On the causality of electrodynamics and the Aharonov-Bohm effect

Richard Kowar

Department of Mathematics, University of Innsbruck,
Technikerstrasse 21a/2, A-6020, Innsbruck, Austria

December 6, 2018

Abstract

This paper presents a *non-instant field model* for electrodynamics that permits a causal explanation of the *Aharonov-Bohm effect* and a *covariant quantization* of the respective Maxwell equations via the *Gupta-Bleuler method*. Our model satisfies the following *correspondence principle*: if A^μ, $\mathbf{E}$, $\mathbf{B}$ denote the four potential, the electric field and the magnetic field of the non-instant field model, then the respective classical quantities are $\mathcal{A}[A^\mu]$, $\mathcal{A}[\mathbf{E}]$, $\mathcal{A}[\mathbf{B}]$, where $\mathcal{A}$ is a covariant time averaging operator. Here $\mathcal{A}[A^\mu]$ is interpreted as the best possible measurement of the four potential A^μ. Although the Lorentz condition is not satisfied for A^μ, it is satisfied for $\mathcal{A}[A^\mu]$. The latter fact means that the Lorentz condition does not hold for the quantized field but for its expectation value (cf. *Gupta-Bleuler method* of quantization). Finally, we derive the energy conservation law of our field model and show that the field energy is quantized.

Keywords: non-instant field model, Aharonov-Bohm effect, causality

1 Introduction

The goal of this paper is to present and discuss a non-instant field model of charged particles that has stronger causality properties than the classical model and can be quantized. A highlight of this paper is the causal explanation of the *Aharonov-Bohm effect*. It is best to start with a shortly review of electrodynamics and discuss the differences and relations to our model without carrying out the details.

The electromagnetic field concept in classical field theory is an instant one and in this sense causality is violated. For example, the four potential

1

$A_{cl}^\mu = (\phi_{cl}, \mathbf{A}_{cl})$ of a charged particle in its restframe S' is modeled by (cf. [4, 8])

$$\phi_{cl}' = \frac{Q_0}{|\mathbf{x}'|} \qquad \text{and} \qquad \mathbf{A}_{cl}' = 0\,,$$

where a prime indicates a quantity observed in the restframe S' of the charge. The charge Q_0 is an invariant and therefore is not primed. We see at once that the Coulomb condition $\nabla' \cdot \mathbf{A}_{cl}' = 0$ as well as the Lorentz condition $\partial_\mu' A_{cl}'^\mu = \frac{1}{c}\frac{\partial \phi_{cl}'}{\partial t'} + \nabla' \cdot \mathbf{A}_{cl}' = 0$ are satisfied. Since the latter condition is covariant (cf. [15, 8]), it holds in all reference frames. In contrast to this model, a non-instant field model must have a time-dependent scalar potential $\phi' = \phi'(\mathbf{x}', t')$ and if $\mathbf{A}' = 0$ is assumed, then the Lorentz condition cannot be satisfied. Consequently, for the non-instant field model presented in this paper

$$q := \partial_\mu A^\mu \quad \text{is a non-vanishing } invariant$$

and the respective field theory depends not only on the electric and magnetic fields $\mathbf{E}$ and $\mathbf{B}$ but also on the *invariant q*. As a consequence, a *Lorentz gauge transformation* of the non-instant four potential A^μ is prohibited. Since gauge transformations are used all the time in physics (cf. [4, 9, 8, 15, 16, 11, 12, 13, 17]), this statement seems to be absurd. But on the other hand, gauge theories might be considered as "unsatisfactory" (cf. Chapter III.4 in [17]). It turns out that this is not a problem, since our model satisfies the following *correspondence principle*.

The respective classical quantities of electrodynamics can be obtained from the non-instant field model via a covariant time averaging operator $\mathcal{A}$. If $\phi' = A'^0$, $\mathbf{E}'$ and $\mathbf{B}'$ denote the scalar potential, the electric field and the magnetic field of a charged particle in its restframe S', then it follows

$$\mathcal{A}[\phi'] = \phi_{cl}'\,, \qquad \mathcal{A}[\mathbf{E}'] = \mathbf{E}_{cl}'\,, \qquad \mathcal{A}[\mathbf{B}'] = \mathbf{B}_{cl}' \qquad \text{and} \qquad \partial_\mu \mathcal{A}[A^\mu] = 0\,,$$

where ϕ_{cl}', $\mathbf{E}_{cl}'$ and $\mathbf{B}_{cl}'$ denote the respective quantities from classical electrodynamics. The last condition means that the Lorentz condition is satisfied for the time averaged non-instant four potential. Here the quantity $\mathcal{A}[A^\mu]$ is interpreted as the best possible measurement of the four potential A^μ. Because the time averaging interval τ' is extremely small, it appears as if the measurement process was performed at a time instant (in S').

Because the classical quantities $\mathcal{A}[\phi']$ and $\mathcal{A}[\mathbf{E}']$ are time independent, it seems as if the electromagnetic field fills out the space instantaneously. On the one hand, calculation with time averaged quantities is very convenient (because we have gauge invariance and we got used to it), but on the other hand, the "time averaged picture" does not always allow to associate a cause to an effect. For example, the non-instant field model predicts that a solenoid sends out an electromagnetic wave of extreme high frequency $2\pi/\tau$

such that the *Aharonov-Bohm effect* (cf. [1, 4, 12, 6, 17, 2]) does not violate causality. However, if we go over to the "time averaged picture" causality seems to be violated, since then

$$\mathbf{E}_{cl} \equiv \mathcal{A}[\mathbf{E}] = \mathbf{0} \qquad \text{and} \qquad \mathbf{B}_{cl} \equiv \mathcal{A}[\mathbf{B}] = \mathbf{0} \qquad \text{outside the solenoid.}$$

The *correspondence principle* can be summarized as follows:

$$\text{relevant \& causal quantity } A^\mu \quad \xrightarrow{\mathcal{A}} \quad \text{measurable quantity } A^\mu_{cl}$$
$$\text{(unknown)} \qquad \text{(filter)} \qquad \text{(known)}$$

For modeling the left side of this mapping there might be only one reason/justification, to find the causes of a measurable effects.

Although this paper is mainly concerned with classical field theory I would like to say a view words about quantum electrodynamics. Since Maxwell's equations of the non-instant four potential read as follows

$$\partial_\nu F^{\mu\nu} - \partial^\mu (\partial_\nu A^\nu) = -4\pi \frac{j^\mu}{c} \qquad (\mu = 0,\, 1, 2\,,\, 3)\,,$$

where j^μ is the four current and $F_j^{\mu\nu} = \partial^\mu A_j^\nu - \partial^\nu A_j^\mu$ is the electromagnetic field tensor (cf. Sections 2 and 3), these equations have the same form as the modified equations used for the quantization of the electromagnetic field via the *Gupta-Bleuler method* (cf. Subsection 5.1 in [11]). That the Lorentz condition is only satisfied for $\mathcal{A}[A^\mu]$ means that the Lorentz condition is satisfied for the expectation value of the quantized four potential, which is (and can be) assumed in quantum electrodynamics. In contrast to quantum electrodynamics we can use the previous equations without assuming the Lorentz condition for the four potential which cannot be quantized.

This paper is organized as follows. Firstly, we introduce our non-instant field model in Section 2. Secondly, we derive the respective Maxwell equatiosn in covariant and non-covariant form in Section 3. Afterwards, in Section 4, the correspondence principle is introduced and discussed for the case of a charged particle and a solenoid. Section 4 is concluded with the causal explanation of the Aharonov-Bohm effect. The conservation law of energy and the quantization of the field energy is derived in Section 5. The paper is concluded with the section "Results" containing the subsection "Conclusions".

2   Non-instant field model of charged particles

2.1   The four potential A^μ

In the following we define a four potential $A^\mu = (\phi, \mathbf{A})$ of a particle with charge Q_0 observed from an inertial frame S_I. We do not assume that the

restframe S' of the particle is an inertial system. In its restframe S' the particle is created at $t' = 0$ and destroyed at $t' = N\tau'$. In this paper

$$\tau' > 0 \text{ is the same constant for each particle in its restframe.}$$

The field A^μ is a superposition of fields A_j^μ that are generated at the time instants $\tau_j' = j\,\tau'$ in S' with $j \in J := \{0, 1, \ldots, N-1\}$. It is convenient to introduce S_j' as the inertial systems such that its origin and axes coincide with those of system S' at time $t = \tau_j$ in S_I.

i) We define the field $A_j'^\mu$ in the inertial system S_j' by

$$A_j'^\mu = (\phi_j', \mathbf{A}_j') \qquad j \in J,$$

with

(1) $\qquad \phi_j'(\mathbf{x}', t') = Q_0 \dfrac{f_0'(\omega'\,(t' - \tau_j') - k'\,|\mathbf{x}'|)}{|\mathbf{x}'|} \qquad$ and $\qquad \mathbf{A}_j' = \mathbf{0}\,,$

where

(2) $\qquad Q_0$ is an invariant, $\qquad \omega' := \dfrac{2\pi}{\tau'}\,, \qquad k' := \dfrac{\omega'}{c}\,,$

the *smooth function* $t' \in \mathbb{R} \mapsto f_0(\omega'\,t') \in \mathbb{R}$ is positive, symmetric with respect to $t' = \tau'/2$ and has support in $[0, \tau']$,[1]

(3) $\qquad \displaystyle\int_\mathbb{R} \dfrac{f_0'(r)}{2\pi}\,\mathrm{d}r = 1 \qquad$ and $\qquad 4\pi\,\alpha \displaystyle\int_\mathbb{R} \left[\dfrac{\mathrm{d}f_0'(r)}{\mathrm{d}r}\right]^2 \mathrm{d}r = 1\,.$

Here c, e and $\alpha = \frac{e^2}{c\hbar}$ denote the *speed of light*, the *charge of the electron* and the *fine structure constant*, respectively. The last two conditions can be satisfied for an appropriate choice of f_0'.

ii) Now we consider the situation from an inertial frame S_I satisfying $t = 0$ if $t' = 0$ in S'. The four potential $(\phi_j, \mathbf{A}_j)$ is defined in S_I by (cf. [4, 8, 9, 15])

$$\phi_j^2(\mathbf{x}, t) - |\mathbf{A}_j(\mathbf{x}, t)|^2 = \phi_j'^2(\mathbf{x}', t')\,,$$

where $\mathbf{x}'(\mathbf{x}, t)$ and $t'(\mathbf{x}, t)$ are the coordinates of S_j' with respect to system S_I. More suggestive

$$\phi_j[\mathbf{v}_j]^2 - |\mathbf{A}_j[\mathbf{v}_j]|^2 = \phi_j'[\mathbf{0}]^2\,,$$

where $\mathbf{v}_j$ denotes the velocity of S_j' with respect to S_I at $t = \tau_j$. The total four potential is given by

(4) $$A^\mu = \sum_{j \in J} A_j^\mu[\mathbf{v}_j]\,.$$

---

[1]We use the Gaussian unit system in this paper, i.e. $\epsilon_0 = \mu_0 = 1$. Both right hand sides in (3) correspond to $1/[\epsilon_0] \equiv 1$.

Remark 1 *We note that Q_0 and $q' = \partial_\mu A'^\mu = \frac{1}{c}\frac{\partial \phi'}{\partial t'}$ (under the assumption $f_0'(-\infty) = 0$) determine the field model of a particle created at time $t = 0$ uniquely, since*

$$\tau' \text{ is the smallest positive real number such that}$$

$$\int\limits_0^{\tau'} q'(\mathbf{x}', r)\, \mathrm{d}r = 0$$

and

$$f_0'(r') - f_0'(-\infty) = -\frac{c^2\, \tau'\, r'}{2\,\pi\, Q_0} \int\limits_{-\infty}^0 q'\left(-\frac{c\,\tau'\,r'}{2\,\pi}, 0, 0, t\right) \mathrm{d}t\,.$$

Here the assumption $f_0'(-\infty) = 0$ is considered as a causality assumption. If the particle is created at time $t = 0$, then $f_0'(t) = 0$ for $t < 0$ is required by causality.

For the classical model in electrodynamics with Lorentz gauge, we have $q = 0$ and thus

$$f_0'(r') - f_0'(-\infty) = 0 \qquad \text{with} \qquad f_0'(-\infty) = 1\,,$$

i.e. causality is not satisfied.

2.2 The electromagnetic field and its tensor

As in classical field theory we define the electromagnetic field of a charged particle via its four potential (cf. [4, 9, 8, 15]). In the inertial frame S_I, we define the total electric and magnetic fields by $\mathbf{E} := \sum_{j\in J} \mathbf{E}_j$ and $\mathbf{B} := \sum_{j\in J} \mathbf{B}_j$ with

(5) $$\mathbf{E}_j = -\nabla\phi_j - \frac{1}{c}\frac{\partial \mathbf{A}_j}{\partial t} \qquad \text{and} \qquad \mathbf{B}_j = \nabla \times \mathbf{A}_j\,.$$

The *electromagnetic field tensor* is defined by $F^{\mu\nu} = \sum_{j\in J} F_j^{\mu\nu}$ with

(6) $$F_j^{\mu\nu} = \partial^\mu A_j^\nu - \partial^\nu A_j^\mu\,,$$

where $\partial_\nu := \left(\frac{1}{c}\frac{\partial}{\partial t}, \nabla\right)$ and $\partial^\nu := \left(\frac{1}{c}\frac{\partial}{\partial t}, -\nabla\right)$. Because A^μ is a four vector, $F_j^{\mu\nu}$ is a tensor and $(F_j)_{\mu\nu}\,(F_j)^{\mu\nu}$, $\epsilon^{\mu\nu\alpha\beta}\,(F_j)_{\mu\nu}\,(F_j)_{\alpha\beta}$ are invariants (cf. [9, 15, 5]).[2]

---

[2]Strictly speaking the second quantity is a pseudoscalar.

2.3 Electromagnetic four force

As usual, the *electromagnetic four force* on a particle with charge Q and velocity $\mathbf{v}$ is given by (cf. [4, 9])

$$\mathbf{f}^\mu := Q \, F^{\mu\,\nu} \, u_\nu \, ,$$

where u^μ denotes the *four velocity* $u^\mu := (1 - v^2/c^2)^{-1/2} \, (c, \mathbf{v})$.

Remark 2 *Consider a charged particle in its restframe. Since its field A^μ oscillates at a fixed point in space, the same is true for the electromagnetic field tensor $F^{\mu\,\nu}$ and hence the four force onto another particle "oscillates" too. But this means that the trajectory of the second particle is quite "chaotic" at a very small scale. However, the respective time averaged trajectory is less chaotic and smoother, since we integrate over time. But this fact will not be investigated closer in this paper. We do not need the concept of force in this paper.*

3 Maxwell's equations for the non-instant field model

We now derive the Maxwell equations for the non-instant field model as defined in Section 2. The most striking result will be - without assuming the Lorentz condition - that these equations (cf. (12) with (6)) are equivalent to the form of the Maxwell equations used in [11] for the quantization of the electromagnetic field via the *Gupta-Bleuler method*.

Covariant form for A^μ

From analysis, it is known that $A_j^{'\mu} = (\phi_j', \mathbf{A}_j')$ defined as in Subsection 2.1 satisfies the wave equations (cf. [14, 3, 7])

$$(7) \quad \nabla^{'2}\phi_j' - \frac{1}{c^2}\frac{\partial^2\phi_j'}{\partial t'^2} = -4\,\pi\,\rho_j' \qquad \text{and} \qquad \nabla^{'2}\mathbf{A}_j' - \frac{1}{c^2}\frac{\partial^2\mathbf{A}_j'}{\partial t'^2} = -4\,\pi\,\frac{\mathbf{j}_j'}{c} \, ,$$

where

$$(8) \qquad \rho_j' = Q_0 \, f_0'(\omega' \, (t' - \tau_j')) \, \delta(\mathbf{x}') \qquad \text{and} \qquad \mathbf{j}_j' = \mathbf{0} \, .$$

We define j_j^μ as the four vector satisfying $j_j^{'\mu} = (c\,\rho_j', \mathbf{j}_j')$ in S_j' and j^μ as the total *four current* $j^\mu := \sum_{j\in J} j_j^\mu$. Because A^μ (cf. (4)) and j^μ are four vectors, it follows from (7) and (8) that (cf. [15]).

$$(9) \qquad \nabla^2 A^\mu - \frac{1}{c^2}\frac{\partial^2 A^\mu}{\partial t^2} = -4\,\pi\,\frac{j^\mu}{c} \qquad (\mu = 0, 1, 2, 3) \, .$$

6

Formally, these equations are equivalent to the classical equations of Maxwell if $\partial_\mu A_j^\mu = 0$, i.e. the Lorentz condition is satisfied (cf. Chapter 25 in [4]). However, ϕ_j' is time dependent for our non-instant field model, i.e.

$$\partial_\mu A_j^\mu = \partial_\mu' A_j'^\mu = \frac{1}{c} \frac{\partial \phi_j'}{\partial t'} \neq 0$$

and thus each

(10) $q_j := \partial_\mu A_j^\mu$ is an invariant that is not identical zero.

Consequently, a *Lorentz gauge transformation* is prohibited for the non-instant field model presented in this paper.

With the electromagnetic field tensor $F^{\mu\nu}$ defined as in (6) and the scalar

$$(11) \qquad q := \sum_{j \in J} q_j = \partial^\nu A^\mu \,,$$

the equations of Maxwell (9) read as follows (cf. [4, 9, 15])

$$(12) \qquad \partial_\nu F^{\mu\nu} - \partial^\mu q = -4\pi \frac{j^\mu}{c} \qquad (\mu = 0,\, 1, 2,\, 3)\,.$$

These are the Maxwell equations used in [11] for the quantization of the electromagnetic field via the *Gupta-Bleuler method*.

Non-covariant form

From (5), (10), (9) and the identities

$$\nabla \times (\nabla \phi) = \mathbf{0} \qquad \text{and} \qquad \nabla \times (\nabla \times \mathbf{A}) = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A}\,,$$

it follows that

$$(13) \qquad
\begin{aligned}
\nabla \cdot \mathbf{E} &= 4\pi \rho - \frac{1}{c} \frac{\partial q}{\partial t}\,, & \nabla \times \mathbf{E} &= -\frac{1}{c} \frac{\partial \mathbf{B}}{\partial t}\,, \\
\nabla \cdot \mathbf{B} &= 0 & \text{and} \qquad \nabla \times \mathbf{B} &= 4\pi \frac{\mathbf{j}}{c} + \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t} + \nabla q\,.
\end{aligned}$$

Here ρ and $\mathbf{j}$ differ from their classical counterparts and the new laws of Gauss and Ampere differ from their classical counterparts, too (cf. [4, 9, 8]). Moreover, we see that Maxwell's equations (13) depend not only on the fields $\mathbf{E}$ and $\mathbf{B}$ but also on the scalar $q = \partial_\mu A^\mu$.

4 Correspondence principle

Now it is time to show how classical electrodynamics can be obtained via a special "time averaging process" from our non-instant field model. As in quantum mechanics we call this the *correspondence principle* of our field model. At the end of this section we examine the *Aharonov-Bohm effect* from the point of view of the non-instant field model together with the correspondence principle.

4.1 Covariant time averaging

For an introduction into the principle of covariance we refer to [15]. We define the covariant time average of the non-instant four potential A^μ by

$$(14) \qquad \mathcal{A}[A^\mu](\mathbf{x}, T) := \int\limits_{T}^{T+\tau} A^\mu(\mathbf{x}, t)\, \frac{\mathrm{d}t}{\tau},$$

where $\tau = \tau'/\sqrt{1 - v^2/c^2}$ and τ' is defined as in Subsection 2.1. This is a four vector, because

- A^μ is a four vector,
- $\mathrm{d}t\,\sqrt{1 - v^2/c^2}$ and $\tau\,\sqrt{1 - v^2/c^2}$ are scalars.

The time averaging of other tensors is defined analogously.

Remark 3 *We note that in general $\mathcal{A}[X(t)\,Y(t)] \neq \mathcal{A}[X(t)]\,\mathcal{A}[Y(t)]$ for two time-dependent quantities $X(t)$ and $Y(t)$. In particular,*

$$\mathcal{A}[X(t)] = 0 \qquad \not\Rightarrow \qquad \mathcal{A}[X(t)\,Y(t)] = 0\,.$$

For example, this fact is important for the calculation of the energy content of A'^μ_j in Section 5.

4.2 Derivation of the classical equations of Maxwell

In the following we assume that all quantities are sufficiently smooth. From (14), it follows at once that

$$(15) \qquad \frac{\partial \mathcal{A}[A^\mu]}{\partial x_j}(\mathbf{x}, t) = \mathcal{A}\left[\frac{\partial A^\mu}{\partial x_j}\right](\mathbf{x}, t) \qquad j = 1,\, 2,\, 3\,.$$

Moreover, from analysis it is known that (cf. [10])

$$\frac{\partial}{\partial T} \int\limits_{T}^{T+\tau} g(t)\, \frac{\mathrm{d}t}{\tau} = \frac{g(T+\tau) - g(T)}{\tau} = \int\limits_{T}^{T+\tau} \frac{\partial g(t)}{\partial t}\, \frac{\mathrm{d}t}{\tau}$$

8

and therefore we have

$$(16) \qquad \frac{\partial \mathcal{A}[A^\mu]}{\partial t}(\mathbf{x}, t) = \mathcal{A}\left[\frac{\partial A^\mu}{\partial t}\right](\mathbf{x}, t).$$

Similar statements hold for other tensors. Hence the application of the operator $\mathcal{A}$ onto the equations of Maxwell (12) yields

$$\partial_\nu \mathcal{A}[F^{\mu\nu}] - \partial^\mu \mathcal{A}[q] = -4\pi \frac{\mathcal{A}[j^\mu]}{c}.$$

In the following subsection, we show that $\mathcal{A}[q] = 0$ holds for a system of charged particles, i.e. the Lorentz condition is satisfied for the time averaged field.

4.3 Classical fields of a charged particle

For simplicity we assume that the restframe S' of the charged particle Q_0 is an inertial system such that $A'^\mu = \sum_{j \in J} A'^\mu_j$ holds in S' (cf. (4)). Moreover, we assume that the particle lives in S' for $t \in (-\infty, \infty)$, i.e. $J = \mathbb{Z}$.

In the restframe S' of the charged particle, the electric and magnetic fields are

$$(17) \qquad \begin{aligned} \mathbf{E}'(\mathbf{x}', t') &= \frac{Q_0}{c} \sum_{j \in \mathbb{Z}} \frac{\frac{\mathrm{d} f_0'}{\mathrm{d} t'}(\omega'(t' - \tau_j') - k'|\mathbf{x}'|)}{|\mathbf{x}'|} \mathbf{e}_{R'} \\ &\quad + Q_0 \sum_{j \in \mathbb{Z}} \frac{f_0'(\omega'(t' - \tau_j') - k'|\mathbf{x}'|)}{|\mathbf{x}'|^2} \mathbf{e}_{R'}, \\ \mathbf{B}'(\mathbf{x}', t') &= 0, \end{aligned}$$

where $\mathbf{e}_R' := \nabla' |\mathbf{x}'|$. From this, (1), (3) and $\mathcal{A}[\frac{\partial f_0'}{\partial t'}] = 0$, we get

$$\mathcal{A}[\phi'](\mathbf{x}', t') = \frac{Q_0}{|\mathbf{x}'|}, \qquad\qquad \mathcal{A}[\mathbf{A}'](\mathbf{x}', t') = \mathbf{0},$$
$$\mathcal{A}[\mathbf{E}'](\mathbf{x}', t') = -\nabla' \mathcal{A}[\phi'](\mathbf{x}') \qquad \text{and} \qquad \mathcal{A}[\mathbf{B}'](\mathbf{x}', t') = \mathbf{0},$$

which are nothing else but the quantities of classical electrodynamics. From these equations, (15), (16) and (8), it follows that

$$\mathcal{A}[q'] = \nabla' \cdot \mathcal{A}[\mathbf{A}'] + \frac{1}{c} \frac{\partial \mathcal{A}[\phi']}{\partial t} = 0 \qquad \text{and} \qquad \mathcal{A}[j'^\mu] := (c\, Q_0\, \delta(\mathbf{x}), \mathbf{0}).$$

Here we have used that $\mathcal{A}[\mathbf{A}'] = \mathbf{0}$ and the classical field $\mathcal{A}[\phi']$ does not depend on time. Thus we interprete $\mathcal{A}[A^\mu]$ as the (ideal) data of the four potential A^μ obtained by a real measurement.

Remark 4 *Assume for the moment that the exact values of A^μ cannot be measured by experiment. Does this mean that it is not reasonable to use the non-instant field A^μ in calculations? If the quantity A^μ is related to reality it must have an effect. The effect is the warranty of causality, i.e. using only $\mathcal{A}[A^\mu]$ yields effects without cause, like the Aharonov-Bohm effect (cf. [1, 4, 12, 6, 17]) but using A^μ gives us the cause. This will be shown in the following two subsections.*

4.4 The field of a solenoid

Consider a solenoid with axis $\{(0, 0, z) \,|\, z \in \mathbb{R}\}$ and radius $R_S > 0$ that was switched on at time $t = -\infty$ (cf. Fig. 1). Our reference frame is the restframe of the solenoid. Because the speed v of the electrons in the solenoid are much smaller than the speed c of light, we model the current in the solenoid by

$$(18) \qquad \mathbf{j} = \sum_{i \in I} \sigma_i(t)\, v\, \frac{\delta(R - R_S)}{R}\, \mathbf{e}_\varphi \,,$$

with the assumptions

$$\mathcal{A}[\sigma_i] = \sigma = const. \qquad \text{and} \qquad v = const.$$

For a fixed $i \in I$ $\sigma_i = \sigma_i(t)$ denotes the surface charge density of the electrons that send out their fields at the same time. In the "time averaged picture" we have the time-independent source term

$$(19) \qquad \mathcal{A}[\mathbf{j}] = \sigma\, v\, \frac{\delta(R - R_S)}{R}\, \mathbf{e}_\varphi \,.$$

Here the important point is that in contrast to (19), the "real" source term (18) is time-dependent.

The symmetry of the solenoid implies $\mathbf{B}(\mathbf{x}, t) = B(|\mathbf{x}|, t)\, \mathbf{e}_z$ which together with $\mathbf{B} = \nabla \times \mathbf{A}$ implies

$$\mathbf{A}(\mathbf{x}, t) = A(|\mathbf{x}|, t)\, \mathbf{e}_\varphi \,.$$

For the solution of (9) with time-dependent source term (18), it follows from pde theory (cf. [14]) that $\mathbf{A}$ does not vanish and is time-dependent outside the solenoid and therefore

$$\mathbf{E}_w := -\frac{1}{c}\, \frac{\partial \mathbf{A}}{\partial t} \qquad \text{(not identical zero)}$$

implies due to the law of induction $\nabla \times \mathbf{E}_w = -\frac{1}{c}\, \frac{\partial \mathbf{B}_w}{\partial t}$ a non-vanishing magnetic field

$$\mathbf{B}_w = \nabla \times \mathbf{A} \,,$$

which is orthogonal to $\mathbf{E}_w$. Indeed, $(\mathbf{E}_w, \mathbf{B}_w)$ is a wave with extrem high frequency $2\pi/\tau$ (cf. (1) and (2)) such that

$$\mathcal{A}[\mathbf{E}_w] = 0 \quad \text{and} \quad \mathcal{A}[\mathbf{B}_w] = 0 \qquad \text{outside the solenoid.}$$

This can be seen as follows. The time-independence of $\mathcal{A}[\mathbf{j}]$ implies the time-independence of $\mathcal{A}[\mathbf{A}]$ and consequently by (16)

$$\mathcal{A}[\mathbf{E}_w] = -\frac{1}{c}\frac{\partial \mathcal{A}[\mathbf{A}]}{\partial t} = 0\,.$$

This and the second equation in (13) together with (15) and (16) imply

$$-\frac{1}{c}\frac{\partial \mathcal{A}[\mathbf{B}_w]}{\partial t} = \nabla \times \mathcal{A}[\mathbf{E}_w] = 0$$

and consequently $\mathcal{A}[\mathbf{B}_w]$ is a constant. That $\mathcal{A}[\mathbf{B}_w]$ vanishes follows from the fact that $\mathcal{A}[\mathbf{B}_w]$ vanishes at infinite.

Since it is very instructive, we give another prove of $\mathcal{A}[\mathbf{B}_w] = 0$ via the law of Ampere (cf. [4]). Consider the situation visualized in Fig. 1. From
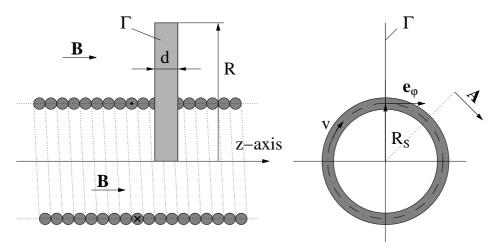

Figure 1: Visualization of a solenoid and its magnetic field.

the law of Ampere (cf. last equation in (13)), the theorem of Stokes (cf. [10]) and the symmetry of the problem, it follows

$$(B(0,t) - B(R,t))\, d = \int_{\partial\Gamma} \mathbf{B}\cdot d\mathbf{x} = \int_{\Gamma} (\nabla \times \mathbf{B})\cdot \mathbf{e}_\varphi\, dS$$

$$= 4\pi \int_{\Gamma}\left(\frac{j}{c} + \nabla q\right) dS + \frac{1}{c}\int_{\Gamma}\frac{\partial \mathbf{E}}{\partial t}\, dS\,.$$

Because of $\mathcal{A}[q] = 0$ and $\mathcal{A}[\frac{\partial \mathbf{E}}{\partial t}] = 0$ (due to $\mathcal{A}[j^\mu] = 0$), application of the time averaging operator $\mathcal{A}$ to our last result yields

(20)
$$\mathcal{A}[B](0,t) - \mathcal{A}[B](R,t) = \frac{4\pi}{cd} \int_\Gamma \mathcal{A}[j] \, dS \, .$$

Here we have used the assumption that the solenoid was switched on at time $t = -\infty$. But (20) implies at once $\mathcal{A}[B](R_1,t) = \mathcal{A}[B](R_2,t)$ for all $R_1, R_2 > R_S$. Because $\mathcal{A}[B](R,t) \overset{R\to\infty}{\longrightarrow} 0$ must hold, we infer

(21)
$$\mathcal{A}[B](R,t) = 0 \qquad \text{for all} \qquad R > R_S \, ,$$

as was to be shown. This calculation shows in which way classical electrodynamics fits into our non-instant field theory.

In summary, the non-instant field model implies that a solenoid emits a wave with extrem high frequency $2\pi/\tau$ and since the time interval τ is extremely small, a measurement of the electromagnetic field outside the solenoid yields the mean zero (correspondence principle).

4.5 The Aharonov-Bohm effect

In this subsection we show that the Aharonov-Bohm effect does not violate causality if the non-instant field model is used and that the classical result follows after time averaging.

First we explain the classical point of view of the Aharonov-Bohm effect. Although measurements show that there is no magnetic field outside a solenoid, an electron passing by is effected if the solenoid was switched on. In mathematical terms this means that due to Stokes theorem and $B_{cl}\, \mathbf{e}_z = \nabla \times \mathbf{A}_{cl}$, the classical quantities $\mathbf{A}_{cl}$ and $\mathbf{B}_{cl}$, which we identify as $\mathcal{A}[\mathbf{A}]$ and $\mathcal{A}[\mathbf{B}]$, satisfy (cf. Fig. 2)

(22)
$$\int_\gamma \mathbf{A}_{cl} \cdot d\mathbf{x} = \int_\Gamma \mathbf{B}_{cl} \cdot d\mathbf{S} = \int_{\Gamma_0} \mathbf{B}_{cl} \cdot d\mathbf{S} = \int_{\gamma_0} \mathbf{A}_{cl} \cdot d\mathbf{x}$$

for every

closed path γ that encloses the simple curve γ_0.

Here the sets Γ and Γ_0 are the closures of γ and γ_0, respectively. Thus it appears as if "information" is transported infinitely fast from the solenoid to the curve $\gamma = \partial\Gamma$ enclosing the solenoid. In other words, the electron passing by feels somehow the magnetic field inside the solenoid and its motion is influence by it. This is the enigma of the Aharonov-Bohm effect.

What happens if we use the non-instant field model? We already know from the previous subsection that there is a wave which transports information with the speed of light c. Let's check the mathematics. From $\mathbf{B} = B\,\mathbf{e}_z$
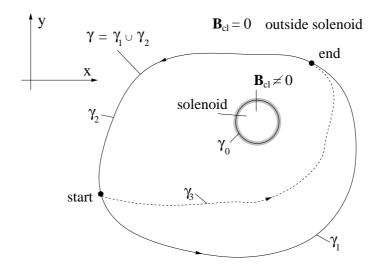
Figure 2: A consequence of the Aharonov-Bohm effect is that $\int_{\gamma} \mathbf{A}_{cl} \cdot d\mathbf{x} = \int_{\gamma_0} \mathbf{A}_{cl} \cdot d\mathbf{x}$ for each γ enclosing γ_0. In particular, $\int_{\gamma_1} \mathbf{A}_{cl} \cdot d\mathbf{x} = \int_{\gamma_3} \mathbf{A}_{cl} \cdot d\mathbf{x}$.

and the theorem of Stokes we get

$$(23) \quad \int_{\partial\Gamma} \mathbf{A} \cdot d\mathbf{x} = \int_{\Gamma} (\nabla \times \mathbf{A}) \cdot \mathbf{e}_z \, dS = \int_{\Gamma} \mathbf{B} \cdot d\mathbf{S} = \int_{\Gamma_0} \mathbf{B} \cdot d\mathbf{S} + \int_{\Gamma\backslash\Gamma_0} \mathbf{B} \cdot d\mathbf{S}$$

in which causality is not violated. Due to $\mathcal{A}[\mathbf{B}] = 0$ outside the solenoid (cf. (21)), the last term satisfies the property

$$\int_{\Gamma\backslash\Gamma_0} \mathcal{A}[\mathbf{B}] \cdot d\mathbf{S} = 0 \,.$$

From this, the assumption that the solenoid was switched on at time $t = -\infty$ and (15), it follows after time averaging of (23) that

$$\int_{\partial\Gamma} \mathcal{A}[\mathbf{A}] \cdot d\mathbf{x} = \int_{\Gamma_0} \mathcal{A}[\mathbf{B}] \cdot d\mathbf{S} = \int_{\partial\Gamma_0} \mathcal{A}[\mathbf{A}] \cdot d\mathbf{x} \,,$$

which is according to our correspondence principle nothing else but (22). Thus if $\mathbf{B}$ is considered as the real magnetic field, then the information is transported with the speed of light but not faster. More to the point, it is causality that "guarantees" the Aharonov-Bohm effect and permits - in the time averaged picture - a Lorentz gauge.

5 Conservation of energy and the quantization of field energy

5.1 The conservation law

The energy conservation law of the non-instant field model reads as follows

$$(24) \qquad \nabla \cdot \mathbf{S} + \frac{1}{c} \frac{\partial u}{\partial t} = 4\pi \left(\rho\, q - \frac{\mathbf{j}}{c} \cdot \mathbf{E} \right),$$

where $q = \partial_\mu A^\mu$ and

$$(25) \qquad u := \frac{1}{2} \left(q^2 + |\mathbf{E}|^2 + |\mathbf{B}|^2 \right) \qquad \text{and} \qquad \mathbf{S} := q\,\mathbf{E} + \mathbf{E} \times \mathbf{B}$$

correspond to the *energy density* and the *energy flux (Poynting vector)* of the electromagnetic field, respectively. Moreover, $j^\mu = (c\,\rho, \mathbf{j})$ denotes the four current and the right hand side models the source of energy.

This conservation law can be derived similarly as in classical field theory (cf. [9]): Scalar multiplication of the second equation in (13) with $\mathbf{B}$ and the last equation in (13) with $\mathbf{E}$ together with identity

$$\nabla \cdot (\mathbf{E} \times \mathbf{B}) = \mathbf{B} \cdot (\nabla \times \mathbf{E}) - \mathbf{E} \cdot (\nabla \times \mathbf{B})$$

yields

$$(26) \quad \frac{1}{c} \frac{\partial}{\partial t} \frac{|\mathbf{E}|^2 + |\mathbf{B}|^2}{2} + 4\pi \frac{\mathbf{j}}{c} \cdot \mathbf{E} + \nabla \cdot (q\,\mathbf{E}) - q\, \nabla \cdot \mathbf{E} + \nabla \cdot (\mathbf{E} \times \mathbf{B}) = 0.$$

Because of the law of Gauss (cf. (13))

$$\nabla \cdot \mathbf{E} = 4\pi\,\rho - \frac{1}{c} \frac{\partial q}{\partial t},$$

equation (26) is equivalent to the *energy conservation law* (24) with u and $\mathbf{S}$ defined as in (25).

5.2 Quantized energy of the field A'^μ_j

In the following we consider a single electron $Q_0 := e^-$ that has life time $N\,\tau'$ ($N \in \mathbb{N}$) in its restframe S'. We show that the energy E'_j of the field A'^μ_j in S'_j is given by

$$(27) \qquad E'_j = \omega'\, \hbar \qquad \text{where} \qquad \omega' = \frac{2\pi}{\tau'}.$$

We recall that S'_j as the inertial systems such that its origin and axes coincide with those of the restframe S' of the electron at time $t' = \tau'_j$ when the field

is sent out (cf. Subsection 2.1). And thus the total energy sent out by the electron in S' during its life time $N \tau'$ is

$$E'_{total} = N \omega' \hbar.$$

Due to the quantization of charges, it follows that the field energy A'^{μ}_j of a charged particle is quantized.

Since $E'_j = E'_0$ for all $j \in J$ we restrict the following calculating to the index $j = 0$. In the frame S'_0 we have $\mathbf{j} = \mathbf{0}$ (vector current) and $\mathbf{B}' = \mathbf{0}$, thus the conservation law simplifies to

(28) $\qquad \dfrac{\partial u'_0}{\partial t'} + c \, \nabla' \cdot \mathbf{S}'_0 = 4 \, \pi \, \rho'_0 \, q'_0 \qquad$ with $\qquad \mathbf{S}'_0 = q'_0 \, \mathbf{E}'_0$

and q'_0, ρ'_0 defined as in (10), (8). From this and the divergence theorem (cf. [10]), it follows that the energy E'_0 is given by

$$E'_0 = c \int\limits_0^{\tau'} \int\limits_{B_{R'}(\mathbf{0})} (\nabla' \cdot \mathbf{S}'_0) \, dV' \, dt' = c \, \tau' \int\limits_{\partial B_{R'}(\mathbf{0})} \mathcal{A}[\mathbf{S}'_0] \cdot \mathbf{e}_{R'} \, dS' \quad \text{with} \quad R' = c \, \tau'.$$

In classical electrodynamics we have $\mathbf{S}'_0 = \mathbf{0}$, but because of

$$\mathcal{A}[q'_0 \, \mathbf{E}'_0] \neq \mathcal{A}[q'_0] \, \mathcal{A}[\mathbf{E}'_0] = 0 \qquad (\mathcal{A}[q'_0] = 0),$$

this is not true for the non-instant field model. From $q'_0 = \frac{1}{c} \frac{\partial \phi'_0}{\partial t'}$ with (1), (2), (17) and (3), it follows that

$$\mathcal{A}[\mathbf{S}'_0](\mathbf{x}', t') \cdot \mathbf{e}_{R'} = \frac{e^2}{\tau' \, c^2 \, |\mathbf{x}'|^2} \int\limits_{\mathbb{R}} \left[\left(\frac{\partial f'_0(\omega' \, t')}{\partial t'} \right)^2 \right] dt' = \frac{\omega' \, \hbar}{4 \, \pi \, \tau' \, c \, |\mathbf{x}'|^2},$$

since

$$\mathcal{A} \left[\frac{\partial f'_0}{\partial t'} (\omega' \, t') \, f'_0(\omega' \, t') \right] = 0 \quad \text{and} \quad \int\limits_{\mathbb{R}} \left[\left(\frac{\partial f'_0(\omega' \, t')}{\partial t'} \right)^2 \right] dt' = \frac{\omega' \, c \, \hbar}{4 \, \pi \, e^2}.$$

From this we get the energy content of A'^{μ}_0 as

$$E'_0 = \int\limits_{\partial B_{R'}(\mathbf{0})} \frac{\omega' \, \hbar}{4 \, \pi \, |\mathbf{x}'|^2} \, dS' = \omega' \, \hbar \qquad (R' = c \, \tau').$$

This proves our claim (27) for the field A'^{μ}_j for $j \in J$.

5.3 Is there an energy problem?

As mentioned in Remark 3, we have

$$\mathcal{A}[X(t)\,Y(t)] \neq \mathcal{A}[X(t)]\,\mathcal{A}[Y(t)]$$

and therefore the correspondence principle cannot be applied to products of physical quantities. Let us look closer at this in the context of field energy. Let u_{cl} and $\mathbf{S}_{cl}$ denote the energy density and Poynting vector used in classical field theory, i.e

$$u_{cl} := \frac{1}{2}\left(\mathcal{A}[E]^2 + \mathcal{A}[B]^2\right) = \mathcal{A}[u_{cl}] \quad \text{and} \quad \mathbf{S}_{cl} := \mathcal{A}[\mathbf{E}] \times \mathcal{A}[\mathbf{B}] = \mathcal{A}[\mathbf{S}_{cl}]\,.$$

Because of

$$\mathcal{A}[q'^2] \neq \mathcal{A}[q']\,\mathcal{A}[q'] = 0 \qquad \text{and} \qquad \mathcal{A}[q'\,\mathbf{E}'] \neq \mathcal{A}[q']\,\mathcal{A}[\mathbf{E}'] = 0\,,$$

it follows that

$$\mathcal{A}[u'] \neq \mathcal{A}[u'_{cl}] \qquad \text{and} \qquad \mathcal{A}[\mathbf{S}'] \neq \mathcal{A}[\mathbf{S}'_{cl}] = 0$$

in the restframe of the charge. Consequently,

$$E'_j = \omega'\,\hbar \neq 0 = c\,\tau' \int\limits_{\partial B_{R'}(\mathbf{0})} \mathcal{A}[\mathbf{S}'_{cl}] \cdot \mathbf{e}_{R'}\,\mathrm{d}S\,,$$

i.e. according to the non-instant field model, a charge sends out quantized energy, which is in contrast to the standard model. But this has to be expected of a non-instant field model.

As a consequence we have to ask: "Is there an energy problem for charged particles?" My point of view is that a charged particle is able to use the energy from its surrounding (e.g. the microwave background radiation) to maintain its field. Besides, there is plenty of energy near our sun. In this way energy is spread over the universe and can be used by other particles. (This is very suggestive, is it not?) Geometric considerations imply that a macroscopic object can maintain its required energy content much longer than a single particle. The worst case is a single particle far away from energy suppliers. So, how does a charged particle react if it does not get enough energy? Is it

- destroyed, or

- does it use its rest mass (or a part of it) to maintain its field or

- does it move without sending out an electromagnetic field?

I consider these questions as very intriguing, but it is beyond the scope of this paper to clarify and solve them.

6 Results

We have presented a non-instant field model for classical field theory that can be quantized with the *Gupta-Bleuler method*. The fact that the non-instant field model does not permit a Lorentz gauge transformation is not an obstacle for quantization and does not contradict the essential results of *classical field theory*. Indeed, our model obeys a correspondence principle, i.e. the classical field quantities $\mathcal{A}_{cl}^{\mu}$, $\mathbf{E}_{cl}$ and $\mathbf{B}_{cl}$ can be obtained via a covariant time averaging operator $\mathcal{A}$ from the respective quantities of the non-instant field model and the Lorentz gauge transformation is satisfied for the time averaged four potential. Although the calculation with the time averaged quantities is very convenient, it does not allow to associate a cause to each effect. As shown in this paper, the non-instant field model predicts that a solenoid sends out an electromagnetic wave of extreme high frequency such that $\mathcal{A}[\mathbf{E}] = \mathbf{0}$ and $\mathcal{A}[\mathbf{B}] = \mathbf{0}$ outside the solenoid, and consequently - according to our correspondence principle - a measurement of the magnetic field yields the value zero outside the solenoid. As a consequence, the *Aharonov-Bohm effect* does not violate causality if the non-instant field model is used. However, if we go over to the "time averaged picture" ($\equiv$ classical electrodynamics), causality seems to be violated.

In addition, we have shown that the field energy of the non-instant field model is quantized.

Conclusions

As we have demonstrated, classical field theory describes the behavior of time averaged field quantities $\mathcal{A}[X]$, $\mathcal{A}[Y]$, ... rather than the non-instant field quantities X, Y, ... themselves. If we grant the quantities X, Y, ... physical significance by using a non-instant field theory, then we are able to describe the *Aharonov-Bohm effect* in a causal way and the quantization of the respective equations of Maxwell is still possible. If you find gauge theories "deeply disturbing and unsatisfactory" (cf. Chapter III.4 in [17]), then the non-instant model for X, Y, ... is the right thing, but if you like gauge theories, then you go over to the time averaged picture $\mathcal{A}[X]$, $\mathcal{A}[Y]$, ... In the latter case, non-causality is not always meaningful. The role of the covariant time averaging operator $\mathcal{A}$ is understood as a causal restriction on the measurement process. That is to say, every measurement of a quantity X will require at least a minimal time period τ and produces the time averaged value $\mathcal{A}[X]$. (Of course this is an idealized picture and the time period τ depends on the observers reference frame.) In other word, even the most cunning experimental setup will not permit to measure X and Y, but $\mathcal{A}[X]$ and $\mathcal{A}[Y]$, and if

$$X \neq Y \qquad \text{and} \qquad \mathcal{A}[X] = \mathcal{A}[Y],$$

then there will be an experimental setup that shows $X \neq Y$. However, if you misinterpret your result, then you conclude a failure of logic or causality.

References

[1] Y. Aharonov and D. Bohm: *Significance of Electromagnetic Potential in the Quantum Theory.* The Physical Reviews, second series, Vol. 115, No. 3, 1959.

[2] L. Bergmann and L. Schäfer: *Lehrbuch der Experimentalphysik. Bd 3: Optik.* Walter de Gruyter, Berlin, 1993.

[3] R. Dautray and J.-L. Lions: *Mathematical Analysis and Numerical Methods for Science and Technology. Volume 5. Springer-Verlag*, New York, 1992.

[4] R. P. Feynman and R. B. Leighton and M. Sands: *The Feynman Lectures on Physics. Vol. 2.* Addison-Wesley Publishing Company, Massachusetts, 6th edition, 1989.

[5] W. Greiner and J. Rafelski: *Spezielle Relativitätstheorie.* Verlag Harri Deutsch, Thun, Frankfurt am Main, 1992.

[6] W. Greiner: *Quantentheorie. Spezielle Kapitel.* Verlag Harri Deutsch, Thun, Frankfurt am Main, 1993.

[7] L. Hörmander: *The Analysis of Linear Partial Differential Operators I.* Springer Verlag, New York, 2nd edition, 2003.

[8] J. D. Jackson: *Classical Electrodynamics.* John Wile & Sons, INC, New York, 1999.

[9] L. Landau and E. Lifschitz: *Klassische Feldtheorie.* Akademie Verlag GmbH, Berlin, 1989.

[10] S. Lang: *Undergraduate Analysis.* Springer-Verlag, New York, 1983.

[11] F. Mandl and G. Shaw: *Quantum Field Theory.* John Wiley & Sons, Ltd, 2nd edition, 2010.

[12] F. Schwabl: *Quantenmechanik.* Springer Verlag, Berlin, 4. Auflage, 1993.

[13] F. Schwabl: *Quantenmechanik für Fortgeschrittene.* Springer Verlag, Berlin, 1997.

[14] F. Treves: *Basic Linear Partial Differential Equations.* Academic Press, INC, London, 1975.

[15] S. Weinberg: *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity.* John Wiley & Sons, New York, 1972.

[16] S. Weinberg: *The Quantum Theory of Field. Vol. I. Foundations.* Cambridge University Press, Cambridge, 1995.

[17] A. Zee: *Quantum Field Theory in a Nutshell.* Princeton University Press, New Jersey, 2003.