clear, the last term represents the derivative of $\mathcal{L}(A_\alpha, A_{\alpha,\beta}, x^\gamma)$ with respect to its final argument keeping A_α and $A_{\alpha,\beta}$ fixed, and is equal to ${j^\alpha}_{,\nu}A_\alpha$ since the only *explicit* functional dependence of $\mathcal{L}$ on $\vec{x}$ is through the current $j^\alpha(\vec{x})$.

Eliminating $\partial\mathcal{L}/\partial A_\alpha$ from this using (57), and using $A_{\alpha,\mu\nu} = \partial_\mu A_{\alpha,\nu}$, the first two terms on the right combine, as usual, to become the derivative of a single product, so

$$\partial_\nu\mathcal{L}(x^\gamma) = \partial_\mu\left(A_{\alpha,\nu}\frac{\partial\mathcal{L}}{\partial A_{\alpha,\mu}}\right) + \frac{\partial\mathcal{L}}{\partial x^\nu} = \partial_\mu\left(A_{\alpha,\nu}F^{\mu\alpha}\right) + {j^\alpha}_{,\nu}A_\alpha. \tag{59}$$

Finally, using $\partial_\nu\mathcal{L}\vec{x} = \delta^\mu_\nu\partial_\mu\mathcal{L}(\vec{x}) = -\delta^\mu_\nu\partial_\mu(\frac{1}{4}F^{\alpha\beta}F_{\alpha\beta} - j^\alpha A_\alpha)$, we obtain the continuity equation

$$T^{\mu}_{{\rm c}\ \nu,\mu} = -j^\mu A_{\mu,\nu} \tag{60}$$

where the *canonical stress tensor for radiation* is

$$\boxed{T^{\mu}_{{\rm c}\ \nu} \equiv F^{\alpha\mu}A_{\alpha,\nu} - \tfrac{1}{4}\delta^\mu_\nu F^{\alpha\beta}F_{\alpha\beta}} \tag{61}$$

3.10.7 The symmetric stress-energy tensor for the radiation

The stress-energy tensor (61) is, like the canonical stress-energy tensor for particles, gauge-dependent. This might not seem surprising since we have obtained this from a Lagrangian density (56) containing a gauge-dependent interaction term. But even if we remove the interaction term, and start with the gauge invariant free-field electromagnetic Lagrangian density $\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$ alone, we still end up with the gauge dependent stress-energy tensor (61). Another unsatisfactory feature of (61) is that the source-term for its divergence in (60) is also gauge-dependent. It is in fact, however, minus the source term for the canonical stress tensor for the particles, so the sum of the canonical stress-energy tensors for the radiation and particles has a vanishing divergence; i.e. the total canonical 4-momentum is conserved. Also, it does not, at first sight, seem to agree with what one might expect from Poynting's theorem, but again that is not surprising as what appears in that theorem is $\mathbf{j}\cdot\mathbf{E}$ which is the rate at which the mechanical, rather than the canonical, energy of the particles is changing.

These unsatisfactory features are readily avoidable. This is because the continuity equation (60), while valid, does not uniquely specify the stress-energy tensor. It is possible to modify the radiation stress-energy tensor without affecting the continuity equation, and, in the process get rid of these problems.

Imagine we were to add to $T_{c\ \nu}^{\mu}$ an additional term of the form $\partial_\alpha(C^{\mu\alpha}{}_\nu)$. Then, when we take the divergence, we get an extra term $\partial_\mu\partial_\alpha(C^{\mu\alpha}{}_\nu)$. If $C^{\mu\alpha}{}_\nu$ is anti-symmetric under $\alpha \leftrightarrow \mu$ then the extra divergence will vanish. Looking at the (gauge-dependent) first term in (61) suggests that we might want to try something like $C^{\mu\alpha}{}_\nu = F^{\mu\alpha}A_\nu$. This is indeed anti-symmetric under $\alpha \leftrightarrow \mu$ and would add to $T_{c\ \nu}^{\mu}$ a divergence-free contribution

$$\partial_\alpha C^{\mu\alpha}{}_\nu = F^{\mu\alpha}A_{\nu,\alpha} + F^{\mu\alpha}{}_{,\alpha}A_\nu = F^{\mu\alpha}A_{\nu,\alpha} - j^\mu A_\nu \tag{62}$$

where we have used the inhomogeneous Maxwell's equations: $F^{\alpha\mu}{}_{,\alpha} = j^\mu$. The first term here is $F^{\mu\alpha}A_{\nu,\alpha} = -F^{\alpha\mu}A_{\nu,\alpha}$ which, when combined with $F^{\alpha\mu}A_{\alpha,\nu}$ in (61), gives the gauge invariant product $F^{\alpha\mu}F_{\alpha\nu}$.

Adding $\partial_\alpha C^{\mu\alpha}{}_\nu + j^\mu A_\nu = F^{\mu\alpha}A_{\nu,\alpha}$ to (61) gives the symmetric stress-tensor

$$\boxed{T_{r\ \nu}^{\mu} \equiv F^{\mu\alpha}F_{\nu\alpha} - \tfrac{1}{4}\delta^\mu_\nu F^{\alpha\beta}F_{\alpha\beta}} \tag{63}$$

while changing its divergence, on the right hand side of (60), from $-j^\mu A_{\mu,\nu}$ to $-j^\mu A_{\mu,\nu} + \partial_\mu(j^\mu A_\nu) = -j^\mu F_{\mu\nu}$ (invoking charge conservation $j^\mu{}_{,\mu} = 0$) so, on raising the index ν,

$$\boxed{T_{r\ \ ,\mu}^{\mu\nu} = -j^\mu F_\mu{}^\nu} \tag{64}$$

So $T_r^{\mu\nu}$ is a symmetric, gauge-invariant tensor that depends only on the radiation field $F_{\mu\nu}$, and has a 4-divergence (64) with a source term which is gauge invariant also. Moreover, this source term is just the opposite to that which sources the divergence of the mechanical stress (46), so

$$\boxed{T_{r\ \ ,\mu}^{\mu\nu} + T_{m\ \ ,\mu}^{\mu\nu} = 0} \tag{65}$$

so whatever energy and momentum is given up by the radiation appears in the stress tensor for the matter and vice versa, the combination $T_r^{\mu\nu} + T_m^{\mu\nu}$ being conserved. The time component of the above equation expresses conservation of total (field plus mechanical) energy – in fact it is just Poynting's theorem – and the spatial components are the expression of Newton's law of action being equal and opposite to reaction.

If we compute the components of $T_r^{\mu\nu}$ in terms of $\mathbf{E}$ and $\mathbf{B}$ (see appendix C), we find

$$T_r^{\alpha\beta} = \begin{bmatrix} \frac{1}{2}(|\mathbf{E}|^2 + |\mathbf{B}|^2) & \mathbf{S} \\ \mathbf{S} & -\boldsymbol{\sigma} \end{bmatrix} \tag{66}$$

where we see $T_r^{tt} = (|\mathbf{E}|^2 + |\mathbf{B}|^2)/2$, the usual expression for the energy density of radiation, $T_r^{ti} = T_r^{it} = S_i$ where $\mathbf{S} \equiv \mathbf{E} \times \mathbf{B}$ is the Poynting energy flux density (also the momentum density), while $\boldsymbol{\sigma} \equiv \mathbf{EE} + \mathbf{BB} - \frac{1}{2}\mathbf{I}(|\mathbf{E}|^2 + |\mathbf{B}|^2)$, with $\mathbf{I}$ the 3-by-3 identity matrix, is the Maxwell stress tensor.

We arrived at (63) and (64) by taking the derivative of the Lagrangian density. This actually led us to the canonical stress tensor, which we then had to massage to obtain the symmetric, gauge-invariant version. A simpler alternative would have been to *postulate* (63) based on Poynting's theorem. Directly taking its derivative gives $T_{r\ \nu,\mu}^{\mu} = F^{\mu\alpha}{}_{,\mu}F_{\nu\alpha} + F^{\mu\alpha}F_{\nu\alpha,\mu} - \frac{1}{2}F^{\alpha\beta}F_{\alpha\beta,\nu}$. The first term on the right is, from Maxwell's equations, equal to $-j^\alpha F_{\alpha\nu}$, so the other terms must vanish. To show this we replace the dummy index μ by α in the second term, so the last two terms become $-\frac{1}{2}F^{\alpha\beta}[2F_{\nu\alpha,\beta} + F_{\alpha\beta,\nu}]$ and invoking the definition of $F_{\alpha\beta} \equiv A_{[\alpha,\beta]}$ this is $-\frac{1}{2}F^{\alpha\beta}[2A_{\nu,\alpha\beta} - (A_{\alpha,\nu\beta} + A_{\beta,\nu\alpha})]$ where by inspection $[\ldots]$ is symmetric under $\alpha \leftrightarrow \beta$ so this, when contracted with the anti-symmetric $F^{\alpha\beta}$ vanishes and we obtain (64).

4 Classical relativistic scalar field electrodynamics

4.1 The free real massive scalar field

Here we review the simplest scalar field; a real field with no coupling to any other fields (a 'free-field'). We start with the Lagrangian, from which we obtain the equation of motion (the Klein-Gordon equation – first obtained by Schrödinger and Klein but as a relativistic version of the S-equation rather than as a classical field. We then derive the continuity equation for the stress-energy tensor that expresses continuity and conservation of energy and momentum. We then discuss the dispersion relation for solutions of the Klein-Gordon equation and how the general free-field solution can be written as a Fourier synthesis (as a sum of positive and negative frequency $\mathbf{k}$-modes) and then we discuss phase- and group-velocities and the properties of wave-packets.

4.1.1 The Lagrangian density

The massive scalar field is mathematically equivalent to a lattice of masses coupled to each other by springs, much like a mattress, and with a harmonic potential well constraining them to their rest positions.

The Lagrangian is the sum of the kinetic energies of the masses minus the sum of the potential energies of all of the springs.

We assume that the masses and springs are all identical, and that we are dealing with wave-like displacements with wavelength much greater than the lattice space, so we can take the *continuum limit*.

The Lagrangian is then the space integral of a *Lagrangian density* $\mathcal{L}$ which is a function of the displacement and and of its time- and space-derivatives: $\mathcal{L} = \mathcal{L}(\phi_{,\mu}, \phi)$. With suitable choices of units, the Lagrangian density is then $\mathcal{L} = \frac{1}{2}(\dot{\phi}^2 - (\nabla\phi)^2 - m^2\phi^2)$, or, in 4-notation,

$$\mathcal{L} = -\tfrac{1}{2}(\phi_{,\mu}\phi^{,\mu} + m^2\phi^2) \tag{67}$$

which is invariant under Lorentz or Lorentz-like transformations. If we were modelling the mattress, the dimensionless velocity parameter β appearing in the transformation matrix would be the physical velocity divided by the asymptotic speed of sound waves in the limit of high spatial frequency. Here we are more interested in modelling scalar fields such as the inflaton that is invoked to drive inflation in the early universe; the quintessence field that is invoked to drive cosmic acceleration in the late-time universe; or the axion or 'fuzzy' dark matter candidates. In that case the Lagrangian density is invariant under the normal relativistic Lorentz transformation.

4.1.2 Units of $\mathcal{L}$, m and ϕ

There is some ambiguity in the dimensions of the field – which lives in some abstract space – and therefore of the Lagrangian density. However, we want the stress tensor obtained from this – which has the same units as $\mathcal{L}$ – to have T^{tt} equal to the energy density, which has units of mass times velocity squared divided by volume or $\mathcal{L} \rightarrow [MT^{-2}L^{-1}]$. The parameter m has units of frequency $m \rightarrow [T^{-1}]$ so the field must then have units $\phi \rightarrow [M^{1/2}L^{-1/2}]$.

If, in addition to having $c = 1$, we choose units such that $\hbar = 1$, this means that we can express a length L as an equivalent inverse mass $M^{-1}(L) = cL/\hbar$, and express a time T as an equivalent inverse mass $M^{-1}(T) = c^2T/\hbar$. That means we can express a value of ϕ as an equivalent mass $M(\phi) = \sqrt{\hbar/c}\phi$ and the frequency m also as an equivalent mass $M(m) = \hbar m/c^2$ (this being the mass for which the Compton frequency is m), so we can speak loosely of ϕ having 'units' of mass and $\mathcal{L}$ having 'units' of mass^4.

4.1.3 Equations of motion

The Lagrangian is the space-integral of the Lagrangian density $L = \int d^3x\mathcal{L}$, while the action is the time-integral of the Lagrangian $S = \int dtL$, so the action is the space-time integral $S = \int dt \int d^3x\mathcal{L}$.

The variation of the action for a variation $\phi(\vec{x}) \rightarrow \phi(\vec{x}) + \delta\phi(\vec{x})$ is

$$\delta S = \int d^4x \left(\frac{\partial\mathcal{L}}{\partial\phi}\delta\phi + \frac{\partial\mathcal{L}}{\partial\phi_{,\mu}}\delta\phi_{,\mu} \right) = \int d^4x\delta\phi \left(\frac{\partial\mathcal{L}}{\partial\phi} - \partial_\mu \frac{\partial\mathcal{L}}{\partial\phi_{,\mu}} \right) \tag{68}$$

where we have integrated by parts and assumed that the field vanishes at spatial infinity (or that the Universe is periodic inside some very large cubical box) and that $\delta\phi$ vanishes on the initial and final time hyper-surfaces.

Requiring that $\delta S = 0$ for arbitrary $\delta\phi$ gives, on performing the partial derivatives of $\mathcal{L}$, the equations of motion

$$\phi^{,\mu}{}_{,\mu} = m^2\phi \tag{69}$$

or equivalently $\Box\phi + m^2\phi = 0$.

4.1.4 The stress-energy tensor and energy-momentum conservation

Partially differentiating the Lagrangian density $\mathcal{L}(\vec{x}) = \mathcal{L}(\phi(\vec{x}), \phi_{,\alpha}(\vec{x}))$ (i.e. considering $\mathcal{L}$ to be a function of space and time for some solution of the field equations) with respect to x^ν gives

$$\frac{\partial\mathcal{L}(\vec{x})}{\partial x^\nu} = \frac{\partial\mathcal{L}}{\partial\phi}\phi_{,\nu} + \frac{\partial\mathcal{L}}{\partial\phi_{,\mu}}\phi_{,\mu\nu}. \tag{70}$$

Using the equations of motion to replace $\partial\mathcal{L}/\partial\phi$ by $\partial_\mu(\partial\mathcal{L}/\partial\phi_{,\mu})$ and noting that $\phi_{,\mu\nu} = \phi_{,\nu\mu}$ the right hand side becomes $\partial_\mu(\phi_{,\nu}(\partial\mathcal{L}/\partial\phi_{,\mu}))$ while writing the left hand side as $\partial\mathcal{L}(\vec{x})/\partial x^\nu = \delta^\mu_\nu \partial_\mu \mathcal{L}$ gives

$$\partial_\mu \left(\delta^\mu_\nu \mathcal{L} - \phi_{,\nu} \frac{\partial \mathcal{L}}{\partial \phi_{,\mu}} \right) = 0. \tag{71}$$

Invoking (67), from which $\partial\mathcal{L}/\partial\phi_{,\mu} = -\phi^{,\mu}$, and raising the index ν gives the continuity equation $T^{\mu\nu}{}_{,\mu} = 0$ where the *stress-energy tensor* is

$$T^{\mu\nu} = \phi^{,\mu}\phi^{,\nu} - \tfrac{1}{2}\eta^{\mu\nu}(\phi_{,\alpha}\phi^{,\alpha} + m^2\phi^2). \tag{72}$$

$T^{\mu\nu}{}_{,\mu} = 0$ is four equations (one for each of $\nu = 0, 1, 2, 3$) and they express the continuity of energy and the components of the 3-momentum. Integrating this equation over all space gives 4 globally conserved quantities $P^\nu = \int d^3x T^{t\nu}$: the total energy and the total 3-momentum.

4.1.5 Dispersion relation, and Fourier synthesis for the free field

The free-field equation of motion $\Box\phi + m^2\phi = 0$ admits plane wave solutions $\phi(\mathbf{x}, t) = \phi_0 \cos(\mathbf{k}\cdot\mathbf{x} - \omega_\mathbf{k} t + \varphi)$, specified by an amplitude ϕ_0 and the phase φ, and where the temporal and spatial frequencies are related by the *dispersion relation*

$$\omega_\mathbf{k} = \sqrt{m^2 + |\mathbf{k}|^2} \tag{73}$$

where we take the positive root in order that a wave with wave-vector $\mathbf{k}$ has crests (or nodes) that advance with time in the direction $+\hat{\mathbf{k}}$.

In 4-notation can write $\phi(\vec{x}) = \phi_0 \cos(k_\mu x^\mu + \varphi)$ where the 4-wave-vector $\vec{k} \to k_\mu = (-\omega_\mathbf{k}, \mathbf{k})$ lies on a hyperboloid in 4D $\vec{k}$-space.

The general solution of (69) can be synthesised as a sum of plane waves with complex coefficients

$$\phi(t, \mathbf{x}) = \tfrac{1}{2} \sum_\mathbf{k} (\phi_\mathbf{k} e^{+i(\mathbf{k}\cdot\mathbf{x} - \omega_\mathbf{k} t)} + \text{c.c.}) \tag{74}$$

where $\phi_\mathbf{k}$ encodes the amplitude and the phase, c.c. denotes complex conjugate, and where we assume that the field is periodic in a large box of side L so the possible spatial frequencies lie on a grid in $\mathbf{k}$-space with spacing $dk = 2\pi/L$.

Alternatively, we can write

$$\phi(\vec{x}) = \tfrac{1}{2} \sum_\mathbf{k} (\phi_\mathbf{k} e^{+ik_\mu x^\mu} + \phi^\star_\mathbf{k} e^{-ik_\mu x^\mu}) \tag{75}$$

where we note that the positive and negative 4-frequencies $\pm\vec{k}$ both correspond to waves moving in the direction $+\hat{\mathbf{k}}$.

The partial derivatives are

$$\phi_{,\mu}(t, \mathbf{x}) = \tfrac{1}{2} \sum_\mathbf{k} ik_\mu (\phi_\mathbf{k} e^{+ik_\mu x^\mu} - \phi^\star_\mathbf{k} e^{-ik_\mu x^\mu}) \tag{76}$$

and specifying ϕ and $\phi_{,t}$ at some initial time are sufficient to determine the complex amplitudes:

$$\phi_\mathbf{k} = \tilde{\phi}(\mathbf{k}) + \tilde{\phi}_{,t}(\mathbf{k})/i\omega_\mathbf{k} \tag{77}$$

where tilde denotes the Fourier transform.

4.1.6 Stress-energy tensor for a plane wave

foo-bar

4.1.7 Wave packets; phase and group velocities

The surfaces of constant phase $\psi = \vec{k} \cdot \vec{x}$ advance with time in the direction parallel to $\mathbf{k}$, with *phase velocity*

$$\mathbf{v}_{\text{phase}} = \omega_{\mathbf{k}} \hat{\mathbf{k}}/k. \tag{78}$$

One may synthesise a *nearly monochromatic wave packet* by summing a large number of waves with similar wave-vectors, the size of the packet being on the order of the inverse of the spread of the wave vectors. Consider a wave that, at $t = 0$, has $\phi_0(\mathbf{x}) = f_0(\mathbf{x}) e^{i\overline{\mathbf{k}} \cdot \mathbf{x}}$ where $f_0(\mathbf{x})$ is the initial *envelope function* and where it is implicitly understood that we are to take the real part. The Fourier transform of this is $\tilde{\phi}_0(\mathbf{k}) = \int d^3x \phi_0(\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}} = \int d^3x f_0(\mathbf{x}) e^{-i(\mathbf{k}-\overline{\mathbf{k}})\cdot\mathbf{x}} = \tilde{f}_0(\mathbf{k} - \overline{\mathbf{k}})$ which is the same as $\phi_{\mathbf{k}}$. If we assume that $\dot{\phi}_0(\mathbf{x})$ corresponds to a wave propagating in the direction $\hat{\mathbf{k}}$ then at a later time t the field is $\phi(\mathbf{x}, t) = \sum_{\mathbf{k}} \phi_{\mathbf{k}} e^{i(\mathbf{k}\cdot\mathbf{x} - \omega_{\mathbf{k}} t)} \rightarrow \int d^3k \tilde{f}_0(\mathbf{k} - \overline{\mathbf{k}}) e^{i(\mathbf{k}\cdot\mathbf{x} - \omega_{\mathbf{k}} t)} = \int d^3p \tilde{f}_0(\mathbf{p}) e^{i[(\overline{\mathbf{k}}+\mathbf{p})\cdot\mathbf{x} - \omega_{\overline{k}+\mathbf{p}} t]}$. If the envelope $f_0(\mathbf{x})$ is very broad (width $\Delta x \gg 1/\overline{k}$) then its transform $\tilde{f}_0(\mathbf{k})$ will be very narrow and we can Taylor expand $\omega_{\overline{k}+\mathbf{p}} = \omega_{\overline{\mathbf{k}}} + \mathbf{p} \cdot (d\omega_{\mathbf{k}}/d\mathbf{k})_{\overline{\mathbf{k}}} + \ldots$ and so $\phi(\mathbf{x}, t) = e^{i(\overline{\mathbf{k}}\cdot\mathbf{x} - \omega_{\overline{\mathbf{k}}} t)} \int d^3p \tilde{f}_0(\mathbf{p}) e^{i\mathbf{p}\cdot(\mathbf{x} - (d\omega_{\mathbf{k}}/d\mathbf{k})t)}$ or $\phi(\mathbf{x}, t) = f_t(\mathbf{x}) e^{i(\overline{\mathbf{k}}\cdot\mathbf{x} - \omega_{\overline{\mathbf{k}}} t)}$ which is a wave with envelope function $f_t(\mathbf{x}) = \int d^3p \tilde{f}_0(\mathbf{p}) e^{i\mathbf{p}\cdot(\mathbf{x} - (d\omega_{\mathbf{k}}/d\mathbf{k})t)} = f_0(\mathbf{x} - (d\omega_{\mathbf{k}}/d\mathbf{k})t)$.

So such a packet evidently moves with *group velocity*

$$\mathbf{v}_{\text{group}} = d\omega_{\mathbf{k}}/d\mathbf{k} \tag{79}$$

which, for the free-field dispersion relation, is $\mathbf{v}_{\text{group}} = \mathbf{k}/\omega_{\mathbf{k}}$.

Putting in the speed of light, and expressing the frequency as $m = Mc^2/\hbar$ (i.e. the Compton frequency for a particle of mass M), we have $\omega_k = \sqrt{(Mc^2)^2/\hbar^2 + c^2k^2}$ and therefore $v_{\text{group}} = c^2k/\omega_k = (\hbar k/M)/\sqrt{1 + (\hbar k/Mc)^2}$ which is sub-luminal (tending to c for $k \to \infty$) and is the same as the speed of a particle of momentum $\hbar k$. The phase-speed, on the other hand, is always super-luminal and is not of any great physical significance.

Taking the Taylor expansion of $\omega_{\overline{\mathbf{k}}+\mathbf{p}}$ to next order reveals that the packet will spread with time. This is because, for a wave packet of finite extend $L \sim N\lambda$, the width of the transform of the envelope function is on the order of $\Delta k \sim 1/L$ and the different modes composing the packet propagate in slightly different directions and at slightly different speeds so as time goes on the packet must broaden. We can estimate the rate for this as follows: If we have a packet of size $\sim N$ wavelengths, so the spread in wave-vectors is $\Delta k \sim \overline{k}/N$, corresponding to a range of angles $\Delta k/\overline{k} \sim 1/N$. We can define a 'spreading time' t_{spread} such that $v_{\text{group}} t_{\text{spread}}$ times the angular range be on the order of the size of the wave-packet $\sim N\lambda$. Thus in a time t_{spread}, the change in size of the packet will be comparable to its inition size. The crossing time t_{cross}, on the other hand, is the packet size divided by the group velocity. It follows that $t_{\text{spread}} \sim N t_{\text{cross}}$, so the spreading of the packet becomes appreciable only after about N crossing times.

4.2 The complex scalar field as electrically charged matter

As discussed in the introduction two independent real scalar fields of identical frequency m possess a conserved 4-current density $j^\mu = a\partial^\mu b - b\partial^\mu a$.

This was obtained directly from the equations of motion. It can also be thought of as arising from invariance of the Lagrangian density $\mathcal{L} = \mathcal{L}_a + \mathcal{L}_b$ which, expressed in terms of $\phi \equiv a + ib$, is

$$\mathcal{L} = -\tfrac{1}{2}(\partial_\mu \phi \partial^\mu \phi^\star + m^2 \phi \phi^\star) \tag{80}$$

under a global phase-shift $\phi \to \phi' = \phi e^{i\theta}$.

The 4-current density is given in terms of ϕ by $j^\mu = (i/2)(\phi \partial^\mu \phi^\star - \phi^\star \partial^\mu \phi)$, whose spatial components are identical in form to the Schrödinger probability flux density.

4.2.1 Coupling to an EM potential: the "gauge covariant derivative"

The S-equation (relativistic or non-relativistic) can be coupled to an EM potential by replacing ∂_μ by *gauge-covariant derivatives* $\partial_\mu - (q/\hbar)A_\mu$ and the same is true for the classical Klein-Gordon equation for a complex scalar field. This is not surprising since the Klein-Gordon equation is identical to the relativistic version of the Schrödinger equation (Developed before the non-relativistic version by Weyl and Klein as well by Schrödinger, but then dropped as a candidate for an equation for a single-particle wave function on the grounds that it did not have a conserved total probability.) To simplify the following, we will define $q' = q/\hbar$

and then drop the prime (so in what follows q denotes physical charge divided by $\hbar$). So the coupling to A_μ is induced by replacing partial derivatives by

$$\partial_\mu \to D_\mu \equiv \partial_\mu - iqA_\mu \tag{81}$$

with complex conjugate $D^\star_\mu \equiv \partial_\mu + iqA_\mu$.

4.2.2 The Lagrangian and EoMs

The Lagrangian density for the matter plus the electromagnetic field then becomes

$$\mathcal{L} = \mathcal{L}_\mathrm{m} + \mathcal{L}_\mathrm{r} = -\tfrac{1}{2}(D_\mu \phi D^{\star\mu} \phi^\star + m^2 \phi\phi^\star) - \tfrac{1}{4} F^{\mu\nu} F_{\mu\nu} \tag{82}$$

where $\mathcal{L}_\mathrm{r} = -F^{\mu\nu}F_{\mu\nu}/4$ as before, and is entirely independent of the field ϕ, whereas $\mathcal{L}_\mathrm{m}$ involves both ϕ and A_μ. This is gauge invariant under $A_\mu \to A_\mu + \xi_{,\mu}$ and $\phi \to e^{iq\xi}\phi$.

The equations of motion for ϕ are the Klein-Gordon equation

$$D_\mu D^\mu \phi = m^2 \phi \quad \text{and} \quad D^\star_\mu D^{\star\mu} \phi^\star = m^2 \phi^\star. \tag{83}$$

These are also gauge invariant in the sense that if $\phi(\vec{x})$ is a solution for some potential A_μ then $\phi' = e^{iq\xi}\phi$ is a solution for a different potential $A'_\mu = A_\mu + \xi_{,\mu}$.

The equations of motion for the EM potential A_μ are

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial A_{\nu,\mu}} = \frac{\partial \mathcal{L}}{\partial A_\nu} \quad \text{or} \quad F^{\mu\nu}{}_{,\mu} = j^\nu \tag{84}$$

which are the inhomogeneous Maxwell equations with source term as before, but where now the current is

$$j^\mu \equiv \frac{iq}{2}(\phi D^{\star\mu}\phi^\star - \phi^\star D^\mu \phi). \tag{85}$$

4.2.3 The conserved 4-current

– Charge conservation from the equation of motion

The divergence of (85) is

$$\begin{aligned} j^\mu{}_{,\mu} &= \partial_\mu \frac{iq}{2}(\phi D^{\star\mu}\phi^\star - \text{c.c.}) = \frac{iq}{2}((\partial_\mu \phi) D^{\star\mu}\phi^\star + \phi \partial_\mu D^{\star\mu}\phi^\star - \text{c.c.}) \\ &= \frac{iq}{2}((D_\mu \phi + iqA_\mu \phi) D^{\star\mu}\phi^\star - \phi(D^\star_\mu + iqA_\mu) D^{\star\mu}\phi^\star - \text{c.c.}) \\ &= \frac{iq}{2}(D_\mu \phi D^{\star\mu}\phi^\star + \phi D^\star_\mu D^{\star\mu}\phi^\star - \text{c.c.}) = \frac{iq}{2}(D_\mu \phi D^{\star\mu}\phi^\star + m^2 \phi\phi^\star - \text{c.c.}) \end{aligned} \tag{86}$$

where, in the last step, we have used the Klein-Gordon equation. But both of the terms in the last expression are real, so

$$j^\mu{}_{,\mu} = 0 \tag{87}$$

and the gauge invariant current (85) is exactly conserved.

– Charge conservation via Noether's theorem

The charge conservation law (87) was obtained (fairly) straightforwardly by applying the field equations, but one can also obtain this from the invariance of the Lagrangian density (82) under the gauge transformation $A_\mu \to A'_\mu = A_\mu + \xi_{,\mu}$ and $\phi \to \phi' = \phi e^{iq\xi}$. Let's say we have some solution $\phi(\vec{x})$ and $A_\mu(\vec{x})$ of the field equations. From these we can evaluate $\phi_{,\nu}(\vec{x})$ and $F_{\mu\nu}$ and use these in (82) to evaluate $\mathcal{L}(\vec{x})$. If instead we substitute $\phi e^{iq\xi}$ for ϕ and $A_\mu + \xi_{,\mu}$ for A_μ, the resulting $\mathcal{L}(\vec{x})$ will be no different. We can say that the differential $d\mathcal{L}(\vec{x})$ under some small change $\xi(\vec{x}) \to \xi(\vec{x}) + d\xi(\vec{x})$ (with associated change in $\xi_{,\mu}$) must vanish.

Applying the chain rule we have

$$d\mathcal{L}(\vec{x}) = 0 = \frac{\partial \mathcal{L}}{\partial \phi} d\phi + \frac{\partial \mathcal{L}}{\partial \phi_{,\mu}} d\phi_{,\mu} + \frac{\partial \mathcal{L}}{\partial \phi^\star} d\phi^\star + \frac{\partial \mathcal{L}}{\partial \phi^\star_{,\mu}} d\phi^\star_{,\mu} + \frac{\partial \mathcal{L}}{\partial A_\mu} dA_\mu + \frac{\partial \mathcal{L}}{\partial A_{\mu,\nu}} dA_{\mu,\nu}. \tag{88}$$

The first thing to note here is that $\partial\mathcal{L}/\partial A_{\mu,\nu} = -F^{\mu\nu}$, which is anti-symmetric, while $dA_{\mu,\nu} = d\xi_{,\mu\nu}$ is, by virtue of the commutation properties of partial derivatives, symmetric, so the last term vanishes.

Next, using the Euler-Lagrange equations to replace $\partial\mathcal{L}/\partial\phi$ by $\partial_\mu(\partial\mathcal{L}/\partial\phi_{,\mu})$ and noting that $d\phi_{,\mu} = \partial_\mu d\phi$ and similarly for the terms involving $\phi^\star$ we can combine the first four terms into a single derivative and we have

$$d\mathcal{L}(\vec{x}) = \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial\phi_{,\mu}}d\phi + \frac{\partial\mathcal{L}}{\partial\phi^\star_{,\mu}}d\phi^\star\right) + \frac{\partial\mathcal{L}}{\partial A_\mu}dA_\mu = 0. \tag{89}$$

Writing $\phi = \phi_0 e^{iq\xi}$ (with ϕ_0 the potential before we make the transformation) we have $d\phi = iq\phi d\xi$ and $d\phi^\star = -iq\phi^\star d\xi$, while, from (82) $\partial\mathcal{L}/\partial\phi_{,\mu} = -D^{\star\mu}\phi^\star/2$ and $\partial\mathcal{L}/\partial\phi^\star_{,\mu} = -D^\mu\phi/2$, so the term in parentheses above is

$$\frac{\partial\mathcal{L}}{\partial\phi_{,\mu}}d\phi + \frac{\partial\mathcal{L}}{\partial\phi^\star_{,\mu}}d\phi^\star = -d\xi\frac{iq}{2}(\phi D^{\star\mu}\phi^\star - \phi^\star D^\mu\phi) = -j^\mu d\xi \tag{90}$$

while, in the last term of (89), $\partial\mathcal{L}/\partial A_\mu = j^\mu$ and $dA_\mu = d\xi_{,\mu}$, so we have, finally,

$$d\mathcal{L}(\vec{x}) = -\partial_\mu(j^\mu d\xi) + j^\mu d\xi_{,\mu} = -j^\mu{}_{,\nu}d\xi = 0 \tag{91}$$

and as this is true for arbitrary $d\xi$ we must have $j^\mu{}_{,\mu} = 0$.

4.2.4 The stress-energy tensor for the matter

Considering only the ϕ-dependent terms in the Lagrangian:

$$\mathcal{L}_\mathrm{m}(\phi, \phi_{,\nu}, \vec{x}) = -\tfrac{1}{2}(D_\mu\phi D^{\star\mu}\phi^\star + m^2\phi\phi^\star) \tag{92}$$

with the explicit $\vec{x}$ dependence coming from the field $A_\mu(\vec{x})$ appearing in the covariant derivatives (considered here as an external influence) and taking the derivative of $\mathcal{L}_\mathrm{m}(\vec{x}) = \mathcal{L}_\mathrm{m}(\phi(\vec{x}), \phi_{,\nu}(\vec{x}), \vec{x})$ with respect to x_ν gives the continuity equation

$$T^\nu_{\mathrm{cm}\mu,\nu} = \partial_\mu\mathcal{L}_\mathrm{m}(\phi, \phi_{,\nu}, \vec{x}) = j^\mu A_{\mu,\nu} \tag{93}$$

where the canonial matter stress is

$$T^\nu_{\mathrm{cm}\mu} \equiv -\phi_{,\mu}\frac{\partial\mathcal{L}_\mathrm{m}}{\partial\phi_{,\nu}} - \phi^\star_{,\mu}\frac{\partial\mathcal{L}_\mathrm{m}}{\partial\phi^\star_{,\nu}} + \delta^\nu_\mu\mathcal{L}_\mathrm{m} \tag{94}$$

or, performing the derivatives,

$$T^\nu_{\mathrm{cm}\mu} = \tfrac{1}{2}(D_\mu\phi D^{\star\nu}\phi^\star + D^\star_\mu\phi^\star D^\nu\phi) + \delta^\nu_\mu\mathcal{L}_\mathrm{m} + A_\mu j^\nu \tag{95}$$

where the last term is non-symmetric and gauge dependent, but if we define the symmetric tensor

$$\begin{aligned} T^\nu_{\mathrm{m}\mu} = T^\nu_{\mathrm{cm}\mu} - A_\mu j^\nu &= \tfrac{1}{2}(D_\mu\phi D^{\star\nu}\phi^\star + D^\star_\mu\phi^\star D^\nu\phi) + \delta^\nu_\mu\mathcal{L}_\mathrm{m} \\ &= \tfrac{1}{2}(D_\mu\phi D^{\star\nu}\phi^\star + D^\star_\mu\phi^\star D^\nu\phi + \delta^\nu_\mu(D_\alpha\phi D^{\star\alpha}\phi^\star + m^2\phi\phi^\star)) \end{aligned} \tag{96}$$

we find, on raising the index μ, that this has continuity equation

$$T^{\nu\mu}_{\mathrm{m}\ ,\nu} = j^\nu F_\nu{}^\mu \tag{97}$$

just as for particles.

4.2.5 The total stress energy tensor

Taking the partial derivative of the total Lagrangian density $\mathcal{L} = \mathcal{L}_\mathrm{m} + \mathcal{L}_\mathrm{r}$ in (82) – but considered as a function of $\vec{x}$ – we find the energy-momentum continuity equation $T^\nu{}_{\mu,\nu} = 0$ where the canonical 4-stress is

$$T^\nu{}_\mu = \left(-\phi_{,\mu}\frac{\partial\mathcal{L}_\mathrm{m}}{\partial\phi_{,\nu}} - \phi^\star_{,\mu}\frac{\partial\mathcal{L}_\mathrm{m}}{\partial\phi^\star_{,\nu}} + \delta^\nu_\mu\mathcal{L}_\mathrm{m}\right) + \left(-A_{\gamma,\mu}\frac{\partial\mathcal{L}_\mathrm{r}}{\partial A_{\gamma,\nu}} + \delta^\nu_\mu\mathcal{L}_\mathrm{r}\right) \tag{98}$$

The second collection of terms is the canonical stress-tensor for the radiation, which is given, in terms of the symmetric radiation stress-tensor by $T_\mathrm{r}{}^\nu{}_\mu + A_{\mu,\gamma}F^{\gamma\nu}$. The first collection of terms is the canonical matter stress-tensor, which, in terms of the symmetric mass stress-tensor, is $T^\nu_{\mathrm{m}\mu} + A_\mu j^\nu = T^\nu_{\mathrm{m}\mu} + A_\mu F^{\gamma\nu}{}_{,\gamma}$. Thus the combination is

$$T^\nu{}_\mu = T^\nu_{\mathrm{m}\mu} + T_\mathrm{r}{}^\nu{}_\mu + \partial_\gamma(A_\mu F^{\gamma\nu}) \tag{99}$$

so there is an extra, generally asymmetric and gauge dependent term in the total stress tensor. This does not affect the law of conservation of 4-momentum since, when we differentiate this with respect to x^ν we get $\partial_\nu\partial_\gamma(A_\mu F^{\gamma\nu})$ which is a symmetric differential operator applied to a tensor that is anti-symmetric under $\gamma \leftrightarrow \nu$.

4.2.6 The matter stress-energy from Hilbert's action

An alternative....

4.3 Charged wave-packets and beams

We now explore the behaviour of wave-packets and beams of complex scalar field matter. The main results are as follows:

- As with a real scalar field, we will see that we can write the general solution of the complex Klein-Gordon equation coupled to a locally constant EM field as the sum over **p**-modes carrying momentum **p** with positive and negative frequencies and with complex amplitudes $\phi^+_{\mathbf{p}}$ and $\phi^-_{\mathbf{p}}$. But a crucial difference, as we shall see, is that while for a real field the amplitudes for the positive and negative frequency modes $\pm\vec{k}$, with $k^\mu = (\omega_{\mathbf{k}}, \mathbf{k}))$ are just complex conjugates of one another, here they describe physically independent waves: Both of these modes transport 4-momentum in the direction $\hat{\mathbf{p}}$ but their currents have opposite sign. Positive frequency waves $\phi \propto e^{+i\vec{k}\cdot\vec{x}}$ carry charge in the same direction as the wave-momentum while negative frequency waves $\phi \propto e^{-i\vec{k}\cdot\vec{x}}$ carry negative charge in that direction.

- In the absence of an EM field, the dispersion relation is as for a real scalar, and the positive and negative frequency waves have identical phase and group velocities. If $\phi^-_{\mathbf{k}} = 0$ we have a positively charged wave for which the real and imaginary parts of the complex field $\phi = a + ib$ rotate clockwise in the (Argand) a, b plane. If $\phi^+_{\mathbf{k}} = 0$ we have a negatively charge wave which rotates in the opposite sense. These are just like circularly polarised EM waves. On the other hand, if $|\phi^+_{\mathbf{k}}| = |\phi^-_{\mathbf{k}}|$ we have the superposition of two waves with the same amplitude but rotating in opposite directions and the resultant field oscillates along a line in the a, b plane; just like a linearly polarised EM wave. Such waves are electrically neutral. In general there are 4 number needed to specify a wave with wave vector **k**: these can be taken to be the amplitude and phase of the positive frequency mode and the amplitude of the negative frequency mode along with the relative phase (which determines the direction of the linearly resultant polarised component). This is just as for EM beams and one can employ the machinery of Stokes's parameters to describe this.

- If we switch on a slowly varying EM field, this modifies the dispersion relation, and the +ve and -ve frequency modes have different phase and group velocities. If we fire a beam with a certain temporal frequency from a free-field region into one with non-zero $\vec{A}$ it is like firing a EM beam into a dispersive plasma. The positive and negative frequency modes have different wave-numbers and so the relative phase changes and one has the phenomenology of Faraday rotation. There is also a change, in general, in the direction of propagation of the two different modes, analogous to what happens in 'birefringent' crystals.

- Another interesting aspect of matter beams and wave-packets is that how phase θ advances as measured by an observer who moves with the packet or beam. The result is that $d\theta/dt = -m/\gamma + q\varphi - q\mathbf{A}\cdot\dot{\mathbf{x}}$. This is the same as the Lagrangian for a charged particle. This allows one to show how the effect of a potential gradient is to cause the phase to advance with position along a beam or in a wave-packet in a way that depends on position, and this in turn is a nice way to show how wave-packets of a complex scalar field coupled to EM get accelerated and deflected.

- Can effects like the Faraday rotation of a neutral beam, or the phase of a charged beam, actually be measured? If the only way one can interact with such waves is via electromagnetism it would seem that it would be difficult - AB effect - ,,,,

4.3.1 The dispersion relation and properties of plane waves

Consider a region of space-time with constant EM potential A^μ. In this region the Faraday tensor and therefore the electric and magnetic field vanish. Applying the operator $D_\mu D^\mu$ to a travelling wave $\phi = \phi_{\mathbf{k}} e^{ik_\mu x^\mu} = \phi_{\mathbf{k}} e^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)}$, where ϕ_k is a complex amplitude, gives $D_\mu D^\mu \phi = (ik_\mu - iqA_\mu)(ik^\mu - iqA^\mu)\phi = ((\omega - q\varphi)^2 - |\mathbf{k} - q\mathbf{A}|^2)\phi$, so such a wave solves the Klein-Gordon equation $D_\mu D^\mu \phi = m^2\phi$ (83) provided the dispersion relation

$$(\omega - q\varphi)^2 = |\mathbf{k} - q\mathbf{A}|^2 + m^2 \tag{100}$$

is satisfied.

The group velocity (the velocity of a packet composed of waves with wave vectors close to $k_\mu = (-\omega, \mathbf{k})$) is

$$\mathbf{v} = d\omega/d\mathbf{k} = (\mathbf{k} - q\mathbf{A})/(\omega - q\varphi). \tag{101}$$

The relativistic γ-factor for this velocity is

$$\gamma = 1/\sqrt{1 - |\mathbf{v}|^2} = |\omega - q\varphi|/m \tag{102}$$

so the 4-momentum for a particle of mass m with velocity $\mathbf{v}$ would be

$$p^\mu = (E, \mathbf{p}) = (\gamma m, \gamma m \mathbf{v}) \tag{103}$$

where

$$\mathbf{p} = |\mathbf{k} - q\mathbf{A}|\hat{\mathbf{v}}. \tag{104}$$

There are two distinct wave-vectors $\vec{k}$ corresponding to a given group velocity $\mathbf{v}$ (or momentum $\mathbf{p}$ - defined above to have the same direction as $\mathbf{v}$): one can have $\mathbf{k} - q\mathbf{A} = +|\omega - q\varphi|\mathbf{v} = +\mathbf{p}$ and $\omega - q\varphi = +\sqrt{|\mathbf{k} - q\mathbf{A}|^2 + m^2} = +E$ or one can change the sign of *both* $\mathbf{k} - q\mathbf{A}$ and $\omega - q\varphi$. These choices correspond to 4-wave-vectors $\vec{k}_\pm = \pm\vec{p} + q\vec{A}$.

A complex scalar wave with momentum $\mathbf{p}$ is a superposition of waves with these two possible wave-vectors:

$$\phi = \phi^+_{\mathbf{p}} e^{i\vec{k}_+\cdot\vec{x}} + \phi^-_{\mathbf{p}} e^{i\vec{k}_-\cdot\vec{x}} = e^{iq\vec{A}\cdot\vec{x}}(\phi^+_{\mathbf{p}} e^{+i\vec{p}\cdot\vec{x}} + \phi^-_{\mathbf{p}} e^{-i\vec{p}\cdot\vec{x}}). \tag{105}$$

with complex amplitudes $\phi^\pm_{\mathbf{p}}$.

This is similar to the situation for a real scalar field. But there the amplitudes for the positive and negative frequency components were complex conjugates of one another. That gave 2 numbers as needed to specify the amplitude and phase for a real wave (or equivalently to specify the field and its rate of change at some initial time). Here we have a complex wave $\phi = a + ib$ so we need 4 numbers to specify this: an amplitude and phase for a and an amplitude and phase for b (perhaps relative to that of a). Or, alternatively, we need 4 degrees of freedom to specify the two components of the field and their rates of change at the initial time. So the two amplitudes $\phi^\pm_{\mathbf{p}}$ for a complex field must be independent of one another.

The current for plane waves: Applying D^μ to ϕ we get

$$D^\mu\phi = (\partial^\mu - iqA^\mu)\phi = ip^\mu e^{iq\vec{A}\cdot\vec{x}}(\phi^+_{\mathbf{p}} e^{+i\vec{p}\cdot\vec{x}} - \phi^-_{\mathbf{p}} e^{-i\vec{p}\cdot\vec{x}}). \tag{106}$$

from which we find that the current is

$$j^\mu = \frac{iq}{2}(\phi D^{\star\mu}\phi^\star - \phi^\star D^\mu\phi) = qp^\mu(|\phi^+_{\mathbf{p}}|^2 - |\phi^-_{\mathbf{p}}|^2) \tag{107}$$

so a wave $\phi \propto e^{i\vec{k}_+\cdot\vec{x}}$ has a positive charge density j^0 (q being a positive parameter here) and a current $\mathbf{j}$ that has the same direction as $\mathbf{p}$ while a wave $\phi \propto e^{i\vec{k}_-\cdot\vec{x}}$ has negative charge which it transports in the same direction (so the current is opposite to $\mathbf{p}$).

Polarisation states for plane waves: A fully positively (negatively) charged wave is one for which $\phi^-_{\mathbf{p}}$ ($\phi^+_{\mathbf{p}}$) vanishes. The field $\phi = a + ib$ at some a fixed position rotates in the Argand (a, b) plane. These are like circularly polarised EM waves (though the field here lives in an abstract space whereas the E_x and E_y components of a planar EM wave live in real space). The rates of rotation for the two charged waves are different (if $A^t \neq 0$). If we set $|\phi^-_{\mathbf{p}}| = |\phi^+_{\mathbf{p}}|$ we get an electrically neutral wave; the sum of two charged waves of equal intensity. If $A^t = 0$ the two waves rotate with equal but opposite frequency and the resultant a and b oscillate in phase with one another along a straight line through the origin of the Argand plane (with the direction determined by the relative phase of the two charged components). This is like a linearly polarised EM wave. But if $A^t \neq 0$, the plane of polarisation rotates.

This reminds one of Faraday rotation of linear polarised radiation propagating through a magnetised plasma where the rate of change of polarisation angle is proportional to electron density times the line-of-sight component of the $\mathbf{B}$-field. But here the polarisation angle is changing at a rate set by the *potential*, not the field, which is different.

In electromagnetism, for a beam along the z-axis, and if we write $E_x = \mathrm{Re}(Xe^{i\omega t})$ (and similarly for E_y), the Stokes parameters are usually defined as

$$\begin{aligned} I &= \langle E_x^2 + E_y^2 \rangle = \tfrac{1}{2}\langle XX^\star + YY^\star \rangle \\ Q &= \langle E_x^2 - E_y^2 \rangle = \tfrac{1}{2}\langle XX^\star - YY^\star \rangle \\ U &= \langle 2E_x E_y \rangle = \tfrac{1}{2}\langle XY^\star + YX^\star \rangle \\ V &= \langle E_x \dot{E}_y - E_y \dot{E}_x \rangle/\omega = \tfrac{i}{2}\langle XY^\star - YX^\star \rangle \end{aligned} \tag{108}$$

Here, for $\mathbf{k} = 0$, and writing the field as $\phi = (a + ib) = \phi_+ e^{+i\omega t} + \phi_- e^{-i\omega t}$, the analogous parameters are

$$\begin{aligned} I &= \langle a^2 + b^2 \rangle = \tfrac{1}{2}\langle \phi_+\phi_+^\star + \phi_-\phi_-^\star \rangle \\ Q &= \langle a^2 - b^2 \rangle = \tfrac{1}{2}\langle \phi_+\phi_- + \phi_+^\star\phi_-^\star \rangle \\ U &= \langle 2ab \rangle = \tfrac{1}{2i}\langle \phi_+\phi_- - \phi_+^\star\phi_-^\star \rangle \\ V &= \langle a\dot{b} - b\dot{a} \rangle/\omega = \tfrac{1}{2}\langle \phi_+\phi_+^\star - \phi_-\phi_-^\star \rangle \end{aligned} \tag{109}$$

where we see that V is proportional to the electric charge density.

Stress tensor for plane waves: To form the matter stress tensor, we need

$$\phi\phi^\star = |\phi_\mathbf{p}^+|^2 + |\phi_\mathbf{p}^-|^2 + (\psi_\mathbf{p}^+\psi_\mathbf{p}^{-\star} e^{2i\vec{p}\cdot\vec{x}} + \text{c.c.}) \tag{110}$$

and

$$D^\mu\phi D^{\star\nu}\phi^\star = p^\mu p^\nu (|\phi_\mathbf{p}^+|^2 + |\phi_\mathbf{p}^-|^2 - (\psi_\mathbf{p}^+\psi_\mathbf{p}^{-\star} e^{2i\vec{p}\cdot\vec{x}} + \text{c.c.})) \tag{111}$$

so

$$\mathcal{L}_\mathrm{m} = -\tfrac{1}{2}(D_\mu\phi D^{\star\mu}\phi^\star + m^2\phi\phi^\star) = m^2(\psi_\mathbf{p}^+\psi_\mathbf{p}^{-\star} e^{2i\vec{p}\cdot\vec{x}} + \text{c.c.}) \tag{112}$$

where we have used the dispersion relation, which in terms of $\vec{p}$ is $p^\mu p_\mu = -m^2$. The Lagrangian density is an oscillating function of space and time and its average vanishes. The symmetric matter stress tensor is

$$\begin{aligned} T_\mathrm{m}^{\mu\nu} &= \tfrac{1}{2}(D^\mu\phi D^{\star\nu}\phi^\star + D^{\star\mu}\phi^\star D^\nu\phi) + \eta^{\mu\nu}\mathcal{L}_\mathrm{m} \\ &= p^\mu p^\nu (|\phi_\mathbf{p}^+|^2 + |\phi_\mathbf{p}^-|^2) - (p^\mu p^\nu - m^2\eta^{\mu\nu})(\psi_\mathbf{p}^+\psi_\mathbf{p}^{-\star} e^{2i\vec{p}\cdot\vec{x}} + \text{c.c.}) \end{aligned} \tag{113}$$

Each of these expressions contain an oscillating factor involving $\psi_\mathbf{p}^+\psi_\mathbf{p}^{-\star} e^{2i\vec{p}\cdot\vec{x}} + \text{c.c.}$. Note, however, that this vanishes for a purely positively (negatively) charged wave for which $\phi_\mathbf{p}^-$ ($\phi_\mathbf{p}^+$) vanishes. It arises from interference between the positively and negatively charged waves. The oscillating term in $T_\mathrm{m}^{\mu\nu}$ can, in principle, be observed as it could generate gravitational waves or cause time varying gravitational lensing, or, most promising, pulse arrival time modulation for pulsars that can be observed by pulsar timing arrays. The steady flux density of momentum is $T_\mathrm{m}^{\mu\nu} = p^\mu p^\nu (|\phi_\mathbf{p}^+|^2 + |\phi_\mathbf{p}^-|^2)$. Note that there is no analogous oscillating term in the current; consequently there is no EM radiation from these planar waves.

If we have the superposition of many plane waves there will also be interference effects, but if the waves are incoherent, then the summed steady stress tensor is $\sum_\mathbf{p} p^\mu p^\mu \langle |\phi_\mathbf{p}^+|^2 + |\phi_\mathbf{p}^-|^2 \rangle$ or, converting the sum to an integral, $T_\mathrm{m}^{\mu\nu} \sim \int (d^3p/\omega_\mathbf{p}) p^\mu p^\nu f(\mathbf{p})$ where the effective phase-space density is $f(\mathbf{p}) = \omega_\mathbf{p}\langle |\phi_\mathbf{p}^+|^2 + |\phi_\mathbf{p}^-|^2 \rangle$ or, equivalently the power-spectrum $\omega_\mathbf{p}^2 \langle |\phi_\mathbf{p}^+|^2 + |\phi_\mathbf{p}^-|^2 \rangle$ divided by frequency.

4.3.2 Wave-packet electrodynamics

We saw in (112) that the matter Lagrangian density $\mathcal{L}_\mathrm{m}$ for a planar complex scalar wave is an oscillating function of $\vec{x}$, so the positive and negative half-waves tend to cancel, and actually vanishes for a fully charged (i.e. circularly polarised) wave. The same is true within a large, nearly monochromatic, wave-packet. This may seem at odds with the fact that the Lagrangian for a charged particle is $L = -m/\gamma + q\dot{x}^\mu A_\mu$. One might perhaps have expected the Lagrangian $\int d^3x \mathcal{L}_\mathrm{m}$ for a wave packet as a function of the position and velocity of the packet to reproduce the particle result, but evidently that is not the case.

Nonetheless, we expect that the dynamics of charged classical scalar wave-packets should reproduce classical particle electrodynamics. One way to develop this is to use the continuity equation for $T_\mathrm{m}^{\mu\nu}$ which, when integrated over space, says that $dP^\nu/dt = d/dt \int d^3x T_\mathrm{m}^{t\nu} = F^{\mu\nu} J_\mu$.

There is another, somewhat illuminating, approach, that allows us to recover the results for classical particles directly from the dispersion relation (100), which itself comes directly from the Klein-Gordon equation (83), thus side-stepping much of the heavy lifting involved with the stress-tensor. This analysis also shows how the particle Lagrangian emerges from classical wave-mechanics.

Consider first a wave packet $\phi(\mathbf{x}, t) = f(\mathbf{x}, t)e^{i\vec{k}\cdot\vec{x}}$ with envelope function $f(\mathbf{x}, t)$. This, as we have seen, moves with velocity $\mathbf{v} = d\omega/d\mathbf{k} = (\mathbf{k} - q\mathbf{A})/(\omega - q\varphi)$, implying a relativistic $\gamma = (\omega - q\varphi)/m$ and momentum $\mathbf{p} = \mathbf{k} - q\mathbf{A} = \gamma m\mathbf{v}$ satisfying the usual energy momentum relation $(\gamma m)^2 = |\mathbf{p}|^2 + m^2$.

Now consider an observer who moves along with the wave packet. That observer sees a time-varying field $\phi \propto e^{i\theta(t)}$ where the phase $\theta(t) = k_\mu x^\mu(t) = (\mathbf{k}\cdot\mathbf{v} - \omega)t$ or, using $\mathbf{k} = \mathbf{p} + q\mathbf{A} = \gamma m\mathbf{v} + q\mathbf{A}$, and $\omega = \gamma m + q\varphi$, $\theta(t) = ((\gamma m\mathbf{v} + q\mathbf{A})\cdot\mathbf{v} - (\gamma m + q\varphi))t$ or, since $1 - \mathbf{v}\cdot\mathbf{v} = 1/\gamma^2$, the rate of change of θ is

$$d\theta/dt = -m/\gamma - q\varphi + q\mathbf{v}\cdot\mathbf{A} \tag{114}$$

which is identical to the particle Lagrangian. So the latter is not equivalent to the integrated Lagrangian density for the packet – which vanishes – it is simply the rate of change of the *phase* of the classical wave-packet.

This should not be surprising. The motivation for Dirac and Feynman's picture of quantum mechanics, in which the amplitude for a path $x(t)$ is $\psi \propto e^{iS/\hbar}$ where $S[x(t)]$ is the classical action, is that they realised that this would explain classical mechanics as the 'geometric optics' limit of wave mechanics, the classical paths being those for which S is stationary and for which neighbouring paths interfere constructively.

The dispersion relation describes plane wave solutions when the $\mathbf{E}$ and $\mathbf{B}$ fields vanish. To see how wave packets are accelerated and deflected by such fields, consider a small fleet of packets; a central packet with velocity and momentum parallel to the z-axis for simplicity, and a set of identical neighbouring copies with identical velocities (and therefore identical momenta $\mathbf{p}$) that are displaced by distances δx_i from the central packet. The purpose of these copies is to sample the way that the phase advances as a function of position.

Let the central packet move for a time t; its phase will change by $\theta = -\int dt(m/\gamma - q\dot{x}^\mu A_\mu(\vec{x}))$ whereas that for another packet will be $\theta = -\int dt(m/\gamma - q\dot{x}^\mu A_\mu(\vec{x} + \vec{\delta x}))$ so there will be a difference in the change of the phase given, to first order in $\vec{\delta x}$ by $\delta\theta(\delta x_i) = q\dot{x}^\mu \delta x_i \int dt A_{\mu,i}$ or, assuming a constant potential gradient, $\delta\theta(\delta x_i) = q\dot{x}^\mu t A_{\mu,i}\delta x_i$. So the potential gradient $A_{\mu,i}$ causes the phase to advance at different rates for the different packets, depending on their position. The same is true within a single wave-packet. For the central packet the spatial wave-vector is initially $\mathbf{k} = \mathbf{p}(t_0) + q\mathbf{A}(t_0)$. This means the phase is, initially, as a function of position $\theta = \text{const} + k_i\delta x_i + \ldots = \text{const} + (p_i(t_0) + qA_i(t_0))\delta x_i$, where we note that we do not need to take account of the fact that A_i is varying across the packet as we are working only at first order in δx_i. After an interval of time t, the phase will vary with position as $\theta = \text{const} + (p_i(t_0) + qA_i(t_0))\delta x_i + q\dot{x}^\mu t A_{\mu,i}\delta x_i$.

This means that the i^{th} component of the wave vector at time $t_0 + t$ is

$$k_i(t_0 + t) = \partial\theta/\partial x_i = p_i(t_0) + qA_i(t_0) + q\dot{x}^\mu t A_{\mu,i} \tag{115}$$

and therefore the momentum at that time is

$$p_i(t_0 + t) = k_i(t_0 + t) - qA_i(t_0 + t) = p_i(t_0) + qA_i(t_0) - qA_i(t_0 + t) + q\dot{x}^\mu t A_{\mu,i}. \tag{116}$$

Only the last two terms are t-dependent, so the rate of change of the momentum is

$$dp_i/dt = -q\dot{A}_i + q\dot{x}^\mu A_{\mu,i} = -q\dot{x}^\mu A_{i,\mu} + q\dot{x}^\mu A_{\mu,i} = q\dot{x}^\mu F_{\mu i} \tag{117}$$

which is the Lorentz force-law. An identical argument provides the rate of change of p^t and we find, for the rate of change of the 4-momentum,

$$dp^\nu/dt = q\dot{x}^\mu F_\mu{}^\nu \tag{118}$$

which again is the same as for a particle.

A The 4-current density in terms of 3, 4 and 6 dimensional particle densities

A.1 The 4-current density in terms of the density in 3D space

The space-density of a collection of point-like particles with label P and trajectories $x_P(t)$ is a sum of 3-dimensional Dirac δ-functions: $n = \sum_P \delta(\mathbf{x} - \mathbf{x}_P(t))$. It has the required property that its integral over some volume is the number of particles contained therein.

For a *fluid* – i.e. a dense collection of particles whose velocities are some function of space $\mathbf{v}(\mathbf{x}, t)$ – the flux of particles across a surface $d\mathbf{A}$ is $n\mathbf{v} \cdot d\mathbf{A}$. The rate of change of the number of particles $\delta N = n\delta V$ in a volume δV with time is the sum of the inward fluxes across the surfaces. This gives $\partial_t n = -\boldsymbol{\nabla} \cdot (n\mathbf{v})$ or $n^\nu{}_{,\nu} = 0$ where $n^\nu \equiv n \times (1, \mathbf{v})$. For a *gas*, where there is, in general, a distribution of velocities at any position, the particle flux is still $n\mathbf{v} \cdot d\mathbf{A}$, but with $\mathbf{v}$ the mean velocity.

The 4-current n^ν (which we will see transforms as a 4-vector) can be defined as

$$n^\nu(\mathbf{x}, t) \equiv \sum_P \dot{x}_P^\nu(t)\delta^{(3)}(\mathbf{x} - \mathbf{x}_P(t)). \tag{119}$$

If we integrate $\mathbf{n}$, the spatial part of this, over some volume δV then we get $\sum\limits_{P \in \delta V} \dot{\mathbf{x}}_P$, while, if we integrate n over the same volume we get $\sum\limits_{P \in \delta V}$. Dividing these gives the mean velocity $\mathbf{v} = \langle \dot{\mathbf{x}} \rangle = \sum\limits_{P \in \delta V} \dot{\mathbf{x}}_P / \sum\limits_{P \in \delta V} = \mathbf{n}/n$, so the spatial components are indeed $\mathbf{n} = n\mathbf{v}$ and, since $\dot{x}^t = dt/dt = 1$, the time component is $n^t \equiv n\dot{x}^t$.

For a collection of particles with electric charges q_P, the charge 4-current-density can be defined similarly as the sum over particles weighted by their charge q_P:

$$j^\nu(\mathbf{x}, t) = \sum_P q_P \dot{x}_P^\nu(t)\delta^{(3)}(\mathbf{x} - \mathbf{x}_P(t)). \tag{120}$$

We could, of course, replace the argument t of $\mathbf{x}_P$ and $\dot{x}_P^\nu$ by some other parameter along the path such as the proper time τ and replace $\dot{x}_P^\nu(t) \rightarrow \dot{x}_P^\nu(\tau) = \dot{x}_P^\nu(\tau(t))$ and $\mathbf{x}_P(t) \rightarrow \mathbf{x}_P(\tau) = \mathbf{x}_P(\tau(t))$.

If we integrate (120) over a spatial volume δV we obtain

$$\delta J^\nu = (\delta Q, \delta \mathbf{J}) = \int\limits_{\delta V} d^3x j^\nu = \sum_{P \in \delta V} q_P \dot{x}_P^\nu = \sum_{P \in \delta V} q_P \times (1, \dot{\mathbf{x}}_P) \tag{121}$$

whose time component is the charge in the volume and whose spatial components are an element of current à la Biot and Savart. The charge/current element δJ^ν may look like a 4-vector, but it isn't. It doesn't transform properly under boosts.

We constructed n^ν so that it would obey $n^\nu{}_{,\nu} = 0$ by appealing to properties of fluids. Alternatively, we can show that $n^\nu{}_{,\nu} = 0$ is implicit in the definition (119) above. Consider the particle with label P. Its contribution to $\vec{n}$ has time component $n_P^t = \delta^{(3)}(\mathbf{x} - \mathbf{x}_P(t))$ which is a function of $\mathbf{x}$ and $\mathbf{x}_P(t)$ so its partial derivative with respect to t at fixed $\mathbf{x}$ is, from the chain rule, $\partial_t n_P^t = \dot{\mathbf{x}}_P \cdot \boldsymbol{\nabla}_{\mathbf{x}_P} \delta^{(3)}(\mathbf{x} - \mathbf{x}_P(t))$. But $\boldsymbol{\nabla}_{\mathbf{x}_P} \delta^{(3)}(\mathbf{x} - \mathbf{x}_P(t)) = -\boldsymbol{\nabla}_{\mathbf{x}} \delta^{(3)}(\mathbf{x} - \mathbf{x}_P(t))$, so $\partial_t n_P^t = -\dot{\mathbf{x}}_P \cdot \boldsymbol{\nabla}_{\mathbf{x}} \delta^{(3)}(\mathbf{x} - \mathbf{x}_P(t))$. The spatial components of this particle's contribution to $\vec{n}$ are $\mathbf{n}_P = \dot{\mathbf{x}}_P(t)\delta^{(3)}(\mathbf{x} - \mathbf{x}_P(t))$. In the divergence of this, $\boldsymbol{\nabla}_{\mathbf{x}} \cdot \mathbf{n}_P$, the partial derivative operator – being the derivatives with respect to components of $\mathbf{x}$ at fixed t – acts only on the factor $\delta^{(3)}(\mathbf{x} - \mathbf{x}_P(t))$, so $\boldsymbol{\nabla}_{\mathbf{x}} \cdot \mathbf{n}_P = \dot{\mathbf{x}}_P(t) \cdot \boldsymbol{\nabla}_{\mathbf{x}} \delta^{(3)}(\mathbf{x} - \mathbf{x}_P(t))$. But this is $-\partial_t n_P^t$. Hence $n_{P,\nu}^\nu = \partial_t n_P^t + \boldsymbol{\nabla} \cdot \mathbf{n}_P = 0$, and summing the over particles gives $n^\nu{}_{,\nu} = 0$. Continuity of the charge current density $j^\nu{}_{,\nu} = 0$ follows in exactly the same way directly from (120).

A.2 The 4-current density and the density in 4D spacetime

A definition of the density in 4-dimensions of a set of particles with world-lines $\vec{x}_P(\tau)$, which is a set of filaments in spacetime that vanishes except on the world-lines, is

$$\rho(\vec{x}) = \sum_P \int d\tau \delta^{(4)}(\vec{x} - \vec{x}_P(\tau)). \tag{122}$$

We can think of this as analogous to the density of beads on strings $\mathbf{x}_S(\lambda)$ in 3D space where the physical spacing $d\lambda$ between the beads is constant (and which we can usefully take to define a unit distance – which we will demand be much less than the radius of curvature of the strings). Here each 'bead' represents the existence of a particle for one unit of proper time.

Or we might think of the density as being analogous to a tangle of garden hoses in 3-dimensions: $\rho(\mathbf{x}) = \sum_S \int d\lambda \delta^{(3)}(\mathbf{x} - \mathbf{x}_S(\lambda))$. Imagine there is water flowing at constant speed $d\lambda/dt = 1$ through these hoses. What is the flux density f_x (amount of water per unit area per unit time) of water flowing through an element of area perpendicular to the x coordinate axes? The answer is $f_x = \sum_S \int d\lambda (dx_S/d\lambda)\delta^{(3)}(\mathbf{x} - \mathbf{x}_S(\lambda))$ where here $dx_S/d\lambda$ is the rate at which x increases with distance along the S^{th} hose.

To see why, think of a slab of thickness Δx. The length of a section of hose passing through this slab is $|d\lambda/dx_S|\Delta x$. So if we multiply the density of hose by $dx_S/d\lambda$ and then integrate along the section we get Δx times ± 1 depending on whether the water is flowing in the positive or negative x-direction.

Evidently f_i (for $i = x, y, z$) are the components of a vector $\mathbf{f} = \sum_S \int d\lambda (d\mathbf{x}_S/d\lambda)\delta^{(3)}(\mathbf{x} - \mathbf{x}_S(\lambda))$ which, when dotted with an element of area $d\mathbf{A}$, gives the flux (amount of water per unit time) of water through that area.

This motivates the definition the $\nu^{\rm th}$ component of the particle 4-current-density as

$$n^\nu(\vec{x}) = \sum_P \int d\tau u_P^\nu(\tau)\delta^{(4)}(\vec{x} - \vec{x}_P(\tau)) \tag{123}$$

and for the charge 4-current-density:

$$j^\nu(\vec{x}) = \sum_P q_P \int d\tau u_P^\nu(\tau)_P \delta^{(4)}(\vec{x} - \vec{x}_P(\tau)) \tag{124}$$

where $\vec{u}_P = d\vec{x}_P/d\tau$.

We now demonstrate that (123) and (124) are equivalent to (119) and (120). The equivalence of (124) and (120) follows from the fact that $d\tau u_P^\nu = dx_P^\nu = \dot{x}_P^\nu dt$, while $\delta^{(4)}(\vec{x} - \vec{x}_P(\tau)) = \delta(t - t_P(\tau))\delta^{(3)}(\mathbf{x} - \mathbf{x}_P(\tau))$, so the integral is $\int dt \dot{x}_P^\nu(\tau)\delta^{(3)}(\mathbf{x} - \mathbf{x}_P(\tau))\delta(t - t_P(\tau)) = \dot{x}_P^\nu(t)\delta^{(3)}(\mathbf{x} - \mathbf{x}_P(t))$, where $\dot{x}_P^\nu(t)$ denotes the value of $\dot{x}_P^\nu(\tau)$ at the proper time τ which is the solution of $t_P(\tau) = t$, and similarly $\mathbf{x}_P(t) = \mathbf{x}_P(\tau)$. Multiplying by q_P and summing over particles gives (120).

Maxwell's equations, in which j^ν appears, are empirical laws based on observations involving quantities like δQ and $\delta \mathbf{J}$, as well as forces and hence measurements of fields. We can thus consider either (120) or (124) to be empirically based definitions of the 4-current-density.

While it is somewhat redundant, one can show that the continuity equation $n^\nu{}_{,\nu} = 0$ follows directly from the definition (123) as follows: The contribution to the particle 4-current density n^ν from an infinitesimal element of proper time $d\tau$ for one of the particles is $dn_P^\nu(\vec{x}) = d\tau u_P^\nu(\tau)\delta^{(4)}(\vec{x} - \vec{x}_P(\tau))$. Its 4-divergence is $dn_{P,\nu}^\nu = d\tau u_P^\nu(\tau)\partial/\partial x^\nu \delta^{(4)}(\vec{x} - \vec{x}_P(\tau)) = -d\tau u_P^\nu(\tau)\partial/\partial x_P^\nu(\delta^{(4)}(\vec{x} - \vec{x}_P(\tau))) = -dx_P^\nu \partial/\partial x_P^\nu(\delta^{(4)}(\vec{x} - \vec{x}_P(\tau)))$. The integral of this is just the boundary term: $n_{P,\nu}^\nu = \int_{\tau_1}^{\tau_2} dn_{P,\nu}^\nu = -[\delta^{(4)}(\vec{x} - \vec{x}_P(\tau))]_{\tau_1}^{\tau_2}$. But real world-lines do not end, or, if they do, it is at $t = \pm\infty$. Everywhere else $n^\nu{}_{,\nu} = 0$.

A.3 The 4-current density in terms of the density in 6D phase-space

The density $f(\mathbf{x}, \mathbf{p}, t)$ of particles in 6-dimensional phase space $(\mathbf{x}, \mathbf{p})$, defined such that the number of particles in a 6-dimensional volume element $d^3x d^3p$ at time t is $d^6N = f(\mathbf{x}, \mathbf{p}, t) d^3x d^3p$, is a sum of 6-dimensional Dirac δ-functions:

$$f(\mathbf{x}, \mathbf{p}, t) = \sum_P \delta^{(3)}(\mathbf{x} - \mathbf{x}_P(t))\delta^{(3)}(\mathbf{p} - \mathbf{p}_P(t)). \tag{125}$$

where $(\mathbf{x}_P(t), \mathbf{p}_P(t))$ is the trajectory of the $P^{\rm th}$ particle.

The space-density $n(\mathbf{x}, t) = \sum_P \delta(\mathbf{x} - \mathbf{x}_P(t))$ is simply $n(\mathbf{x}, t) = \int d^3p f(\mathbf{x}, \mathbf{p}, t)$. The mean velocity is $\mathbf{v} = \int d^3p \dot{\mathbf{x}} f(\mathbf{x}, \mathbf{p}, t) / \int d^3p f(\mathbf{x}, \mathbf{p}, t)$, from which if follows that the 4-current-density for a set of particles of equal charge q is

$$j^\nu(\vec{x}) = q n^\nu(\vec{x}) = q \int d^3p \, \dot{x}^\nu f(\mathbf{x}, \mathbf{p}, t) = q \int \frac{d^3p}{p^t} p^\nu f(\mathbf{x}, \mathbf{p}, t) \tag{126}$$

where, in the last form, d^3p/p^t and $f(\mathbf{x}, \mathbf{p}, t)$ are both Lorentz scalars. This is for particles of equal charge and mass. For such particles, Hamilton's equations tell us that the 6-dimensional velocity $(\dot{\mathbf{x}}, \dot{\mathbf{p}})$ is only a function of 6-dimensional position $(\mathbf{x}, \mathbf{p})$. I.e. these particles are like a *fluid* in phase-space (unlike a gas in 3-dimensional space where, at any position $\mathbf{x}$, there is a range of velocities $\dot{\mathbf{x}}$). If we have different types of particles with different charges or charge-to-mass ratios, such as particles and their anti-particles, or electrons and ions in a plasma, then we need to sum over the different types as we then have a superposition of phase-space fluids.

B Continuity of 4-momentum in terms of 3, 4 and 6 dimensional particle densities

B.1 Continuity equation in terms of the 3D density

We now show how the continuity equation (46) follows from the expression $T^{\mu\nu} = m\sum_P \dot{x}_P^\mu(t)u_P(t)\delta^{(3)}(\mathbf{x}-\mathbf{x}_P(t))$ in terms of the 3-D density (44). Consider the P^{th} particle, using $\dot{x}^t = 1$ and $mu^\nu = p_P^\nu$ its contribution to the time-time part of the 4-divergence $T^{t\nu}{}_{,t}$ is $T_P^{t\nu}{}_{,t} = \partial_t(p_P^\nu(t)\delta^{(3)}(\mathbf{x}-\mathbf{x}_P(t))) = \dot{p}_P^\nu\delta^{(3)}(\mathbf{x}-\mathbf{x}_P(t)) + p_P^\nu\partial_t\delta^{(3)}(\mathbf{x}-\mathbf{x}_P(t))$. Its contribution to the spatial divergence is $T_P^{i\nu}{}_{,i} = \boldsymbol{\nabla}_{\mathbf{x}}\cdot(\mathbf{x}_P(t)p_P^\nu(t)\delta^{(3)}(\mathbf{x}-\mathbf{x}_P(t))) = p_P^\nu(t)\mathbf{x}_P(t)\cdot\boldsymbol{\nabla}_{\mathbf{x}}\delta^{(3)}(\mathbf{x}-\mathbf{x}_P(t)) = -p_P^\nu(t)\mathbf{x}_P(t)\cdot\boldsymbol{\nabla}_{\mathbf{x}_P}\delta^{(3)}(\mathbf{x}-\mathbf{x}_P(t)) = -p_P^\nu\partial_t\delta^{(3)}(\mathbf{x}-\mathbf{x}_P(t))$. This is minus the second term in $T_P^{t\nu}{}_{,t}$ giving $T_P^{\mu\nu}{}_{,\mu} = T_P^{t\nu}{}_{,t} + T_P^{i\nu}{}_{,i} = \dot{p}_P^\nu\delta^{(3)}(\mathbf{x}-\mathbf{x}_P(t))$. Finally, using $\dot{p}_P^\nu = q_P\dot{x}_P^\mu F_\mu{}^\nu$ gives $T_P^{\mu\nu}{}_{,\mu} = q_P F_\mu{}^\nu\dot{x}_P^\mu\delta^{(3)}(\mathbf{x}-\mathbf{x}_P(t))$, and summing over particles gives (46).

B.2 Continuity equation in terms of the 4D density

The continuity equation (46) can also be obtained directly from the second expression in (44) as follows: The partial derivative ∂_μ acts only on the 4D δ-function so (again considering first the contribution from the P^{th} particle) $T_P^{\mu\nu}{}_{,\mu} = m\int d\tau u_P^\mu(\tau)u_P^\nu(\tau)\partial/\partial_{x^\mu}\delta^{(4)}(\vec{x}-\vec{x}_P(\tau)) = -\int dx_P^\mu p_P^\nu(\tau)\partial/\partial_{x_P^\mu}\delta^{(4)}(\vec{x}-\vec{x}_P(\tau))$ where we have used $d\tau u_P^\mu = dx_P^\mu$ and $mu_P^\nu = p_P^\nu$. Integrating by parts gives $T_P^{\mu\nu}{}_{,\mu} = \int dx_P^\mu\delta^{(4)}(\vec{x}-\vec{x}_P(\tau))dp_P^\nu(\tau)/dx_P^\mu$ plus a boundary term at the end of the particle's world-line, which we may ignore, so $T_P^{\mu\nu}{}_{,\mu} = \int d\tau\delta^{(4)}(\vec{x}-\vec{x}_P(\tau))dp_P^\nu(\tau)/d\tau$ and using $dp_P^\nu(\tau)/d\tau = q_P u_P^\mu F_\mu{}^\nu$ and summing gives $T_P^{\mu\nu}{}_{,\mu} = F_\mu{}^\nu\sum_P q_P\int d\tau u_P^\mu\delta^{(4)}(\vec{x}-\vec{x}_P(\tau)) = F_\mu{}^\nu j_P^\mu$.

B.3 Continuity equation in terms of the 6D density

Taking the divergence of the last expression for $T^{\mu\nu}$ in (44) gives

$$T^{\mu\nu}{}_{,\mu} = \partial_\mu\int d^3p f(\mathbf{x},\mathbf{p},t)\dot{x}^\mu(\mathbf{p})p^\nu(\mathbf{p}) = \int d^3p\, p^\nu\dot{x}^\mu\partial_\mu f(\mathbf{x},\mathbf{p},t) = -\int d^3p\, p^\nu\boldsymbol{\nabla}_{\mathbf{p}}\cdot(f\dot{\mathbf{p}}) \tag{127}$$

where we have used Liouville's theorem $df/dt = \partial_t f + \dot{\mathbf{x}}\cdot\boldsymbol{\nabla}_{\mathbf{x}}f + \dot{\mathbf{p}}\cdot\boldsymbol{\nabla}_{\mathbf{p}}f = \dot{x}^\mu\partial_\mu f + \dot{\mathbf{p}}\cdot\boldsymbol{\nabla}_{\mathbf{p}}f = 0$ and $\boldsymbol{\nabla}_{\mathbf{p}}\dot{\mathbf{p}} = 0$.

Integrating by parts, and assuming that f tends to zero at infinity, we have

$$T^{\mu\nu}{}_{,\mu} = \int d^3p f\dot{\mathbf{p}}\cdot\boldsymbol{\nabla}_{\mathbf{p}}p^\nu(\mathbf{p}) = qF_{\mu i}\int d^3p f\dot{x}^\mu\partial_{p_i}p^\nu(\mathbf{p}) \tag{128}$$

where we have used the Lorentz force law $\dot{p}_i = qF_{\mu i}\dot{x}^\mu$.

The $\nu = t$ component of $\partial_{p_i}p^\nu(\mathbf{p})$ is, from $p^t = \sqrt{m^2+|\mathbf{p}|^2}$, given by $\partial_{p_i}p^t = \dot{x}^i$, while $\partial_{p_i}p^j = \delta_{ij}$. Using the latter in (128) we obtain

$$T^{\mu i}{}_{,\mu} = F_{\mu i}q\int d^3p f\dot{x}^\mu = j^\mu F_\mu{}^i. \tag{129}$$

And using the former (in the penultimate expression for $T_{\text{m}}^{\mu\nu}{}_{,\mu}$ in (128)) gives

$$T^{\mu t}{}_{,\mu} = \int d^3p f\dot{p}_i\dot{x}_i = \int\frac{d^3p}{p^t}f\dot{p}_i p_i = -\int\frac{d^3p}{p^t}f\dot{p}^t p_t = \int d^3p f\dot{p}^t = F_\mu{}^t q\int d^3p f\dot{x}^\mu = j^\mu F_\mu{}^t \tag{130}$$

where, in the third step, we used the fact that $p^\mu p_\mu = -m^2$ is constant, so (half) its derivative $\dot{p}^\mu p_\mu = \dot{p}^t p_t + \dot{p}_i p_i = 0$, and in the fifth we used the work equation $\dot{p}^t = qF_\mu{}^t\dot{x}^\mu$. Thus again we have $T^{\mu\nu}{}_{,\mu} = j^\mu F_\mu{}^\nu$ just as we found at the outset.

C The radiation Lagrangian and stress tensor in terms of E and B

The Lagrangian for the radiation is

$$\mathcal{L}_{\text{r}} = -\tfrac{1}{4}F^{\mu\nu}F_{\mu\nu} \tag{131}$$

or equivalently

$$\mathcal{L}_{\text{r}} = \tfrac{1}{4}\text{Tr}(F^{\alpha\nu}F_{\nu\beta}). \tag{132}$$

Using

$$F^{\mu\nu} = \begin{bmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{bmatrix} \quad \text{and} \quad F_{\mu\nu} = \begin{bmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -B_z & B_y \\ -E_y & B_z & 0 & -B_x \\ -E_z & -B_y & B_x & 0 \end{bmatrix} \tag{133}$$

we obtain their product

$$F^{\alpha\nu}F_{\nu\beta} = \begin{bmatrix} |\mathbf{E}|^2 & E_zB_y - E_yB_z & E_xB_z - E_zB_x & E_yB_x - E_xB_y \\ E_yB_z - E_zB_y & E_x^2 - B_z^2 - B_y^2 & E_xE_y + B_xB_y & E_xE_z + B_xB_z \\ E_zB_x - E_xB_z & E_xE_y + B_xB_y & E_y^2 - B_z^2 - B_x^2 & E_yE_z + B_yB_z \\ E_xB_y - E_yB_x & E_xE_z + B_xB_z & E_yE_z + B_yB_z & E_z^2 - B_x^2 - B_y^2 \end{bmatrix} \tag{134}$$

Its trace is $2(|\mathbf{E}|^2 - |\mathbf{B}|^2)$ so

$$\mathcal{L}_{\mathrm{r}} = -\tfrac{1}{4}F^{\mu\nu}F_{\mu\nu} = \tfrac{1}{2}(|\mathbf{E}|^2 - |\mathbf{B}|^2). \tag{135}$$

We can write this product (134) more simply as

$$F^{\alpha\nu}F_{\nu\beta} = \begin{bmatrix} |\mathbf{E}|^2 & -\mathbf{S} \\ \mathbf{S} & \boldsymbol{\sigma} + \tfrac{1}{2}\mathbf{I}(|\mathbf{E}|^2 - |\mathbf{B}|^2) \end{bmatrix} \tag{136}$$

where $\mathbf{S} \equiv \mathbf{E} \times \mathbf{B}$ is the *Poynting energy flux density* and where $\mathbf{I} \to \delta_{ij} = \mathrm{diag}(1,1,1)$ is the 3-by-3 identity matrix and where $\boldsymbol{\sigma}$ is the 3 by 3 symmetric *Maxwell stress tensor* is defined by

$$\boldsymbol{\sigma} \equiv \mathbf{EE} + \mathbf{BB} - \tfrac{1}{2}\mathbf{I}(|\mathbf{E}|^2 + |\mathbf{B}|^2). \tag{137}$$

Raising the index β we obtain the symmetric matrix

$$F^{\alpha\nu}F_{\nu}{}^{\beta} = \begin{bmatrix} -|\mathbf{E}|^2 & -\mathbf{S} \\ -\mathbf{S} & \boldsymbol{\sigma} + \tfrac{1}{2}\mathbf{I}(|\mathbf{E}|^2 - |\mathbf{B}|^2) \end{bmatrix} \tag{138}$$

from which we obtain

$$T_{\mathrm{r}}^{\alpha\beta} = F^{\alpha\nu}F^{\beta}{}_{\nu} + \eta^{\alpha\beta}\mathcal{L}_{\mathrm{r}} = \begin{bmatrix} \tfrac{1}{2}(|\mathbf{E}|^2 + |\mathbf{B}|^2) & \mathbf{S} \\ \mathbf{S} & -\boldsymbol{\sigma} \end{bmatrix} \tag{139}$$

in which we recognise the usual expressions for the energy density in T_{r}^{tt} and the energy flux density and momentum density in T_{r}^{it} and T_{r}^{ti} respectively (these being equal).

D Stress, energy and momentum for beams and wave-packets

In the expression for $T^{\nu}{}_{\mu}$ (72) we have factors like $\phi_{,\mu}\phi^{,\nu}$. Using (76) this is a double sum over wave modes

$$\phi_{,\mu}\phi^{,\nu} = -\tfrac{1}{4}\sum_{\mathbf{k}}\sum_{\mathbf{k}'} k_\mu k'^\nu (\phi_{\mathbf{k}} e^{+i\vec{k}\cdot\vec{x}} - \phi^\star_{\mathbf{k}} e^{-i\vec{k}\cdot\vec{x}})(\phi_{\mathbf{k}'} e^{+i\vec{k}'\cdot\vec{x}} - \phi^\star_{\mathbf{k}'} e^{-i\vec{k}'\cdot\vec{x}}) \tag{140}$$

If we integrate this over all space we will get non-zero contribution from the terms with $\mathbf{k}' = \mathbf{k}$, for which the terms like $\phi_{\mathbf{k}} e^{+i\vec{k}\cdot\vec{x}}\phi^\star_{\mathbf{k}'} e^{-i\vec{k}'\cdot\vec{x}} \to \phi_{\mathbf{k}}\phi^\star_{\mathbf{k}}$ which has no space- or time-oscillations. There are also non-vanishing contributions from terms with $\mathbf{k}' = -\mathbf{k}$. For example, $\phi_{\mathbf{k}} e^{+i\vec{k}\cdot\vec{x}}\phi_{\mathbf{k}'} e^{+i\vec{k}'\cdot\vec{x}}$ has no spatial oscillation for $\mathbf{k}' = -\mathbf{k}$ and so contributes, but it has a rapid time-oscillation $\sim e^{2i\omega_{\mathbf{k}}t}$ which will drops out if we average over a multiple of half-periods. Neglecting the latter we have

$$\overline{\phi_{,\mu}\phi^{,\nu}} \equiv \frac{1}{L^3}\int d^3x \phi_{,\mu}\phi^{,\nu} = \tfrac{1}{2}\sum_{\mathbf{k}} k_\mu k^\nu \phi_{\mathbf{k}}\phi^\star_{\mathbf{k}} \tag{141}$$

Similarly, we find

$$\overline{\phi^2} = \tfrac{1}{2}\sum_{\mathbf{k}} \phi_{\mathbf{k}}\phi^\star_{\mathbf{k}} \tag{142}$$

It follows that the average of the Lagrangian density is

$$-\tfrac{1}{2}\overline{\phi_{,\mu}\phi^{,\mu}} - \tfrac{1}{2}m^2\overline{\phi^2} = \tfrac{1}{4}\sum_{\mathbf{k}}(k_\mu k^\mu + m^2)\phi_{\mathbf{k}}\phi^\star_{\mathbf{k}} \tag{143}$$

which vanishes, by virtue for the dispersion relation (physically representing the fact that for an oscillator, the time average of the kinetic and potential energies are equal, so their difference - the Lagrangian - vanishes) hence the average of the stress tensor is

$$\overline{T^{\nu}{}_{\mu}} = \overline{\phi_{,\mu}\phi^{,\nu}} \equiv \frac{1}{L^3}\int d^3x \phi_{,\mu}\phi^{,\nu} = \frac{1}{2}\sum_{\mathbf{k}} k_\mu k^\nu \phi_{\mathbf{k}}\phi^\star_{\mathbf{k}} \tag{144}$$

or, with $\Delta k = 2\pi/L$,

$$\overline{T^{\nu\mu}} = \frac{1}{2}\frac{L^3}{2\pi}\sum_{\mathbf{k}}(\Delta k)^3 k^\mu k^\nu \phi_{\mathbf{k}}\phi^\star_{\mathbf{k}} \rightarrow \frac{L^3}{2(2\pi)^3}\int \frac{d^3k}{\omega_{\mathbf{k}}} k^\mu k^\nu (\omega_{\mathbf{k}}\phi_{\mathbf{k}}\phi^\star_{\mathbf{k}}) \tag{145}$$

This is identical in form to the canonical stress energy tensor for particles $T^{\mu\nu} = \int (d^3p/p^t)p^\mu p^\nu f(\mathbf{p})$ (in the absence of an EM field) if we identify $\mathbf{p}$ with $\mathbf{k}$, $\omega_{\mathbf{k}}$ with p^t, and $\omega_{\mathbf{k}}\phi_{\mathbf{k}}\phi^\star_{\mathbf{k}}$ with the phase space density of particles $f(\mathbf{p})$.

It follows that for a nearly monochromatic wave packet for which the mode amplitudes are appreciable only close to $\mathbf{k} = \overline{\mathbf{k}}$, the total 4-momentum is

$$p^\mu = \overline{T^{t\mu}} = \overline{k}^\mu \sum_{\mathbf{k}} \omega_{\mathbf{k}}\phi_{\mathbf{k}}\phi^\star_{\mathbf{k}} \tag{146}$$

and the components of $p^\mu = (E, \mathbf{p})$ satisfy the same energy momentum relation $E^2 - P^2 =$ constant as a relativistic particle.

The same is true of the total 4-momentum of a broad beam (which can be thought of as a wave-packet that is extended along the direction parallel to $\overline{\mathbf{k}}$).