EJTP **6, No. 21** (2009) 39–52 *Electronic Journal of Theoretical Physics*

Underdeterminacy and Redundance in Maxwell's Equations. I. The Origin of Gauge Freedom

Peter Enders*

Ahornallee 11, 15754 Senzig, Germany

Received 1 December 2008, Accepted 25 April 2009, Published 5 May 2009

Abstract: The gauge freedom in the electromagnetic potentials indicates an underdeterminacy in Maxwell's theory. This underdeterminacy will be found in Maxwell's (1864) original set of equations by means of Helmholtz's (1858) decomposition theorem. Since it concerns only the longitudinal electric field, it is intimately related to charge conservation, on the one hand, and to the transversality of free electromagnetic waves, on the other hand (as will be discussed in Pt. II). Exploiting the concept of Newtonian and Laplacian vector fields, the role of the static longitudinal component of the vector potential being *not* determined by Maxwell's equations, but important in quantum mechanics (Aharonov-Bohm effect) is elucidated. These results will be exploited in Pt.III for formulating a manifest gauge invariant canonical formulation of Maxwell's theory as input for developing Dirac's (1949) approach to special-relativistic dynamics.

Keywords: Electromagnetic Fields; Maxwell Equations; Helmholtz Decomposition; Gauge Theory; Aharonov-Bohm Effect
PACS (2008): 03.50.De; 41.20.-q; 03.50.-z; 03.50.Kk; 11.15.-q; 11.10.-z; 73.23.-b

1. Introduction

Traditionally, there are two main approaches to classical electromagnetism, *viz*,

(1) the experimental one going from the phenomena to the rationalized Maxwell equations (*eg*, Maxwell 1873, Mie 1941, Jackson 1999, Feynman, Leighton & Sands 2001);
(2) the deductive one deriving the phenomena from the rationalized Maxwell equations (*eg*, Hertz 1889, Lorentz 1909, Sommerfeld 2001).

* enders@dekasges.de

"Rationalized Maxwell equations" (Poynting 1884, Heaviside 1892) means the macroscopic Gauss' laws for the magnetic and electric field and the induction and flux laws (SI units).

$$\nabla \cdot \vec{B}(\vec{r},t) = 0 \tag{1}$$

$$\frac{\partial}{\partial t}\vec{B}(\vec{r},t) = -\nabla \times \vec{E}(\vec{r},t) \tag{2}$$

$$\nabla \cdot \vec{D}(\vec{r},t) = \rho(\vec{r},t) \tag{3}$$

$$\frac{\partial}{\partial t}\vec{D}(\vec{r},t) = \nabla \times \vec{H}(\vec{r},t) - \vec{j}(\vec{r},t) \tag{4}$$

For moving charges in vacuo, they can be simplified via $\vec{D}(\vec{r},t) = \varepsilon_0 \vec{E}(\vec{r},t)$, $\vec{B}(\vec{r},t) = \mu_0 \vec{H}(\vec{r},t)$ to the microscopic Maxwell equations (Lorentz 1892).

$$\nabla \cdot \vec{B}(\vec{r},t) = 0 \tag{5}$$

$$\frac{\partial}{\partial t}\vec{B}(\vec{r},t) = -\nabla \times \vec{E}(\vec{r},t) \tag{6}$$

$$\nabla \cdot \vec{E}(\vec{r},t) = \frac{1}{\varepsilon_0}\rho(\vec{r},t) \tag{7}$$

$$\frac{\partial}{\partial t}\vec{E}(\vec{r},t) = \frac{1}{\varepsilon_0\mu_0}\nabla \times \vec{B}(\vec{r},t) - \frac{1}{\varepsilon_0}\vec{j}(\vec{r},t) \tag{8}$$

For both sets, two fundamental problems have to be clarified, *viz*,

(1) the origin of the gauge freedom in the potentials, and

(2) the origin of the transversality of free (unbounded) electromagnetic waves.

For this, I will – following the recommendation by Boltzmann (2001) – return to Maxwell's (1864) *original* set of equations. Using the Helmholtz (1858) decomposition of 3D vector fields into transverse and longitudinal components, I will show that this set is both underdetermined and redundant (but not inconsistent). Remarkably enough, both deficiencies are related to *longitudinal* vector components.

Thus, this Pt. I of a series of papers starts with an exposition of the Helmholtz decomposition. Special attention is paid to the various kinds of fields occurring in electromagnetism, notably to fields like the possible static longitudinal component of the vector potential, $\vec{A}_L(\vec{r})$, which is not accounted for in any variant of Maxwell equations, but is determined by boundary conditions. In Section 3, Maxwell's (1864) original set of equations is rewritten in terms of the transverse and longitudinal components of *all* fields, and a revised set being *free* of underdeterminacy and redundance is proposed. Section 4 considers the role of $\vec{A}_L(\vec{r},t)$ for the gauge freedom both in electromagnetism and in Schrödinger's (1926) wave mechanics, where the latter provides a short-cut to a gauge *invariant* Hamiltonian. The results are summarized and discussed in Section 5.

2. Helmholtz Decomposition of 3D Vector Fields

In order to apply Helmholtz's decomposition theorem appropriately, one has carefully to discriminate between certain types of vector fields, *viz*, Newtonian, Laplacian and vector fields in multiply connected domains.

2.1 Newtonian Vector Fields

Newtonian vector fields are vector fields in unbounded domains with a given distribution of sources and vortices (Schwab 2002). The classical example is Newton's force of gravity. They are the actual subject of

Helmholtz's decomposition theorem: Any sufficiently well-behaving 3D vector field, $\vec{f}(\vec{r})$, can uniquely be decomposed into a transverse or solenoidal, $\vec{f}_T(\vec{r})$, a longitudinal or irrotational, $\vec{f}_L(\vec{r})$, and a constant components (which I will omit in what follows).

$$\begin{aligned}\vec{f}(\vec{r}) = \iiint_V \vec{f}(\vec{r}')\delta(\vec{r}-\vec{r}')dV'; \quad \vec{r} \in V\backslash\partial V \\ = -\frac{1}{4\pi}\iiint_V \vec{f}(\vec{r}')\Delta\frac{1}{|\vec{r}-\vec{r}'|}dV' \\ = \frac{1}{4\pi}\nabla\times\nabla\times\iiint_V \frac{\vec{f}(\vec{r}')}{|\vec{r}-\vec{r}'|}dV' - \frac{1}{4\pi}\nabla\nabla\cdot\iiint_V \frac{\vec{f}(\vec{r}')}{|\vec{r}-\vec{r}'|}dV' \\ = \vec{f}_T(\vec{r}) + \vec{f}_L(\vec{r})\end{aligned} \quad (9)$$

(These notions of longitudinal and transverse fields should not be confused with the notions of longitudinal and transverse waves in waveguides!)

It is thus most useful to introduce scalar, $\phi_f(\vec{r})$, and vector potentials, $\vec{a}_f(\vec{r})$, as

$$\vec{f}_T(\vec{r}) = \nabla\times\vec{a}_{\vec{f}}(\vec{r}); \quad \vec{f}_L(\vec{r}) = -\nabla\phi_{\vec{f}}(\vec{r}) \quad (10)$$

The minus sign is chosen to follow the definitions of the mechanical potential energy and the scalar potential in the electric field strength. $\vec{a}_{\vec{f}}$ is *sourceless*; otherwise, one would increases the number of independent field variables.

As a consequence, each such vector field is uniquely determined by its sources, $\phi_{\vec{f}}$, and source*less* vertices, $\vec{j}_{\vec{f}}$.

$$\nabla\times\vec{f}(\vec{r}) = \nabla\times\vec{f}_T(\vec{r}) = \nabla\times\nabla\times\vec{a}_{\vec{f}}(\vec{r}) = -\Delta\vec{a}_{\vec{f}}(\vec{r}) = \vec{j}_{\vec{f}}(\vec{r}) \quad (11)$$

$$\nabla\cdot\vec{f}(\vec{r}) = \nabla\cdot\vec{f}_L(\vec{r}) = -\Delta\phi_{\vec{f}}(\vec{r}) = \rho_{\vec{f}}(\vec{r}) \quad (12)$$

Including the surface terms (Oughstun 2006, Appendix A), the potentials follow as

$$\begin{aligned}\phi_{\vec{f}}(\vec{r}) = \frac{1}{4\pi}\nabla\cdot\iiint_V \frac{\vec{f}(\vec{r}')}{|\vec{r}-\vec{r}'|}dV' \\ = \frac{1}{4\pi}\iiint_V \frac{\rho_{\vec{f}}(\vec{r}')}{|\vec{r}-\vec{r}'|}dV' - \frac{1}{4\pi}\oint_{\partial V} \frac{\vec{f}(\vec{r}')}{|\vec{r}-\vec{r}'|}\cdot d\vec{\sigma}'\end{aligned} \quad (13)$$

$$\begin{aligned}\vec{a}_{\vec{f}}(\vec{r}) = \frac{1}{4\pi}\nabla\times\iiint_V \frac{\vec{f}(\vec{r}')}{|\vec{r}-\vec{r}'|}dV' \\ = \frac{1}{4\pi}\iiint_V \frac{\vec{j}_{\vec{f}}(\vec{r}')}{|\vec{r}-\vec{r}'|}dV' + \frac{1}{4\pi}\oint_{\partial V} \frac{\vec{f}(\vec{r}')}{|\vec{r}-\vec{r}'|}\times d\vec{\sigma}'\end{aligned} \quad (14)$$

For the balance equations I will also need the
Orthogonality theorem: Integrals over mixed scalar products vanishe,

$$\iiint_V \vec{f}_T \cdot \vec{g}_L dV = -\iiint_V \nabla \times \vec{a}_{\vec{f}} \cdot \nabla \phi_{\vec{g}} dV$$
$$= -\iiint_V \nabla\left(\phi_{\vec{g}} \nabla \times \vec{a}_{\vec{f}}\right) dV = -\iiint_V \nabla\left(\vec{a}_{\vec{f}} \times \nabla \phi_{\vec{g}}\right) dV$$
$$= -\oint_{\partial V} \phi_{\vec{g}} \nabla \times \vec{a}_{\vec{f}} \cdot d\vec{\sigma} = -\oint_{\partial V} \vec{a}_{\vec{f}} \times \nabla \phi_{\vec{g}} \cdot d\vec{\sigma} = 0 \quad (15)$$

if the surface, ∂V, lies infinitely away from the sources of the fields (*cf* Stewart 2008), or if the fields obey appropriate periodic boundary conditions on ∂V (Heitler 1954, I.6.3).

If the Orthogonality theorem holds true, the integrals over the scalar products of two vectors separates as

$$\iiint_V \vec{f}(\vec{r}) \cdot \vec{g}(\vec{r}) dV = \iiint_V \vec{f}_T(\vec{r}) \cdot \vec{g}_T(\vec{r}) dV + \iiint_V \vec{f}_L(\vec{r}) \cdot \vec{g}_L(\vec{r}) dV \quad (16)$$

In particular, both the Joule power and the electric field energy decompose into the contributions of the transverse and longitudinal components of the (di)electric field vectors.

The validity of this theorem will be assumed throughought this series of papers.

Notice that the electromagnetic vector potential, $\vec{A}$, is a vector potential in the sense of Helmholtz's theorem only w.r.t. the magnetic induction, $\vec{B}$, not, however, w.r.t. the electric field strength, $\vec{E}$. As a consequence, both its transverse and longitudinal components are physically significant. The clue is thus to Helmholtz-decompose $\vec{A}$, too.

It should also be noted that the longitudinal and transverse components of a *localized* vector field are *spread over the whole* volume of definition. For instance, for a point-like body of charge q moving along the trajectory $\vec{r}(t)$,

$$\rho(\vec{r}, t) = q\delta(\vec{r} - \vec{r}(t)); \quad \vec{j}(\vec{r}, t) = q\vec{v}(t)\delta(\vec{r} - \vec{r}(t)) \quad (17)$$

one has ($\vec{r}_t \equiv \vec{r}(t)$; here, t is merely a parameter)

$$\rho_{\vec{j}} = \nabla \vec{j} = q\vec{v} \cdot \nabla\delta(\vec{r} - \vec{r}(t)) = -\frac{\partial \rho}{\partial t} \quad (18a)$$
$$\vec{a}_{\vec{j}} = \nabla \times \vec{j} = q\nabla\delta(\vec{r} - \vec{r}_t) \times \vec{v} \quad (18b)$$

$$\phi_{\vec{j}} = \frac{q}{4\pi} \nabla \cdot \frac{\vec{v}}{|\vec{r} - \vec{r}_t|} = -\frac{q}{4\pi} \frac{\vec{v} \cdot (\vec{r} - \vec{r}_t)}{|\vec{r} - \vec{r}_t|^3} \quad (19a)$$
$$\vec{a}_{\vec{j}} = \frac{q}{4\pi} \nabla \times \frac{\vec{v}}{|\vec{r} - \vec{r}_t|} = -\frac{q}{4\pi} \frac{(\vec{r} - \vec{r}_t) \times \vec{v}}{|\vec{r} - \vec{r}_t|^3} \quad (19b)$$

$$\vec{j}_L = \frac{q}{4\pi} \nabla \frac{\vec{v} \cdot (\vec{r} - \vec{r}_t)}{|\vec{r} - \vec{r}_t|^3}; \quad \vec{j}_T = -\frac{q}{4\pi} \nabla \times \frac{(\vec{r} - \vec{r}_t) \times \vec{v}}{|\vec{r} - \vec{r}_t|^3} \quad (20)$$

2.2 Laplacean Vector Fields

Laplacean vector fields are vector fields outside any sources and vortices, they (or their potentials) satisfy the Laplace equation and are essentially determined by the (inhomogeneous) boundary conditions (Schwab 2002). A typical example is the electric field between electrodes.

Since both their divergence and curl vanishe identically, Helmholtz's theorem is not really useful for them.

2.3 Vector Fields in Multiply Connected Domains

Vector fields in multiply connected domains assume an 'androgyne' position in that they (or their potentials) satisfy the Laplace equation in certain, bounded domains, but not globally. A well known example is the magnetic field strength, $\vec{H}$, of a constant current, I, through an infinite straight conductor in vacuo. The 'magnetic ring voltage', $\oint \vec{H} \cdot d\vec{s}$, vanishes identically, as long as the path of integration lies *entirely outside* the conductor, so that no current flows through the area bounded by it. But it equals

$$n \iint_{\sigma} \left(\nabla \times \vec{H} \right) \cdot d\vec{\sigma} = n \iint_{\sigma} \vec{j} \cdot d\vec{\sigma} = nI \tag{21}$$

if the path *surrounds* the conductor n times (n integer). That means, that *inside* the conductor, $\vec{H}$ is a *vortex* field: $\nabla \times \vec{H} = \vec{j} \neq \vec{0}$, while *outside* the conductor, $\vec{H}$ is a *gradient* field: $\nabla \times \vec{H} = \vec{0}$. Obviously, Helmholtz's theorem is only conditionally applicable, since the integral rather than the differential form of Ampère's flux law is appropriate.

An analogous example is the vector potential in the Aharonov-Bohm (1959) setup. A constant current through an ideal straight infinite coil in vacuo with no spacing between its windings creates a magnetic field strength and induction being constant inside and vanishing outside the coil. However, by virtue of its continuity, the vector potential does *not* vanishe outside the coil, but represents a gradient field there. I will return to this issue in Section 4.

3. Maxwell's (1864) Original Equations Revisited

"He [Maxwell] would not have been so often misunderstood, if one would have started the study not with the treatise, while the specific Maxwellian method occours much more clearly in his earlier essays." (Boltzmann 2001; *cf* also Sommerfeld 2001, §1) For this, let us return to Maxwell's (1864) original set of "20 equations for the 20 variables" $(F, G, H) = \vec{A}$, $(\alpha, \beta, \gamma) = \vec{H}$, $(P, Q, R) = \vec{E}$, $(p, q, r) = \vec{j}$, $(f, g, h) = \vec{D}$, $(p', q', r') = \vec{J}$, $e = \rho$, $\psi = \Phi$. I will rewrite them in modern notation (the r.h.s. of the foregoing relations), SI units and together with their Helmholtz decomposition. For easier reference, Maxwell's equation numbering is applied. In place of his eqs. (D) for moving conductors his eqs. (35) for conductors at rest is used. The signs in his eqs. (30) and (33) are changed according to the nowadays use.

3.1 Helmholtz Decomposition

A) The total current density, $\vec{J}$, is the sum of electric (conduction, convection) current density, $\vec{j}$, and displacement ('total polarization') current density, $\partial\vec{D}/\partial t$.

$$\vec{J}(\vec{r},t) = \vec{j}(\vec{r},t) + \frac{\partial}{\partial t}\vec{D}(\vec{r},t) \tag{22}$$

This is Maxwell's famous and crucial step to generalize Ampère's flux law to open circuits and to convective currents. The time derivative is a precondition to obtain wave equations for the field variables.

The Helmholtz decomposition of this equation is obvious.

$$\vec{J}_{T,L}(\vec{r},t) = \vec{j}_{T,L}(\vec{r},t) + \frac{\partial}{\partial t}\vec{D}_{T,L}(\vec{r},t) \tag{23}$$

B) The "magnetic force" (induction, flux density), $\mu\vec{H}$, is the vortex of the vector potential, $\vec{A}$: $\mu\vec{H} = \nabla\times\vec{A}$. Hence, it has got no longitudinal component.

$$(\mu\vec{H})_T(\vec{r},t) = \nabla\times\vec{A}_T(\vec{r},t) \tag{24a}$$
$$(\mu\vec{H})_L(\vec{r},t) \equiv \vec{0} \tag{24b}$$

C) The total current density, $\vec{J}$, is the vortex of the magnetic field strength, $\vec{H}$.

$$\nabla\times\vec{H}(\vec{r},t) = \vec{J}(\vec{r},t) \tag{C}$$

Hence, it has got no longitudinal component, too.

$$\vec{J}_T(\vec{r},t) = \nabla\times\vec{H}_T(\vec{r},t) \tag{25a}$$
$$\vec{J}_L(\vec{r},t) = \vec{0} \tag{25b}$$

D) The "electromotive force" (electric field strength) equals

$$\vec{E}(\vec{r},t) = -\frac{\partial}{\partial t}\vec{A}(\vec{r},t) - \nabla\Phi(\vec{r},t) \tag{26}$$

Therefore,

$$\vec{E}_T(\vec{r},t) = -\frac{\partial}{\partial t}\vec{A}_T(\vec{r},t) \tag{27a}$$
$$\vec{E}_L(\vec{r},t) = -\frac{\partial}{\partial t}\vec{A}_L(\vec{r},t) - \nabla\Phi(\vec{r},t) \underset{def}{=} -\nabla\phi_{\vec{E}}(\vec{r},t) \tag{27b}$$

The longitudinal component consists of *two* terms, for which, however, there is *no other* equation. This makes the whole set to be *underdetermined* and is the origin of the gauge freedom in the potentials $\vec{A}$ and Φ. Due to the redundancy in some equations below, it is not inconsistent, however.

This underdeterminacy is overcome, if one can work solely with $\phi_E(\vec{r},t)$, the 'total scalar potential of $\vec{E}(\vec{r},t)$', or if one finds an additional equation for $\vec{A}_L$ and Φ, respectively. An example is the boundary conditions in the Aharonov-Bohm (1959) setup, which determine $\vec{A}_L$ outside the coil.

E) Electric field strength and dielectric displacement are related through the "equation of electric elasticity".

$$\vec{E}(\vec{r},t) = \frac{1}{\varepsilon}\vec{D}(\vec{r},t) \tag{28}$$

Thus,

$$\vec{E}_{T,L}(\vec{r},t) = \frac{1}{\varepsilon}\vec{D}_{T,L}(\vec{r},t) \tag{29}$$

if ε is a scalar constant.

F) Electric field strength and electric current density are related through the "equation of electric resistance" (σ being the specific conductivity).

$$\vec{E}(\vec{r},t) = \frac{1}{\sigma}\vec{j}(\vec{r},t) \tag{30}$$

Thus,

$$\vec{E}_{T,L}(\vec{r},t) = \frac{1}{\sigma}\vec{j}_{T,L}(\vec{r},t) \tag{31}$$

if σ is a scalar constant.

For N point-like charges $\{q_a\}$ in vacuo ($\sigma = 0$), eq. (30) is to be replaced with

$$\sum_{a=1}^{N} q_a \vec{v}_a(t)\delta(\vec{r} - \vec{r}_a(t)) = \vec{j}(\vec{r},t) \tag{32}$$

G) The "free" charge density is related to the dielectric displacement through the "equation of free electricity".

$$\rho(\vec{r},t) - \nabla\vec{D}(\vec{r},t) = 0 \tag{33}$$

Obviously, it concerns the longitudinal component of $\vec{D}$ only.

$$\rho(\vec{r},t) - \nabla\vec{D}_L(\vec{r},t) = 0 \tag{34}$$

H) In a conductor, there is – in analogy to hydrodynamics – "another condition", the "equation of continuity".

$$\frac{\partial}{\partial t}\rho(\vec{r},t) + \nabla\vec{j}(\vec{r},t) = 0 \tag{35}$$

It concerns the longitudinal component of $\vec{j}$ only.

$$\frac{\partial}{\partial t}\rho(\vec{r},t) + \nabla\vec{j}_L(\vec{r},t) = 0 \tag{36}$$

At once, by virtue of eq.(25b), it is merely a consequence of eq.(34). Here is the redundancy mentioned above.

With

$$\vec{j} = \vec{j}_T + \vec{j}_L = \nabla\times\vec{a}_{\vec{j}} - \nabla\phi_{\vec{j}} \tag{37}$$

one obtains the continuity equation in the form

$$-\Delta\phi_{\vec{j}}(\vec{r},t) + \frac{\partial}{\partial t}\rho(\vec{r},t) = 0 \tag{38}$$

It has the advantage of being a single equation relating *two* scalar quantities one to another rather than four, as in its usual form (35).

3.2 Elimination of Underdeterminacy and Redundance

The underdeterminacy and redundance in Maxwell's original set can be eliminiated through removing $(\mu\vec{H})_L$, Φ and $\vec{A}_L$ from the set of field variables, but retaining $\phi_{\vec{E}} = -\partial\phi_{\vec{A}}/\partial t + \Phi$. I also remove the total current in view of its merely historical relevance. Then, it remains 18 equations for the 18 variables $(\mu\vec{H})_{(T)} = \vec{B}$, $\vec{H}$, $\vec{A}_T$, $\vec{D}$, $\vec{E}$, $\phi_{\vec{E}}$, $\vec{j}$ and ρ.

B') The magnetic induction (flux density), $\mu\vec{H}$, is solenoidal, since it is the vortex of the transverse component of the vector potential.

$$\mu\vec{H} = (\mu\vec{H})_T = \nabla \times \vec{A}_T \tag{39}$$

C') The transverse components of the conduction/convection and displacement current densities build the vortex of the transverse component, $\vec{H}_T$, of the magnetic field strength, $\vec{H}$.

$$\nabla \times \vec{H}_T = \vec{j}_T + \frac{\partial}{\partial t}\vec{D}_T \tag{40}$$

D') The electric field strength equals (and Helmholtz decomposes as)

$$\vec{E} = -\frac{\partial}{\partial t}\vec{A}_T - \nabla\phi_{\vec{E}} \tag{41}$$

E) Electric field strength and dielectric displacement are related through the "equation of electric elasticity" (28).

F) Electric field strength and electric current density are related through the "equation of electric resistance" (30).

G') The "free" charge density is related to the longitudinal component of the dielectric displacement through the "equation of free electricity".

$$\rho(\vec{r},t) - \nabla\vec{D}_L(\vec{r},t) = 0 \tag{42}$$

H') The conservation of charge is expressed through the equation of continuity.

$$\frac{\partial}{\partial t}\rho(\vec{r},t) + \nabla\vec{j}_L(\vec{r},t) = 0 \tag{43}$$

Therefore, the redundance is removed in the flux law rather than eliminating the continuity equation from the set of basic equations. The continuity equation is retained, because it is a direct consequence of the fact, that – within this approach – the charge of a point-like body is a given, invariant property of it (like its mass). This also allows for an immediate explanation of the transversality of free electromagnetic waves, as will be shown in Pt.II.

4. Gauge Freedom and the Role of $\vec{A}_L$

4.1 Classical Gauge Freedom

As mentioned after eq.(27b) above, there is only *one* equation for the *two* fields $\partial\vec{A}_L/\partial t$ and Φ. Hence, any change of the scalar and vector potentials such, that the expression

$(-\partial\phi_{\vec{A}}/\partial t+\Phi) = \phi_{\vec{E}}$ remains unchanged, is without any physical effect within Maxwell's theory.

In fact, the Helmholtz components and potentials of vector potential, $\vec{A}$, and electrical field strength, $\vec{E}$,

$$\vec{A} = \vec{A}_T + \vec{A}_L = \nabla\times\vec{a}_{\vec{A}} - \nabla\phi_{\vec{A}} \tag{44a}$$
$$\vec{E} = \vec{E}_T + \vec{E}_L = \nabla\times\vec{a}_{\vec{E}} - \nabla\phi_{\vec{E}} \tag{44b}$$

are known to be interrelated as

$$\vec{E}_T = -\frac{\partial}{\partial t}\vec{A}_T; \quad \vec{a}_{\vec{E}} = -\frac{\partial}{\partial t}\vec{a}_{\vec{A}} \tag{45}$$
$$\vec{E}_L = -\frac{\partial}{\partial t}\vec{A}_L - \nabla\Phi; \quad \phi_{\vec{E}} = -\frac{\partial}{\partial t}\phi_{\vec{A}} + \Phi \tag{46}$$

Hence, the gauge transformation,

$$\vec{A} = \vec{A}' - \nabla\chi; \quad \Phi = \Phi' + \frac{\partial\chi}{\partial t} \tag{47}$$

actually concerns only the scalar potential, $\phi_{\vec{A}}$, of $\vec{A}$ as

$$\phi_{\vec{A}} = \phi'_{\vec{A}} + \chi \tag{48}$$

but not the vector potential, $\vec{a}_{\vec{A}}$, of $\vec{A}$.

In the Lorenz (1867) gauge used in Lorentz covariant formulations of the theory, one has

$$\nabla\vec{A} = -\Delta\phi_{\vec{A}} = -\frac{\partial\Phi}{\partial t} \tag{49}$$

while in the Coulomb (transverse, radiation) gauge being popular in quantum electrodynamics,

$$\nabla\vec{A} = -\Delta\phi_{\vec{A}} = 0 \tag{50}$$

This all suggests to avoid the gauge indeterminacy at all through working solely with $\vec{A}_T$ and $\phi_{\vec{E}}$. If necessary, $\vec{A}_L$ can be determined as boundary value problem.

4.2 Quantum Gauge Freedom (Schrödinger Theory)

Although this series of papers deals with classical electromagnetism, it is enlightening and pedagogically useful to sidestep for looking at gauge freedom within Schrödinger wave mechanics.

In order to be independent of the interpretation of the quantum mechanical formalism, let me proceed a follows (Enders 2006, 2008a,b).

$|\psi|^2$ and $<\psi|\hat{H}|\psi>$ are 'Newtonian state functions' of a non-relativistic quantum system as they are time-independent in stationary states and as their time-dependence is governed by solely the time-dependent part of the Hamiltonian. This suggests to extend Helmholtz's (1847, 1911) explorations about the relationships between forces and energies to the question, which 'external influences' leave $|\psi|^2$ and $<\psi|\hat{H}|\psi>$ unchanged?

Obviously, $|\psi|^2$ is unchanged, if an external influence, w, affects only the phase, φ, of ψ. (Dirac 1931 required the phase to be independent of the state.)

$$\psi_w = \psi_0 e^{i\varphi(w)}; \quad \varphi(0) = 0 \tag{51}$$

Then, if $\psi_0(\vec{r}, t)$ obeys the Schrödinger equation

$$i\hbar \frac{\partial}{\partial t}\psi_0 = \frac{\hat{p}^2}{2m}\psi_0 + V\psi_0 \tag{52}$$

$\psi_w(\vec{r}, t)$ obeys the Schrödinger equation

$$i\hbar \frac{\partial}{\partial t}\psi_w = \hat{H}_w\psi_w = \frac{1}{2m}\left(\hat{p} - \hbar\nabla\varphi\right)^2\psi_w + \left(V - \hbar\frac{\partial\varphi}{\partial t}\right)\psi_0 \tag{53}$$

Consequently, in stationary states, $< \psi_w|\hat{H}_w|\psi_w >$ is independent of w, because $i\hbar\frac{\partial}{\partial t}\psi_w = E\psi_w$, where – by the very definition of w – E is independent of w. This is essentially the gauge invariance of the Schrödinger (Pauli 1926) and Dirac equations (Fock 1929) (see also Weyl 1929, 1931).

For influences caused by external electromagnetic fields, this quite general arguing leads to the following important observation, which will be exploited in Pt.III.

The common quasi-classical Schrödinger equation for a point-like charge, q, in an electromagnetic field reads

$$i\hbar \frac{\partial}{\partial t}\psi_{\vec{A},\Phi}(\vec{r}, t) = \left[\frac{1}{2m}\left(\hat{p} - q\vec{A}(\vec{r}, t)\right)^2 + q\Phi(\vec{r}, t)\right]\psi_{\vec{A},\Phi}(\vec{r}, t) \tag{54}$$

Thus, the wave function

$$\psi_{\vec{A}_T,\phi_{\vec{E}}}(\vec{r}, t) = \psi_{\vec{A},\Phi}(\vec{r}, t)e^{i\frac{q}{\hbar}\phi_{\vec{A}}(\vec{r},t)} \tag{55}$$

obeys a Schrödinger equation with a *manifest gauge invariant* Hamiltonian.

$$i\hbar \frac{\partial}{\partial t}\psi_{\vec{A}_T,\phi_{\vec{E}}}(\vec{r}, t) = \left[\frac{1}{2m}\left(\hat{p} - q\vec{A}_T(\vec{r}, t)\right)^2 + q\phi_{\vec{E}}(\vec{r}, t)\right]\psi_{\vec{A}_T,\phi_{\vec{E}}}(\vec{r}, t) \tag{56}$$

This suggests that manifest gauge invariant theories can be obtained through replacing $\vec{A}$ with $\vec{A}_T$ and Φ with $\phi_{\vec{E}}$.

It is noteworthy, that in *multiply* connected domains, notably outside an infinite coil, where the $\vec{B}$-field vanishes, $\phi_{\vec{A}}$ is not globally integrabel. The phase of the wave function can aquire physical significance, as in the Aharonov-Bohm (1959) effect. This underpins the physical significance of the Helmholtz decomposition of the field variables.

Thus, the longitudinal component of a static vector potential, $\vec{A}_L(\vec{r})$, is classically *not* observable, because it does not contribute to the Maxwell-Lorentz force (Maxwell 1864, Lorentz 1892),

$$q\vec{E} + q\vec{v} \times \vec{B} = q\left(-\frac{\partial\vec{A}}{\partial t} - \nabla\Phi + \vec{v} \times \nabla \times \vec{A}\right) \tag{57}$$

This suggests to remove $\vec{A}_L(\vec{r})$ from the classical theory altogether and to consider it to be a 'quantum potential' being proportional to Planck's quantum of action, h. On the other hand, if one requires – for good reasons – $\vec{A}(\vec{r})$ to be continuous, $\vec{A}_L(\vec{r})$ can be finite even in the classical (limit) case.

Eq.(55) suggests to incorporate other non-dynamical fields not entering the Hamiltonian and being determined by Laplacean boundary-value problems, by means of appropriate phase factors, too.

"As emphasized by Yang [1974] the vector potential is an over complete specification of the physics of a gauge theory but the gauge covariant field strength underspecifies the content of a gauge theory. The Bohm-Aharanov [1959] effect is the most striking example of this, wherein there exist physical effects on charged particles in a region where the field strength vanishes. The complete and minimal set of variables necessary to capture all the physics are the non-integrable phase factors." (Gross 1992, II.4) Because there are no such phase factors within classical electromagnetism, their classical limit is rather unclear. The complete and minimal set of classical variables obtained in Pts.II and III, respectively, is only loosely related to those. It is thus hoped that the gauge-free representation presented in Pt.III will narrow this gap between classical and quantum theory.

5. Summary and Discussion

The gauge freedom in classical electromagnetism roots in an underdeterminacy in Maxwell's (1864) original set of equations. Maxwell's representation (26) of the electric field strength in terms of the potentials contains *two* contributions to its longitudinal component, one from the vector potential, $-\frac{\partial}{\partial t}\vec{A}_L$, and one from the scalar potential, $-\nabla\Phi$. For these two, no other equation is established. Consequently, Maxwell's (1864) set is actually *not* "20 equations for 20 variables" (§70), but only 19 equations for 20 variables. It is not inconsistent, however, because the equations being related to charge conservation are redundant.

In the rationalized and microscopic Maxwell equations, this underdeterminacy is hidden, because only $\vec{E}(\vec{r},t)$ occurs. It returned in the canonical theory, as will be discussed in Pt. III of this series, where manifest gauge invariant Lagrangians and Hamiltonians will be proposed. The redundancy of the 1864 set is absent in those equations, because the continuity equation has not been retained. The advantage of this concentration on the field equations is their Lorentz invariance. A disadvantage consists in that the experimentally observed transversality of free electromagnetic waves does not naturally emerges out of the theory. (Within quantum theory, it consists in the artifact of the occurrence of not observable photons.)

Both deficiencies, underdeterminacy and redundance, are absent in the revised set of equations proposed in this paper. Their 'rationalization' leads to separate Poynting theorems for both, the propagating transverse and the non-propagating longitudinal field variables, respectively, and to a complete and minimal set of field variables (see Pts.II and III).

Littlejohn (2008) has stressed correctly, that the gauge transformation changes only the longitudinal component of the vector potential. His conclusion, however, that this is the "nonphysical" part, while the transverse component is the physical one (Sect. 34.8), overlooks its role in the Aharanov-Bohm effect. Such contradictions will be avoided through, (i), working with combinations of Φ and $\vec{A}$, in which those "nonphysical parts", if present, cancel each another and, (ii), treating this gauge invariant combination separately from the dynamics of the other field components.

It is perhaps no accident that the history of the electromagnetic potentials is even more curvilineal than that of the field strengths. Maxwell (1861, 1862, 1864) saw the vector potential to represent Faraday's "electrotonic state" and the electromagnetic field momentum, respectively. Later, the potentials were considered to be superfluous or merely mathematical tools for solving the rationalized Maxwell's equations. This mistake lived for a surprisingly long time, in spite of their appearance in the principle of least action (Schwartzschild 1903), in the Hamiltonian (Pauli 1926, Fock 1929) and, last but not least, in the Aharonov-Bohm (1959) effect. The double role of the vector potential, $\vec{A}$, in the electric field strength, $\vec{E}$, where $\partial\vec{A}/\partial t$ contributes to both the transverse and the longitudinal components, has surely hindered the clarification.

The approach presented here benefits from the methodological advantages of the treatments by Newton, Euler and Helmholtz (see Enders 2006, 2008, 2009). In particular, the subject under investigation (moving charged bodies and the electromagnetic fields created by them and acting back onto them) is defined *before* the mathematical formalism is developed. This keeps the latter physically clear.

Acknowledgement

I feel highly indebted to Prof. O. Keller and Dr. E. Stefanovich for numerous enlightening explanations, and to various posters in the moderated Usenet group 'sci.physics.foundations' for discussing this issue. I'm also indebted to a referee for his proposals to make this paper more straight and to remove not essential associations to related topics. Last but not least I like to thank Prof. J. López-Bonilla and the Leopoldina for encouraging this work.

References

[1] Y. Aharonov & D. Bohm, *Significance of Electromagnetic Potentials in the Quantum Theory*, Phys. Rev. 115 (1959) 485-491; Reprint in: A. Shapere & F. Wilczek, *Geometric Phases in Physics*, Singapore etc.: World Scientific 1989, paper [2.6]

[2] L. Boltzmann, *Comments*, in: J. C. Maxwell, *Über Faradays Kraftlinien*, Frankfurt: Deutsch ^{3}1995/2001 (Ostwalds Klassiker 69), p.99

[3] P. A. M. Dirac, *Quantized singularities in the electromagnetic field*, Proc. Roy. Soc. (L.) A133 (1931) 60-72

[4] P. A. M. Dirac, *Forms of Relativistic Dynamics*, Rev. Mod. Phys. 21 (1949) 392-399

[5] P. Enders, *Von der klassischen Physik zur Quantenphysik. Eine historisch-kritische deduktive Ableitung mit Anwendungsbeispielen aus der Festkörperphysik*, Berlin · Heidelberg: Springer 2006

[6] P. Enders, *The Mechanical Roots of Bopp's Principles of Electromagnetism*, Adv. Stud. Theor. Phys. 2 (2008) 199-214

[7] P. Enders, *Towards the Unity of Classical Physics*, Apeiron 16, No. 1 (2009) 22-44

[8] R. P. Feynman, R. B. Leighton & M. Sands, The Feynman Lectures on *Physics. Vol. II: Mainly Electromagnetism and Matter*, Reading (MA): Addison-Wesley ^{3}2001

[9] V. Fock, *Geometrisierung der Diracschen Theorie des Elektrons*, Z. Phys. 57 (1929) 261-275

[10] D. J. Gross, *Gauge Theory – Past, Present, and Future?*, Chin. J. Phys. 30 (1992) 7, 955-972

[11] O. Heaviside, *On the Forces, Stresses and Fluxes of Energy in the Electromagnetic Field*, Phil. Trans. Roy. Soc. 183A (1892) 423ff.

[12] W. Heitler, *The Quantum Theory of Radiation*, Clarendon, Oxford, ^{3}1954; Reprint: New York: Dover 1984

[13] H. Helmholtz, *Über die Erhaltung der Kraft*, Berlin: Reimer 1847

[14] H. Helmholtz, *Über Integrale der hydrodynamischen Gleichungen, welche den Wirbelbewegungen entsprechen*, J. Reine Angew. Math.55 (1858) 25-55

[15] H. v. Helmholtz, *Vorlesungen über die Dynamik discreter Massenpunkte*, Leipzig: Barth ^{2}1911

[16] H. Hertz, *Die Kräfte electrischer Schwingungen, behandelt nach der Maxwell'schen Theorie*, Ann. Phys 36 (1889) 1-22

[17] R. Littlejohn, *The Classical Electromagnetic Field Hamiltonian*, in: *Online lecture notes*, Physics 221B, Spring 2008, Notes 34, http://bohr.physics.berkeley.edu/classes/221/0708/notes/hamclassemf.pdf

[18] H. A. Lorentz, *La théorie électromagnétique de Maxwell et son application aux corps mouvants*, 1892; in: *Collected Papers*, The Hague 1936, Vol.2, 164-343

[19] H. A. Lorentz, *The Theory of Electrons and its Applications to the Phenomena of Light and Radiant Heat*, Leipzig: Teubner 1909; Reprint of 2^{nd} Ed.: New York: Dover 1952

[20] L. Lorenz, *On the identity of the vibrations of light with electrical currents*, Phil. Mag. 34 (1867) 287-301

[21] J. C. Maxwell, *On Physical Lines of Force*, Phil. Mag. [4] 21 (1861) 161-175, 281-291, 338-348; 23 (1862) 12-24, 85-95; *Scient. Papers* I, 451ff.

[22] J. C. Maxwell, *A Dynamical Theory of the Electromagnetic Field*, Phil. Trans. Roy. Soc. CLV (1865) 459-512 (article accompanying the Dec. 8, 1864 presentation to the Royal Society)

[23] J. Mehra & H. Rechenberg, *The Historical Development of Quantum Theory*, New York: Springer 1999ff.

[24] G. Mie, *Untersuchungen zum Problem der Quantenelektrik*, Ann. Phys. [4] 85 (1928) 711-729

[25] K. E. Oughstun, *Electromagnetic and Optical Pulse Propagation 1. Spectral Representations in Temporally Disperse Media*, New York: Springer 2006 (Springer Series in Optical Sciences 125)

[26] W. Pauli, *Letter to E. Schrödinger dated 12.12.1926* (after Mehra & Rechenberg, Vol. 6, Pt. I, p.218)

[27] J. H. Poynting, *On the transfer of energy in the electromagnetic field*, London Trans. 175 (1884) 343ff.

[28] A. J. Schwab, *Begriffswelt der Feldtheorie. Elektromagnetische Felder, Maxwellsche Gleichungen, Gradient, Rotation, Divergenz*, Berlin etc.: Springer ^{6}2002

[29] K. Schwartzschild, *Zur Elektrodynamik [I.]. Zwei Formen des Princips der kleinsten Action in der Elektronentheorie*, Nachr. Königl. Ges. Wiss. Göttingen. Math.-phys. Klasse (1903) 126-131

[30] A. Sommerfeld, *Vorlesungen über theoretische Physik*, *Bd. III Elektrodynamik*, Frankfurt a. Main: Deutsch ^{4}2001

[31] A. M. Stewart, *Longitudinal and transverse components of a vector field*, http://arxiv.org/pdf/0801.0335v1

[32] H. Weyl, *Elektron und Gravitation*, Z. Phys. 56 (1929) 330-352

[33] H. Weyl, *Gruppentheorie und Quantenmechanik*, Leipzig: Hirzel ^{2}1931; *The Theory of Groups and Quantum Mechanics*, London · New York: Methuen and Dover 1931