General U(N) gauge transformations in the realm of covariant Hamiltonian field theory

Jürgen Struckmeier and Hermine Reichau

Abstract A consistent, local coordinate formulation of covariant Hamiltonian field theory is presented. While the covariant canonical field equations are equivalent to the Euler-Lagrange field equations, the covariant canonical transformation theory offers more general means for defining mappings that preserve the action functional — and hence the form of the field equations — than the usual Lagrangian description. Similar to the well-known canonical transformation theory of point dynamics, the canonical transformation rules for fields are derived from generating functions. As an interesting example, we work out the generating function of type F_2 of a general local U(N) gauge transformation and thus derive the most general form of a Hamiltonian density $\mathscr{H}_3$ that is *form-invariant* under *local* U(N) gauge transformations. As a result, a generalized gauge-invariant Dirac-Lagrangian $\mathscr{L}_3$ is obtained that includes the description of Pauli-coupling of an N-tuple of fermions with the set of bosonic gauge fields.

1 Covariant Hamiltonian density

In field theory, the usual definition of a Hamiltonian density emerges from a Legendre transformation of a Lagrangian density $\mathscr{L}$ that only maps the time derivative $\partial_t \phi$ of a field $\phi(t,x,y,z)$ into a corresponding canonical momentum variable, π_t. Taking then the spatial integrals, we obtain a description of the field dynamics that corresponds to that of point dynamics. In contrast, a fully covariant Hamiltonian description treats space and time variables on equal footing[1, 2]. If $\mathscr{L}$ is a Lorentz scalar, this property is passed to the *covariant Hamiltonian*. Moreover, this descrip-

Jürgen Struckmeier
Frankfurt Institute for Advanced Studies (FIAS)
Ruth-Moufang-Str. 1, 60438 Frankfurt am Main, Germany
e-mail: `struckmeier@fias.uni-frankfurt.de`

tion enables us to derive a consistent theory of canonical transformations in the realm of classical field theory.

1.1 Covariant canonical field equations

The transition from particle dynamics to the dynamics of a *continuous* system is based on the assumption that a *continuum limit* exists for the given physical problem[3]. This limit is defined by letting the number of particles involved in the system increase over all bounds while letting their masses and distances go to zero. In this limit, the information on the location of individual particles is replaced by the *value* of a smooth function $\phi(x)$ that is given at a spatial location x^1, x^2, x^3 at time $t \equiv x^0$. The differentiable function $\phi(x)$ is called a *field*. In this notation, the index μ runs from 0 to 3, hence distinguishes the four independent variables of space-time $x^\mu \equiv (x^0, x^1, x^2, x^3) \equiv (t, x, y, z)$, and $x_\mu \equiv (x_0, x_1, x_2, x_3) \equiv (t, -x, -y, -z)$. We furthermore assume that the given physical problem can be described in terms of a set of $I = 1, \ldots, N$ — possibly interacting — scalar fields $\phi_I(x)$ or vector fields $\boldsymbol{A}_I = (A_I^0, A_I^1, A_I^2, A_I^3)$, with the index "$I$" enumerating the individual fields. In order to clearly distinguish scalar quantities from vector quantities, we denote the latter with boldface letters. Throughout the article, the summation convention is used. Whenever no confusion can arise, we omit the indexes in the argument list of functions in order to avoid the number of indexes to proliferate.

The Lagrangian description of the dynamics of a continuous system is based on the Lagrangian density function $\mathscr{L}$ that is supposed to carry the complete information on the given physical system. In a first-order field theory, the Lagrangian density $\mathscr{L}$ is defined to depend on the ϕ_I, possibly on the vector of independent variables x, and on the four first derivatives of the fields ϕ_I with respect to the independent variables, i.e., on the 1-forms (covectors)

$$\boldsymbol{\partial}\phi_I \equiv (\partial_t\phi_I, \partial_x\phi_I, \partial_y\phi_I, \partial_z\phi_I).$$

The Euler-Lagrange field equations are then obtained as the zero of the variation δS of the action integral

$$S = \int \mathscr{L}(\phi_I, \boldsymbol{\partial}\phi_I, x)\, d^4x \tag{1}$$

as[3]

$$\frac{\partial}{\partial x^\alpha} \frac{\partial \mathscr{L}}{\partial(\partial_\alpha \phi_I)} - \frac{\partial \mathscr{L}}{\partial \phi_I} = 0. \tag{2}$$

To derive the equivalent *covariant* Hamiltonian description of continuum dynamics, we first define for each field $\phi_I(x)$ a 4-vector of conjugate momentum fields $\pi_I^\mu(x)$. Its components are given by

$$\pi_I^\mu = \frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi_I)} \equiv \frac{\partial \mathscr{L}}{\partial\left(\frac{\partial \phi_I}{\partial x^\mu}\right)}. \tag{3}$$

The 4-vector $\boldsymbol{\pi}_I$ is thus induced by the Lagrangian $\mathscr{L}$ as the *dual counterpart* of the 1-form $\boldsymbol{\partial}\phi_I$. For the entire set of N scalar fields $\phi_I(x)$, this establishes a set of N conjugate 4-vector fields. With this definition of the 4-vectors of canonical momenta $\boldsymbol{\pi}_I(x)$, we can now define the Hamiltonian density $\mathscr{H}(\phi_I,\boldsymbol{\pi}_I,x)$ as the covariant Legendre transform of the Lagrangian density $\mathscr{L}(\phi_I,\boldsymbol{\partial}\phi_I,x)$

$$\mathscr{H}(\phi_I,\boldsymbol{\pi}_I,x) = \pi_J^\alpha \frac{\partial \phi_J}{\partial x^\alpha} - \mathscr{L}(\phi_I,\boldsymbol{\partial}\phi_I,x). \tag{4}$$

In order for the Hamiltonian $\mathscr{H}$ to be valid, we must require the Legendre transformation to be *regular*, which means that for each index "I" the Hesse matrices $(\partial^2\mathscr{L}/\partial(\partial^\mu\phi_I)\,\partial(\partial_\nu\phi_I))$ are non-singular. This ensures that by means of the Legendre transformation, the Hamiltonian $\mathscr{H}$ takes over the complete information on the given dynamical system from the Lagrangian $\mathscr{L}$. The definition of $\mathscr{H}$ by Eq. (4) is referred to in literature as the "De Donder-Weyl" Hamiltonian density.

Obviously, the dependencies of $\mathscr{H}$ and $\mathscr{L}$ on the ϕ_I and the x^μ only differ by a sign,

$$\left.\frac{\partial \mathscr{H}}{\partial x^\mu}\right|_{\text{expl}} = -\left.\frac{\partial \mathscr{L}}{\partial x^\mu}\right|_{\text{expl}}, \qquad \frac{\partial \mathscr{H}}{\partial \phi_I} = -\frac{\partial \mathscr{L}}{\partial \phi_I} = -\frac{\partial}{\partial x^\alpha}\frac{\partial \mathscr{L}}{\partial(\partial_\alpha \phi_I)} = -\frac{\partial \pi_I^\alpha}{\partial x^\alpha}.$$

These variables thus do not take part in the Legendre transformation of Eqs. (3), (4). Thus, with respect to this transformation, the Lagrangian density $\mathscr{L}$ represents a function of the $\partial_\mu\phi_I$ only and does *not depend* on the canonical momenta π_I^μ, whereas the Hamiltonian density $\mathscr{H}$ is to be considered as a function of the π_I^μ only and does not depend on the derivatives $\partial_\mu\phi_I$ of the fields. In order to derive the second canonical field equation, we calculate from Eq. (4) the partial derivative of $\mathscr{H}$ with respect to π_I^μ,

$$\frac{\partial \mathscr{H}}{\partial \pi_I^\mu} = \delta_{IJ}\,\delta_\mu^\alpha\,\frac{\partial \phi_J}{\partial x^\alpha} = \frac{\partial \phi_I}{\partial x^\mu} \qquad \Longleftrightarrow \qquad \frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi_I)} = \pi_J^\alpha\,\delta_{JI}\,\delta_\alpha^\mu = \pi_I^\mu.$$

The complete set of covariant canonical field equations is thus given by

$$\frac{\partial \mathscr{H}}{\partial \pi_I^\mu} = \frac{\partial \phi_I}{\partial x^\mu}, \qquad \frac{\partial \mathscr{H}}{\partial \phi_I} = -\frac{\partial \pi_I^\alpha}{\partial x^\alpha}. \tag{5}$$

This pair of first-order partial differential equations is equivalent to the set of second-order differential equations of Eq. (2). We observe that in this formulation of the canonical field equations, all coordinates of space-time appear symmetrically — similar to the Lagrangian formulation of Eq. (2). Provided that the Lagrangian density $\mathscr{L}$ is a Lorentz scalar, the dynamics of the fields is invariant with respect to Lorentz transformations. The covariant Legendre transformation (4) passes this property to the Hamiltonian density $\mathscr{H}$. It thus ensures *a priori* the relativistic invariance of the fields that emerge as integrals of the canonical field equations if $\mathscr{L}$ — and hence $\mathscr{H}$ — represents a Lorentz scalar.

2 Canonical transformations in covariant Hamiltonian field theory

The covariant Legendre transformation (4) allows us to derive a canonical transformation theory in a way similar to that of point dynamics. The main difference is that now the generating function of the canonical transformation is represented by a *vector* rather than by a scalar function. The main benefit of this formalism is that we are not dealing with plain transformations. Instead, we restrict ourselves *right from the beginning* to those transformations that preserve the form of the action functional. This ensures all eligible transformations to be *physical*. Furthermore, with a generating function, we not only define the transformations of the fields but also pinpoint simultaneously the corresponding transformation law of the canonical momentum fields.

2.1 Generating functions of type $\boldsymbol{F}_1(\boldsymbol{\phi},\boldsymbol{\Phi},x)$

Similar to the canonical formalism of point mechanics, we call a transformation of the fields $(\boldsymbol{\phi},\boldsymbol{\pi})\mapsto(\boldsymbol{\Phi},\boldsymbol{\Pi})$ *canonical* if the form of the variational principle that is based on the action functional (1) is maintained,

$$\delta\int_R\left(\pi_I^\alpha\frac{\partial\phi_I}{\partial x^\alpha}-\mathscr{H}(\boldsymbol{\phi},\boldsymbol{\pi},x)\right)d^4x\stackrel{!}{=}\delta\int_R\left(\Pi_I^\alpha\frac{\partial\Phi_I}{\partial x^\alpha}-\mathscr{H}'(\boldsymbol{\Phi},\boldsymbol{\Pi},x)\right)d^4x. \tag{6}$$

Equation (6) tells us that the *integrands* may differ by the divergence of a vector field F_1^μ, whose variation vanishes on the boundary ∂R of the integration region R within space-time

$$\delta\int_R\frac{\partial F_1^\alpha}{\partial x^\alpha}d^4x=\delta\oint_{\partial R}F_1^\alpha dS_\alpha\stackrel{!}{=}0.$$

The immediate consequence of the form invariance of the variational principle is the form invariance of the covariant canonical field equations (5)

$$\frac{\partial\mathscr{H}'}{\partial\Pi_I^\mu}=\frac{\partial\Phi_I}{\partial x^\mu},\qquad\frac{\partial\mathscr{H}'}{\partial\Phi_I}=-\frac{\partial\Pi_I^\alpha}{\partial x^\alpha}.$$

For the integrands of Eq. (6) — hence for the Lagrangian densities $\mathscr{L}$ and $\mathscr{L}'$ — we thus obtain the condition

$$\mathscr{L}=\mathscr{L}'+\frac{\partial F_1^\alpha}{\partial x^\alpha}$$
$$\pi_I^\alpha\frac{\partial\phi_I}{\partial x^\alpha}-\mathscr{H}(\boldsymbol{\phi},\boldsymbol{\pi},x)=\Pi_I^\alpha\frac{\partial\Phi_I}{\partial x^\alpha}-\mathscr{H}'(\boldsymbol{\Phi},\boldsymbol{\Pi},x)+\frac{\partial F_1^\alpha}{\partial x^\alpha}. \tag{7}$$

With the definition $F_1^\mu \equiv F_1^\mu(\boldsymbol{\phi}, \boldsymbol{\Phi}, x)$, we restrict ourselves to a function of exactly those arguments that now enter into transformation rules for the transition from the original to the new fields. The divergence of F_1^μ writes, explicitly,

$$\frac{\partial F_1^\alpha}{\partial x^\alpha} = \frac{\partial F_1^\alpha}{\partial \phi_I}\frac{\partial \phi_I}{\partial x^\alpha} + \frac{\partial F_1^\alpha}{\partial \Phi_I}\frac{\partial \Phi_I}{\partial x^\alpha} + \left.\frac{\partial F_1^\alpha}{\partial x^\alpha}\right|_{\text{expl}}. \tag{8}$$

The rightmost term denotes the sum over the *explicit* dependence of the generating function F_1^μ on the x^ν. Comparing the coefficients of Eqs. (7) and (8), we find the local coordinate representation of the field transformation rules that are induced by the generating function F_1^μ

$$\pi_I^\mu = \frac{\partial F_1^\mu}{\partial \phi_I}, \qquad \Pi_I^\mu = -\frac{\partial F_1^\mu}{\partial \Phi_I}, \qquad \mathscr{H}' = \mathscr{H} + \left.\frac{\partial F_1^\alpha}{\partial x^\alpha}\right|_{\text{expl}}. \tag{9}$$

The transformation rule for the Hamiltonian density implies that summation over α is to be performed. In contrast to the transformation rule for the Lagrangian density $\mathscr{L}$ of Eq. (7), the rule for the Hamiltonian density is determined by the *explicit* dependence of the generating function F_1^μ on the x^ν. Hence, if a generating function does not explicitly depend on the independent variables, x^ν, then the *value* of the Hamiltonian density is not changed under the particular canonical transformation emerging thereof.

Differentiating the transformation rule for π_I^μ with respect to Φ_J, and the rule for Π_J^μ with respect to ϕ_I, we obtain a symmetry relation between original and transformed fields

$$\frac{\partial \pi_I^\mu}{\partial \Phi_J} = \frac{\partial^2 F_1^\mu}{\partial \phi_I \partial \Phi_J} = -\frac{\partial \Pi_J^\mu}{\partial \phi_I}.$$

The emerging of symmetry relations is a characteristic feature of *canonical* transformations. As the symmetry relation directly follows from the second derivatives of the generating function, is does not apply for arbitrary transformations of the fields that do not follow from generating functions.

2.2 Generating functions of type $\boldsymbol{F}_2(\boldsymbol{\phi}, \boldsymbol{\Pi}, x)$

The generating function of a canonical transformation can alternatively be expressed in terms of a function of the original fields ϕ_I and of the new *conjugate* fields Π_I^μ. To derive the pertaining transformation rules, we perform the covariant Legendre transformation

$$F_2^\mu(\boldsymbol{\phi}, \boldsymbol{\Pi}, x) = F_1^\mu(\boldsymbol{\phi}, \boldsymbol{\Phi}, x) + \Phi_J \Pi_J^\mu, \qquad \Pi_I^\mu = -\frac{\partial F_1^\mu}{\partial \Phi_I}. \tag{10}$$

By definition, the functions F_1^μ and F_2^μ agree with respect to their ϕ_I and x^μ dependencies

$$\frac{\partial F_2^\mu}{\partial \phi_I} = \frac{\partial F_1^\mu}{\partial \phi_I} = \pi_I^\mu, \qquad \left.\frac{\partial F_2^\alpha}{\partial x^\alpha}\right|_{\text{expl}} = \left.\frac{\partial F_1^\alpha}{\partial x^\alpha}\right|_{\text{expl}} = \mathscr{H}' - \mathscr{H}.$$

The variables ϕ_I and x^μ thus do not take part in the Legendre transformation from Eq. (10). Therefore, the two F_2^μ-related transformation rules coincide with the respective rules derived previously from F_1^μ. As F_1^μ does not depend on the Π_I^μ whereas F_2^μ does not depend on the the Φ_I, the new transformation rule thus follows from the derivative of F_2^μ with respect to Π_J^ν as

$$\frac{\partial F_2^\mu}{\partial \Pi_I^\nu} = \Phi_J \frac{\partial \Pi_J^\mu}{\partial \Pi_I^\nu} = \Phi_J\, \delta_{IJ}\, \delta_\nu^\mu .$$

We thus end up with set of transformation rules

$$\pi_I^\mu = \frac{\partial F_2^\mu}{\partial \phi_I}, \qquad \Phi_I\, \delta_\nu^\mu = \frac{\partial F_2^\mu}{\partial \Pi_I^\nu}, \qquad \mathscr{H}' = \mathscr{H} + \left.\frac{\partial F_2^\alpha}{\partial x^\alpha}\right|_{\text{expl}}, \tag{11}$$

which is equivalent to the set (9) by virtue of the Legendre transformation (10) if the matrices $(\partial^2 F_1^\mu / \partial \phi_I \partial \Phi_J)$ are non-singular for all indexes "μ". From the second partial derivations of F_2^μ one immediately derives the symmetry relation

$$\frac{\partial \pi_I^\mu}{\partial \Pi_J^\nu} = \frac{\partial^2 F_2^\mu}{\partial \phi_I \partial \Pi_J^\nu} = \frac{\partial \Phi_J}{\partial \phi_I}\, \delta_\nu^\mu ,$$

whose existence characterizes the transformation to be canonical.

3 Examples for Hamiltonian densities in covariant field theory

We present some simple examples Hamiltonian densities as they emerge from Lagrangian densities of classical Lagrangian field theory. It is shown that resulting canonical field equations are equivalent to the corresponding Euler-Lagrange equations.

3.1 Klein-Gordon Hamiltonian density for complex fields

We first consider the Klein-Gordon *Lagrangian density* $\mathscr{L}_{\text{KG}}$ for a *complex* scalar field ϕ that is associated with mass m (see, for instance, Ref. [4]):

$$\mathscr{L}_{\mathrm{KG}}\left(\phi,\phi^*,\partial^\mu\phi,\partial_\mu\phi^*\right)=\frac{\partial\phi^*}{\partial x^\alpha}\frac{\partial\phi}{\partial x_\alpha}-m^2\,\phi^*\phi.$$

Herein, ϕ^* denotes complex conjugate field of ϕ. Both quantities are to be treated as independent. With $[L]$ denoting the dimension of "length," we have with $\hbar=c=1$, i.e. in "natural units", $[\mathscr{L}]=[L]^{-4}$, $[m]=[L]^{-1}$, and $[\partial_\mu]=[L]^{-1}$ so that $[\phi]=[L]^{-1}$. The Euler-Lagrange equations (2) for ϕ and ϕ^* follow from this Lagrangian density as

$$\frac{\partial^2}{\partial x_\alpha\partial x^\alpha}\phi^*=-m^2\,\phi^*,\qquad\frac{\partial^2}{\partial x_\alpha\partial x^\alpha}\phi=-m^2\,\phi.\tag{12}$$

As a prerequisite for deriving the corresponding Hamiltonian density $\mathscr{H}_{\mathrm{KG}}$ we must first define from $\mathscr{L}_{\mathrm{KG}}$ the conjugate momentum fields,

$$\pi^\mu=\frac{\partial\mathscr{L}_{\mathrm{KG}}}{\partial\left(\partial_\mu\phi^*\right)}=\frac{\partial\phi}{\partial x_\mu},\qquad\pi^*_\mu=\frac{\partial\mathscr{L}_{\mathrm{KG}}}{\partial(\partial^\mu\phi)}=\frac{\partial\phi^*}{\partial x^\mu},$$

which means that $[\pi^\mu]=[L]^{-2}$. The determinant of the Hesse matrix does not vanish for the actual Lagrangian $\mathscr{L}_{\mathrm{KG}}$ since

$$\det\left(\frac{\partial^2\mathscr{L}_{\mathrm{KG}}}{\partial(\partial^\mu\phi)\partial\left(\partial_\nu\phi^*\right)}\right)=\det\left(\frac{\partial\pi^*_\mu}{\partial\left(\partial_\nu\phi^*\right)}\right)=\det\left(\delta_\mu^\nu\right)=1.$$

This condition is always satisfied if the Lagrangian density $\mathscr{L}$ is *quadratic* in the derivatives of the fields. The Hamiltonian density $\mathscr{H}$ then follows as the Legendre transform of the Lagrangian density

$$\mathscr{H}(\pi^\mu,\pi^*_\mu,\phi,\phi^*)=\pi^*_\alpha\frac{\partial\phi}{\partial x_\alpha}+\frac{\partial\phi^*}{\partial x^\alpha}\pi^\alpha-\mathscr{L}(\partial^\mu\phi,\partial_\mu\phi^*,\phi,\phi^*),$$

thus $[\mathscr{H}]=[\mathscr{L}]=[L]^{-4}$. The Klein-Gordon *Hamiltonian density* $\mathscr{H}_{\mathrm{KG}}$ is then given by

$$\mathscr{H}_{\mathrm{KG}}(\pi_\mu,\pi^*_\mu,\phi,\phi^*)=\pi^*_\alpha\pi^\alpha+m^2\,\phi^*\phi.\tag{13}$$

For the Hamiltonian density (13), the canonical field equations (5) provide the following set of coupled first order partial differential equations

$$\frac{\partial\phi^*}{\partial x^\mu}=\frac{\partial\mathscr{H}_{\mathrm{KG}}}{\partial\pi^\mu}=\pi^*_\mu,\qquad\qquad\frac{\partial\phi}{\partial x_\mu}=\frac{\partial\mathscr{H}_{\mathrm{KG}}}{\partial\pi^*_\mu}=\pi^\mu$$

$$-\frac{\partial\pi^*_\alpha}{\partial x_\alpha}=\frac{\partial\mathscr{H}_{\mathrm{KG}}}{\partial\phi}=m^2\phi^*,\qquad\qquad-\frac{\partial\pi^\alpha}{\partial x^\alpha}=\frac{\partial\mathscr{H}_{\mathrm{KG}}}{\partial\phi^*}=m^2\phi.$$

In the first row, the canonical field equations for the scalar fields ϕ and ϕ^* reproduce the definitions of the momentum fields π^μ and π^*_μ from the Lagrangian density $\mathscr{L}_{\mathrm{KG}}$. Eliminating the π^μ, π^*_μ from the canonical field equations then yields the Euler-Lagrange equations of Eq. (12).

3.2 Maxwell's equations as canonical field equations

The Lagrangian density $\mathscr{L}_{\mathrm{M}}$ of the electromagnetic field is given by

$$\mathscr{L}_{\mathrm{M}}(\boldsymbol{a},\boldsymbol{\partial a},x) = -\tfrac{1}{4} f_{\alpha\beta} f^{\alpha\beta} - j^{\alpha}(x)\, a_{\alpha}, \quad f_{\mu\nu} = \frac{\partial a_{\nu}}{\partial x^{\mu}} - \frac{\partial a_{\mu}}{\partial x^{\nu}}. \tag{14}$$

Herein, the four components a^{μ} of the 4-vector potential $\boldsymbol{a}$ now take the place of the scalar fields $\phi_I \equiv a^{\mu}$ in the notation used so far. The Lagrangian density (14) thus entails a set of *four* Euler-Lagrange equations, i.e., an equation for each component a_{μ}. The source vector $\boldsymbol{j} = (\rho, j_x, j_y, j_z)$ denotes the 4-vector of electric currents combining the usual current density vector (j_x, j_y, j_z) of configuration space with the charge density ρ. In a local Lorentz frame, i.e., in Minkowski space, the Euler-Lagrange equations (2) take on the form,

$$\frac{\partial}{\partial x^{\alpha}} \frac{\partial \mathscr{L}_{\mathrm{M}}}{\partial(\partial_{\alpha} a_{\mu})} - \frac{\partial \mathscr{L}_{\mathrm{M}}}{\partial a_{\mu}} = 0, \qquad \mu = 0, \ldots, 3. \tag{15}$$

With $\mathscr{L}_{\mathrm{M}}$ from Eq. (14), we obtain directly

$$\frac{\partial f^{\mu\alpha}}{\partial x^{\alpha}} + j^{\mu} = 0. \tag{16}$$

In Minkowski space, this is the tensor form of the inhomogeneous Maxwell equation. In order to formulate the equivalent Hamiltonian description, we first define, according to Eq. (3), the canonically field components $p^{\mu\nu}$ as the conjugate objects of the derivatives of the 4-vector potential $\boldsymbol{a}$

$$p^{\mu\nu} = \frac{\partial \mathscr{L}_{\mathrm{M}}}{\partial(\partial_{\nu} a_{\mu})} \equiv \frac{\partial \mathscr{L}_{\mathrm{M}}}{\partial a_{\mu,\nu}} \tag{17}$$

With the particular Lagrangian density (14), Eq. (17) means

$$\begin{aligned}
f_{\alpha\beta} &= a_{\beta,\alpha} - a_{\alpha,\beta} \\
p^{\mu\nu} &= -\tfrac{1}{4} \left(\frac{\partial f_{\alpha\beta}}{\partial a_{\mu,\nu}} f^{\alpha\beta} + \frac{\partial f^{\alpha\beta}}{\partial a_{\mu,\nu}} f_{\alpha\beta} \right) = -\tfrac{1}{2} \frac{\partial f_{\alpha\beta}}{\partial a_{\mu,\nu}} f^{\alpha\beta} \\
&= -\tfrac{1}{2} \left(\delta_{\beta}^{\mu} \delta_{\alpha}^{\nu} - \delta_{\alpha}^{\mu} \delta_{\beta}^{\nu} \right) f^{\alpha\beta} = \tfrac{1}{2} (f^{\mu\nu} - f^{\nu\mu}) \\
&= f^{\mu\nu}.
\end{aligned}$$

The tensor $p^{\mu\nu}$ thus matches exactly the electromagnetic field tensor $f^{\mu\nu}$ from Eq. (14) and hence inherits the skew-symmetry of $f^{\mu\nu}$ because of the particular dependence of $\mathscr{L}_{\mathrm{M}}$ on the $a_{\mu,\nu} \equiv \partial a_{\mu} / \partial x^{\nu}$.

As the Lagrangian density (14) now describes the dynamics of a *vector field*, a_{μ}, rather than a set of scalar fields ϕ_I, the canonical momenta $p^{\mu\nu}$ now constitute a second rank *tensor* rather than a vector. The Legendre transformation corresponding to

Eq. (4) then comprises the product $p^{\alpha\beta}\partial_\beta a_\alpha$. The skew-symmetry of the momentum tensor $p^{\mu\nu}$ picks out the skew-symmetric part of $\partial_\nu a_\mu$ as the symmetric part of $\partial_\nu a_\mu$ vanishes identically calculating the product $p^{\alpha\beta}\partial_\beta a_\alpha$

$$p^{\alpha\beta}\frac{\partial a_\alpha}{\partial x^\beta} = \tfrac{1}{2}p^{\alpha\beta}\underbrace{\left(\frac{\partial a_\alpha}{\partial x^\beta} - \frac{\partial a_\beta}{\partial x^\alpha}\right)}_{=f_{\beta\alpha}} + \tfrac{1}{2}\underbrace{p^{\alpha\beta}\left(\frac{\partial a_\alpha}{\partial x^\beta} + \frac{\partial a_\beta}{\partial x^\alpha}\right)}_{\equiv 0}.$$

For a skew-symmetric momentum tensor $p^{\mu\nu}$, we thus obtain the Hamiltonian density $\mathscr{H}_\mathrm{M}$ as the Legendre-transformed Lagrangian density $\mathscr{L}_\mathrm{M}$

$$\mathscr{H}_\mathrm{M}(\boldsymbol{a},\boldsymbol{p},x) = \tfrac{1}{2}p^{\alpha\beta}f_{\alpha\beta} - \mathscr{L}_\mathrm{M}(\boldsymbol{a},\boldsymbol{\partial a},x).$$

From this (non-standard) Legendre transformation prescription and the corresponding Euler-Lagrange equations (15), the canonical field equations are immediately obtained as

$$\begin{aligned}
\frac{\partial \mathscr{H}_\mathrm{M}}{\partial p^{\mu\nu}} &= -\frac{1}{2}f_{\mu\nu} = \frac{1}{2}\left(\frac{\partial a_\mu}{\partial x^\nu} - \frac{\partial a_\nu}{\partial x^\mu}\right)\\
\frac{\partial \mathscr{H}_\mathrm{M}}{\partial a_\mu} &= -\frac{\partial \mathscr{L}_\mathrm{M}}{\partial a_\mu} = -\frac{\partial}{\partial x^\alpha}\frac{\partial \mathscr{L}_\mathrm{M}}{\partial(\partial_\alpha a_\mu)} = -\frac{\partial p^{\mu\alpha}}{\partial x^\alpha}\\
\frac{\partial \mathscr{H}_\mathrm{M}}{\partial x^\nu} &= -\frac{\partial \mathscr{L}_\mathrm{M}}{\partial x^\nu}.
\end{aligned}$$

The Hamiltonian density for the Lagrangian density (14) follows as

$$\begin{aligned}
\mathscr{H}_\mathrm{M}(\boldsymbol{a},\boldsymbol{p},x) &= -\tfrac{1}{2}p^{\alpha\beta}p_{\alpha\beta} + \tfrac{1}{4}p^{\alpha\beta}p_{\alpha\beta} + j^\alpha(x)\,a_\alpha\\
&= -\tfrac{1}{4}p^{\alpha\beta}p_{\alpha\beta} + j^\alpha(x)\,a_\alpha.
\end{aligned} \tag{18}$$

The first canonical field equation follows from the derivative of the Hamiltonian density (18) with respect to $p^{\mu\nu}$ and $p_{\mu\nu}$

$$\frac{1}{2}\left(\frac{\partial a_\mu}{\partial x^\nu} - \frac{\partial a_\nu}{\partial x^\mu}\right) = \frac{\partial \mathscr{H}_\mathrm{M}}{\partial p^{\mu\nu}} = -\tfrac{1}{2}p_{\mu\nu}, \quad \frac{1}{2}\left(\frac{\partial a^\mu}{\partial x_\nu} - \frac{\partial a^\nu}{\partial x_\mu}\right) = \frac{\partial \mathscr{H}_\mathrm{M}}{\partial p_{\mu\nu}} = -\tfrac{1}{2}p^{\mu\nu}, \tag{19}$$

which reproduces the definition of $p_{\mu\nu}$ and $p^{\mu\nu}$ from Eq. (17).

The second canonical field equation is obtained calculating the derivative of the Hamiltonian density (18) with respect to a_μ

$$-\frac{\partial p^{\mu\alpha}}{\partial x^\alpha} = \frac{\partial \mathscr{H}_\mathrm{M}}{\partial a_\mu} = j^\mu.$$

Inserting the first canonical equation, the second order field equation for the a_μ is thus obtained for the Maxwell Hamiltonian density (18) as

$$\frac{\partial f^{\mu\alpha}}{\partial x^{\alpha}} + j^{\mu} = 0,$$

which agrees, as expected, with the corresponding Euler-Lagrange equation (16).

3.3 The Proca Hamiltonian density

In relativistic quantum field theory, the dynamics of particles of spin 1 and mass m is derived from the Proca Lagrangian density $\mathscr{L}_{\mathrm{P}}$,

$$\mathscr{L}_{\mathrm{P}} = -\tfrac{1}{4} f^{\alpha\beta} f_{\alpha\beta} + \tfrac{1}{2} m^2 a^{\alpha} a_{\alpha}, \qquad f_{\mu\nu} = \frac{\partial a_{\nu}}{\partial x^{\mu}} - \frac{\partial a_{\mu}}{\partial x^{\nu}}.$$

We observe that the kinetic term of $\mathscr{L}_{\mathrm{P}}$ agrees with that of the Lagrangian density $\mathscr{L}_{\mathrm{M}}$ of the electromagnetic field of Eq. (14). Therefore, the field equations emerging from the Euler-Lagrange equations (15) are similar to those of Eq. (16)

$$\frac{\partial f^{\mu\alpha}}{\partial x^{\alpha}} - m^2 a^{\mu} = 0. \tag{20}$$

Thus $[\mathscr{L}] = [L]^{-4}$, $[m] = [L]^{-1}$, and $[\partial_{\mu}] = [L]^{-1}$ entails a dimension of the 4-vector fields $[\boldsymbol{a}] = [L]^{-1}$ and $[\boldsymbol{f}] = [L]^{-2}$ in natural units. The transition to the corresponding Hamilton description is performed by defining on the basis of the actual Lagrangian $\mathscr{L}_{\mathrm{P}}$ the canonical momentum field tensors $p^{\mu\nu}$ as the conjugate objects of the derivatives of the 4-vector potential $\boldsymbol{a}$

$$p^{\mu\nu} = \frac{\partial \mathscr{L}_{\mathrm{P}}}{\partial \left(\partial_{\nu} a_{\mu}\right)} \equiv \frac{\partial \mathscr{L}_{\mathrm{P}}}{\partial a_{\mu,\nu}}.$$

Similar to the preceding section, we find

$$p^{\mu\nu} = f^{\mu\nu}, \qquad p_{\mu\nu} = f_{\mu\nu}, \qquad [\boldsymbol{p}] = [\boldsymbol{f}] = [L]^{-2},$$

because of the particular dependence of $\mathscr{L}_{\mathrm{P}}$ on the derivatives of the a^{μ}. With $p^{\alpha\beta}$ being skew-symmetric in α, β, the product $p^{\alpha\beta} a_{\alpha,\beta}$ picks out the skew-symmetric part of the partial derivative $\partial a_{\alpha}/\partial x^{\beta}$ as the product with the symmetric part vanishes identically. Denoting the skew-symmetric part by $a_{[\alpha,\beta]}$, the Legendre transformation prescription

$$\begin{aligned}\mathscr{H}_{\mathrm{P}} &= p^{\alpha\beta} a_{\alpha,\beta} - \mathscr{L}_{\mathrm{P}} = p^{\alpha\beta} a_{[\alpha,\beta]} - \mathscr{L}_{\mathrm{P}} \\ &= \tfrac{1}{2} p^{\alpha\beta} \left(\frac{\partial a_{\alpha}}{\partial x^{\beta}} - \frac{\partial a_{\beta}}{\partial x^{\alpha}} \right) - \mathscr{L}_{\mathrm{P}},\end{aligned}$$

leads to the Proca Hamiltonian density by following the path of Eq. (18)

$$\mathscr{H}_{\mathrm{P}} = -\tfrac{1}{4} p^{\alpha\beta} p_{\alpha\beta} - \tfrac{1}{2} m^2 a^{\alpha} a_{\alpha}. \tag{21}$$

The canonical field equations emerge as

$$\begin{aligned} a_{[\mu,\nu]} \equiv \frac{1}{2}\left(\frac{\partial a_\mu}{\partial x^\nu} - \frac{\partial a_\nu}{\partial x^\mu}\right) = \frac{\partial \mathscr{H}_{\mathrm{P}}}{\partial p^{\mu\nu}} &= -\tfrac{1}{2} p_{\mu\nu} \\ -\frac{\partial p^{\mu\alpha}}{\partial x^\alpha} = \frac{\partial \mathscr{H}_{\mathrm{P}}}{\partial a_\mu} &= -m^2 a^\mu . \end{aligned}$$

By means of eliminating $p^{\mu\nu}$, this coupled set of first order equations can be converted into second order equations for the vector field $\boldsymbol{a}(x)$,

$$\frac{\partial}{\partial x_\alpha}\left(\frac{\partial a_\mu}{\partial x^\alpha} - \frac{\partial a_\alpha}{\partial x^\mu}\right) - m^2 a_\mu = 0.$$

As expected, this equation coincides with the Euler-Lagrange equation (20).

3.4 The Dirac Hamiltonian density

The dynamics of particles with spin $\frac{1}{2}$ and mass m is described by the Dirac equation. With γ^i, $i = 1,\ldots,4$ denoting the 4×4 Dirac matrices, and ψ a four component Dirac spinor, the Dirac Lagrangian density $\mathscr{L}_{\mathrm{D}}$ is given by

$$\mathscr{L}_{\mathrm{D}} = i\overline{\psi}\gamma^{\alpha}\frac{\partial \psi}{\partial x^{\alpha}} - m\overline{\psi}\psi, \tag{22}$$

wherein $\overline{\psi} \equiv \psi^{\dagger}\gamma^0$ denotes the adjoint spinor of ψ. In the following we summarize some fundamental relations that apply for the Dirac matrices γ^μ, and their duals, γ_μ,

$$
\begin{aligned}
\{\gamma^\mu, \gamma^\nu\} &\equiv \gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2\eta^{\mu\nu} \mathbb{1} \\
\gamma^\alpha \gamma_\alpha &= \gamma_\alpha \gamma^\alpha = 4\,\mathbb{1} \\
[\gamma^\mu, \gamma^\nu] &\equiv \gamma^\mu \gamma^\nu - \gamma^\nu \gamma^\mu \equiv -2i\,\sigma^{\mu\nu} \\
[\gamma_\mu, \gamma_\nu] &\equiv \gamma_\mu \gamma_\nu - \gamma_\nu \gamma_\mu \equiv -2i\,\sigma_{\mu\nu} \\
\det \sigma^{\mu\nu} &= 1, \qquad \mu \neq \nu \\
\tau_{\mu\alpha} \sigma^{\alpha\nu} &= \sigma^{\nu\alpha} \tau_{\alpha\mu} = \delta^\nu_\mu \mathbb{1} \\
\gamma^\alpha \tau_{\alpha\mu} &= \tau_{\mu\alpha} \gamma^\alpha = -\frac{i}{3}\, \gamma_\mu \\
\gamma_\alpha \sigma^{\alpha\mu} &= \sigma^{\mu\alpha} \gamma_\alpha = 3i\, \gamma^\mu \\
\gamma^\alpha \tau_{\alpha\beta} \gamma^\beta &= -\frac{4i}{3} \mathbb{1} \\
\gamma_\alpha \sigma^{\alpha\beta} \gamma_\beta &= 12i\, \mathbb{1}, \qquad \sigma^{\alpha\beta}\, \sigma_{\alpha\beta} = 12\, \mathbb{1} \\
3\tau_{\mu\nu} + \sigma_{\mu\nu} &= 2i\, \eta_{\mu\nu}\, \mathbb{1}.
\end{aligned} \tag{23}
$$

Herein, the symbol $\mathbb{1}$ stands for the 4×4 unit matrix, and the real numbers $\eta^{\mu\nu}, \eta_{\mu\nu} \in \mathbb{R}$ for an element of the Minkowski metric $(\eta^{\mu\nu}) = (\eta_{\mu\nu})$. The matrices $(\sigma^{\mu\nu})$ and $(\tau_{\mu\nu})$ are to be understood as 4×4 block matrices, with each block $\sigma^{\mu\nu}$, $\tau_{\mu\nu}$ representing a 4×4 matrix of complex numbers. Thus, $(\sigma^{\mu\nu})$ and $(\tau_{\mu\nu})$ are actually 16×16 matrices of complex numbers.

Natural units are defined by setting $\hbar = c = 1$. Denoting "the dimension of" by the symbol "[]", we then have for the dimension of the mass m, length L, time T, and energy E

$$
[m] = [L]^{-1} = [T]^{-1} = [E].
$$

Then

$$
[\mathscr{L}_\mathrm{D}] = [L]^{-4}, \qquad [\psi] = [L]^{-3/2}, \qquad [\partial_\mu] = [m] = [L]^{-1}.
$$

The Dirac Lagrangian density $\mathscr{L}_\mathrm{D}$ can be rendered symmetric by combining the Lagrangian density Eq. (22) with its adjoint, which leads to

$$
\mathscr{L}_\mathrm{D} = \frac{i}{2}\left(\overline{\psi} \gamma^\alpha \frac{\partial \psi}{\partial x^\alpha} - \frac{\partial \overline{\psi}}{\partial x^\alpha} \gamma^\alpha \psi \right) - m \overline{\psi} \psi. \tag{24}
$$

The resulting Euler-Lagrange equations are identical to those derived from Eq. (22),

$$
\begin{aligned}
i\gamma^\alpha \frac{\partial \psi}{\partial x^\alpha} - m\psi &= 0 \\
i\frac{\partial \overline{\psi}}{\partial x^\alpha} \gamma^\alpha + m\overline{\psi} &= 0.
\end{aligned} \tag{25}
$$

As both Lagrangians (22) and (24) are *linear* in the derivatives of the fields, the determinant of the Hessian vanishes,

$$\det\left[\frac{\partial^2 \mathscr{L}_{\mathrm{D}}}{\partial\left(\partial_\mu \psi\right)\partial\left(\partial_\nu \overline{\psi}\right)}\right] = 0. \tag{26}$$

Therefore, Legendre transformations of the Lagrangian densities (22) and (24) are irregular. Nevertheless, as a Lagrangian density is determined only up to the divergence of an arbitrary vector function F^μ according to Eq. (7), one can construct an equivalent Lagrangian density $\mathscr{L}_{\mathrm{D}}'$ that yields identical Euler-Lagrange equations while yielding a regular Legendre transformation. The additional term[5] emerges as the divergence of a vector function F^μ, which may be expressed in symmetric form as

$$F^\mu = \frac{i}{6\tilde{m}}\left(\overline{\psi}\,\sigma^{\mu\alpha}\frac{\partial\psi}{\partial x^\alpha} + \frac{\partial\overline{\psi}}{\partial x^\alpha}\sigma^{\alpha\mu}\,\psi\right), \qquad [\boldsymbol{F}] = [L]^{-3}.$$

The “gauge-fixing parameter” $\tilde{m}$ must have the natural dimension of mass in order to match the dimensions correctly. Explicitly, the additional term is given by

$$\begin{aligned}\frac{\partial F^\beta}{\partial x^\beta} &= \frac{i}{6\tilde{m}}\left(\partial_\beta\overline{\psi}\sigma^{\beta\alpha}\partial_\alpha\psi + \overline{\psi}\sigma^{\beta\alpha}\partial_\beta\partial_\alpha\psi + \partial_\beta\partial_\alpha\overline{\psi}\sigma^{\alpha\beta}\psi + \partial_\alpha\overline{\psi}\sigma^{\alpha\beta}\partial_\beta\psi\right)\\ &= \frac{\partial\overline{\psi}}{\partial x^\alpha}\frac{i\sigma^{\alpha\beta}}{3\tilde{m}}\frac{\partial\psi}{\partial x^\beta}.\end{aligned}$$

Note that the double sums $\sigma^{\beta\alpha}\partial_\beta\partial_\alpha\psi$ and $\partial_\beta\partial_\alpha\overline{\psi}\sigma^{\alpha\beta}$ vanish identically, as we sum over a symmetric ($\partial_\mu\partial_\nu\psi = \partial_\nu\partial_\mu\psi$) and a skew-symmetric ($\sigma^{\mu\nu} = -\sigma^{\nu\mu}$) factor. Following Eq. (7), the equivalent Lagrangian density is given by $\mathscr{L}_{\mathrm{D}}' = \mathscr{L}_{\mathrm{D}} + \partial F^\beta/\partial x^\beta$, which means, explicitly,

$$\mathscr{L}_{\mathrm{D}}' = \frac{i}{2}\left(\overline{\psi}\gamma^\alpha\frac{\partial\psi}{\partial x^\alpha} - \frac{\partial\overline{\psi}}{\partial x^\alpha}\gamma^\alpha\psi\right) + \frac{\partial\overline{\psi}}{\partial x^\alpha}\frac{i\sigma^{\alpha\beta}}{3\tilde{m}}\frac{\partial\psi}{\partial x^\beta} - m\,\overline{\psi}\psi. \tag{27}$$

Due to the skew-symmetry of the $\sigma^{\mu\nu}$, the Euler-Lagrange equations (2) for $\mathscr{L}_{\mathrm{D}}'$ yield again the Dirac equations (25). We remark that the regularized Dirac Lagrangian (27) can equivalently be written as

$$\mathscr{L}_{\mathrm{D}}' = \left(\frac{\partial\overline{\psi}}{\partial x^\alpha} - \frac{i\tilde{m}}{2}\overline{\psi}\gamma_\alpha\right)\frac{i\sigma^{\alpha\beta}}{3\tilde{m}}\left(\frac{\partial\psi}{\partial x^\beta} + \frac{i\tilde{m}}{2}\gamma_\beta\psi\right) + (\tilde{m}-m)\,\overline{\psi}\psi.$$

This representation of the Dirac Lagrangian will be recognized as the analogue of the Dirac Hamiltonian $\mathscr{H}_{\mathrm{D}}$ to be derived in Eq. (31).

As desired, the Hessian of $\mathscr{L}_{\mathrm{D}}'$ is not singular,

$$\det\left[\frac{\partial^2\mathscr{L}_{\mathrm{D}}'}{\partial\left(\partial_\mu\overline{\psi}\right)\partial\left(\partial_\nu\psi\right)}\right] = \det\frac{i\sigma^{\mu\nu}}{3\tilde{m}} \neq 0 \quad \text{since} \quad \det\sigma^{\mu\nu} = 1,\ \nu\neq\mu. \tag{28}$$

Thus, the Legendre transformation of the Lagrangian density $\mathscr{L}_{\mathrm{D}}'$ is now *regular*. It is remarkable that it is exactly a term which does *not* contribute to the Euler-Lagrange equations that makes the Legendre transformation of $\mathscr{L}_{\mathrm{D}}'$ *regular* and thus transfers the information on the dynamical system that is contained in the Lagrangian to the Hamiltonian description. The canonical momenta follow as

$$
\begin{aligned}
\overline{\pi}^{\mu} &= \frac{\partial \mathscr{L}_{\mathrm{D}}'}{\partial\left(\partial_{\mu}\psi\right)} = \quad \frac{i}{2}\overline{\psi}\gamma^{\mu} + \frac{\partial\overline{\psi}}{\partial x^{\alpha}}\,\frac{i\sigma^{\alpha\mu}}{3\tilde{m}} \\
\pi^{\mu} &= \frac{\partial \mathscr{L}_{\mathrm{D}}'}{\partial\left(\partial_{\mu}\overline{\psi}\right)} = -\frac{i}{2}\gamma^{\mu}\psi + \frac{i\sigma^{\mu\alpha}}{3\tilde{m}}\,\frac{\partial\psi}{\partial x^{\alpha}},
\end{aligned} \tag{29}
$$

which states that $[\pi^{\mu}] = [\psi] = [L]^{-3/2}$. The Legendre transformation can now be worked out, yielding

$$
\begin{aligned}
\mathscr{H}_{\mathrm{D}} &= \overline{\pi}^{\alpha}\frac{\partial\psi}{\partial x^{\alpha}} + \frac{\partial\overline{\psi}}{\partial x^{\alpha}}\pi^{\alpha} - \mathscr{L}_{\mathrm{D}}' \\
&= \frac{\partial\overline{\psi}}{\partial x^{\alpha}}\,\frac{i\sigma^{\alpha\beta}}{3\tilde{m}}\,\frac{\partial\psi}{\partial x^{\beta}} + m\,\overline{\psi}\,\psi \\
&= \left(\overline{\pi}^{\beta} - \frac{i}{2}\overline{\psi}\gamma^{\beta}\right)\frac{\partial\psi}{\partial x^{\beta}} + m\,\overline{\psi}\,\psi,
\end{aligned}
$$

thus $[\mathscr{H}_{\mathrm{D}}] = [\mathscr{L}_{\mathrm{D}}] = [L]^{-4}$. As the Hamiltonian density must always be expressed in terms of the canonical momenta rather then by the velocities, we must solve Eq. (29) for $\partial_{\mu}\psi$ and $\partial_{\mu}\overline{\psi}$. To this end, we multiply $\overline{\pi}^{\mu}$ by $\tau_{\mu\nu}$ from the right, and π^{μ} by $\tau_{\nu\mu}$ from the left,

$$
\begin{aligned}
\frac{\partial\overline{\psi}}{\partial x^{\nu}} &= \frac{3\tilde{m}}{i}\left(\overline{\pi}^{\alpha} - \frac{i}{2}\overline{\psi}\gamma^{\alpha}\right)\tau_{\alpha\nu} \\
\frac{\partial\psi}{\partial x^{\nu}} &= \frac{3\tilde{m}}{i}\,\tau_{\nu\beta}\left(\pi^{\beta} + \frac{i}{2}\gamma^{\beta}\psi\right).
\end{aligned} \tag{30}
$$

The Dirac Hamiltonian density is then finally obtained as

$$
\mathscr{H}_{\mathrm{D}} = \left(\overline{\pi}^{\alpha} - \frac{i}{2}\overline{\psi}\gamma^{\alpha}\right)\frac{3\tilde{m}\tau_{\alpha\beta}}{i}\left(\pi^{\beta} + \frac{i}{2}\gamma^{\beta}\psi\right) + m\,\overline{\psi}\psi. \tag{31}
$$

We may expand the products in Eq. (31) using Eqs. (23) to find

$$
\mathscr{H}_{\mathrm{D}} = i\tilde{m}\left(\frac{1}{2}\overline{\psi}\,\gamma_{\alpha}\pi^{\alpha} - \frac{1}{2}\overline{\pi}^{\alpha}\gamma_{\alpha}\psi - 3\overline{\pi}^{\alpha}\tau_{\alpha\beta}\pi^{\beta}\right) + (m - \tilde{m})\,\overline{\psi}\psi. \tag{32}
$$

In order to show that the Hamiltonian density $\mathscr{H}_{\mathrm{D}}$ describes the same dynamics as $\mathscr{L}_{\mathrm{D}}$ from Eq. (22), we set up the canonical equations from Eq. (32)

$$
\begin{aligned}
\frac{\partial\overline{\psi}}{\partial x^{\nu}} &= \frac{\partial\mathscr{H}_{\mathrm{D}}}{\partial\pi^{\nu}} = \quad i\tilde{m}\left(\tfrac{1}{2}\overline{\psi}\,\gamma_{\nu} - 3\overline{\pi}^{\alpha}\tau_{\alpha\nu}\right) \\
\frac{\partial\psi}{\partial x^{\mu}} &= \frac{\partial\mathscr{H}_{\mathrm{D}}}{\partial\overline{\pi}^{\mu}} = -i\tilde{m}\left(\tfrac{1}{2}\gamma_{\mu}\psi + 3\tau_{\mu\beta}\pi^{\beta}\right).
\end{aligned}
$$

Obviously, these equations reproduce the definition of the canonical momenta from Eqs. (29) in their inverted form given by Eqs. (30). The second set of canonical equations follows from the ψ and $\overline{\psi}$ dependence of the Hamiltonian $\mathscr{H}_{\mathrm{D}}$,

$$
\begin{aligned}
\frac{\partial \overline{\pi}^{\alpha}}{\partial x^{\alpha}} = -\frac{\partial \mathscr{H}_{\mathrm{D}}}{\partial \psi} &= \frac{i\tilde{m}}{2}\overline{\pi}^{\beta}\gamma_{\beta} - (m-\tilde{m})\,\overline{\psi} \\
&= \frac{i\tilde{m}}{2}\left(\frac{i}{2}\overline{\psi}\gamma^{\beta} + \frac{\partial \overline{\psi}}{\partial x^{\alpha}}\frac{i\sigma^{\alpha\beta}}{3\tilde{m}}\right)\gamma_{\beta} - (m-\tilde{m})\,\overline{\psi} \\
&= -\frac{i}{2}\frac{\partial \overline{\psi}}{\partial x^{\alpha}}\gamma^{\alpha} - m\overline{\psi} \\
\frac{\partial \pi^{\alpha}}{\partial x^{\alpha}} = -\frac{\partial \mathscr{H}_{\mathrm{D}}}{\partial \overline{\psi}} &= -\frac{i\tilde{m}}{2}\gamma_{\beta}\pi^{\beta} - (m-\tilde{m})\,\psi \\
&= -\frac{i\tilde{m}}{2}\gamma_{\beta}\left(-\frac{i}{2}\gamma^{\beta}\psi + \frac{i\sigma^{\beta\alpha}}{3\tilde{m}}\frac{\partial \psi}{\partial x^{\alpha}}\right) - (m-\tilde{m})\,\psi \\
&= \frac{i}{2}\gamma^{\alpha}\frac{\partial \psi}{\partial x^{\alpha}} - m\psi.
\end{aligned}
$$

The divergences of the canonical momenta follow equally from the derivatives of the first canonical equations, or, equivalently, from the derivatives of Eqs. (29),

$$
\begin{aligned}
\frac{\partial \overline{\pi}^{\alpha}}{\partial x^{\alpha}} &= \quad \frac{i}{2}\frac{\partial \overline{\psi}}{\partial x^{\alpha}}\gamma^{\alpha} + \cancel{\frac{\partial^{2}\overline{\psi}}{\partial x^{\alpha}\partial x^{\beta}}\frac{i\sigma^{\alpha\beta}}{3\tilde{m}}} = \quad \frac{i}{2}\frac{\partial \overline{\psi}}{\partial x^{\alpha}}\gamma^{\alpha} \\
\frac{\partial \pi^{\alpha}}{\partial x^{\alpha}} &= -\frac{i}{2}\gamma^{\alpha}\frac{\partial \psi}{\partial x^{\alpha}} - \cancel{\frac{i\sigma^{\alpha\beta}}{3\tilde{m}}\frac{\partial^{2}\psi}{\partial x^{\alpha}\partial x^{\beta}}} = -\frac{i}{2}\gamma^{\alpha}\frac{\partial \psi}{\partial x^{\alpha}}.
\end{aligned}
$$

The terms containing the second derivatives of ψ and $\overline{\psi}$ vanish due to the skew-symmetry of $\sigma^{\mu\nu}$. Equating finally the expressions for the divergences of the canonical momenta, we encounter, as expected, the Dirac equations (25)

$$
\begin{aligned}
\frac{i}{2}\frac{\partial \overline{\psi}}{\partial x^{\alpha}}\gamma^{\alpha} &= -m\overline{\psi} - \frac{i}{2}\frac{\partial \overline{\psi}}{\partial x^{\alpha}}\gamma^{\alpha} \\
-\frac{i}{2}\gamma^{\alpha}\frac{\partial \psi}{\partial x^{\alpha}} &= -m\psi + \frac{i}{2}\gamma^{\alpha}\frac{\partial \psi}{\partial x^{\alpha}}.
\end{aligned}
$$

It should be mentioned that this section is similar to the derivation of the Dirac Hamiltonian density in Ref. [6]. We note that the additional term in the Dirac Lagrangian density $\mathscr{L}_{\mathrm{D}}'$ from Eq. (27) — as compared to the Lagrangian $\mathscr{L}_{\mathrm{D}}$ from Eq. (24) — entails additional terms in the energy-momentum tensor, namely,

$$
T_{\mu}^{\nu\,\prime} - T_{\mu}^{\nu} \equiv j_{\mu}^{\nu}(x) = \frac{i}{3\tilde{m}}\left(\partial_{\alpha}\overline{\psi}\sigma^{\alpha\nu}\partial_{\mu}\psi + \partial_{\mu}\overline{\psi}\sigma^{\nu\alpha}\partial_{\alpha}\psi - \delta_{\mu}^{\nu}\partial_{\alpha}\overline{\psi}\sigma^{\alpha\lambda}\partial_{\lambda}\psi\right).
$$

We easily convince ourselves by direct calculation that the divergences of $T_{\mu}^{\nu\,\prime}$ and T_{μ}^{ν} coincide,

$$
\begin{aligned}
\frac{\partial j_\mu^\beta}{\partial x^\beta} &= \frac{i}{3\tilde{m}}\Big(\cancel{\partial_\beta\partial_\alpha\overline{\psi}\sigma^{\alpha\beta}\partial_\mu\psi} + \partial_\alpha\overline{\psi}\sigma^{\alpha\beta}\partial_\beta\partial_\mu\psi + \partial_\beta\partial_\mu\overline{\psi}\sigma^{\beta\alpha}\partial_\alpha\psi \\
&\quad + \cancel{\partial_\mu\overline{\psi}\sigma^{\beta\alpha}\partial_\beta\partial_\alpha\psi} - \delta_\mu^\beta\,\partial_\beta\partial_\alpha\overline{\psi}\sigma^{\alpha\lambda}\partial_\lambda\psi - \delta_\mu^\beta\,\partial_\alpha\overline{\psi}\sigma^{\alpha\lambda}\partial_\beta\partial_\lambda\psi\Big) \\
&= \frac{i}{3\tilde{m}}\Big(\partial_\alpha\overline{\psi}\sigma^{\alpha\beta}\partial_\beta\partial_\mu\psi + \partial_\beta\partial_\mu\overline{\psi}\sigma^{\beta\alpha}\partial_\alpha\psi \\
&\qquad - \partial_\mu\partial_\alpha\overline{\psi}\sigma^{\alpha\beta}\partial_\beta\psi - \partial_\alpha\overline{\psi}\sigma^{\alpha\beta}\partial_\mu\partial_\beta\psi\Big) \\
&\equiv 0,
\end{aligned}
$$

which means that both energy-momentum tensors represent the same physical system. For each index μ, $j_\mu^\nu(x)$ represents a conserved current vector which are all associated with the transformation from $\mathscr{L}_\mathrm{D}$ to $\mathscr{L}'_\mathrm{D}$.

4 Examples of canonical transformations in covariant Hamiltonian field theory

The formalism of canonical transformations that was worked out in Sect. 2 is now shown to yield a generalized representation of Noether's theorem. Furthermore, a generalized theory of $\mathrm{U}(N)$ gauge transformations is outlined.

4.1 Generalized Noether theorem

Canonical transformations are defined by Eq. (6) as the particular subset of general transformations of the fields ϕ_I and their conjugate momentum vector fields $\boldsymbol{\pi}_I$ that preserve the action functional (6). Such a transformation depicts a symmetry transformation that is associated with a conserved four-current vector, hence with a vector whose space-time divergence vanishes[7]. In the following, we shall work out the correlation of this conserved current by means an *infinitesimal* canonical transformation of the field variables. The generating function F_2^μ of an *infinitesimal* transformation differs from that of an *identical* transformation by a infinitesimal parameter $\varepsilon \neq 0$ times an as yet arbitrary function $g^\mu(\phi_I, \boldsymbol{\pi}_I, x)$,

$$
F_2^\mu(\phi_I, \boldsymbol{\Pi}_I, x) = \phi_J\,\Pi_J^\mu + \varepsilon\,g^\mu(\phi_I, \boldsymbol{\pi}_I, x). \tag{33}
$$

To first order in ε, the subsequent transformation rules follow from the general rules (11) as

$$\pi_I^\mu = \frac{\partial F_2^\mu}{\partial \phi_I} = \Pi_I^\mu + \varepsilon \frac{\partial g^\mu}{\partial \phi_I}, \qquad \Phi_I \, \delta_\nu^\mu = \frac{\partial F_2^\mu}{\partial \Pi_I^\nu} = \phi_I \, \delta_\nu^\mu + \varepsilon \frac{\partial g^\mu}{\partial \pi_I^\nu},$$
$$\mathcal{H}' = \mathcal{H} + \left.\frac{\partial F_2^\alpha}{\partial x^\alpha}\right|_{\mathrm{expl}} = \mathcal{H} + \varepsilon \left.\frac{\partial g^\alpha}{\partial x^\alpha}\right|_{\mathrm{expl}},$$

hence

$$\delta\pi_I^\mu = -\varepsilon \frac{\partial g^\mu}{\partial \phi_I}, \qquad \delta\phi_I \, \delta_\nu^\mu = \varepsilon \frac{\partial g^\mu}{\partial \pi_I^\nu}, \qquad \delta\mathcal{H}\big|_{\mathrm{CT}} = \varepsilon \left.\frac{\partial g^\alpha}{\partial x^\alpha}\right|_{\mathrm{expl}}. \tag{34}$$

As the transformation does not change the independent variables, x^μ, both the original as well as the transformed fields refer to the same space-time event x, hence $\delta x^\mu = 0$. Making use of the canonical field equations (5), the variation of $\mathcal{H}$ due to the variations (34) of the canonical field variables ϕ_I and π_I^μ emerges as

$$\begin{aligned}
\delta\mathcal{H} &= \frac{\partial \mathcal{H}}{\partial \phi_I} \, \delta\phi_I + \frac{\partial \mathcal{H}}{\partial \pi_I^\alpha} \, \delta\pi_I^\alpha \\
&= -\frac{\partial \pi_I^\beta}{\partial x^\alpha} \, \delta_\beta^\alpha \, \delta\phi_I + \frac{\partial \phi_I}{\partial x^\alpha} \, \delta\pi_I^\alpha \\
&= -\varepsilon \left(\frac{\partial g^\alpha}{\partial \pi_I^\beta} \frac{\partial \pi_I^\beta}{\partial x^\alpha} + \frac{\partial g^\alpha}{\partial \phi_I} \frac{\partial \phi_I}{\partial x^\alpha} \right) \\
&= -\varepsilon \left(\frac{\partial g^\alpha}{\partial x^\alpha} - \left.\frac{\partial g^\alpha}{\partial x^\alpha}\right|_{\mathrm{expl}} \right) \\
&= -\varepsilon \frac{\partial g^\alpha}{\partial x^\alpha} + \delta\mathcal{H}\big|_{\mathrm{CT}}.
\end{aligned} \tag{35}$$

If and only if the infinitesimal transformation rule $\delta\mathcal{H}|_{\mathrm{CT}}$ for the Hamiltonian from Eqs. (34) coincides with the variation $\delta\mathcal{H}$ at $\delta x^\mu = 0$ from Eq. (35), then the set of infinitesimal transformation rules is consistent and actually defines a *canonical* transformation. We thus have

$$\delta\mathcal{H}\big|_{\mathrm{CT}} \stackrel{!}{=} \delta\mathcal{H} \quad \Longleftrightarrow \quad \frac{\partial g^\alpha}{\partial x^\alpha} \stackrel{!}{=} 0. \tag{36}$$

Thus, the divergence of the characteristic function $g^\mu(x)$ in the generating function (33) must vanish in order for the transformation (34) to be *canonical*, and hence to preserve the form of the action functional (6). The $g^\mu(x)$ then define a conserved four-current vector, commonly referred to as *Noether current*. The canonical transformation rules then furnish the corresponding infinitesimal one-parameter group of symmetry transformations

$$\frac{\partial g^{\alpha}(x)}{\partial x^{\alpha}} = 0 \tag{37}$$

$$\delta\pi_I^{\mu} = -\varepsilon\,\frac{\partial g^{\mu}}{\partial \phi_I}, \qquad \delta\phi_I\,\delta_{\nu}^{\mu} = \varepsilon\,\frac{\partial g^{\mu}}{\partial \pi_I^{\nu}}, \qquad \delta\mathscr{H} = \varepsilon\,\frac{\partial g^{\alpha}}{\partial x^{\alpha}}\bigg|_{\mathrm{expl}}.$$

We can now formulate the generalized Noether theorem and its inverse in the realm of covariant Hamiltonian field theory as:

Theorem 1 (generalized Noether). *The characteristic vector function* $g^{\mu}(\phi_I, \boldsymbol{\pi}_I, x)$ *in the generating function* F_2^{μ} *from Eq. (33) must have* zero divergence *in order to define a canonical transformation. The subsequent transformation rules (37) then define an infinitesimal one-parameter group of symmetry transformations that preserve the form of the action functional (6).*

Conversely, if a one-parameter symmetry transformation is known to preserve the form of the action functional (6), then the transformation is canonical *and hence can be derived from a generating function. The characteristic* 4*-vector function* $g^{\mu}(\phi_I, \boldsymbol{\pi}_I, x)$ *in the corresponding* infinitesimal *generating function (33) then represents a conserved current, hence* $\partial g^{\alpha}/\partial x^{\alpha} = 0$.

In contrast to the usual derivation of this theorem in the Lagrangian formalism, we are not restricted to point transformations as the g^{μ} may be *any* divergence-free 4-vector function of the given dynamical system. In this sense, we have found a generalization of Noether's theorem.

4.1.1 Gauge invariance of the electromagnetic 4-potential

For the Maxwell Hamiltonian $\mathscr{H}_{\mathrm{M}}$ from Eq. (18), the correlation of the 4-vector potential a^{μ} with the conjugate fields $p_{\mu\nu}$ is determined by the first field equation (19) as the generalized curl of $\boldsymbol{a}$. This means on the other hand that the correlation between $\boldsymbol{a}$ and the $p_{\mu\nu}$ is *not unique*. Defining a transformed 4-vector potential $\boldsymbol{A}$ according to

$$A_{\mu} = a_{\mu} + \frac{\partial\chi(x)}{\partial x^{\mu}}, \tag{38}$$

with $\chi = \chi(x)$ an arbitrary differentiable function of the independent variables. This means for the transformation of the $p_{\mu\nu}$

$$p_{\mu\nu} = \frac{\partial a_{\nu}}{\partial x^{\mu}} - \frac{\partial a_{\mu}}{\partial x^{\nu}} = \frac{\partial A_{\nu}}{\partial x^{\mu}} - \cancel{\frac{\partial^2\chi(x)}{\partial x^{\nu}\partial x^{\mu}}} - \frac{\partial A_{\mu}}{\partial x^{\nu}} + \cancel{\frac{\partial^2\chi(x)}{\partial x^{\mu}\partial x^{\nu}}} = P_{\mu\nu}. \tag{39}$$

The transformations (38) and (39) can be regarded as a canonical transformation, whose generating function F_2^{μ} is given by

$$F_2^{\mu}(\boldsymbol{a}, \boldsymbol{P}, x) = a_{\alpha}P^{\alpha\mu} + \frac{\partial}{\partial x^{\alpha}}\left(P^{\alpha\mu}\chi(x)\right). \tag{40}$$

For a vector field $\boldsymbol{a}$ and its set of canonical conjugate fields $\boldsymbol{p}^{\mu}$, the general transformation rules (11) are rewritten as

$$p^{\nu\mu} = \frac{\partial F_2^{\mu}}{\partial a_{\nu}}, \qquad A_{\nu}\,\delta_{\beta}^{\mu} = \frac{\partial F_2^{\mu}}{\partial P^{\nu\beta}}, \qquad \mathscr{H}' = \mathscr{H} + \left.\frac{\partial F_2^{\alpha}}{\partial x^{\alpha}}\right|_{\text{expl}}, \tag{41}$$

which yield for the particular generating function of Eq. (40) the transformation prescriptions

$$\begin{aligned}
p^{\nu\mu} &= \frac{\partial a_{\alpha}}{\partial a_{\nu}} P^{\alpha\mu} = \delta_{\alpha}^{\nu} P^{\alpha\mu} = P^{\nu\mu} \\
A_{\nu}\,\delta_{\beta}^{\mu} &= a_{\alpha}\delta_{\nu}^{\alpha}\delta_{\beta}^{\mu} + \delta_{\nu}^{\alpha}\delta_{\beta}^{\mu}\,\frac{\partial\chi(x)}{\partial x^{\alpha}} \\
\Rightarrow \quad A_{\nu} &= a_{\nu} + \frac{\partial\chi(x)}{\partial x^{\nu}} \\
\mathscr{H}' - \mathscr{H} &= \frac{\partial^2 p^{\alpha\beta}}{\partial x^{\alpha}\partial x^{\beta}}\chi(x) + \frac{\partial p^{\alpha\beta}}{\partial x^{\alpha}}\frac{\partial\chi(x)}{\partial x^{\beta}} + p^{\alpha\beta}\,\frac{\partial^2\chi(x)}{\partial x^{\alpha}\partial x^{\beta}} \\
&= -\frac{\partial p^{\alpha\beta}}{\partial x^{\beta}}\frac{\partial\chi(x)}{\partial x^{\alpha}}.
\end{aligned}$$

The canonical transformation rules coincide with the correlations of Eqs. (38) and (39) defining the Lorentz gauge. The last equation holds because of the skew-symmetry of the canonical momentum tensor $p^{\nu\mu} = -p^{\mu\nu}$.

In order to determine the conserved Noether current that is associated with the canonical point transformation generated by $\boldsymbol{F}_2$ from Eq. (40), we need the generator of the corresponding *infinitesimal* canonical point transformation,

$$F_2^{\mu}(\boldsymbol{a},\boldsymbol{P},x) = a_{\alpha}P^{\alpha\mu} + \varepsilon g^{\mu}(\boldsymbol{p},x), \qquad g^{\mu} = \frac{\partial}{\partial x^{\alpha}}\left[p^{\alpha\mu}\chi(x)\right].$$

Herein, $\varepsilon \neq 0$ denotes a small parameter. The pertaining infinitesimal canonical transformation rules are

$$\begin{aligned}
p^{\nu\mu} &= \frac{\partial F_2^{\mu}}{\partial a_{\nu}} = P^{\nu\mu}, \qquad A_{\nu} = a_{\nu} + \varepsilon\frac{\partial\chi(x)}{\partial x^{\nu}} \\
\delta\mathscr{H}\big|_{\text{CT}} &= \left.\frac{\partial F_2^{\alpha}}{\partial x^{\alpha}}\right|_{\text{expl}} = \mathscr{H}_{\text{M}}' - \mathscr{H}_{\text{M}} = -\varepsilon\,\frac{\partial p^{\alpha\beta}}{\partial x^{\beta}}\frac{\partial\chi(x)}{\partial x^{\alpha}}.
\end{aligned}$$

The coordinate transformation rules agree with Eqs. (38) and (39) in the finite limit. Because of $\delta p^{\nu\mu} \equiv P^{\nu\mu} - p^{\nu\mu} = 0$, the variation $\delta\mathscr{H}$ due to the variation of the canonical variables reduces to the term proportional to $\delta a_{\nu} \equiv A_{\nu} - a_{\nu}$,

$$\delta\mathscr{H} = \frac{\partial\mathscr{H}_{\text{M}}}{\partial a_{\alpha}}\,\delta a_{\alpha} = -\varepsilon\,\frac{\partial p^{\alpha\beta}}{\partial x^{\beta}}\,\frac{\partial\chi(x)}{\partial x^{\alpha}}.$$

Hence, $\delta\mathscr{H}$ coincides with the corresponding canonical transformation rule $\delta\mathscr{H}|_{\mathrm{CT}}$, as required for the transformation to be canonical. With the requirement (36) fulfilled, the characteristic function $g^{\mu}(\boldsymbol{p},x)$ in the infinitesimal generating function F_2^{μ} then directly yields the conserved 4-current $\boldsymbol{j}_{\mathrm{N}}(x)$, $j_{\mathrm{N}}^{\mu} = g^{\mu}$ according to Noether's theorem from Eq. (37)

$$\frac{\partial j_{\mathrm{N}}^{\alpha}(x)}{\partial x^{\alpha}} = 0, \qquad j_{\mathrm{N}}^{\mu}(x) = \frac{\partial}{\partial x^{\alpha}}\big(p^{\alpha\mu}\chi(x)\big).$$

By calculating its divergence, we verify directly that $\boldsymbol{j}_{\mathrm{N}}(x)$ is indeed the conserved Noether current that corresponds to the symmetry transformation (38)

$$\begin{aligned}\frac{\partial j_{\mathrm{N}}^{\beta}(x)}{\partial x^{\beta}} &= \frac{\partial}{\partial x^{\beta}}\left(\frac{\partial p^{\alpha\beta}}{\partial x^{\alpha}}\chi + p^{\alpha\beta}\frac{\partial \chi}{\partial x^{\alpha}}\right)\\ &= \frac{\partial^2 p^{\alpha\beta}}{\partial x^{\alpha}\partial x^{\beta}}\chi + \left(\frac{\partial p^{\beta\alpha}}{\partial x^{\beta}} + \frac{\partial p^{\alpha\beta}}{\partial x^{\beta}}\right)\frac{\partial \chi}{\partial x^{\alpha}} + p^{\alpha\beta}\frac{\partial^2 \chi}{\partial x^{\alpha}\partial x^{\beta}}.\end{aligned}$$

As $\chi(x)$ represents by assumption an *arbitrary* function of x, a zero divergence of the Noether current j_{N}^{β} means that the coefficients associated with χ and its first and second derivative must separately vanish. This is equally ensured for all three terms if $p^{\nu\mu}$ is a skew-symmetric tensor

$$p^{\nu\mu} = -p^{\mu\nu}.$$

4.2 General local $U(N)$ gauge transformation

As an interesting example of a canonical transformation in the covariant Hamiltonian description of classical fields, the general local $\mathrm{U}(N)$ gauge transformation is treated in this section. The main feature of the approach is that the terms to be added to a given Hamiltonian $\mathscr{H}$ in order to render it *locally* gauge invariant only depends on the *type of fields* contained in the Hamiltonian $\mathscr{H}$ and not on the particular form of the original Hamiltonian itself. The only precondition is that $\mathscr{H}$ must be invariant under the corresponding *global* gauge transformation, hence a transformation *not* depending explicitly on x.

4.2.1 External gauge field

We consider a system consisting of a vector of N complex fields ϕ_I, $I = 1,\ldots,N$, and the adjoint field vector, $\overline{\boldsymbol{\phi}}$,

$$\boldsymbol{\phi} = \begin{pmatrix} \phi_1 \\ \vdots \\ \phi_N \end{pmatrix}, \qquad \overline{\boldsymbol{\phi}} = \left(\overline{\phi}_1 \cdots \overline{\phi}_N\right).$$

A general local linear transformation may be expressed in terms of a dimensionless complex matrix $U(x) = (u_{IJ}(x))$ and its adjoint, $U^{\dagger}$ that may depend explicitly on the independent variables, x^{μ}, as

$$\begin{aligned} \boldsymbol{\Phi} &= U\,\boldsymbol{\phi}, & \overline{\boldsymbol{\Phi}} &= \overline{\boldsymbol{\phi}}\,U^{\dagger} \\ \Phi_I &= u_{IJ}\,\phi_J, & \overline{\Phi}_I &= \overline{\phi}_J\,u^{*}_{JI}, \qquad [u_{IJ}] = 1. \end{aligned} \tag{42}$$

With this notation, ϕ_I may stand for a set of $I = 1,\ldots,N$ complex scalar fields ϕ_I or Dirac spinors. In other words, U is supposed to define an isomorphism within the space of the ϕ_I, hence to linearly map the ϕ_I into objects of the same type. The uppercase Latin letter indexes label the field or spinor number. Their transformation in iso-space are not associated with any metric. We, therefore, do not use superscripts for these indexes as there is not distinction between covariant and contravariant components. In contrast, Greek indexes are used for those components that *are* associated with a metric — such as the derivatives with respect to a space-time variable, x^{μ}. As usual, summation is understood for indexes occurring in pairs.

We restrict ourselves to transformations that preserve the norm $\overline{\boldsymbol{\phi}}\boldsymbol{\phi}$

$$\begin{aligned} \overline{\boldsymbol{\Phi}}\boldsymbol{\Phi} &= \overline{\boldsymbol{\phi}}\,U^{\dagger}U\,\boldsymbol{\phi} = \overline{\boldsymbol{\phi}}\boldsymbol{\phi} & &\Longrightarrow & U^{\dagger}U &= \mathbb{1} = UU^{\dagger} \\ \overline{\Phi}_I\Phi_I &= \overline{\phi}_J u^{*}_{JI}\,u_{IK}\phi_K = \overline{\phi}_K\phi_K & &\Longrightarrow & u^{*}_{JI}\,u_{IK} &= \delta_{JK} = u_{JI}\,u^{*}_{IK}. \end{aligned}$$

This means that $U^{\dagger} = U^{-1}$, hence that the matrix U is supposed to be *unitary*. The transformation (42) follows from a generating function that — corresponding to $\mathscr{H}$ — must be a real-valued function of the generally complex fields $\boldsymbol{\phi}$ and their canonical conjugates, $\boldsymbol{\pi}^{\mu}$,

$$\begin{aligned} F_2^{\mu}(\boldsymbol{\phi},\overline{\boldsymbol{\phi}},\boldsymbol{\Pi}^{\mu},\overline{\boldsymbol{\Pi}}^{\mu},x) &= \overline{\boldsymbol{\Pi}}^{\mu}\,U\,\boldsymbol{\phi} + \overline{\boldsymbol{\phi}}\,U^{\dagger}\,\boldsymbol{\Pi}^{\mu} \\ &= \overline{\Pi}^{\mu}_K\,u_{KJ}\,\phi_J + \overline{\phi}_K\,u^{*}_{KJ}\,\Pi^{\mu}_J. \end{aligned} \tag{43}$$

According to Eqs. (11) the set of transformation rules follows as

$$\begin{aligned} \overline{\pi}^{\mu}_I &= \frac{\partial F_2^{\mu}}{\partial \phi_I} = \overline{\Pi}^{\mu}_K u_{KJ}\delta_{IJ}, & \overline{\Phi}_I\delta^{\mu}_{\nu} &= \frac{\partial F_2^{\mu}}{\partial \Pi^{\nu}_I} = \overline{\phi}_K u^{*}_{KJ}\delta^{\mu}_{\nu}\delta_{IJ} \\ \pi^{\mu}_I &= \frac{\partial F_2^{\mu}}{\partial \overline{\phi}_I} = \delta_{IK}u^{*}_{KJ}\Pi^{\mu}_J, & \Phi_I\delta^{\mu}_{\nu} &= \frac{\partial F_2^{\mu}}{\partial \overline{\Pi}^{\nu}_I} = \delta^{\mu}_{\nu}\,\delta_{IK}u_{KJ}\phi_J. \end{aligned}$$

The complete set of transformation rules and their inverses then read in component notation

$$\begin{aligned}
\Phi_I &= u_{IJ}\,\phi_J, & \overline{\Phi}_I &= \overline{\phi}_J\,u^*_{JI}, & \Pi^\mu_I &= u_{IJ}\,\pi^\mu_J, & \overline{\Pi}^\mu_I &= \overline{\pi}^\mu_J\,u^*_{JI} \\
\phi_I &= u^*_{IJ}\,\Phi_J, & \overline{\phi}_I &= \overline{\Phi}_J u_{JI}, & \pi^\mu_I &= u^*_{IJ}\,\Pi^\mu_J, & \overline{\pi}^\mu_I &= \overline{\Pi}^\mu_J\,u_{JI}.
\end{aligned} \tag{44}$$

We assume the Hamiltonian $\mathscr{H}$ to be *form-invariant* under the *global* gauge transformation (42), which is given for $U = \mathrm{const}$, hence for all u_{IJ} *not* depending on the independent variables, x^μ. In contrast, if $U = U(x)$, the transformation (44) is referred to as a *local* gauge transformation. The transformation rule for the Hamiltonian is then determined by the explicitly x^μ-dependent terms of the generating function F_2^μ according to

$$\begin{aligned}
\mathscr{H}' - \mathscr{H} = \left.\frac{\partial F_2^\alpha}{\partial x^\alpha}\right|_{\mathrm{expl}} &= \overline{\Pi}^\alpha_I\,\frac{\partial u_{IJ}}{\partial x^\alpha}\,\phi_J + \overline{\phi}_I\,\frac{\partial u^*_{IJ}}{\partial x^\alpha}\,\Pi^\alpha_J \\
&= \overline{\pi}^\alpha_K\,u^*_{KI}\,\frac{\partial u_{IJ}}{\partial x^\alpha}\,\phi_J + \overline{\phi}_I\,\frac{\partial u^*_{IJ}}{\partial x^\alpha}\,u_{JK}\,\pi^\alpha_K \\
&= \overline{\pi}^\alpha_K\,u^*_{KI}\,\frac{\partial u_{IJ}}{\partial x^\alpha}\,\phi_J + \overline{\phi}_K\,\frac{\partial u^*_{KI}}{\partial x^\alpha}\,u_{IJ}\,\pi^\alpha_J \\
&= \left(\overline{\pi}^\alpha_K\,\phi_J - \overline{\phi}_K\,\pi^\alpha_J\right)u^*_{KI}\,\frac{\partial u_{IJ}}{\partial x^\alpha}.
\end{aligned} \tag{45}$$

In the last step, the identity

$$\frac{\partial u^*_{KI}}{\partial x^\mu}\,u_{IJ} + u^*_{KI}\,\frac{\partial u_{IJ}}{\partial x^\mu} = \frac{\partial}{\partial x^\mu}\left(u^*_{KI}u_{IJ}\right) = \frac{\partial}{\partial x^\mu}\,\delta_{KJ} = 0$$

was inserted. If we want to set up a Hamiltonian $\mathscr{H}_1$ that is *form-invariant* under the *local*, hence x^μ-dependent transformation generated by (43), then we must compensate the additional terms (45) that emerge from the explicit x^μ-dependence of the generating function (43). The only way to achieve this is to *adjoin* the Hamiltonian $\mathscr{H}$ of our system with terms that correspond to (45) with regard to their dependence on the canonical variables, $\boldsymbol{\phi}, \overline{\boldsymbol{\phi}}, \boldsymbol{\pi}^\mu, \overline{\boldsymbol{\pi}}^\mu$. With a *unitary* matrix U, the u_{IJ}-dependent terms in Eq. (45) are *skew-hermitian*,

$$\overline{\left(u^*_{KI}\,\frac{\partial u_{IJ}}{\partial x^\mu}\right)} = \frac{\partial u^*_{JI}}{\partial x^\mu}\,u_{IK} = -u^*_{JI}\,\frac{\partial u_{IK}}{\partial x^\mu}, \qquad \overline{\left(\frac{\partial u_{KI}}{\partial x^\mu}\,u^*_{IJ}\right)} = u_{JI}\,\frac{\partial u^*_{IK}}{\partial x^\mu} = -\frac{\partial u_{JI}}{\partial x^\mu}\,u^*_{IK},$$

or in matrix notation

$$\left(U^\dagger\,\frac{\partial U}{\partial x^\mu}\right)^\dagger = \frac{\partial U^\dagger}{\partial x^\mu}U = -U^\dagger\,\frac{\partial U}{\partial x^\mu}, \qquad \left(\frac{\partial U}{\partial x^\mu}U^\dagger\right)^\dagger = U\,\frac{\partial U^\dagger}{\partial x^\mu} = -\frac{\partial U}{\partial x^\mu}U^\dagger.$$

The u-dependent terms in Eq. (45) can thus be compensated by a *Hermitian* matrix $(\boldsymbol{a}_{KJ})$ of "4-vector gauge fields", with each off-diagonal matrix element, $\boldsymbol{a}_{KJ}$, $K \neq J$, a complex 4-vector field with components $a_{KJ\mu}$, $\mu = 0, \ldots, 3$

$$a_{KJ\mu} = a^*_{JK\mu}.$$

The number of independent gauge fields thus amount to N^2 real 4-vectors. The amended Hamiltonian $\mathscr{H}_1$ thus reads

$$\mathscr{H}_1 = \mathscr{H} + \mathscr{H}_\mathrm{a}, \qquad \mathscr{H}_\mathrm{a} = ig\left(\overline{\pi}^\alpha_K \phi_J - \overline{\phi}_K \pi^\alpha_J\right) a_{KJ\alpha}. \tag{46}$$

With the real coupling constant g, the interaction Hamiltonian $\mathscr{H}_\mathrm{a}$ is thus real. Usually, g is defined to be dimensionless. We then infer the dimension of the gauge fields $\boldsymbol{a}_{KJ}$ to be

$$[g] = 1, \qquad [\boldsymbol{a}_{KJ}] = [L]^{-1} = [m] = [\partial_\mu].$$

In contrast to the given system Hamiltonian $\mathscr{H}$, the *amended* Hamiltonian $\mathscr{H}_1$ is supposed to be *invariant in its form* under the canonical transformation, hence

$$\mathscr{H}_1' = \mathscr{H}' + \mathscr{H}_\mathrm{a}', \qquad \mathscr{H}_\mathrm{a}' = ig\left(\overline{\Pi}^\alpha_K \Phi_J - \overline{\Phi}_K \Pi^\alpha_J\right) A_{KJ\alpha}. \tag{47}$$

Submitting the amended Hamiltonian $\mathscr{H}_1$ from Eq. (46) to the canonical transformation generated by Eq. (43), the new Hamiltonian $\mathscr{H}_1'$ emerges with Eqs. (45) and (47) as

$$\begin{aligned}
\mathscr{H}_1' &= \mathscr{H}_1 + \left.\frac{\partial F_2^\alpha}{\partial x^\alpha}\right|_\mathrm{expl} = \mathscr{H} + \mathscr{H}_\mathrm{a} + \left.\frac{\partial F_2^\alpha}{\partial x^\alpha}\right|_\mathrm{expl} \\
&= \mathscr{H} + \left(\overline{\pi}^\alpha_K \phi_J - \overline{\phi}_K \pi^\alpha_J\right)\left(ig\, a_{KJ\alpha} + u^*_{KI}\frac{\partial u_{IJ}}{\partial x^\alpha}\right) \\
&\overset{!}{=} \mathscr{H}' + \left(\overline{\Pi}^\alpha_K \Phi_J - \overline{\Phi}_K \Pi^\alpha_J\right) ig A_{KJ\alpha}.
\end{aligned}$$

The original base fields, $\phi_J, \overline{\phi}_K$ and their conjugates can now be expressed in terms of the transformed ones according to the rules (44), which yields, after index relabeling, the conditions

$$\begin{aligned}
\mathscr{H}'(\boldsymbol{\Phi}, \overline{\boldsymbol{\Phi}}, \boldsymbol{\Pi}^\mu, \overline{\boldsymbol{\Pi}}^\mu, x^\mu) &\overset{\text{global GT}}{=} \mathscr{H}(\boldsymbol{\phi}, \overline{\boldsymbol{\phi}}, \boldsymbol{\pi}^\mu, \overline{\boldsymbol{\pi}}^\mu, x^\mu) \\
\left(\overline{\Pi}^\alpha_K \Phi_J - \overline{\Phi}_K \Pi^\alpha_J\right) ig A_{KJ\alpha} &= \left(\overline{\Pi}^\alpha_K \Phi_J - \overline{\Phi}_K \Pi^\alpha_J\right)\left(ig\, u_{KL}\, a_{LI\alpha}\, u^*_{IJ} + \frac{\partial u_{KI}}{\partial x^\alpha} u^*_{IJ}\right).
\end{aligned}$$

This means that the system Hamiltonian must be invariant under the *global* gauge transformation defined by Eq. (44), whereas the gauge fields $A_{IJ\mu}$ must satisfy the transformation rule

$$A_{KJ\mu} = u_{KL}\, a_{LI\mu}\, u^*_{IJ} + \frac{1}{ig}\,\frac{\partial u_{KI}}{\partial x^\mu}\, u^*_{IJ}. \tag{48}$$

We observe that for any type of canonical field variables ϕ_I and for any Hamiltonian system $\mathscr{H}$, the transformation of the 4-vector gauge fields $\boldsymbol{a}_{IJ}(x)$ is uniquely determined according to Eq. (48) by the transformation matrix $U(x)$ for the N fields ϕ_I. In the notation of the 4-vector gauge fields $\boldsymbol{a}_{KJ}(x),\ K,J = 1,\ldots,N$, the transformation rule is equivalently expressed as

$$\boldsymbol{A}_{KJ} = u_{KL}\boldsymbol{a}_{LI}\,u^*_{IJ} + \frac{1}{ig}\,\frac{\partial u_{KI}}{\partial x}\,u^*_{IJ},$$

or, in matrix notation

$$\hat{A}_\mu = U\hat{a}_\mu U^\dagger + \frac{1}{ig}\,\frac{\partial U}{\partial x^\mu}U^\dagger, \qquad \hat{\boldsymbol{A}} = U\,\hat{\boldsymbol{a}}\,U^\dagger + \frac{1}{ig}\,\frac{\partial U}{\partial x}\,U^\dagger, \tag{49}$$

with $\hat{a}_\mu$ denoting the $N \times N$ matrices of the μ-components of the 4-vectors $\boldsymbol{A}_{IK}(x)$, and, finally, $\hat{\boldsymbol{a}}$ the $N \times N$ matrix of gauge 4-vectors $\boldsymbol{a}_{IK}(x)$. The matrix $U(x)$ is *unitary*, and thus constitutes a member of the group $\mathrm{U}(N)$

$$U^\dagger(x) = U^{-1}(x), \qquad |\det U(x)| = 1.$$

For $\det U(x) = +1$, the matrix $U(x)$ is a member of the group $\mathrm{SU}(N)$.

Inserting the transformation rule for the base fields, $\boldsymbol{\Phi} = U\,\boldsymbol{\phi}$, into Eq. (49), we immediately find the *homogeneous* transformation condition

$$\frac{\partial \boldsymbol{\Phi}}{\partial x^\mu} - ig\hat{A}_\mu \boldsymbol{\Phi} = U\left(\frac{\partial \boldsymbol{\phi}}{\partial x^\mu} - ig\,\hat{a}_\mu \boldsymbol{\phi}\right).$$

We identify this “amended” partial derivative as the covariant derivative that defines the minimum coupling rule for our gauge transformation.

Equation (49) is the general transformation law for gauge bosons. U and $\hat{a}_\mu$ do not commute if $N > 1$, hence if U is a unitary matrix rather than a complex number of modulus 1. We are then dealing with a non-Abelian gauge theory. As the matrices $\hat{a}_\mu$ are Hermitian, the number of independent gauge 4-vectors $\boldsymbol{a}_{IK}$ amounts to N real vectors on the main diagonal, and $(N^2 - N)/2$ independent complex off-diagonal vectors, which corresponds to a total number of N^2 independent real gauge 4-vectors for a $\mathrm{U}(N)$ symmetry transformation, and hence $N^2 - 1$ real gauge 4-vectors for a $\mathrm{SU}(N)$ symmetry transformation.

4.2.2 Including the gauge field dynamics

With the knowledge of the required transformation rule for the gauge fields from Eq. (48), it is now possible to redefine the generating function (43) to also describe the gauge field transformation. This simultaneously defines the transformation of the canonical conjugates, $p^{\mu\nu}_{JK}$, of the gauge fields $a_{JK\mu}$. Furthermore, the redefined generating function yields additional terms in the transformation rule for the Hamiltonian. Of course, in order for the Hamiltonian to be invariant under local gauge transformations, the additional terms must be invariant as well. The transformation rules for the fields $\boldsymbol{\phi}$ and the gauge field matrices $\hat{\boldsymbol{a}}$ (Eq. (49)) can be regarded as a canonical transformation that emerges from an explicitly x^μ-dependent and real-valued generating function vector of type $F_2^\mu = F_2^\mu(\boldsymbol{\phi}, \overline{\boldsymbol{\phi}}, \boldsymbol{\Pi}, \overline{\boldsymbol{\Pi}}, \boldsymbol{a}, \boldsymbol{P}, x)$,

$$F_2^\mu = \overline{\Pi}_K^\mu u_{KJ}\,\phi_J + \overline{\phi}_K u_{KJ}^* \Pi_J^\mu + P_{JK}^{\alpha\mu}\left(u_{KL}\,a_{LI\alpha}\,u_{IJ}^* + \frac{1}{ig}\frac{\partial u_{KI}}{\partial x^\alpha}\,u_{IJ}^*\right). \tag{50}$$

Accordingly, the subsequent transformation rules for canonical variables $\boldsymbol{\phi}, \overline{\boldsymbol{\phi}}$ and their conjugates, $\boldsymbol{\pi}^\mu, \overline{\boldsymbol{\pi}}^\mu$, agree with those from Eqs. (44). The rule for the gauge fields $a_{IK\alpha}$ emerges as

$$A_{KJ\alpha}\,\delta_\nu^\mu = \frac{\partial F_2^\mu}{\partial P_{JK}^{\alpha\nu}} = \delta_\nu^\mu\left(u_{KL}\,a_{LI\alpha}\,u_{IJ}^* + \frac{1}{ig}\frac{\partial u_{KI}}{\partial x^\alpha}u_{IJ}^*\right),$$

which obviously coincides with Eq. (48), as demanded. The transformation of the momentum fields is obtained from the generating function (50) as

$$p_{IL}^{\alpha\mu} = \frac{\partial F_2^\mu}{\partial a_{LI\alpha}} = u_{IJ}^* P_{JK}^{\alpha\mu}\,u_{KL}. \tag{51}$$

It remains to work out the difference of the Hamiltonians that are submitted to the canonical transformation generated by (50). Hence, according to the general rule from Eq. (11), we must calculate the divergence of the explicitly x^μ-dependent terms of F_2^μ

$$\begin{aligned}\left.\frac{\partial F_2^\alpha}{\partial x^\alpha}\right|_{\text{expl}} &= \overline{\Pi}_I^\alpha\,\frac{\partial u_{IJ}}{\partial x^\alpha}\,\phi_J + \overline{\phi}_I\,\frac{\partial u_{IJ}^*}{\partial x^\alpha}\,\Pi_J^\alpha \\ &\quad + P_{JK}^{\alpha\beta}\left(\frac{\partial u_{KL}}{\partial x^\beta}a_{LI\alpha}u_{IJ}^* + u_{KL}a_{LI\alpha}\frac{\partial u_{IJ}^*}{\partial x^\beta} + \frac{1}{ig}\frac{\partial u_{KI}}{\partial x^\alpha}\frac{\partial u_{IJ}^*}{\partial x^\beta} + \frac{1}{ig}\frac{\partial^2 u_{KI}}{\partial x^\alpha \partial x^\beta}u_{IJ}^*\right).\end{aligned} \tag{52}$$

We are now going to replace all u_{IJ}-dependencies in (52) by canonical variables making use of the canonical transformation rules. The first two terms on the right-hand side of Eq. (52) can be expressed in terms of the canonical variables by means of the transformation rules (44), (48), and (51) that all follow from the generating function (50)

$$\begin{aligned}\overline{\Pi}_I^\alpha\frac{\partial u_{IJ}}{\partial x^\alpha}\phi_J + \overline{\phi}_I\frac{\partial u_{IJ}^*}{\partial x^\alpha}\Pi_J^\alpha &= \overline{\Pi}_I^\alpha\frac{\partial u_{IJ}}{\partial x^\alpha}u_{JK}^*\Phi_K + \overline{\Phi}_K u_{KI}\frac{\partial u_{IJ}^*}{\partial x^\alpha}\Pi_J^\alpha \\ &= \overline{\Pi}_I^\alpha\frac{\partial u_{IJ}}{\partial x^\alpha}u_{JK}^*\Phi_K - \overline{\Phi}_K\frac{\partial u_{KI}}{\partial x^\alpha}u_{IJ}^*\Pi_J^\alpha \\ &= ig\overline{\Pi}_I^\alpha\left(A_{IK\alpha} - u_{IL}a_{LJ\alpha}u_{JK}^*\right)\Phi_K \\ &\qquad - ig\overline{\Phi}_K\left(A_{KJ\alpha} - u_{KL}a_{LI\alpha}u_{IJ}^*\right)\Pi_J^\alpha \\ &= ig\left(\overline{\Pi}_K^\alpha\Phi_J - \overline{\Phi}_K\Pi_J^\alpha\right)A_{KJ\alpha} - ig\left(\overline{\pi}_K^\alpha\phi_J - \overline{\phi}_K\pi_J^\alpha\right)a_{KJ\alpha}.\end{aligned}$$

The second derivative term in Eq. (52) is *symmetric* in the indexes α and β. If we split $P_{JK}^{\alpha\beta}$ into a symmetric $P_{JK}^{(\alpha\beta)}$ and a skew-symmetric part $P_{JK}^{[\alpha\beta]}$ in α and β

$$P_{JK}^{\alpha\beta} = P_{JK}^{(\alpha\beta)} + P_{JK}^{[\alpha\beta]}, \qquad P_{JK}^{[\alpha\beta]} = \tfrac{1}{2}\left(P_{JK}^{\alpha\beta} - P_{JK}^{\beta\alpha}\right), \qquad P_{JK}^{(\alpha\beta)} = \tfrac{1}{2}\left(P_{JK}^{\alpha\beta} + P_{JK}^{\beta\alpha}\right),$$

then the second derivative term vanishes for $P_{JK}^{[\alpha\beta]}$,

$$P_{JK}^{[\alpha\beta]}\frac{\partial^2 u_{KI}}{\partial x^\alpha \partial x^\beta}=0.$$

By inserting the transformation rules for the gauge fields from Eqs. (48), the remaining terms of (52) for the skew-symmetric part of $P_{JK}^{\alpha\beta}$ are converted into

$$\begin{aligned}
&P_{JK}^{[\alpha\beta]}\left(\frac{\partial u_{KL}}{\partial x^\beta}a_{LI\alpha}\,u_{IJ}^* + u_{KL}\,a_{LI\alpha}\frac{\partial u_{IJ}^*}{\partial x^\beta}+\frac{1}{ig}\frac{\partial u_{KI}}{\partial x^\alpha}\frac{\partial u_{IJ}^*}{\partial x^\beta}\right)\\
&= ig\,p_{JK}^{[\alpha\beta]}\,a_{KI\alpha}\,a_{IJ\beta} - ig\,P_{JK}^{[\alpha\beta]}A_{KI\alpha}A_{IJ\beta}\\
&= \tfrac{1}{2}ig\left(p_{JK}^{\alpha\beta}-p_{JK}^{\beta\alpha}\right)a_{KI\alpha}\,a_{IJ\beta} - \tfrac{1}{2}ig\left(P_{JK}^{\alpha\beta}-P_{JK}^{\beta\alpha}\right)A_{KI\alpha}A_{IJ\beta}\\
&= \tfrac{1}{2}ig\,p_{JK}^{\alpha\beta}\left(a_{KI\alpha}\,a_{IJ\beta}-a_{KI\beta}\,a_{IJ\alpha}\right) - \tfrac{1}{2}ig\,P_{JK}^{\alpha\beta}\left(A_{KI\alpha}A_{IJ\beta}-A_{KI\beta}A_{IJ\alpha}\right).
\end{aligned}$$

For the symmetric part of $P_{JK}^{\alpha\beta}$, we obtain

$$\begin{aligned}
&P_{JK}^{(\alpha\beta)}\left(\frac{\partial u_{KL}}{\partial x^\beta}a_{LI\alpha}\,u_{IJ}^* + u_{KL}a_{LI\alpha}\frac{\partial u_{IJ}^*}{\partial x^\beta}+\frac{1}{ig}\frac{\partial u_{KI}}{\partial x^\alpha}\frac{\partial u_{IJ}^*}{\partial x^\beta}+\frac{1}{ig}\frac{\partial^2 u_{KI}}{\partial x^\alpha \partial x^\beta}u_{IJ}^*\right)\\
&= P_{JK}^{(\alpha\beta)}\left(\frac{\partial A_{KJ\alpha}}{\partial x^\beta}-u_{KL}\frac{\partial a_{LI\alpha}}{\partial x^\beta}u_{IJ}^*\right)\\
&= \tfrac{1}{2}P_{JK}^{\alpha\beta}\left(\frac{\partial A_{KJ\alpha}}{\partial x^\beta}+\frac{\partial A_{KJ\beta}}{\partial x^\alpha}\right) - \tfrac{1}{2}p_{JK}^{\alpha\beta}\left(\frac{\partial a_{KJ\alpha}}{\partial x^\beta}+\frac{\partial a_{KJ\beta}}{\partial x^\alpha}\right).
\end{aligned}$$

In summary, by inserting the transformation rules into Eq. (52), the divergence of the explicitly x^μ-dependent terms of F_2^μ — and hence the difference of transformed and original Hamiltonians — can be expressed completely in terms of the canonical variables as

$$\begin{aligned}
\left.\frac{\partial F_2^\alpha}{\partial x^\alpha}\right|_{\mathrm{expl}} = ig\Big[&\left(\overline{\Pi}_K^\alpha\Phi_J-\overline{\Phi}_K\Pi_J^\alpha\right)A_{KJ\alpha}-\left(\overline{\pi}_K^\alpha\phi_J-\overline{\phi}_K\pi_J^\alpha\right)a_{KJ\alpha}\\
&-\tfrac{1}{2}P_{JK}^{\alpha\beta}\left(A_{KI\alpha}A_{IJ\beta}-A_{KI\beta}A_{IJ\alpha}\right)+\tfrac{1}{2}p_{JK}^{\alpha\beta}\left(a_{KI\alpha}\,a_{IJ\beta}-a_{KI\beta}\,a_{IJ\alpha}\right)\Big]\\
&+\tfrac{1}{2}P_{JK}^{\alpha\beta}\left(\frac{\partial A_{KJ\alpha}}{\partial x^\beta}+\frac{\partial A_{KJ\beta}}{\partial x^\alpha}\right)-\tfrac{1}{2}p_{JK}^{\alpha\beta}\left(\frac{\partial a_{KJ\alpha}}{\partial x^\beta}+\frac{\partial a_{KJ\beta}}{\partial x^\alpha}\right).
\end{aligned}$$

We observe that *all* u_{IJ}-dependencies of Eq. (52) were expressed *symmetrically* in terms of the original and transformed complex scalar fields ϕ_J,Φ_J and 4-vector gauge fields $\boldsymbol{a}_{JK},\boldsymbol{A}_{JK}$, in conjunction with their respective canonical momenta. Consequently, an amended Hamiltonian $\mathscr{H}_2$ of the form

$$\begin{aligned}
\mathscr{H}_2 = \mathscr{H}(\boldsymbol{\pi},\boldsymbol{\phi},x)&+ig\left(\overline{\pi}_K^\alpha\phi_J-\overline{\phi}_K\pi_J^\alpha\right)a_{KJ\alpha}\\
&-\tfrac{1}{2}ig\,p_{JK}^{\alpha\beta}\left(a_{KI\alpha}\,a_{IJ\beta}-a_{KI\beta}\,a_{IJ\alpha}\right)+\tfrac{1}{2}p_{JK}^{\alpha\beta}\left(\frac{\partial a_{KJ\alpha}}{\partial x^\beta}+\frac{\partial a_{KJ\beta}}{\partial x^\alpha}\right) \qquad (53)
\end{aligned}$$

is then transformed according to the general rule (11)

$$\mathscr{H}_2' = \mathscr{H}_2 + \left.\frac{\partial F_2^\alpha}{\partial x^\alpha}\right|_{\mathrm{expl}}$$

into the new Hamiltonian

$$\begin{aligned}\mathscr{H}_2' = \mathscr{H}(\boldsymbol{\Pi},\boldsymbol{\Phi},x) + ig\left(\overline{\Pi}_K^\alpha \Phi_J - \overline{\Phi}_K \Pi_J^\alpha\right) A_{KJ\alpha} \\ - \tfrac{1}{2} ig P_{JK}^{\alpha\beta}\left(A_{KI\alpha}A_{IJ\beta} - A_{KI\beta}A_{IJ\alpha}\right) + \tfrac{1}{2} P_{JK}^{\alpha\beta}\left(\frac{\partial A_{KJ\alpha}}{\partial x^\beta} + \frac{\partial A_{KJ\beta}}{\partial x^\alpha}\right).\end{aligned} \tag{54}$$

The entire transformation is thus *form-conserving* provided that the original Hamiltonian $\mathscr{H}(\boldsymbol{\pi},\boldsymbol{\phi},x)$ is also form-invariant if expressed in terms of the new fields, $\mathscr{H}(\boldsymbol{\Pi},\boldsymbol{\Phi},x)$, according to the transformation rules (44). In other words, $\mathscr{H}(\boldsymbol{\pi},\boldsymbol{\phi},x)$ must be form-invariant under the corresponding *global* gauge transformation.

In order for the presented transformation theory to be *physically consistent*, we must ensure that the *canonical field equations* for the derivatives of the gauge fields that follow from the final form-invariant amended Hamiltonians, $\mathscr{H}_3$ and $\mathscr{H}_3'$, coincide with the derivatives of the transformation rules for the gauge fields from Eq. (48). As it turns out, the form-invariant Hamiltonians $\mathscr{H}_2$ from Eq. (53) and $\mathscr{H}_2'$ from Eq. (54) must be further amended by terms $\mathscr{H}_{\mathrm{dyn}}(\boldsymbol{p})$ and $\mathscr{H}_{\mathrm{dyn}}'(\boldsymbol{P})$ that describe the dynamics of the free 4-vector gauge fields, $\boldsymbol{a}_{KJ}$ and $\boldsymbol{A}_{KJ}$, respectively

$$\begin{aligned}\mathscr{H}_3' = \mathscr{H}(\boldsymbol{\Pi},\boldsymbol{\Phi},x) + \mathscr{H}_{\mathrm{dyn}}'(\boldsymbol{P}) + ig\left(\overline{\Pi}_K^\alpha \Phi_J - \overline{\Phi}_K \Pi_J^\alpha\right) A_{KJ\alpha} \\ - \tfrac{1}{2} ig P_{JK}^{\alpha\beta}\left(A_{KI\alpha}A_{IJ\beta} - A_{KI\beta}A_{IJ\alpha}\right) + \tfrac{1}{2} P_{JK}^{\alpha\beta}\left(\frac{\partial A_{KJ\alpha}}{\partial x^\beta} + \frac{\partial A_{KJ\beta}}{\partial x^\alpha}\right).\end{aligned}$$

Of course, $\mathscr{H}_{\mathrm{dyn}}'(\boldsymbol{P})$ must be form-invariant as well in order to ensure the form-invariance of the *final* amended Hamiltonians, $\mathscr{H}_3$ and $\mathscr{H}_3'$. To derive $\mathscr{H}_{\mathrm{dyn}}'$, we set up the first canonical equation

$$\frac{\partial A_{KJ\mu}}{\partial x^\nu} = \frac{\partial \mathscr{H}_3'}{\partial P_{JK}^{\mu\nu}} = \frac{\partial \mathscr{H}_{\mathrm{dyn}}'}{\partial P_{JK}^{\mu\nu}} - \tfrac{1}{2} ig\left(A_{KI\mu}A_{IJ\nu} - A_{KI\nu}A_{IJ\mu}\right) + \tfrac{1}{2}\left(\frac{\partial A_{KJ\mu}}{\partial x^\nu} + \frac{\partial A_{KJ\nu}}{\partial x^\mu}\right).$$

Applying now the transformation rules (48), for the gauge fields $\boldsymbol{A}_{KJ}$, we find after straightforward calculation

$$\begin{aligned}\frac{\partial \mathscr{H}_{\mathrm{dyn}}'}{\partial P_{JK}^{\mu\nu}} &= \tfrac{1}{2}\left(\frac{\partial A_{KJ\mu}}{\partial x^\nu} - \frac{\partial A_{KJ\nu}}{\partial x^\mu}\right) + \tfrac{1}{2} ig\left(A_{KI\mu}A_{IJ\nu} - A_{KI\nu}A_{IJ\mu}\right) \\ &= \tfrac{1}{2} u_{KL}\left[\frac{\partial a_{LN\mu}}{\partial x^\nu} - \frac{\partial a_{LN\nu}}{\partial x^\mu} + ig\left(a_{LI\mu}a_{IN\nu} - a_{LI\nu}a_{IN\mu}\right)\right] u_{NJ}^* \\ &= u_{KL}\,\frac{\partial \mathscr{H}_{\mathrm{dyn}}}{\partial p_{NL}^{\mu\nu}}\,u_{NJ}^*.\end{aligned}$$

The derivatives of $\mathscr{H}_{\mathrm{dyn}}$ and $\mathscr{H}'_{\mathrm{dyn}}$ obviously transform like the canonical momenta, as stated in Eq. (51). Consequently, these expressions must be identified with $p_{KJ\nu\mu}$ and $P_{KJ\nu\mu}$, respectively

$$\frac{\partial \mathscr{H}'_{\mathrm{dyn}}}{\partial P_{JK}^{\mu\nu}} = -\tfrac{1}{2} P_{KJ\mu\nu}, \qquad \frac{\partial \mathscr{H}_{\mathrm{dyn}}}{\partial p_{JK}^{\mu\nu}} = -\tfrac{1}{2} p_{KJ\mu\nu}.$$

This means, in turn, that $\mathscr{H}'_{\mathrm{dyn}}$ and thus $\mathscr{H}_{\mathrm{dyn}}$ are given by

$$\mathscr{H}'_{\mathrm{dyn}}(\boldsymbol{P}) = -\tfrac{1}{4} P_{JK}^{\alpha\beta} P_{KJ\alpha\beta}, \qquad \mathscr{H}_{\mathrm{dyn}}(\boldsymbol{p}) = -\tfrac{1}{4} p_{JK}^{\alpha\beta} p_{KJ\alpha\beta}. \tag{55}$$

We conclude that Eq. (55) is the only choice for the free dynamics term of the gauge fields in order for the entire gauge transformation formalism to be consistent. Thus, the amended Hamiltonian $\mathscr{H}_3$ given by

$$\begin{aligned}
\mathscr{H}_3 =& \mathscr{H} + \mathscr{H}_{\mathrm{g}} \\
\mathscr{H}_{\mathrm{g}} =& ig\left(\overline{\pi}_K^{\alpha}\phi_J - \overline{\phi}_K \pi_J^{\alpha}\right) a_{KJ\alpha} - \tfrac{1}{4} p_{JK}^{\alpha\beta}\, p_{KJ\alpha\beta} \\
& - \tfrac{1}{2} ig\, p_{JK}^{\alpha\beta}\left(a_{KI\alpha}\, a_{IJ\beta} - a_{KI\beta}\, a_{IJ\alpha}\right) + \tfrac{1}{2} p_{JK}^{\alpha\beta}\left(\frac{\partial a_{KJ\alpha}}{\partial x^{\beta}} + \frac{\partial a_{KJ\beta}}{\partial x^{\alpha}}\right).
\end{aligned} \tag{56}$$

4.2.3 Inserting the gauge-invariant Hamiltonian $\mathscr{H}_3$ into the action integral

With gauge fields $a_{KJ\mu}$ and their conjugates, $p_{JK}^{\mu\nu}$, the additional dynamical quantities of the locally gauge-invariant system, the amended action integral from Eq. (6) reads

$$S = \int_R \left(\pi_I^{\beta} \frac{\partial \phi_I}{\partial x^{\beta}} + p_{JK}^{\alpha\beta} \frac{\partial a_{KJ\alpha}}{\partial x^{\beta}} - \mathscr{H}_3 \right) d^4x. \tag{57}$$

Inserting the explicit representation of $\mathscr{H}_3$ from Eq. (56) then yields the following non-standard form of the action integral

$$S = \int_R \left[\pi_I^{\beta} \frac{\partial \phi_I}{\partial x^{\beta}} + \tfrac{1}{2} p_{JK}^{\alpha\beta} \left(\frac{\partial a_{KJ\alpha}}{\partial x^{\beta}} - \frac{\partial a_{KJ\beta}}{\partial x^{\alpha}} \right) - \mathscr{H}_4 \right] d^4x, \tag{58}$$

with

$$\begin{aligned}
\mathscr{H}_4 = \mathscr{H} &+ ig\left(\overline{\pi}_K^{\alpha}\phi_J - \overline{\phi}_K \pi_J^{\alpha}\right) a_{KJ\alpha} - \tfrac{1}{4} p_{JK}^{\alpha\beta}\, p_{KJ\alpha\beta} \\
&- \tfrac{1}{2} ig\, p_{JK}^{\alpha\beta}\left(a_{KI\alpha}\, a_{IJ\beta} - a_{KI\beta}\, a_{IJ\alpha}\right)
\end{aligned} \tag{59}$$

We observe in Eq. (58) that only the *skew-symmetric* part of $p_{JK}^{\alpha\beta}$ in α,β contributes to the action S. In this form, the action integral is manifestly *form-invariant* under a local $\mathrm{U}(N)$ symmetry transformation (42) of the fields $\boldsymbol{\phi}, \overline{\boldsymbol{\phi}}$, and $\boldsymbol{a}$. The canonical equation for the derivative of the gauge fields is now obtained directly from (58) as

$$\frac{1}{2}\left(\frac{\partial a_{KJ\alpha}}{\partial x^{\beta}}-\frac{\partial a_{KJ\beta}}{\partial x^{\alpha}}\right)=\frac{\partial \mathscr{H}_4}{\partial p_{JK}^{\alpha\beta}}. \tag{60}$$

With $\mathscr{H}_4$ from Eq. (59), this reads in explicit form

$$\frac{1}{2}\left(\frac{\partial a_{KJ\alpha}}{\partial x^{\beta}}-\frac{\partial a_{KJ\beta}}{\partial x^{\alpha}}\right)=-\tfrac{1}{2}p_{KJ\alpha\beta}-\tfrac{1}{2}ig\left(a_{KI\alpha}\,a_{IJ\beta}-a_{KI\beta}\,a_{IJ\alpha}\right),$$

hence

$$p_{KJ\mu\nu}=\frac{\partial a_{KJ\nu}}{\partial x^{\mu}}-\frac{\partial a_{KJ\mu}}{\partial x^{\nu}}+ig\left(a_{KI\nu}\,a_{IJ\mu}-a_{KI\mu}\,a_{IJ\nu}\right). \tag{61}$$

We observe that $p_{KJ\mu\nu}$ occurs to be skew-symmetric in the indices μ,ν. Here, this feature emerges from the canonical formalism and does not need to be postulated. Yet, the information on the actual form of the action integral (58), hence on the skew-symmetry of $p_{KJ\mu\nu}$ must be supplemented in addition to the specification of the final form of the locally gauge-invariant Hamiltonian $\mathscr{H}_4$

$$\begin{aligned}
&\mathscr{H}_4=\mathscr{H}+\mathscr{H}_{\mathrm{g}}, \qquad p_{JK}^{\mu\nu}=-p_{JK}^{\nu\mu}\\
&\mathscr{H}_{\mathrm{g}}=-\tfrac{1}{4}p_{JK}^{\alpha\beta}\,p_{KJ\alpha\beta}+ig\left(\overline{\pi}_K^{\alpha}\,a_{KJ\alpha}\,\phi_J-\overline{\phi}_K\,a_{KJ\alpha}\,\pi_J^{\alpha}-p_{JK}^{\alpha\beta}\,a_{KI\alpha}\,a_{IJ\beta}\right).
\end{aligned} \tag{62}$$

Thus, $\mathscr{H}_{\mathrm{g}}$ describes the dynamics of *massless* 4-vector fields $\boldsymbol{a}_{IK}$, namely, their couplings to the base fields ϕ_I as well as their self-couplings. This is the final result of the general local U(N) gauge transformation theory in the Hamiltonian formalism.

From the locally gauge-invariant Hamiltonian (62), the canonical equation for the base fields ϕ_I is given by

$$\begin{aligned}
\left.\frac{\partial \phi_I}{\partial x^{\mu}}\right|_{\mathscr{H}_4}&=\frac{\partial \mathscr{H}_4}{\partial \overline{\pi}_I^{\mu}}=\frac{\partial \mathscr{H}}{\partial \overline{\pi}_I^{\mu}}+ig\,a_{IJ\mu}\,\phi_J\\
&=\left.\frac{\partial \phi_I}{\partial x^{\mu}}\right|_{\mathscr{H}}+ig\,a_{IJ\mu}\,\phi_J.
\end{aligned}$$

This is exactly the so-called "minimum coupling rule", which is also referred to as the "gauge covariant derivative". Remarkably, in the canonical formalism this result is *derived*, hence does not need to be postulated. It is commonly assumed that the quantities $a_{IJ\mu}$ exhibit *elementary* fields themselves, hence that the $a_{IJ\mu}$ are not compositions of elementary fields.

4.3 Locally gauge-invariant Lagrangian

4.3.1 Legendre transformation for a general system Hamiltonian

The equivalent gauge-invariant Lagrangian $\mathscr{L}_3$ is derived by Legendre-transforming the gauge-invariant Hamiltonian $\mathscr{H}_3$, defined in Eqs. (56)

$$\mathscr{L}_3 = \overline{\pi}_K^\alpha \frac{\partial \phi_K}{\partial x^\alpha} + \frac{\partial \overline{\phi}_K}{\partial x^\alpha} \pi_K^\alpha + p_{JK}^{\alpha\beta} \frac{\partial a_{KJ\alpha}}{\partial x^\beta} - \mathscr{H}_3, \qquad \mathscr{H}_3 = \mathscr{H} + \mathscr{H}_\mathrm{g}.$$

With $p_{JK}^{\mu\nu}$ from Eq. (61) and $\mathscr{H}_\mathrm{g}$ from Eq. (56), we thus have

$$\begin{aligned} p_{JK}^{\alpha\beta} \frac{\partial a_{KJ\alpha}}{\partial x^\beta} - \mathscr{H}_\mathrm{g} &= \tfrac{1}{2} p_{JK}^{\alpha\beta} \left(\frac{\partial a_{KJ\alpha}}{\partial x^\beta} - \frac{\partial a_{KJ\beta}}{\partial x^\alpha} \right) + \tfrac{1}{2} p_{JK}^{\alpha\beta} \left(\frac{\partial a_{KJ\alpha}}{\partial x^\beta} + \frac{\partial a_{KJ\beta}}{\partial x^\alpha} \right) - \mathscr{H}_\mathrm{g} \\ &= -\tfrac{1}{2} p_{JK}^{\alpha\beta}\, p_{KJ\alpha\beta} - \tfrac{1}{2} ig\, p_{JK}^{\alpha\beta} \left(a_{KI\alpha}\, a_{IJ\beta} - a_{KI\beta}\, a_{IJ\alpha} \right) \\ &\quad + \tfrac{1}{2} p_{JK}^{\alpha\beta} \left(\frac{\partial a_{KJ\alpha}}{\partial x^\beta} + \frac{\partial a_{KJ\beta}}{\partial x^\alpha} \right) - \mathscr{H}_\mathrm{g} \\ &= ig \left(\overline{\pi}_K^\alpha \phi_J - \overline{\phi}_K \pi_J^\alpha \right) a_{KJ\alpha} - \tfrac{1}{4} p_{JK}^{\alpha\beta}\, p_{KJ\alpha\beta}. \end{aligned}$$

The locally gauge-invariant Lagrangian $\mathscr{L}_3$ for any given globally gauge-invariant system Hamiltonian $\mathscr{H}(\overline{\phi}_I, \phi_I, \overline{\boldsymbol{\pi}}_I, \boldsymbol{\pi}_I, x)$ is then

$$\begin{aligned} \mathscr{L}_3 &= -\tfrac{1}{4} p_{JK}^{\alpha\beta}\, p_{KJ\alpha\beta} - ig \left(\overline{\pi}_K^\alpha \phi_J - \overline{\phi}_K \pi_J^\alpha \right) a_{KJ\alpha} + \overline{\pi}_K^\alpha \frac{\partial \phi_K}{\partial x^\alpha} + \frac{\partial \overline{\phi}_K}{\partial x^\alpha} \pi_K^\alpha - \mathscr{H} \qquad (63) \\ &= -\tfrac{1}{4} p_{JK}^{\alpha\beta}\, p_{KJ\alpha\beta} + \overline{\pi}_K^\alpha \left(\frac{\partial \phi_K}{\partial x^\alpha} - ig\, a_{KJ\alpha} \phi_J \right) + \left(\frac{\partial \overline{\phi}_K}{\partial x^\alpha} + ig\, \overline{\phi}_J a_{JK\alpha} \right) \pi_K^\alpha - \mathscr{H}. \end{aligned}$$

As implied by the Lagrangian formalism, the dynamical variables are given by both the fields, $\overline{\phi}_K$, ϕ_J, and $a_{KJ\alpha}$, in conjunction with their respective partial derivatives with respect to the independent variables, x^μ. Therefore, the $\boldsymbol{p}_{KJ}$ in $\mathscr{L}_3$ from Eq. (63) are now merely abbreviations for a combination of the Lagrangian dynamical variables. Independently of the given system Hamiltonian $\mathscr{H}$, the correlation of the $\boldsymbol{p}_{KJ}$ with the gauge fields $\boldsymbol{a}_{KJ}$ and their derivatives is given by the first canonical equation (61).

The correlation of the momenta $\boldsymbol{\pi}_I, \overline{\boldsymbol{\pi}}_I$ to the base fields $\phi_I, \overline{\phi}_I$ and their derivatives are derived from Eq. (63) for the given system Hamiltonian $\mathscr{H}$ via

$$\frac{\partial \mathscr{H}}{\partial \overline{\pi}_I^\mu} = \frac{\partial \phi_I}{\partial x^\mu} - ig\, a_{IJ\mu} \phi_J, \qquad \frac{\partial \mathscr{H}}{\partial \pi_I^\mu} = \frac{\partial \overline{\phi}_I}{\partial x^\mu} + ig\, \overline{\phi}_J\, a_{JI\mu}. \qquad (64)$$

Thus, for any *globally* gauge-invariant system Hamiltonian $\mathscr{H}(\overline{\phi}_I, \phi_I, \overline{\boldsymbol{\pi}}_I, \boldsymbol{\pi}_I, x)$, the amended Lagrangian $\mathscr{L}_3$ from Eq. (63) with the $\overline{\boldsymbol{\pi}}_I, \boldsymbol{\pi}_I$ to be determined from Eqs. (64) describes in the Lagrangian formalism the associated physical system that is invariant under *local* gauge transformations.

4.3.2 Klein-Gordon system Hamiltonian

The generalized Klein-Gordon Hamiltonian $\mathscr{H}_\mathrm{KG}$ describing N complex scalar fields ϕ_I that are associated with equal masses m is

$$\mathscr{H}_{\mathrm{KG}}(\boldsymbol{\pi}_\mu, \boldsymbol{\pi}^{*\mu}, \boldsymbol{\phi}, \boldsymbol{\phi}^*) = \pi^*_{I\alpha}\pi^\alpha_I + m^2\,\phi^*_I\phi_I.$$

This Hamiltonian is clearly form-invariant under the global gauge-transformation defined by Eqs. (44). Following Eqs. (56) and (62), the corresponding locally gauge-invariant Hamiltonian $\mathscr{H}_{3,\mathrm{KG}}$ is then

$$\begin{aligned}\mathscr{H}_{3,\mathrm{KG}} = {} & \pi^*_{I\alpha}\pi^\alpha_I + m^2\,\phi^*_I\phi_I - \tfrac{1}{4}p^{\alpha\beta}_{JK}\,p_{KJ\alpha\beta} \\ & + ig\left(\pi^{*\alpha}_K\,a_{KJ\alpha}\,\phi_J - \phi^*_K\,a_{KJ\alpha}\,\pi^\alpha_J - p^{\alpha\beta}_{JK}\,a_{KI\alpha}\,a_{IJ\beta}\right), \qquad p^{\mu\nu}_{JK} \stackrel{!}{=} -p^{\nu\mu}_{JK}.\end{aligned}$$

To derive the equivalent locally gauge-invariant Lagrangian $\mathscr{L}_{3,\mathrm{KG}}$, we set up the first canonical equation for the gauge-invariant Hamiltonian $\mathscr{H}_{3,\mathrm{KG}}$ of our actual example

$$\frac{\partial\phi_I}{\partial x^\mu} = \frac{\partial\mathscr{H}_{3,\mathrm{KG}}}{\partial\pi^{*\mu}_I} = \pi_{I\mu} + ig\,a_{IJ\mu}\phi_J, \qquad \frac{\partial\phi^*_I}{\partial x^\mu} = \frac{\partial\mathscr{H}_{3,\mathrm{KG}}}{\partial\pi^\mu_I} = \pi^*_{I\mu} - ig\,\phi^*_J\,a_{JI\mu}.$$

Inserting $\partial\phi_I/\partial x^\mu$ and $\partial\phi^*_I/\partial x^\mu$ into Eq. (63), we directly encounter the *locally* gauge-invariant Lagrangian $\mathscr{L}_{3,\mathrm{KG}}$ as

$$\mathscr{L}_{3,\mathrm{KG}} = \pi^*_{I\alpha}\pi^\alpha_I - m^2\,\phi^*_I\phi_I - \tfrac{1}{4}p^{\alpha\beta}_{JK}\,p_{KJ\alpha\beta},$$

with the abbreviations

$$\begin{aligned}\pi_{I\mu} &= \frac{\partial\phi_I}{\partial x^\mu} - ig\,a_{IJ\mu}\phi_J, \qquad \pi^*_{I\mu} = \frac{\partial\phi^*_I}{\partial x^\mu} + ig\,\phi^*_J\,a_{JI\mu} \\ p_{KJ\mu\nu} &= \frac{\partial a_{KJ\nu}}{\partial x^\mu} - \frac{\partial a_{KJ\mu}}{\partial x^\nu} + ig\left(a_{KI\nu}\,a_{IJ\mu} - a_{KI\mu}\,a_{IJ\nu}\right).\end{aligned}$$

In a more explicit form, $\mathscr{L}_{3,\mathrm{KG}}$ is thus given by

$$\mathscr{L}_{3,\mathrm{KG}} = \left(\frac{\partial\phi^*_I}{\partial x^\alpha} + ig\,\phi^*_J\,a_{JI\alpha}\right)\left(\frac{\partial\phi_I}{\partial x_\alpha} - ig\,a^\alpha_{IJ}\phi_J\right) - m^2\,\phi^*_I\phi_I - \tfrac{1}{4}p^{\alpha\beta}_{JK}\,p_{KJ\alpha\beta}$$

The expressions in the parentheses represent the "minimum coupling rule," which appears here as the transition from the *kinetic* momenta to the *canonical* momenta. By inserting $\mathscr{L}_{3,\mathrm{KG}}$ into the Euler-Lagrange equations, and $\mathscr{H}_{3,\mathrm{KG}}$ into the canonical equations, we may convince ourselves that the emerging field equations for ϕ^*_I, ϕ_I, and $\boldsymbol{a}_{JK}$ agree. This means that $\mathscr{H}_{3,\mathrm{KG}}$ and $\mathscr{L}_{3,\mathrm{KG}}$ describe the *same physical system*.

4.3.3 Dirac system Hamiltonian

The generalized Dirac Hamiltonian (31) describing N spin-$\frac{1}{2}$ fields, each of them being associated with the same mass m,

$$\mathscr{H}_{\mathrm{D}} = \left(\overline{\pi}_I^{\alpha} - \frac{i}{2}\overline{\psi}_I\gamma^{\alpha}\right)\frac{3\tilde{m}\tau_{\alpha\beta}}{i}\left(\pi_I^{\beta} + \frac{i}{2}\gamma^{\beta}\psi_I\right) + m\overline{\psi}_I\psi_I, \quad \tau_{\mu\alpha}\sigma^{\alpha\nu} = \delta_{\mu}^{\nu}\mathbb{1}$$

is form-invariant under global gauge transformations (44) since

$$\begin{aligned}
\mathscr{H}_{\mathrm{D}}' &= \left(\overline{\Pi}_K^{\alpha} - \frac{i}{2}\overline{\Psi}_K\gamma^{\alpha}\right)\frac{3\tilde{m}\tau_{\alpha\beta}}{i}\underbrace{u_{KI}u_{IJ}^{*}}_{=\delta_{KJ}}\left(\Pi_J^{\beta} + \frac{i}{2}\gamma^{\beta}\Psi_J\right) + m\overline{\Psi}_K\underbrace{u_{KI}u_{IJ}^{*}}_{=\delta_{KJ}}\Psi_J \\
&= \left(\overline{\Pi}_K^{\alpha} - \frac{i}{2}\overline{\Psi}_K\gamma^{\alpha}\right)\frac{3\tilde{m}\tau_{\alpha\beta}}{i}\left(\Pi_K^{\beta} + \frac{i}{2}\gamma^{\beta}\Psi_K\right) + m\overline{\Psi}_K\Psi_K.
\end{aligned}$$

Again, the corresponding locally gauge-invariant Hamiltonian $\mathscr{H}_{3,\mathrm{D}}$ is found by adding the gauge Hamiltonian $\mathscr{H}_{\mathrm{g}}$ from Eq. (62)

$$\begin{aligned}
\mathscr{H}_{3,\mathrm{D}} = &\left(\overline{\pi}_I^{\alpha} - \frac{i}{2}\overline{\psi}_I\gamma^{\alpha}\right)\frac{3\tilde{m}\tau_{\alpha\beta}}{i}\left(\pi_I^{\beta} + \frac{i}{2}\gamma^{\beta}\psi_I\right) + m\overline{\psi}_I\psi_I \\
&- \tfrac{1}{4}p_{JK}^{\alpha\beta}\,p_{KJ\alpha\beta} + ig\left(\overline{\pi}_K^{\alpha}\,\psi_J - \overline{\psi}_K\,\pi_J^{\alpha} + p_{JI}^{\alpha\beta}\,a_{IK\beta}\right)a_{KJ\alpha}.
\end{aligned} \tag{65}$$

The correlation of the canonical momenta $\overline{\pi}_I^{\mu}, \pi_I^{\mu}$ with the base fields $\overline{\psi}_I, \psi_I$ and their derivatives follows again from first canonical equation for $\mathscr{H}_{3,\mathrm{D}}$

$$\begin{aligned}
\frac{\partial\psi_I}{\partial x^{\mu}} &= \frac{\partial\mathscr{H}_{3,\mathrm{D}}}{\partial\overline{\pi}_I^{\mu}} = \frac{3\tilde{m}\tau_{\mu\beta}}{i}\left(\pi_I^{\beta} + \frac{i}{2}\gamma^{\beta}\psi_I\right) + ig\,a_{IJ\mu}\,\psi_J \\
\frac{\partial\overline{\psi}_I}{\partial x^{\mu}} &= \frac{\partial\mathscr{H}_{3,\mathrm{D}}}{\partial\pi_I^{\mu}} = \left(\overline{\pi}_I^{\alpha} - \frac{i}{2}\overline{\psi}_I\gamma^{\alpha}\right)\frac{3\tilde{m}\tau_{\alpha\mu}}{i} - ig\,\overline{\psi}_J\,a_{JI\mu}.
\end{aligned} \tag{66}$$

Inserting $\partial\psi_I/\partial x^{\mu}$ and $\partial\overline{\psi}_I/\partial x^{\mu}$ into Eq. (63), we encounter the related *locally* gauge-invariant Lagrangian $\mathscr{L}_{3,\mathrm{D}}$ in the intermediate form

$$\mathscr{L}_{3,\mathrm{D}} = -\tfrac{1}{4}p_{JK}^{\alpha\beta}\,p_{KJ\alpha\beta} + \overline{\pi}_I^{\alpha}\frac{3\tilde{m}\tau_{\alpha\beta}}{i}\pi_I^{\beta} - (m-\tilde{m})\,\overline{\psi}_I\psi_I, \tag{67}$$

with the momenta $\overline{\pi}_I^{\alpha}, \pi_I^{\beta}$ determined by Eqs. (66). We can finally eliminate the momenta of the base fields in order to express $\mathscr{L}_{3,\mathrm{D}}$ completely in Lagrangian variables. To this end, we solve Eqs. (66) for the momenta

$$\begin{aligned}
\frac{3\tilde{m}\tau_{\alpha\beta}}{i}\pi_I^{\beta} &= \frac{\partial\psi_I}{\partial x^{\alpha}} - ig\,a_{IK\alpha}\,\psi_K + \frac{i\tilde{m}}{2}\gamma_{\alpha}\,\psi_I \\
\overline{\pi}_I^{\alpha} &= \left(\frac{\partial\overline{\psi}_I}{\partial x^{\beta}} + ig\,\overline{\psi}_J a_{JI\beta} - \frac{i\tilde{m}}{2}\overline{\psi}_I\gamma_{\beta}\right)\frac{i\sigma^{\beta\alpha}}{3\tilde{m}}.
\end{aligned}$$

Then

$$\begin{aligned}
&\overline{\pi}_I^{\alpha}\,\frac{3\tilde{m}\tau_{\alpha\beta}}{i}\,\pi_I^{\beta}\\
&\quad=\left(\frac{\partial\overline{\psi}_I}{\partial x^{\alpha}}+ig\,\overline{\psi}_J a_{JI\alpha}-\frac{i\tilde{m}}{2}\overline{\psi}_I\gamma_{\alpha}\right)\frac{i\sigma^{\alpha\beta}}{3\tilde{m}}\left(\frac{\partial\psi_I}{\partial x^{\beta}}-ig\,a_{IK\beta}\,\psi_K+\frac{i\tilde{m}}{2}\gamma_{\beta}\,\psi_I\right).
\end{aligned}$$

Inserting this expression into (67) yields the final form of the locally gauge-invariant Dirac Lagrangian

$$\begin{aligned}
\mathscr{L}_{3,\mathrm{D}}=&\left(\frac{\partial\overline{\psi}_I}{\partial x^{\alpha}}+ig\,\overline{\psi}_J a_{JI\alpha}-\frac{i\tilde{m}}{2}\overline{\psi}_I\gamma_{\alpha}\right)\frac{i\sigma^{\alpha\beta}}{3\tilde{m}}\left(\frac{\partial\psi_I}{\partial x^{\beta}}-ig\,a_{IK\beta}\,\psi_K+\frac{i\tilde{m}}{2}\gamma_{\beta}\,\psi_I\right)\\
&-\tfrac{1}{4}p_{JK}^{\alpha\beta}\,p_{KJ\alpha\beta}-(m-\tilde{m})\,\overline{\psi}_I\psi_I.
\end{aligned}$$

After expanding, this Lagrangian writes equivalently

$$\begin{aligned}
\mathscr{L}_{3,\mathrm{D}}=&\frac{i}{2}\overline{\psi}_I\gamma^{\alpha}\left(\frac{\partial\psi_I}{\partial x^{\alpha}}-ig\,a_{IK\alpha}\psi_K\right)-\frac{i}{2}\left(\frac{\partial\overline{\psi}_I}{\partial x^{\alpha}}+ig\,\overline{\psi}_J a_{JI\alpha}\right)\gamma^{\alpha}\psi_I-m\,\overline{\psi}_I\psi_I\\
&+\left(\frac{\partial\overline{\psi}_I}{\partial x^{\alpha}}+ig\,\overline{\psi}_J a_{JI\alpha}\right)\frac{i\sigma^{\alpha\beta}}{3\tilde{m}}\left(\frac{\partial\psi_I}{\partial x^{\beta}}-ig\,a_{IK\beta}\,\psi_K\right)-\tfrac{1}{4}p_{JK}^{\alpha\beta}\,p_{KJ\alpha\beta}.
\end{aligned}\tag{68}$$

The sums in parentheses can be regarded as a generalized "minimum coupling rule" for the actual case of a Dirac Lagrangian describing an N-tuple of spinors ψ_I. This also applies for the term involving $\sigma^{\alpha\beta}$ in Eq. (68) that emerges *in addition* to the conventional gauge-invariant Lagrangian if we start from the "regularized" Lagrangian from Eq. (27). This term is easily shown to be separately form-invariant under the combined local gauge transformation that is defined by Eqs. (42) and (48). Since the bilinear covariant $\overline{\psi}_J\sigma^{\alpha\beta}\psi_I$ transforms as a $(2,0)$-tensor, it is in particular also Lorentz-invariant. Physically, the term describes Pauli-coupling of the N-tuple of fermions ψ_I with the matrix of bosonic 4-vector gauge fields $a_{IK\mu}$.

The $\boldsymbol{p}_{KJ}$ stand for the combinations of the Lagrangian dynamical variables of the gauge fields from Eq. (61) that apply to all systems

$$p_{KJ\alpha\beta}=\frac{\partial a_{KJ\beta}}{\partial x^{\alpha}}-\frac{\partial a_{KJ\alpha}}{\partial x^{\beta}}+ig\left(a_{KI\beta}\,a_{IJ\alpha}-a_{KI\alpha}\,a_{IJ\beta}\right).$$

In order to set up the Euler-Lagrange equations for the locally gauge-invariant Lagrangian $\mathscr{L}_{3,\mathrm{D}}$ from Eq. (68), we first calculate the derivatives

$$\begin{aligned}
\frac{\partial}{\partial x^{\alpha}}\frac{\partial\mathscr{L}_{3,\mathrm{D}}}{\partial\left(\partial_{\alpha}\overline{\psi}_I\right)}&=-\frac{i}{2}\gamma^{\alpha}\frac{\partial\psi_I}{\partial x^{\alpha}}+\frac{i\sigma^{\alpha\beta}}{3\tilde{m}}\left(\cancel{\frac{\partial^2\psi_I}{\partial x^{\alpha}\partial x^{\beta}}}-ig\frac{\partial a_{IK\beta}}{\partial x^{\alpha}}\psi_K-ig\,a_{IK\beta}\frac{\partial\psi_K}{\partial x^{\alpha}}\right)\\
\frac{\partial\mathscr{L}_{3,\mathrm{D}}}{\partial\overline{\psi}_I}&=\frac{i}{2}\gamma^{\alpha}\frac{\partial\psi_I}{\partial x^{\alpha}}-m\psi_I+g\,a_{IK\alpha}\gamma^{\alpha}\psi_K\\
&\quad+\frac{i\sigma^{\alpha\beta}}{3\tilde{m}}\,ig\left(a_{IK\alpha}\frac{\partial\psi_K}{\partial x^{\beta}}-ig\,a_{IJ\alpha}a_{JK\beta}\,\psi_K\right)
\end{aligned}$$

and

$$\frac{\partial}{\partial x^\beta}\frac{\partial \mathscr{L}_{3,\mathrm{D}}}{\partial\left(\partial_\beta \psi_I\right)} = \frac{i}{2}\frac{\partial \overline{\psi}_I}{\partial x^\beta}\gamma^\beta + \left(\cancel{\frac{\partial^2 \overline{\psi}_I}{\partial x^\alpha \partial x^\beta}} + ig\,\frac{\partial \overline{\psi}_K}{\partial x^\beta} a_{KI\alpha} + ig\,\overline{\psi}_K \frac{\partial a_{KI\alpha}}{\partial x^\beta}\right)\frac{i\sigma^{\alpha\beta}}{3\tilde{m}}$$

$$\begin{aligned}\frac{\partial \mathscr{L}_{3,\mathrm{D}}}{\partial \psi_I} &= -\frac{i}{2}\frac{\partial \overline{\psi}_I}{\partial x^\alpha}\gamma^\alpha - m\overline{\psi}_I + g\,\overline{\psi}_K \gamma^\alpha a_{KI\alpha} \\ &\quad + ig\left(\frac{\partial \overline{\psi}_K}{\partial x^\beta} a_{KI\alpha} - ig\,\overline{\psi}_K a_{KJ\alpha} a_{JI\beta}\right)\frac{i\sigma^{\alpha\beta}}{3\tilde{m}}.\end{aligned}$$

The second derivative terms drop out due to the skew-symmetry of $\sigma^{\alpha\beta}$. The Euler-Lagrange equations thus finally emerge as

$$\begin{aligned} i\gamma^\alpha \frac{\partial \psi_I}{\partial x^\alpha} + g\gamma^\alpha a_{IK\alpha}\psi_K - m\,\psi_I + \frac{g}{6\tilde{m}} p_{IK\alpha\beta}\,\sigma^{\alpha\beta}\psi_K &= 0 \\ i\frac{\partial \overline{\psi}_I}{\partial x^\alpha}\gamma^\alpha - g\,\overline{\psi}_K a_{KI\alpha}\gamma^\alpha + m\,\overline{\psi}_I - \frac{g}{6\tilde{m}}\overline{\psi}_K \sigma^{\alpha\beta} p_{KI\alpha\beta} &= 0. \end{aligned} \tag{69}$$

We observe that our gauge-invariant Dirac equation contains an additional term that is proportional to $p_{IK\alpha\beta}$, hence to the canonical momenta of the gauge fields $a_{IK\alpha}$. This term is separately gauge invariant. We thus encounter the description of the coupling of the anomalous magnetic moments of the fermions to the gauge bosons, ie., a spin-gauge field coupling.

For the case of a system with a single spinor ψ representing a fermion of mass m, hence for the $\mathrm{U}(1)$ gauge group, we may set $\tilde{m} = m$. The locally gauge-invariant Dirac equation reduces to

$$i\gamma^\alpha \frac{\partial \psi}{\partial x^\alpha} + g\,\gamma^\alpha a_\alpha \psi - m\,\psi + \frac{\mu}{3}\left(\frac{\partial a_\beta}{\partial x^\alpha} - \frac{\partial a_\alpha}{\partial x^\beta}\right)\sigma^{\alpha\beta}\psi = 0,$$

with $\mu = g/2m$ the particle's magneton. The equation is obviously invariant under the combined gauge transformation of base and gauge fields

$$a_\mu(x) \mapsto A_\mu(x) = a_\mu(x) + \frac{1}{g}\frac{\partial \Lambda(x)}{\partial x^\mu}, \qquad \psi(x) \mapsto \Psi(x) = \psi(x)\,e^{i\Lambda(x)},$$

with the spin-gauge field coupling term being separately gauge invariant. Here, the additional term corresponds to a coupling of the electromagnetic field with the spin-induced magnetic moment of the fermion represented by ψ, commonly referred to as "Pauli-coupling" term. It is remarkable that Pauli interaction necessarily emerges in the context of the Hamiltonian formulation of gauge theory. In the Lagrangian description, we encounter this term only if the minimum coupling rule is applied to the *regularized* Lagrangian from Eq. (27).

4.3.4 Comparison with Pauli's amended Lagrangian

In this context, we remark that the Pauli-coupling term in the field equations (69) equally follows from the amended Dirac Lagrangian

$$\begin{aligned}\mathscr{L}_{3,\text{Pauli}} &= \frac{i}{2}\overline{\psi}_I\gamma^\alpha\left(\frac{\partial\psi_I}{\partial x^\alpha} - ig\,a_{IK\alpha}\psi_K\right) - \frac{i}{2}\left(\frac{\partial\overline{\psi}_I}{\partial x^\alpha} + ig\,\overline{\psi}_J a_{JI\alpha}\right)\gamma^\alpha\psi_I - m\overline{\psi}_I\psi_I \\ &\quad \pm \tfrac{1}{2}\ell\,\overline{\psi}_J p_{JK\alpha\beta}\sigma^{\alpha\beta}\psi_K - \tfrac{1}{4}p_{JK}^{\alpha\beta}p_{KJ\alpha\beta} \qquad (70)\end{aligned}$$

if we identify the coupling constant $\ell[L]$ with $\ell = g/m$. The addition of the term proportional to ℓ was proposed by Pauli[8]. Setting up the field equation for the charge conjugate solution $\overline{\psi}_I$, the sign of ℓ must taken to be negative. We may directly convince ourselves that the gauge-invariant Lagrangian from Eq. (68) and the amended Lagrangian (70) yield the same Pauli-coupling contributions to the *classical* field equations for both the $\psi_I, \overline{\psi}_I$ as well as for the gauge fields $a_{JK\mu}$

$$\begin{aligned}\mathscr{L}_{\text{int,Pauli}} &= \pm\ell\,\overline{\psi}_I\left(\frac{\partial a_{IJ\beta}}{\partial x^\alpha} + ig\,a_{IK\beta}\,a_{KJ\alpha}\right)\sigma^{\alpha\beta}\psi_J \\ \mathscr{L}_{\text{int}} &= -\frac{\ell}{ig}\left(\frac{\partial\overline{\psi}_I}{\partial x^\alpha} + ig\,\overline{\psi}_J a_{JI\alpha}\right)\sigma^{\alpha\beta}\left(\frac{\partial\psi_I}{\partial x^\beta} - ig\,a_{IK\beta}\psi_K\right).\end{aligned}$$

The interaction Lagrangian $\mathscr{L}_{\text{int,Pauli}}$ defines a non-minimal coupling. In contrast, with the locally gauge-invariant Lagrangian $\mathscr{L}_{3,\text{D}}$ from Eq. (68) containing the term $\mathscr{L}_{\text{int}}$, we have derived a description of Pauli coupling that conforms with the minimal-coupling rule. While both Lagrangians yield the same contributions to classical field equations, the subsequent interaction vertex factors are *different*. As the Pauli-coupling term $\mathscr{L}_{\text{int}}$ obeys the minimum coupling rule and follows from canonical gauge theory rather than being postulated, we may expect the interaction Lagrangian $\mathscr{L}_{\text{int}}$ to be the correct one. This is essential for the description of Pauli-type coupling effects in both QED as well as in QCD, where strong interactions of the colorless baryons and mesons arise from their nature being composed of colored quarks.

5 Conclusions

With the present paper, we have worked out a consistent local inertial frame description of the canonical formalism in the realm of covariant Hamiltonian field theory. On that basis, the Noether theorem as well as the idea of gauge theory — to amend the Hamiltonian of a given system in order to render the resulting system locally gauge invariant — could elegantly and most generally be formulated as particular canonical transformations.

Acknowledgements To the memory of my (J.S.) colleague and friend Dr. Claus Riedel (GSI), who contributed vitally to this work. Furthermore, the authors are indebted to Prof. Dr. Dr. hc. mult. Walter Greiner from the *Frankfurt Institute of Advanced Studies* (FIAS) for his long-standing hospitality, his critical comments and encouragement.

References

1. Th. De Donder, *Théorie Invariantive Du Calcul des Variations*, (Gaulthier-Villars & Cie., Paris, 1930).
2. H. Weyl, Geodesic Fields in the Calculus of Variation for Multiple Integrals, in *Annals of Mathematics* **36** 607 (1935).
3. cf, for instance: J. V. José and E. J. Saletan, *Classical Dynamics*, Cambridge University Press, Cambridge, 1998.
4. W. Greiner, B. Müller, and J. Rafelski, *Quantum Electrodynamics of Strong Fields*, (Springer-Verlag, Berlin, 1985).
5. S. Gasiorowicz, *Elementary particle physics*, (Wiley, New York, 1966).
6. J. von Rieth, The Hamilton-Jacobi theory of De Donder and Weyl applied to some relativistic field theories, in *J. Math. Phys.* **25**, 1102 (1984).
7. E. Noether, Nachr. Ges. Wiss. Göttingen, Math.-Phys. Kl. **57**, 235 (1918).
8. W. Pauli, *Rev. Mod. Phys.* **13**, 203–232 (1941).