Particles and Fields

Holger Gies

FSU Jena

Lecture Notes, TPI

1. Introduction

1.1 Why particles & fields?

This course is meant to be a preparatory course for an in depth lecture course on quantum field theory (QFT). In fact, QFT has become the language of modern physics. Most prominently, QFT describes the physics of elementary particles and their interactions at the most fundamental level that is currently accessible to observations in the laboratory (e.g. at colliders) or in astrophysical or cosmological data. QFT even has the potential to describe systems to arbitrarily short-distance or arbitrarily high-energy scales (in contrast to classical mechanics, electrodynamics or quantum mechanics). Moreover, QFT provides also for useful tools for the description of condensed-matter systems, many-body physics, critical phenomena, statistical systems, phase transitions, etc.

It is therefore not astonishing that QFT exhibits a deep level of structural and technical complexity, challenging both students and teachers in a compact lecture course.

The purpose of this course hence is to remove a large part of this complexity by ignoring quantization. The remaining body of classical field theory still offers a comprehensive playground where many concepts and actually real physics can be learned and understood.

Though the mathematics of this course deals with classical field theory, the goal (behind the horizon) is QFT and its application to particle physics. Hence, some applications and discussions center around elementary particle physics. As QFT supercedes the point-particle concept, the word "particle" in the title does not allude to classical point particles, but to the modern understanding of particles as quantized excitations of fields. As we stay within the realm of classical physics in this course, a "particle" should be thought of as a classical excitation of a field, such as a localized propagating wave.

1.2 Examples of classical field theories

In classical field theory, each point in spacetime $x \equiv (t, \vec{x})$ is associated with a function ϕ

$$x \rightarrow \phi(x) \qquad (1.1)$$

Depending on the system ϕ could be a real or a complex number, $\phi \in \mathbb{R}$ or $\phi \in \mathbb{C}$, or an N-tuple of such numbers ϕ^a, $a = 1, \dots, N$.

Examples are given by the electrostatic potential $\varphi(x) \in \mathbb{R}$ in classical electrostatics, or the vector potential $\vec{A}(x)$ consisting of 3 components, giving rise to a magnetic field $\vec{B}(x) = \vec{\nabla} \times \vec{A}(x)$.

We typically assume $\phi(x)$ to be sufficiently smooth and differentiable (e.g. $\phi \in C^2$) such that its dynamics can be governed by a differential equation, the <u>field equation</u> or <u>equation of motion</u> (EOM).

This abstract notion is already familiar from classical electrodynamics, being a paradigmatic example for a classical field theory.

The field equations for the electric and magnetic field components, $\vec{E}(x)$ and $\vec{B}(x)$, are given by the Maxwell equations, which in vacuum read

$$\vec{\nabla}\cdot\vec{E} = 0 \qquad\qquad \vec{\nabla}\times\vec{B} = \frac{\partial\vec{E}}{\partial t}$$
$$\vec{\nabla}\cdot\vec{B} = 0 \qquad\qquad \vec{\nabla}\times E + \frac{\partial\vec{B}}{\partial t} = 0 \qquad (1.2)$$

Here, we have already used the convention $c = 1$ (i.e. all velocity-like quantities are measured in fractions of light speed, or lengths are measured by the time that light takes to propagate this distance.)

Mathematically, the field equations are (coupled) partial differential equations (PDEs), the solution of which requires suitable boundary conditions or/and initial data.

The Maxwell equations form a rather peculiar example, as the information encoded in the 6 functions $E_1(x)$, $E_2(x)$, $E_3(x)$, $B_1(x)$, $B_2(x)$, $B_3(x)$ can also be parametrized by the above mentioned 4 auxiliary functions of the electrostatic potential $\varphi(x)$ and the vector potential $\vec{A}$, where

$$\vec{B}(x) = \vec{\nabla} \times \vec{A}(x) \, ,$$
$$\vec{E}(x) = -\vec{\nabla}\varphi(x) - \frac{\partial}{\partial t}\vec{A}(x) \, . \tag{1.3}$$

Inserting (1.3) into (1.2), and using $\vec{\nabla} \times \vec{\nabla}\varphi = 0$ and $\vec{\nabla} \cdot \vec{\nabla} \times \vec{A} = 0$ (for smooth φ and $\vec{A}$), the 2nd line of (1.2) is automatically satisfied, while the 1st line boils down to

$$\vec{\nabla}^2 \varphi + \frac{\partial}{\partial t}(\vec{\nabla}\cdot\vec{A}) = 0$$
$$\vec{\nabla}^2 \vec{A} - \frac{\partial^2}{\partial t^2}\vec{A} - \vec{\nabla}\left(\vec{\nabla}\cdot\vec{A} + \frac{\partial}{\partial t}\varphi\right) = 0 \tag{1.4}$$

forming 4 PDE's for the 4 unknown fields $\varphi(x), \vec{A}(x)$.

This parametrization in terms of the potentials φ and $\vec{A}$ is even more peculiar, as the choice of φ and $\vec{A}$ is not unique. For instance, if φ and $\vec{A}$ are shifted according to

$$\varphi(x) \to \varphi'(x) = \varphi(x) - \frac{\partial}{\partial t}\lambda(x)$$
$$\vec{A}(x) \to \vec{A}'(x) = \vec{A}(x) + \vec{\nabla}\lambda(x) \tag{1.5}$$

with an arbitrary function $\lambda(x) (\in C^2)$, the $\vec{E}$ and $\vec{B}$ fields in (1.3) remain the same. While $\vec{E}$ and $\vec{B}$ can be measured in terms of forces acting on

(moving) charged particles, the values of $\varphi(x)$ and $\vec{A}(x)$ at a given point x can be shifted by (1.5) to any value and thus have no locally observable meaning.
This invariance under local shifts a la (1.5) is called a <u>gauge symmetry</u> and characterizes a very special (and very important) class of field theories.

For our present purpose, it is useful to choose $\lambda(x)$ in such a way that φ' and $\vec{A}'$ satisfy the following auxiliary condition (Lorenz gauge condition)

$$\vec{\nabla}\cdot\vec{A}' + \frac{\partial}{\partial t}\varphi' = 0. \tag{1.6}$$

If so, the field equations (1.4) for φ' and $\vec{A}'$ simplify to

$$\begin{aligned} \vec{\nabla}^2\varphi' - \frac{\partial^2}{\partial t^2}\varphi' &= 0 \\ \vec{\nabla}^2\vec{A}' - \frac{\partial}{\partial t^2}\vec{A}' &= 0 \end{aligned} \tag{1.7}$$

Or simply $\Box\varphi' = 0$, $\Box\vec{A}' = 0$, where $\Box = -\vec{\nabla}^2 + \frac{\partial}{\partial t^2}$ is the d'Alembert operator.

Eqs. (1.7) are a wave equations for all 4 field functions which hence admit plane wave solutions:

$$\varphi', \vec{A}' \sim e^{-i\omega t + i\vec{k}\cdot\vec{x}}, \quad \text{with } \omega^2 = \vec{k}^2. \tag{1.8}$$

(for complexified fields, or Re / Im $e^{-i\omega t + i\vec{k}\cdot\vec{x}}$ for real fields)

In addition to gauge invariance, Maxwell's equations also have an invariance with respect to the choice of coordinate systems. The corresponding invariance is a relativistic invariance, and the corresponding transformations between coordinate systems moving relative to each other at constant speed $\beta = \frac{v}{c} \equiv v$ are the Lorentz transformations. For instance, if two coordinate systems move relative to each other along their common x direction, the Lorentz transformation reads

$$
\begin{aligned}
t' &= \gamma (t - \beta x) \qquad & y' &= y \\
x' &= \gamma (x - \beta t) \qquad & z' &= z
\end{aligned}
\qquad (1.9)
$$

$$\gamma = \frac{1}{\sqrt{1-\beta^2}}$$

Summarizing the spacetime coordinates in a ("contravariant") 4-vector $x^\mu = (\underset{x^0}{t}, \underset{x^1}{x}, \underset{x^2}{y}, \underset{x^3}{z})$. This transformation can be written in a matrix form

$$x^{\mu\prime} = \Lambda^\mu{}_\nu \, x^\nu \qquad \text{(summation over } \nu \text{ is implicitly understood)} \qquad (1.10)$$

where

$$\Lambda = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \qquad (1.11)$$

Of course, by suitably applying rotation matrices,

$$\vec{x}' = R\vec{x} \quad , R^T R = \mathbb{1} \ , \ R \in SO(3) \ ,$$

the Lorentz transformations generalize to "boosts" along any other direction $\vec{\beta}$, as well as to coordinate systems spatially rotated relative to each other (as in classical mechanics).

Recall that (1.9) follows from Einstein's postulate that the wave front of a flash of light starting at a common origin of the coordinate systems propagates at the same speed as measured in both systems. The position of such a (spherical) wave front after time t (t') is at

$$0 = t^2 - (x^2 + y^2 + z^2) \quad , 0 = t'^2 - (x'^2 + y'^2 + z'^2) \qquad (1.12)$$

respectively. This suggests to introduce the Minkowski metric

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} , \qquad (1.13)$$

to write the propagation distance of the wave front in both systems as

$$0 = x^\mu g_{\mu\nu} x^\nu = x^{\mu'} g_{\mu\nu} x^{\nu'} \qquad (1.14)$$

Using (1.10), we get

$$x^\mu g_{\mu\nu} x^\nu = \Lambda^\mu{}_\kappa x^\kappa g_{\mu\nu} \Lambda^\nu{}_\lambda x^\lambda$$

$$\underset{\nu \leftrightarrow \lambda}{\overset{\mu \leftrightarrow \kappa}{=}} x^\mu \Lambda^\kappa{}_\mu g_{\kappa\lambda} \Lambda^\lambda{}_\nu x^\nu \qquad (1.15)$$

Note that in this context x^μ is not just any position in spacetime but a <u>vector</u> specifying the distance of the wave front from the origin. From (1.15) we read off that Lorentz transformations Λ of such vectors satisfy

$$g_{\mu\nu} = g_{\kappa\lambda} \Lambda^\kappa{}_\mu \Lambda^\lambda{}_\nu . \qquad (1.16)$$

It is straightforward to verify that (1.11) satisfies this condition.

More generally, we call any 4-by-4 matrix Λ that satisfies (1.16) for the metric (1.13) a Lorentz transformation. Hence, (1.16) has the same status for Lorentz transformations, as $R^T R = \mathbb{1}$ ($\delta_{ij} = \delta_{kl} R^k{}_i R^l{}_j$) has for rotations. The corresponding matrix group is $SO(3,1)$. We will discuss this group in more detail below.

Any 4-tuple v^μ, $\mu = 0,1,2,3$, that transforms under changes of the Lorentz system as

$$v^{\mu\prime} = \Lambda^\mu{}_\nu \, v^\nu \qquad (1.17)$$

is called a Lorentz 4-vector. Correspondingly, objects $T^{\mu_1 \mu_2 \dots \mu_n}$ that transform as

$$T^{\mu_1\mu_2\mu_3\dots\mu_n\prime} = \Lambda^{\mu_1}{}_{\nu_1} \Lambda^{\mu_2}{}_{\nu_2} \dots \Lambda^{\mu_n}{}_{\nu_n} T^{\nu_1 \dots \nu_n} \qquad (1.18)$$

are called Lorentz tensors of rank n.

It is useful to introduce "covariant" vectors by

$$x_\mu := g_{\mu\nu} \, x^\nu = (t, -\vec{x}) \, . \qquad (1.19)$$

With this notation, the light-front position discussed above can be written as $0 = x_\mu x^\mu = x'_\mu x^{\mu\prime}$, which makes it obvious that expressions with pair-wise contracted upper & lower indices are Lorentz invariant.

For instance, the argument of the plane wave in (1.8) can be written as

$$-i\omega t + i \vec{k} \cdot \vec{x} = -i \, k^\mu x_\mu \qquad (1.20)$$

where $k^\mu = (\omega, \vec{k})$, [NB: the fact that ω and $\vec{k}$ indeed transform as components of a 4-vector is a manifestation of the relativistic Doppler effect.] Hence, the plane wave form of (1.8) is a relativistic invariant. This translates into the invariance of the corresponding wave operator

$$\Box = \frac{\partial^2}{\partial t^2} - \vec{\nabla}^2 . \qquad (1.21)$$

The trivial fact that

$$\frac{\partial}{\partial x^\mu} x^\nu = \begin{cases} 1 & \text{for } \mu=\nu \\ 0 & \text{otherwise} \end{cases} \qquad (1.22)$$

$$\Rightarrow \quad \frac{\partial}{\partial x^\mu} x^\mu = 4$$

holds in any Lorentz frame suggests to interpret $\frac{\partial}{\partial x^\mu}$ as a covariant vector: ∂_μ

$$\partial_\mu x^\mu = 4 \quad , \quad \partial_\mu = \left(\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) \qquad (1.23)$$

The corresponding contravariant vector operator is

$$\partial^\mu = g^{\mu\nu} \partial_\nu \quad , \quad \partial^\mu = \left(\frac{\partial}{\partial t}, -\frac{\partial}{\partial x}, -\frac{\partial}{\partial y}, -\frac{\partial}{\partial z}\right), \qquad (1.24)$$

where $g^{\mu\nu}$ denotes the inverse of $g_{\mu\nu}$. Obviously, we have $(g^{-1})^{\mu\nu} = g_{\mu\nu}$ component-wise. We write

$$g^{-1} g = \mathbb{1} \text{, or in components } g^{\mu\nu} g_{\nu\kappa} = \delta^\mu_\kappa \qquad (1.25)$$

With this notation, we have

$$\Box = \frac{\partial^2}{\partial t^2} - \vec{\nabla}^2 = \partial_\mu \partial^\mu \tag{1.26}$$

which makes Lorentz invariance manifest.

To conclude the discussion of classical electrodynamics, the form invariance of Maxwell's equations under Lorentz transformations becomes manifest by noticing that $\varphi(x)$ and $\vec{A}(x)$ also transform as components of a 4 vector

$$A^\mu_{(x)} = (\varphi(x), \vec{A}(x)) . \tag{1.27}$$

The Lorenz gauge condition (1.6) is hence Lorentz invariant

$$\partial_\mu A^\mu = 0 \tag{1.28}$$

From (1.3) it is clear that $\vec{E}$ and $\vec{B}$ cannot be arranged into 4-vectors. Instead, their components can be arranged into a Lorentz tensor, the field strength tensor

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$$

$$(F)^{\mu\nu} = \begin{pmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & -B_3 & B_2 \\ E_2 & B_3 & 0 & -B_1 \\ E_3 & -B_2 & B_1 & 0 \end{pmatrix}^{\mu\nu} \tag{1.29}$$

such that the 1st line of the Maxwell equations read

$$\partial_\mu F^{\mu\nu} = 0 \; . \tag{1.30}$$

This is a set of 4 equations, $\nu = 0,1,2,3$, that transform as a 4-vector under Lorentz transformations.
In order to write the 2nd line of (1.2) into 4-notation, it is useful to introduce the Minkowskian analogue of the Levi-Civita symbol

$$\varepsilon^{\mu\nu\kappa\lambda} = \begin{cases} +1 & \text{for } \mu=0,\ \nu=1,\ \kappa=2,\ \lambda=3 \\ & \text{and even permutations} \\ -1 & \text{for odd permutations} \\ 0 & \text{if two indices are equal} \end{cases} \; . \tag{1.31}$$

This allows to introduce the dual field strength tensor

$$\tilde{F}^{\mu\nu} = \frac{1}{2}\, \varepsilon^{\mu\nu\kappa\lambda}\, F_{\kappa\lambda} \tag{1.32}$$

where $F_{\kappa\lambda} = g_{\kappa\mu} F^{\mu\nu} g_{\nu\lambda}$

More explicitly,

$$\left(\tilde{F}\right)^{\mu\nu} = \begin{pmatrix} 0 & -B_1 & -B_2 & -B_3 \\ B_1 & 0 & E_3 & -E_2 \\ B_2 & -E_3 & 0 & E_1 \\ B_3 & E_2 & -E_1 & 0 \end{pmatrix} \tag{1.33}$$

By construction, we have

antisymmetric

$$0 = \partial_\mu \tilde{F}^{\mu\nu} = \frac{1}{2}\,\varepsilon^{\mu\nu\kappa\lambda}\,\partial_\mu\left(\partial_\kappa A_\lambda - \partial_\lambda A_\kappa\right) \tag{1.34}$$

symmetric

which is also called the Bianchi identity, which reproduces the 2nd line of (1.2). We close this section on electrodynamics by noting that the whole formalism can be generalized to non-vanishing charges and currents. Combining the charge density ρ and the current density $\vec{j}$ into a 4-vector $j^\mu = (\rho, \vec{j})$, the Maxwell equation (1.30) reads (in Heaviside-Lorentz units)

$$\partial_\mu F^{\mu\nu} = j^\nu , \tag{1.35}$$

while (1.34) remains as it is. Since $F^{\mu\nu}$ (as well as $\tilde{F}^{\mu\nu}$) is antisymmetric by construction, $F^{\mu\nu} = -F^{\nu\mu}$, current conservation is manifest:

antisym.

$$0 = \partial_\nu \partial_\mu F^{\mu\nu} = \partial_\nu j^\nu = \frac{\partial}{\partial t}\rho + \vec{\nabla}\cdot\vec{j} \tag{1.36}$$

sym.

Classical electrodynamics is obviously an example for a classical field theory with a high degree of structure both due to gauge symmetry as well as the vector and tensor nature of the field variables.

With this insight, we can "guess" a much simpler field theory that satisfies relativistic invariance:

$$\Box \, \phi(x) = 0 \qquad (1.37)$$

where $\phi(x)$ is a scalar field that transforms trivially under Lorentz transformations $\phi(x) \to \phi'(x') \equiv \phi(x)$.

In fact, (1.37) is identical to the Klein-Gordon equation

$$\left(\Box + m^2\right) \phi(x) = 0 \qquad (1.38)$$

for the special case of vanishing mass m.
(Here, we use also the convention that $\hbar = 1$.)

From our advanced quantum mechanics course, we know that the Klein-Gordon equation also admits plane wave solutions,

$$\phi \sim e^{-i\omega t + i\vec{k}\cdot\vec{x}} = e^{-ik^\mu x_\mu} \tag{1.39}$$

where $$k_\mu k^\mu = m^2 . \tag{1.40}$$

The last equation is equivalent to

$$\omega^2 = \vec{k}^2 + m^2 \tag{1.41a}$$

which according to our conventions is identical to

$$E^2 = \vec{p}^2 c^2 + (mc^2)^2 , \tag{1.41b}$$

being the relativistic energy - momentum relation (dispersion relation) of a relativistic point - particle. Of course, in the quantum mechanics course, the Klein- Gordon equation has been motivated by the relativistic dispersion relation (1.41) with the wave equation (1.38) being a consequence of the correspondence principle $E \to i\partial_t$, $\vec{p} \to -i\vec{\partial}_x$.

From the viewpoint of field theory, the logic is reversed: we have written down the simplest relativistic field equations in (1.37) and (1.38) which turn out to support wave excitations that obey the dispersion relation of a relativistic point particle.

[NB: in fact, leaving relativity and quantum mechanics aside, the Klein-Gordon equation also appears in continuum mechanics: it describes the propagation of longitudinal waves of (the continuum limit of) a chain or net of oscillators 
with $\phi(x)$ corresponding to the amplitude of an oscillator at point x; the speed c is related to the spring constants, and m is a measure for a harmonic force pulling each oscillator back to its rest position.]

Comparing the dispersion relation (1.41 a/b) to that found for waves in electrodynamics in (1.8), the latter appear to correspond to massless relativistic particles satisfying $\omega^2 = k^2$ or $E = |\vec{p}|c$, the quantized version of which will be the photons.

Having obtained the (quantum mechanical) Klein-Gordon equation from field theory considerations, it is a perfectly legitimate viewpoint, to interpret even the Schrödinger equation (at least mathematically) as a wave equation of a classical field theory,

$$i \, \partial_t \, \Psi(x) = -\frac{1}{2m} \vec{\nabla}^2 \Psi(x) + V(x) \, \Psi(x) . \qquad (1.42)$$

Obviously, the Schrödinger equation is not invariant under Lorentz transformations; instead it is Galilei invariant (as Newton's classical mechanics). Correspondingly, its excitations give rise to dispersion relations of a non-relativistic point particle, $E = \frac{p^2}{2m} + \dots$.

One may justifiably object that there is still a clear distinction between field theories such as electrodynamics on the one hand side, and quantum mechanical field equations on the other hand side, because the quantum mechanical wave functions have a probabilistic interpretation, $P(x) = |\Psi(x)|^2$, i.e. first, one needs to square the amplitude, and second, the result is a probability not a fully deterministic prediction for a single measurement.

However, this distinction becomes less meaningful, if we keep in mind that a typical observable for electromagnetic waves is the intensity, $I \sim |\vec{E}|^2, |\vec{B}|^2$, which is also related to the square of the field amplitude.

Moreover, when we approach the regime of very small intensities (and system sizes with actions of the order $S \sim \hbar$), we expect quantum effects to set in.

Interestingly, it is not the Maxwell equations which break down in this regime, but it is the interpretation of the amplitudes that breaks down: the intensity then is related to the probability of measuring radiation (a photon).

An important difference between the quantum mechanical and the field theory viewpoint is the following: in QM, we first lift space coordinates and momenta to operators $\vec{x}, \vec{p} \rightarrow \hat{\vec{x}}, \hat{\vec{p}}$ with non-trivial commutation relations, and only later when we formulate the Schrödinger equation in position space, the coordinates become "c-numbers" again. In this manner, there is a fundamental difference between space and time, as the latter t always remains a parameter.
By contrast, both time and space remain parameters on an equal footing in field theory. This holds also

true in QFT, where $(t, \vec{x})$ remain parameters; instead, the fields themselves are lifted to operators.

All of the examples of field theories mentioned so far are special in the sense that their field equations are linear in the amplitude $\Phi(x)$ (or $F^{\mu\nu}$, A^μ, ψ). As a consequence, the superposition principle holds: if two solutions $\Phi_1(x)$ and $\Phi_2(x)$ exist, then also

$$\Phi(x) = \alpha \Phi_1(x) + \beta \Phi_2(x) \tag{1.43}$$

is a solution (with $\alpha, \beta = \text{const}$).

This is generally no longer true if we consider non-linear theories. A famous example is Einstein's theory of general relativity, where the field variable is a now dynamical metric $g_{\mu\nu}(x)$ and the field equation reads (in vacuum without cosmological constant)

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = 0 \,. \tag{1.44}$$

Here, the Ricci tensor $R_{\mu\nu}$ and Ricci scalar R depend in a nonlinear way on $g_{\mu\nu}$ (and its inverse) and derivatives thereof.

1.3 The action principle for classical field theories

All of the above given examples for field equations can be derived from an action principle in much the same way as Hamilton's principle gives rise to equations of motion in classical mechanics. The corresponding actions turn out to be of the form

$$S[\Phi] = \int_V d^4x \, \mathcal{L}(\Phi, \partial_\mu \Phi) \qquad (1.45)$$

Here, the action S is considered to be a functional of the field Φ. The integration measure d^4x over spacetime is Lorentz invariant, as the Jacobian of the transformation, $d^4x \to d^4x' = |\det \Lambda| \, d^4x$, involves the modulus of the determinant of Λ, which by virtue of (1.16) satisfies $(\det \Lambda)^2 = 1$. If $\mathcal{L}$ transforms as a scalar, S is a Lorentz invariant number for any field Φ. The integration volume V may be finite or extend over full Minkowski space. Since (1.45) involves a

volume integration, $\mathcal{L}$ is called the Lagrange density.
We assume it to be a function of the field ϕ and its first derivative $\partial_\mu \phi$, since the above given field equations are of second order. As in classical mechanics, we could also allow for higher derivatives at the expense of higher order field equations.
We look for those field configurations that extremize the action S. As in classical mechanics, we assume that the general field can be written as

$$\phi(x,\alpha) = \phi(x) + \alpha\, \eta(x) \qquad (1.46)$$

where $\phi(x)$ is the extremizing solution, α is a parameter, and $\eta(x)$ is an arbitrary field variation that vanishes on the boundary of V:

$$\eta(x)\Big|_{x\in\partial V} = 0 \qquad (1.47)$$

(i.e. if the general field has to satisfy specific boundary conditions on ∂V, these b.c.'s are completely carried by $\phi(x)$, i.e. by the extremizing field)
With these assumptions, S has to be stationary at $\alpha = 0$:

$$0 = \frac{\partial S[\phi]}{\partial\alpha}\Bigg|_{\alpha=0} = \int_V d^4x \left[\frac{\partial\mathcal{L}}{\partial\phi}\,\eta + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\,\partial_\mu\eta\right]_{\alpha=0}$$

Integrating the second term by parts, yields

$$0 = \int_V d^4x \left\{ \left[\frac{\partial \mathcal{L}}{\partial \varphi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)} \right] \eta(x) \right\}_{\alpha=0} + \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)} \eta \right]_{\partial V, \, \alpha=0} \quad (1.48)$$

The last term is a surface term (to be evaluated along the normal of the surface) which vanishes because of (1.47).
Since the first term has to vanish for any $\eta(x)$, we conclude that

$$\frac{\partial \mathcal{L}}{\partial \varphi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)} = 0 \qquad (1.49)$$

This is the field theory version of the Euler-Lagrange equation, representing a necessary condition for $\varphi(x)$ to be a local extremum of the action functional $S[\varphi]$. Note that we have not specified the nature of the field φ any further. If φ represents a multi-component field φ^a, $a = 1, \dots, N$ where a can be any kind of index, we correspondingly obtain N Euler-Lagrange equations

$$\frac{\partial \mathcal{L}}{\partial \varphi^a} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi^a)} = 0 \qquad (1.50)$$

Let us start with the simplest example of a single-component real scalar field $\phi(x) \in \mathbb{R}$. Since $\mathcal{L}$ must be a Lorentz scalar, the simplest term involving $\partial_\mu \phi$ which we can write down is $\sim (\partial_\mu \phi)(\partial^\mu \phi)$. Because of the necessary pairing of Lorentz indices, this term is invariant under the additional symmetry $\phi \Rightarrow -\phi$ (a $\mathbb{Z}_2$ symmetry, a transformation group consisting of the elements $\mathbb{Z}_2 \hat{=} \{-1, 1\}$). If we wish to maintain this symmetry also for the ϕ-dependent parts, the simplest Lagrange density takes the form

$$\mathcal{L} = \frac{1}{2} (\partial_\mu \phi)(\partial^\mu \phi) - \frac{1}{2} m^2 \phi^2, \qquad (1.51)$$

where the factors of $\frac{1}{2}$ are pure convention and the parameter m has been introduced to let the second term have the same "dimensionality" (units) as the first term. Inserting (1.51) into (1.49), we find

$$\frac{\partial \mathcal{L}}{\partial \phi} = - m^2 \phi \qquad (1.52a)$$

and with

$$(\partial_\kappa \phi)(\partial^\kappa \phi) = g^{\kappa\lambda}(\partial_\kappa \phi)(\partial_\lambda \phi)$$

and $$\frac{\partial(\partial_\kappa \phi)}{\partial(\partial_\mu \phi)} = \delta^\mu_\kappa \, ,$$ we get

$$\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} = \frac{1}{2}\frac{\partial}{\partial(\partial_\mu \phi)} g^{\kappa\lambda}(\partial_\kappa \phi)(\partial_\lambda \phi)$$

$$= \frac{1}{2} g^{\kappa\lambda} \delta^\mu_\kappa \partial_\lambda \phi + \frac{1}{2} g^{\kappa\lambda} \partial_\kappa \phi \, \delta^\mu_\lambda$$

$$= \partial^\mu \phi \qquad (1.52b)$$

$$\Rightarrow \quad \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} = \partial_\mu \partial^\mu \phi = \Box \phi \qquad (1.53)$$

In other words, the Euler-Lagrange equation reads

$$(\Box + m^2)\,\phi = 0 \qquad (1.54)$$

being identical to the Klein-Gordon equation.
We conclude that (1.51) corresponds to the Lagrange density of Klein-Gordon theory.

Several comments are in order:

1) We have arrived at (1.51) using <u>symmetry</u> arguments (Lorentz, $\mathbb{Z}_2$) and <u>simplicity</u>. While symmetry is a clearly defined criterion, simplicity (or beauty) is rather vague. While classical field theory has not much to offer as an alternative argument, quantum field theory does have another consistency criterion that can (at least partly) replace simplicity, it goes under the name of "renormalizability" which sounds (and at first sight is) technical, but goes to the very heart of the existence, origin or emergence of quantum field theories (see my lecture notes on "Physics of Scales"). To <u>zeroth</u> approximation, renormalizability is related to dimensionality, see below.

2) Disregarding $\mathbb{Z}_2$ symmetry, an even simpler term would be a linear term $\sim + j\phi$ with a parameter j or a function $j(x)$. The resulting field equation would be

$$(\Box + m^2)\,\phi(x) = j(x) \qquad (1.55)$$

Such a linear term hence would have the meaning of a <u>source term</u>. Note, however, that such a source term would break Z_2 symmetry.

3) Let us clarify the notion of units or dimensionality in our conventions where $\hbar = c = 1$. For instance, from the dispersion relation (1.41a), it is clear that energy, momentum and mass all carry the same units which can be expressed in terms of an arbitrary unit scale. In high-energy physics, the typical choice is the energy unit of electron Volts eV with a GeV (giga eV) corresponding approximately to the mass $\underset{\text{rest energy}}{\overset{}{\underset{\wr\wr\wr}{\text{of}}}}$ the proton. Solely counting mass or energy dimensions, we write

$$[E] = [\omega] = [p_i] = [m] = 1 \qquad (1.56a)$$

Since the action carries the same unit as $\hbar = 1$, the action itself is dimensionless,

$$[S] = 0 \qquad (1.56b)$$

Since position times momentum has the unit of an action (as well as angular momentum), we have

$$[x\, p] = 0$$

With (1.56a) this implies that position carries an inverse mass dimension:

$$[x] = -1 \qquad (1.56c)$$

Consequently, we have

$$[d^4x] = -4 \qquad (1.56d)$$

and thus with (1.56b):

$$[\mathcal{L}] = 4 \qquad (1.56e)$$

in four spacetime dimensions

From (1.56c) we deduce that

$$[\partial_\mu] = \left[\frac{\partial}{\partial x^\mu}\right] = 1 \,. \qquad (1.56f)$$

Combining these findings with the form of $\mathcal{L}$ in (1.51), we see that the field amplitude itself must carry a mass dimension

$$[\phi] = 1 \qquad (1.56g)$$

(Exercise: generalize these considerations to a Klein Gordon field in D dimensional spacetime)

4) The linearity of the resulting field equation is in one-to-one correspondence with the fact that the action / Lagrangian (1.51) is quadratic in the fields. It is straightforward to construct more general non-linear theories, e.g. by generalizing the mass term to a full function,

$$\mathcal{L} = \frac{1}{2} (\partial_\mu \varphi)(\partial^\mu \varphi) - V(\varphi) ; \tag{1.57}$$

in analogy to classical mechanics, we call $V(\varphi)$ a potential. Note, however, that $V(\varphi)$ generically does not give preference for a particle / excitation to be at a certain position in space(time), but for the field to have a certain amplitude. Correspondingly, the first term $\sim (\partial_\mu \varphi)(\partial^\mu \varphi)$ is called a kinetic term. Analogously to mechanics, it is a measure for how much action is stored in variations of the field in time <u>and</u> space.

$\mathbb{Z}_2$ symmetry is preserved if the potential satisfies $V(\varphi) = V(-\varphi)$. Considering its Taylor expansion about the origin in field space

$$V(\varphi) = \frac{1}{2} m^2 \varphi^2 + \underbrace{\frac{\lambda}{4!}}_{\text{convention}} \varphi^4 + \dots \tag{1.58}$$

we encounter a quartic term which, on the level of the equations of motion, turns into a cubic interaction,

$$(\Box + m^2)\phi + \frac{\lambda}{3}\phi^3 + \ldots = 0 \qquad (1.59)$$

The parameter λ is dimensionless $[\lambda]=0$ and serves as a measure for the interaction of the field with itself. For small $\lambda \ll 1$, the dispersion relation of small amplitude fluctuations remains essentially unmodified, and we expect approximate plane wave excitations of mass m. For large couplings and/or large amplitudes, the nonlinearity will lead to sizable modifications both of the wave form as well as the dispersion relation.

We close this section by listing the actions that give rise to the field equations discussed in the previous section:

1) Maxwell's electrodynamics:

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu}F^{\mu\nu} - J_\mu A^\mu \qquad (1.60)$$

in presence of a current J_μ. The signs are chosen such that the above given conventions are met.

2) Klein-Gordon theory for a complex field $\phi \in \mathbb{C}$:

$$\mathcal{L} = (\partial_\mu \phi^*)(\partial^\mu \phi) - m^2 \phi^* \phi \qquad (1.61)$$

with the decomposition into two real fields

$$\phi = \frac{1}{\sqrt{2}} (\phi_1 + i \phi_2) \quad , \phi_{1/2} \in \mathbb{R} , \qquad (1.62)$$

(1.61) splits into two copies of (1.51).

3) Schrödinger theory for $\psi(x) \in \mathbb{C}$:

$$\mathcal{L} = \psi^* i\partial_t \psi - \frac{1}{2m} (\vec{\nabla} \psi^*)(\vec{\nabla} \psi) - V(x) \psi^* \psi \qquad (1.63)$$

The explicit verification of the corresponding field equations is left as an exercise to the reader.

1.4 Functional differenciation

The variational calculus, introducing a variation parameter and an arbitrary variation $\eta(x)$, can be most conveniently formulated in terms of functional differenciation. The latter is a directional derivative of a c-number valued functional taken "into the direction of a function" in function space. Its precise mathematical definition requires a careful discussion of function spaces (see, e.g. Courant, Hilbert '53). For our purposes, it suffices to work with the (mostly) algebraic rules following from its definition (which can equally well be worked out from the variational calculus used above): a functional derivative is linear.

$$\frac{\delta}{\delta\varphi(x)}\Big(\alpha F_1[\varphi] + \beta F_2[\varphi]\Big) = \alpha\,\frac{\delta F_1[\varphi]}{\delta\varphi(x)} + \beta\,\frac{\delta F_2[\varphi]}{\delta\varphi(x)} \tag{1.64}$$

and obeys a Leipniz rule.

$$\frac{\delta}{\delta\varphi(x)}\big(F_1[\varphi]\,F_2[\varphi]\big) = \frac{\delta F_1[\varphi]}{\delta\varphi(x)}\,F_2[\varphi] + F_1[\varphi]\,\frac{\delta F_2[\varphi]}{\delta\varphi(x)}. \tag{1.65}$$

The "most elementary" nontrivial derivative is

$$\frac{\delta\varphi(y)}{\delta\varphi(x)} = \delta^{(D)}(y-x) \tag{1.66}$$

where D is the number of space time dimensions, and $\delta^{(D)}$ is the δ distribution on the considered function space.

With this tool, let us verify that the extrema of the action $S[\varphi]$ satisfy the Euler-Lagrange equations:

$$0 = \frac{\delta S}{\delta \varphi(x)} = \int d^4y \, \frac{\delta}{\delta \varphi(x)} \, \mathcal{L}(\varphi, \partial_\mu \varphi; y)$$

$\mathcal{L}$ is evaluated at y

$$= \int d^4y \left(\frac{\delta \varphi(y)}{\delta \varphi(x)} \frac{\partial \mathcal{L}}{\partial \varphi}(y) + \frac{\delta \partial_\mu^y \varphi(y)}{\delta \varphi(x)} \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)}(y) \right)$$

$$= \int d^4y \left(\delta^{(4)}(y-x) \frac{\partial \mathcal{L}}{\partial \varphi}(y) + \partial_\mu^y \delta^{(4)}(y-x) \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)}(y) \right)$$

$$\stackrel{\text{i.b.p.}}{=} \int d^4y \left[\delta^{(4)}(y-x) \left(\frac{\partial \mathcal{L}}{\partial \varphi}(y) - \partial_\mu^y \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)}(y) \right) \right]$$

$$= \frac{\partial \mathcal{L}}{\partial \varphi(x)} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi(x))} \, . \qquad (1.67)$$

Note that $\mathcal{L}$ is a function of the field and its derivatives and thus only partial derivatives of $\mathcal{L}$ have to be evaluated. The surface term of the partial integration (i.b.p.) does not contribute for obvious reasons as long as x is not on the boundary of the integration volume. If it was, the functional directional derivative would correspond to a change or variation of the boundary conditions imposed on the fields, which we do not want to consider here. This restriction is equivalent to choosing $\eta(x)\big|_{\partial V} = 0$ in the variational calculus.

Further examples of functional differenciation are discussed in the exercises.