HAL
open science

Gauge invariance and relativistic effects in photon absorption and scattering by matter

Nadejda Bouldi, Christian Brouder

To cite this version:

Nadejda Bouldi, Christian Brouder. Gauge invariance and relativistic effects in photon absorption and scattering by matter. 2016. hal-01383508v1

HAL Id: hal-01383508

https://hal.science/hal-01383508v1

Preprint submitted on 18 Oct 2016 (v1), last revised 6 Apr 2017 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Gauge invariance and relativistic effects in photon absorption and scattering by matter

Nadejda Bouldi and Christian Brouder[1]

[1]*Sorbonne Universités, UPMC Univ Paris 06, UMR CNRS 7590, Muséum National d'Histoire Naturelle, IRD UMR 206, Institut de Minéralogie, de Physique des Matériaux et de Cosmochimie, 4 place Jussieu, F-75005 Paris, France*

(Dated: October 18, 2016)

Several semi-relativistic Hamiltonians have been proposed to describe the interaction of light with magnetic matter. In the calculation of x-ray magnetic circular dichroism, these different Hamiltonians lead to different results. Gauge invariance is supposed to help choosing the proper Hamiltonian, but we show that the semi-classical formulation of the absorption cross-section is not gauge invariant in a commonly accepted sense.

To avoid the problem of choosing the proper semi-relativistic Hamiltonian, we start from the fully relativistic absorption and scattering cross-sections given by quantum electrodynamics and we derive a semi-relativistic expansion of the cross-sections by formulating a new many-body Foldy-Wouthuysen transformation of the wavefunctions in the transition matrix elements.

In the process, a new light-matter interaction term emerges, called the spin-position interaction, that contributes significantly to the magnetic x-ray circular dichroism of transition metals. In the case of absorption, the final formula agrees with the result obtained from one of the semi-relativistic Hamiltonians. However, the correct scattering cross-section is not given by any of the semi-relativistic Hamiltonians.

PACS numbers:

I. INTRODUCTION

This paper deals with the description of magnetic effects in spectroscopy. In molecular and condensed matter physics, the effects of relativity are often described by using a one-body Foldy-Wouthuysen Hamiltonian.[1,2] In x-ray scattering spectroscopy, another (many-body) Hamiltonian, due to Blume,[3–5] is almost universally used. In fact, we found in the literature different semi-relativistic Hamiltonians and it is important to determine which one (if any) is correct.

It turns out that the situation is rather confused. We present four non-relativistic Hamiltonians giving different results. Since only gauge-invariant observables are physically meaningful, it seems reasonable to choose a Hamiltonian that gives a gauge-invariant cross-section. However, if we use the semi-classical approach where the incident photon is described by an external field, then the transition probabilities are not gauge invariant. In particular, it is not enough to use a Hamiltonian where $\mathbf{p}$ is replaced by $\mathbf{p} - e\mathbf{A}$. We review the proposed solutions to this problem and we conclude that, to obtain a gauge-invariant result, we need to work in the framework of quantum electrodynamics (QED) where the photon is represented by a state and not by a time-dependent external potential.

This leads to an alternative route to magnetic effects in spectroscopy. Instead of including relativistic effects in a semi-relativistic Hamiltonian, we start from the fully relativistic absorption and scattering cross-sections given by quantum electrodynamics and we apply a many-body Foldy-Wouthuysen transformation to the Dirac wavefunctions in the relativistic matrix elements.

As a result, we find that an additional spin-position term must be including in both absorption and scattering cross-sections.

We describe now the outline of this paper. Section 2 presents four different non-relativistic Hamiltonians that are found in the literature. Section 3 reviews the question of gauge invariance of transition probabilities generated by incident photons. In section 4, we calculate the fully relativistic matrix elements in QED and we derive their multipole expansion up to quadrupole approximation that contains an additional term which is not usually considered. In section 5 we derive a general time-independent Foldy-Wouthuysen-Eriksen transformation which is also valid in the many-body case, and we use it to express dipole and multipole transitions. Section 6 and 7 apply the previous result to the absorption and scattering cross sections, respectively.

II. INTERACTION OF LIGHT WITH MAGNETIC MATTER

In this section, we describe four semi-relativistic Hamiltonians: the one proposed by Blume, the "gauge-invariant" Foldy-Wouthuysen one, the textbook Foldy-Wouthuysen one and the effective Hamiltonian derived in non-relativistic QED (NRQED). To help comparing these Hamiltonians, we express them in a one-particle framework.

The fact that different semi-relativistic Hamiltonians were used in spectroscopy was already noticed.[6,7]

A. The Blume Hamiltonian

Blume discussed the interaction of light with magnetic matter by starting from the Hamiltonian:[3,4]

$$H^{\mathrm{B}} = \frac{\boldsymbol{\pi}^2}{2m} + eV - \frac{e\hbar}{2m}\boldsymbol{\sigma}\cdot\mathbf{B} - \frac{e\hbar}{4m^2c^2}\boldsymbol{\sigma}\cdot(\mathbf{E}\times\boldsymbol{\pi}), \quad (1)$$

where $\boldsymbol{\pi} = \mathbf{p} - e\mathbf{A}$. This Hamiltonian is the sum of four terms: (i) the kinetic energy of the electrons, (ii) the Coulomb interaction of the electrons with the nuclei and the other electrons, (iii) the Zeeman interaction between electrons and magnetic field and (iv) the spin-orbit interaction (because, for a spherical V and a static $\mathbf{A}$, $\boldsymbol{\sigma}\cdot(\mathbf{E}\times\mathbf{p}) = \frac{-1}{r}\frac{dV}{dr}\boldsymbol{\sigma}\cdot(\mathbf{r}\times\mathbf{p}) = \frac{-1}{r}\frac{dV}{dr}\boldsymbol{\sigma}\cdot\mathbf{L}$).

There are several differences between our notation and Blume's: he considers a many-body Hamiltonian (involving sums over electrons) and writes $\sum_{ij} V(r_{ij})$ for our eV, he adds the Hamiltonian H_γ of the free photons, he uses $\mathbf{A}/c$, $\nabla\times\mathbf{A}/c$ and $\mathbf{s}$ where we use $\mathbf{A}$, $\mathbf{B}$ and $\boldsymbol{\sigma}/2$, finally, his Zeeman term is wrong by a factor of 2 in his first two papers on the subject,[3,4] but this was corrected in the third one.[5] In this third paper, Blume also replaces $\mathbf{E}$ by $-\dot{\mathbf{A}}$. This is not compatible with his quantized description of the photon field. Indeed, the time-derivative $\dot{\mathbf{A}}$ is present in the Lagrangian but, after the Legendre transformation leading to the Hamiltonian, $\dot{\mathbf{A}}$ is replaced by its canonical momentum $-\mathbf{E}$. Note that Blume does not sketch any derivation of his Hamiltonian.

B. Foldy-Wouthuysen Hamiltonian

We consider now the so-called "gauge-invariant" Foldy-Wouthuysen Hamiltonian for positive-energy states up to order $1/(mc)^2$:[8]

$$H^{\mathrm{FW}} = H^{\mathrm{B}} + mc^2 - \frac{e\hbar^2}{8m^2c^2}\nabla\cdot\mathbf{E} - \frac{ie\hbar^2}{8m^2c^2}\boldsymbol{\sigma}\cdot(\nabla\times\mathbf{E}).$$

The difference between the Foldy-Wouthuysen and the Blume Hamiltonians consists of three terms: the rest energy mc^2 of positive-energy eigenstates, the Darwin term proportional to $\nabla\cdot\mathbf{E}$ and a last term, proportional to $\boldsymbol{\sigma}\cdot(\nabla\times\mathbf{E})$ and called the *curl-term*, that we discuss now. A basic difference between H^{B} and H^{FW} must first be stressed: the former is a QED expression where the quantum fields $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{E}$ are independent of time because they are written in the Schrödinger representation, while the latter was derived under the assumption that $\mathbf{A}$ and V are external time-dependent potentials. In particular, the curl-term disappears if the external field $\mathbf{A}$ is independent of time.[9] In the semi-classical treatment of light-matter interaction, the photons are represented by an external time-dependent potential and this term is present.

All these Hamiltonians can be written $H(\mathbf{A},\Phi)$, where the total vector potential $\mathbf{A}$ and scalar potential Φ are written as a sum $\mathbf{A} = \mathbf{a}_0+\mathbf{a}$, $\Phi = \phi_0+\phi$, of static external potentials $\mathbf{a}_0$ and ϕ_0 (representing the static internal and external fields) perturbed by dynamical potentials $\mathbf{a}$ and ϕ representing the incident electromagnetic wave. We write the interaction Hamiltonian as $H_I = H(\mathbf{A},\Phi) - H(\mathbf{a}_0,\phi_0)$. The two Hamiltonians H^{B} and H^{FW} lead to two different interactions: $H_I^{\mathrm{B}} = h_1+h_2+h_3+h_4+h_5+h_6$ and $H_I^{\mathrm{FW}} = H_I^{\mathrm{B}} + h_7$, where

$$\begin{aligned}
h_1 &= \frac{e^2}{2m}\mathbf{a}^2,\\
h_2 &= -\frac{e}{m}\mathbf{a}\cdot\boldsymbol{\pi}_0,\\
h_3 &= -\frac{e\hbar}{2m}\boldsymbol{\sigma}\cdot(\nabla\times\mathbf{a}),\\
h_4 &= \frac{e^2\hbar}{4m^2c^2}\boldsymbol{\sigma}\cdot(\mathbf{e}\times\mathbf{a}),\\
h_5 &= -\frac{e\hbar}{4m^2c^2}\boldsymbol{\sigma}\cdot(\mathbf{e}\times\boldsymbol{\pi}_0),\\
h_6 &= \frac{e^2\hbar}{4m^2c^2}\boldsymbol{\sigma}\cdot(\mathbf{e}_0\times\mathbf{a}),\\
h_7 &= -\frac{ie\hbar^2}{8m^2c^2}\boldsymbol{\sigma}\cdot(\nabla\times\mathbf{e}),
\end{aligned}$$

with $\boldsymbol{\pi}_0 = \mathbf{p} - e\mathbf{a}_0$. The *curl-term* in H^{FW} is the origin of the presence of h_7 in H_I^{FW}. The Darwin term gives no contribution to the interaction because $\nabla\cdot\mathbf{e}$ is zero for the electromagnetic wave. The terms h_5 and h_6 were omitted by Blume, who considered them to be small.[3] We shall see that h_5 is the source of a spin-position term which is not negligible in x-ray magnetic circular dichroism (XMCD) spectra.[10]

C. Textbook Foldy-Wouthuysen Hamiltonian

Standard textbooks often derive a Foldy-Wouthuysen Hamiltonian H^{TFW} which is the same as H^{FW}, except for the fact that $\boldsymbol{\pi}$ is replaced by $\mathbf{p}$ in the spin-orbit term.[11,12] A mass-velocity term $-(\mathbf{p}\cdot\mathbf{p})^2/8m^3c^2$ is often added[11] but its contribution to the radiation-matter interaction is zero. The difference with the Foldy-Wouthuysen Hamiltonians is a term in $\boldsymbol{\sigma}\cdot\mathbf{E}\times\mathbf{A}$. This results in the absence of h_4 and h_6 in the perturbation Hamiltonian, which changes the transition probabilities.

D. Foldy-Wouthuysen subtelties

In this section, we would like to stress that the Foldy-Wouthuysen transformation must be used with great care to calculate matrix elements.

The first subtelty was noticed by Nieto:[13,14] If $|\Psi\rangle$ is a solution of the time-dependent Dirac equation $(i\hbar\partial_t - H)|\Psi\rangle = 0$, then the Foldy-Wouthuysen transformation turns it into $|\psi\rangle = U|\Psi\rangle$, where U is a unitary time-dependent operator and $|\psi\rangle$ is a solution of the time-dependent Schrödinger equation with the time-dependent Foldy-Wouthuysen Hamiltonian $H' =$

$UHU^{-1}+i\hbar(\partial_t U)U^{-1}$. The curl-term in $H^{\rm FW}$ originates from $i\hbar(\partial_t U)U^{-1}$. In the following, an uppercase letter refers to a solution of the Dirac equation and a lowercase letter to its Foldy-Wouthuysen.

As a consequence, a matrix element $\langle\Phi|H|\Psi\rangle$ is not equal to $\langle\phi|H'|\psi\rangle$, but to $\langle\phi|H'-i\hbar(\partial_t U)U^{-1}|\psi\rangle$. In other words, H' has to be used to calculate the states $|\phi\rangle$ and $|\psi\rangle$ but not to calculate the matrix elements of the Hamiltonian.

The second subtelty was observed by Yang.[15] A transition probability is calculated as $|\langle\phi_n|\psi(t)\rangle|^2$, where $|\phi_n\rangle$ is a solution of the time-independent Schrödinger equation $H(\mathbf{a}_0,\phi_0)|\psi_n\rangle = E_n|\phi_n\rangle$, while $|\psi(t)\rangle$ is a solution of the time-dependent Schrödinger equation with Hamiltonian $H(\mathbf{A},\Phi)$. Since the Foldy-Wouthuysen operator U depends explicitly on the potentials in the Hamiltonian, the Foldy-Wouthuysen transformations to be applied to $|\phi_n\rangle$ and $|\psi(t)\rangle$ are different because they correspond to different Hamiltonians.

The last example is a development of the previous one: the Foldy-Wouthuysen interaction Hamiltonian $H'_I = H'(\mathbf{A},\Phi) - H'(\mathbf{a}_0,\phi_0) \neq UH_IU^{-1}$ because the Foldy-Wouthuysen transformation corresponding to $H'(\mathbf{A},\Phi)$ and $H'(\mathbf{a}_0,\phi_0)$ are different. As a consequence, $\langle\phi_f|H'_I|\psi_i\rangle$ is not equal to $\langle\Phi_f|H_I|\Psi_i\rangle$.

E. NRQED

To deal with QED calculations involving bound states, Caswell and Lepage proposed an alternative approach to relativistic effects, called non-relativistic QED (NRQED), which turned out to be highly successful.[16] They wrote the most general gauge-invariant non-relativistic Lagrangian terms and fitted the coefficients of these terms to known QED processes.[17]

The corresponding NRQED Hamiltonian is the same as $H^{\rm FW}$ up to order c^{-2}, but its interpretation is different.[17] Indeed, NRQED is a quantum field theory, and the fields are independent of time in the Schrödinger representation. However, the curl-term is present in time-independent NRQED although it is generated by a time-dependence in $H^{\rm FW}$. In particular, the curl-term must not be removed from the Hamiltonian to calculate matrix elements of the Hamiltonian operator, in contrast to the example of section II D.

Besides these four different Hamiltonians, we consider an additional source of discrepancies between authors: the commutators.

F. Commutators

To derive the multipole expansion of the matrix element of H_I, it is useful to replace $\boldsymbol{\pi}$ by a commutator with $H_0 = H(\mathbf{a}_0,\phi_0)$. The derivations that start from Blume's interaction Hamiltonian usually use the relation:[6,18]

$$\mathbf{p} = \frac{mi}{\hbar}[H_0,\mathbf{r}]. \tag{2}$$

However, if one considers the static Hamiltonian given by Blume (1), its commutator with $\mathbf{r}$ is:

$$[H_0^{\rm B},\mathbf{r}] = -\frac{i\hbar}{m}\boldsymbol{\pi}_0 + \frac{e\hbar}{4m^2c^2}(i\hbar)(\boldsymbol{\sigma}\times\mathbf{e}_0),$$

which is different from Eq. (2) because $\mathbf{p}$ is replaced by $\boldsymbol{\pi}_0 = \mathbf{p} - e\mathbf{a}_0$ and because of the term proportional to c^{-2}. The commutator of $\mathbf{r}$ with $H_0^{\rm TFW}$ and $H_0^{\rm FW}$ are the same. In $H_I^{\rm FW}$ and $H_I^{\rm B}$, when $\boldsymbol{\pi}_0$ in h_2 is rewritten as a function of the commutator, the extra relativistic term leads to the cancellation of h_6, which is important in XMCD. On the other hand, it leads to a contribution $\frac{e^2\hbar}{4m^2c^2}\boldsymbol{\sigma}\cdot[\nabla v_0\times\mathbf{a}]$ in $H_I^{\rm TFW}$.

If the mass-velocity term $-(\mathbf{p}\cdot\mathbf{p})^2$ is present in H_0, the additional contribution to the commutator, $\frac{i\hbar(\mathbf{p}\cdot\mathbf{p})\mathbf{p}}{2m^3c^2}$ is small compared to $\frac{i\hbar}{m}\mathbf{p}$ if the order of magnitude of the kinetic energy of the core state satisfies $E_k << mc^2$.

For all the Hamiltonians presented here, using the relation $[\mathbf{p},v_0] = i\hbar\nabla v_0$, the electric field in matter writes at zeroth order in c^{-2} as a function of the commutator of $\boldsymbol{\pi}_0$ with H_0:

$$\mathbf{e}_0 = -\nabla v_0 = \frac{-i}{e\hbar}[\boldsymbol{\pi}_0,H_0].$$

In the case of absorption, the commutator transforms into a factor $-\hbar\omega$ in the cross section so that h_5 and h_6 lead to the same contribution to the matrix element:

$$\frac{-ie\hbar\omega}{4m^2c^2}\boldsymbol{\sigma}\cdot(\mathbf{a}\times\boldsymbol{\pi}_0),$$

which we call *spin-position* interaction. We observed that this contribution can appear two times, one time or cancel completely, according to which Hamiltonian and which commutator was used. These different results illustrate the importance of a proper choice of the starting semi-relativistic Hamiltonian to compute the transition rate. In this paper, we show that the Hamiltonian $H^{\rm B}$ gives the correct result for absorption (but not for scattering). Since physically meaningful observables are gauge invariant, we first check the gauge invariance of the cross-section.

III. GAUGE INVARIANCE

The gauge invariance of the absorption and scattering cross-sections of light is a long-standing problem. It started in 1952 when Willis Lamb calculated the spectrum of Hydrogen in two gauges leading to different results[19], and gave rise to a long series of papers up to this day.[15,20–69]

In 1987, the same Willis Lamb (then Nobel prize winner in physics) still considered it as "one of the outstanding problems of modern quantum optics."[52]

A. The principle of gauge invariance

The two homogeneous Maxwell equations $\nabla\times\mathbf{E}+\dot{\mathbf{B}}=0$ and $\nabla\cdot\mathbf{B}=0$, where the dot denotes time derivative, imply the local existence of a vector potential $\mathbf{A}$ and a scalar potential Φ such that $\mathbf{B}=\nabla\times\mathbf{A}$ and $\mathbf{E}=-\nabla\Phi-\dot{\mathbf{A}}$. We denote $A=(\Phi,\mathbf{A})$. The same $\mathbf{E}$ and $\mathbf{B}$ are obtained from the potentials $A'=(\Phi-\dot{\Lambda},\mathbf{A}+\nabla\Lambda)$, that we also denote $A'=A-\partial\Lambda$, where Λ is any smooth function of space and time. In classical electromagnetism, the gauge invariance means that the physics described by A and A' is the same.

In quantum mechanics, consider a non-relativistic Hamiltonian

$$H_A \;=\; \frac{(\mathbf{p}-e\mathbf{A})^2}{2m}+e\Phi,$$

or a relativistic Hamiltonian

$$H_A \;=\; c\boldsymbol{\alpha}\cdot(\mathbf{p}-e\mathbf{A})+mc^2\beta+e\Phi,$$

where $\boldsymbol{\alpha}=(\alpha_x,\alpha_y,\alpha_z)$ and β are the Dirac matrices. Both Hamiltonians are of the form $H_A=f(\mathbf{p}-e\mathbf{A})+e\Phi$, where f is some function. For such Hamiltonians it can be checked that $M_\Lambda^\dagger(i\hbar\partial_t-H_{A'})M_\Lambda=i\hbar\partial_t-H_A$, where $M_\Lambda=e^{ie\Lambda/\hbar}$. As a consequence, if ψ is a solution of the time-dependent Schrödinger equation $i\hbar\dot{\psi}=H_A\psi$, then $\psi'=M_\Lambda\psi$ is a solution of $i\hbar\dot{\psi}'=H_{A'}\psi'$.

In quantum mechanics, gauge transformation consists in both a change of the potentials and a change in the phase of the wavefunctions. An observable O_A depending on the electromagnetic potential A is said to be *gauge invariant* if $M_\Lambda^\dagger O_{A'}M_\Lambda=O_A$ for every function $\Lambda(t,\mathbf{r})$. An observable must be gauge invariant to be considered a *true physical quantity*.[70]

The principle of gauge invariance has become a cornerstone of particle physics. Since general relativity can also be considered as a gauge theory[71], it may be safely said that gauge invariance was the guiding principle of most of the fundamental physics of the twentieth century. Therefore, we need to check that the cross-section formulas are gauge invariant to ensure their true physical nature.

Note that the time-dependent Dirac or Schrödinger equations are always gauge invariant but the time-independent ones are not because H_A is not gauge invariant due to the the scalar potential Φ. Indeed, under a gauge transformation Φ becomes $\Phi-\dot{\Lambda}$ and the term $\dot{\Lambda}$ cannot be compensated for in the absence of a time derivative.

B. Gauge invariance of transition probabilities

In time-dependent perturbation theory, a system is assumed to be the ground state $|\phi_g\rangle$ of a time-independent Hamiltonian H_{a_0}. Then, at time t_0, an electromagnetic wave represented by the time-dependent potential a is added to the system (with total potential $A=a_0+a$), which is represented at time t by the state $|\psi(t)\rangle$. A good way to take both the initial state and the dynamics into account is to use the evolution operator $U_A(t,t_0)$, which is the solution of $i\hbar\partial_t U_A(t,t_0)=H_A(t)U_A(t,t_0)$ with the boundary condition $U_A(t_0,t_0)=1$. Thus, $|\psi(t)\rangle=U_A(t,t_0)|\phi_g\rangle$. The probability of a transition to the eigenstate $|\phi_n\rangle$ of H_{a_0} at time t is

$$P_{ng}(t) \;=\; |\langle\phi_n|\psi(t)\rangle|^2=|\langle\phi_n|U_A(t,t_0)|\phi_g\rangle|^2. \quad (3)$$

If we carry out a gauge transformation from a to $a'=a-\partial\Lambda$, then the evolution operator becomes[48]

$$U_{A'}(t,t_0)=M_\Lambda(t)U_A(t,t_0)M_\Lambda^\dagger(t_0), \quad (4)$$

where $A'=a_0+a'$.

In general, $P'_{ng}(t)=|\langle\phi_n|U_{A'}(t,t_0)|\phi_g\rangle|^2\neq P_{ng}(t)$ and the transition probabilities calculated in the two gauge are different. Many papers describe the discrepancy between the probability calculated in two different gauges, which is generally not small.[19,54,65] By properly choosing Λ, the discrepancy can even be made arbitrary large.[64]

The absence of gauge invariance is due to the fact that the operator is transformed but not the states. This is called a *hybrid transformation* in the literature.[29]

C. Proposed solutions

Several solutions to this alarming problem have been proposed. Since no clear consensus appears to have emerged,[61] we present a critical review of the subject.

The first solution is based on the observation that, if instead of gauge-transforming a we transform the potential a_0 of the initial Hamiltonian to get $H_{a'_0}$, where $a'_0=a_0-\partial\Lambda$, then the evolution operator becomes again $U_{A'}(t,t_0)$ (because $a'_0+a=a_0+a-\partial\Lambda=A'$) but the eigenstates are also changed $|\phi'_g\rangle=M_\Lambda(t_0)|\phi_g\rangle$ and $|\phi'_n\rangle=M_\Lambda(t)|\phi_n\rangle$. Therefore, the transition probability is now conserved. In other words, gauge invariance is lost if we subtract $\partial\Lambda$ from the perturbation but not if we subtract it from the unperturbed Hamiltonian $H_0=H_{a_0}$.

Starting with Forney and coll.[26] and Epstein,[21] many authors proposed to solve the problem by using the so-called *consistent procedure*, where both the evolution operator and the eigenstates are modified whenever a gauge transformation is made. Although this is not consistent with standard time-dependent perturbation theory, this ensures that, if we start from an initial gauge and make any gauge transformation, then the transformed transition probabilities are the same as in the initial gauge.

However, as noticed by Yang,[15] this does not really solve the problem because, if we start the calculation with the initial potential H_{a_0} and the perturbation a', the transition probability is $P'_{ng}(t)$. If we then use the consistent procedure to come back to the perturbation a, then we still find $P'_{ng}(t)$ and we do not recover the result $P_{ng}(t)$. In other words, the transition probability is now

gauge invariant (in the sense that a change of gauge does not modify the result) but it is gauge-dependent (in the sense that the result depends on the gauge we use to start the calculation). This gauge dependence would be a serious problem because we would have to select the "true" physical gauge.

A second solution was proposed by Grant[23], who states that gauge invariance of transition matrix elements is ensured if the potential is coupled to a conserved current. However, his proof is wrong (the first term of his first equation on p. 1472 does not vanish because $\partial_\mu(\Lambda j^\mu)$ is not integrated over time).

A third solution appeared in a series of papers starting in 1976,[15,24,31,32,36,38,39,44,48,50,55], where Yang and collaborators proposed to define a gauge invariant transition probability. His idea is to start from the gauge-invariant (but time-dependent) initial Hamiltonian

$$H_0(t) \;=\; \frac{(\mathbf{p} - e\mathbf{a}_0 - e\mathbf{a}(t))^2}{2m} + \mathbf{e}V, \qquad (5)$$

where V describes the electron-electron and electron-nuclear interactions so that $H = H_0 + e\phi$: the perturbation is only the scalar potential ϕ. Then, the Hamiltonian $H_0(t)$ is diagonalized at every time t: $H_0(t)|\phi_n(t)\rangle = E_n(t)|\phi_n(t)\rangle$ and the transition are calculated between the time-dependent states $|\phi_n(t)\rangle$. The corresponding transition probabilities are indeed gauge invariant. This solution has been widely used up to this day,[53,67,69] although it was also strongly criticized.[22,33,34,40–42,45–47,72,73] The main arguments against Yang's interpretation are: (i) the quantity $E_n(t)$ is not physical because you cannot measure an energy at a given time with arbitrary precision; (ii) the time-dependent states $|\phi_n(t)\rangle$ can be neither prepared nor detected; (iii) the term V in Eq. (5) should be removed from $H_0(t)$ because it is a scalar potential and, as such, not gauge invariant. But if V is removed, then $H_0(t)$ is so far from the true Hamiltonian that perturbation theory is no longer valid.

Following Goldman,[14] Feuchtwang, Kazes and coll. proposed the following alternative solution.[40,41,46,47,74] They started from the well-known fact that the equations of motion of a Lagrangian are not modified by the addition of the total time derivative of a function.[70] Thus, two Lagrangians that differ by a total time derivative are equivalent.[75] Then, they remark that the addition of a total time derivative $e\dot{\Lambda}$ to the Lagrangian induces a gauge transformation $A \to A - \partial\Lambda$ of the Hamiltonian.[62,70,74] Finally, they use such a total derivative to compensate for the electric potential that is the cause of the gauge variance of the Hamiltonian. However, it is difficult to distinguish this procedure from picking up a specific gauge, namely the Weyl or temporal gauge where the scalar potential vanishes.

We can conclude this short review by stating that no solution was found fully satisfactory. To determine when gauge invariance can be achieved, we consider a Dirac Hamiltonian in two gauges A and $A' = A - \partial\Lambda$ and we calculate the difference

$$\langle\psi|H_A - H_{A'}|\psi'\rangle \;=\; e\langle\psi|c\boldsymbol{\alpha}\cdot\nabla\Lambda + \dot{\Lambda}|\psi'\rangle.$$

The advantage of the Dirac Hamiltonian is that the difference $H_A - H_{A'}$ does not depend on A, but a similar calculation can be carried out in the non-relativistic case.[73] Then, we notice that $c\boldsymbol{\alpha}\cdot\nabla\Lambda = (i/\hbar)[H_D, \Lambda]$ for any Dirac Hamiltonian H_D. Thus, if $|\psi\rangle$ and $|\psi'\rangle$ are eigenstates of H_D with energy E and E', we obtain

$$\langle\psi|H_A - H_{A'}|\psi'\rangle = e\langle\psi|\dot{\Lambda}|\psi'\rangle + ie\frac{E - E'}{\hbar}\langle\psi|\Lambda|\psi'\rangle. \quad (6)$$

If we consider the absorption cross-section (up to first order) of a photon of energy $\hbar\omega$, then energy conservation implies that $E' = E + \hbar\omega$. Thus, if Λ satisfies $\dot{\Lambda} = -i\omega\Lambda$, then $\langle\phi|H_A - H_{A'}|\phi'\rangle = 0$.[55,73] In other words, by restricting the gauge transformation to those satisfying $\dot{\Lambda} = -i\omega\Lambda$, the absorption cross-section, calculated up to first order in perturbation theory, is gauge invariant. However, in the resonant scattering cross-section, energy conservation does not apply to the transition involving intermediate states, and the cross-section is not gauge invariant even for those gauges.[25,55,65]

Equation (6) shows that the matrix elements are also gauge invariant for a time-independent gauge transformation and energy conserving processes (i.e. $E' = E$). However, the gauge invariance principle is not supposed to restrict to gauges satisfying specific constraints such as $\dot{\Lambda} = -i\omega\Lambda$ or $\dot{\Lambda} = 0$.

This rapid overview shows that, in the semi-classical approach where the photon is represented by an external potential, the transition probabilities are not gauge-invariant and no proposed solution has reached general exceptance. Therefore, we turn now to a framework where both electrons and photons are quantized: quantum electrodynamics (QED).

D. Quantum electrodynamics

In QED the incident light is no longer described by an external electromagnetic field but by a photon, i.e. a state in a bosonic Fock space. Therefore, a scattering experiment is now described by the transition from an initial state involving both the electronic system in its ground state and the incident photon, to a final state involving both the electronic system in its (possibly) excited state and the scattered photon. Thus, the energy of the inital and final states is the same and, in the Schrödinger picture, the gauge transformation can be expressed in terms of time-independent operators. The remark at the end of the previous paragraph lets us think that transition probabilities, which are now described through the so-called S-matrix, are now gauge invariant.

However, a review of the literature on quantum electrodynamics is rather confusing. For standard textbooks "the S-matrix is gauge invariant by construction".[76] For

more mathematically-minded authors, "an even approximately complete solution [of the gauge invariance problem] does not exit."[77] The difficulty comes from the fact that a gauge transformation also implies a change of the space of states. For example, in the Coulomb gauge, only the transverse degrees of freedom are quantized and the photon states form a Hilbert space built by acting on the vacuum with creation operators of left and right polarized photons, while in the Lorenz gauge four degrees of freedom are quantized and the states form a Krein space built by acting on the vacuum with creation operators of the left, right, longitudinal and scalar photons. In the Lorenz gauge, the Lorenz condition cannot be satisfied as an operator equation,[78] it becomes a subsidiary condition used to determine a subspace of physical states.

In other words, the state spaces of the Coulomb and Lorenz gauges have a quite different nature and the relation between them is difficult to control, although it was possible to devise a common framework for a certain class of gauges.[56] Moreover, the gauge-invariance can only be expected for the renormalized S-matrix[79–81] and renormalization in a general gauge is not clear.[56]

However, the case of a general *infinitesimal* gauge transformation is well established within the Becchi-Rouet-Stora-Tyutin (BRST) approach: matrix elements of gauge-invariant operators between physical states are independent of the choice of the gauge-fixing functional if and only if the physical states $|\alpha\rangle$ satisfy $Q|\alpha\rangle = 0$, where Q is the BRST charge.[82,83]The case of finite (i.e. not infinitesimal) gauge transformations is in progress.[84,85]

To summarize the discussion, the gauge invariance of the renormalized S-matrix is established for infinitesimal gauge transformations and for reasonably large classes of gauges.[29,56,80,86–91] In other words, it is proved at the physicist level of rigour.

The most studied gauges are the Lorenz and Coulomb gauges. Renormalization is perfectly established for the Lorenz gauge, but in most practical calculations the subsidiary condition (Gauss' law) is not enforced.[92] Although it was proved that the S-matrix elements are often the same with and without the subsidiary condition,[22,29,41,93,94] this fails when the Hamiltonian is suddenly changed,[95] as in the sudden creation of a core hole in photoemission or x-ray absorption[96,97]. In that case, Gauss' law has to be imposed in the Lorentz gauge and the Coulomb gauge result is recovered.[95] To be on the safe side, we shall use the Coulomb gauge.

IV. RELATIVISTIC MATRIX ELEMENTS

After having clarified the issue of gauge invariance, we can calculate the relativistic matrix elements that will be used for establishing x-ray scattering and absorption cross-sections in the framework of QED in the Coulomb gauge.

A. The Hamiltonian

The quantum field Hamiltonian describing the interaction of light with matter in the Coulomb gauge is:[29,70,98,99]

$$H = H_e + H_\gamma + H_{e\gamma},$$

where

$$H_e = \int d\mathbf{r}\psi^\dagger(\mathbf{r})\big(c\boldsymbol{\alpha}\cdot(-i\hbar\nabla - e\mathbf{a}) + \beta mc^2 + e\phi\big)\psi(\mathbf{r}) + \int d\mathbf{r}d\mathbf{r}'\frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{8\pi\epsilon_0|\mathbf{r}-\mathbf{r}'|},$$

where ϕ is a time-independent scalar external potential (for instance the nuclear potential), $\mathbf{a}$ is a time-independent vector potential (describing an external magnetic field) and ψ are fermion field operators . Normal ordering is implicit in H_e. It is the QED form of the Dirac Hamiltonian in the Coulomb gauge. The many-body version of this Hamiltonian is

$$H_N = \sum_{n=1}^{N} c\boldsymbol{\alpha}_n\cdot(-i\hbar\nabla_n - e\mathbf{a}(\mathbf{r}_n)) + \beta_n mc^2 + e\phi(\mathbf{r}_n) + \sum_{m\neq n}\frac{e^2}{8\pi\epsilon_0}\frac{1}{|\mathbf{r}_m-\mathbf{r}_n|},$$

where $\boldsymbol{\alpha}_n$ and β_n act on the nth Dirac electron. It can be given a well-defined mathematical meaning if the electronic system is described with respect to the Dirac sea,[100] although the physical validity of the Dirac sea is sometimes disputed.[101]

The photon Hamiltonian is

$$H_\gamma = \frac{\epsilon_0}{2}\int d\mathbf{r}|\mathbf{E}^\perp|^2 + c^2|\mathbf{B}|^2 = \sum_{\mathbf{k},l}\hbar\omega_{\mathbf{k},l}a^\dagger_{\mathbf{k},l}a_{\mathbf{k},l},$$

where l stands for the polarization of a mode (there are two independent directions for a given wavevector $\mathbf{k}$) and

$$H_{e\gamma} = -ec\int d\mathbf{r}\psi^\dagger(\mathbf{r})\boldsymbol{\alpha}\cdot\mathbf{A}(\mathbf{r})\psi(\mathbf{r}),$$

describes the photon-matter interaction and $\mathbf{A}$ satisfies the Coulomb gauge condition $\nabla\cdot\mathbf{A} = 0$. According to Bialynicki-Birula, the Hamiltonian H also describes the dynamics of gauge-invariant states in any gauge,[99] but the concept of a gauge-invariant state is still controversial.[102,103] The many-body version of this interaction Hamiltonian is

$$H_I = -ec\sum_{n=1}^{N}\boldsymbol{\alpha}_n\cdot\mathbf{A}(\mathbf{r}_n).$$

B. S-matrix elements

Since we saw that the S-matrix is gauge invariant, we calculate its matrix-elements. We recall that

$$S \;=\; \lim_{\epsilon\to 0} T(e^{-\frac{i}{\hbar}\int_{-\infty}^{\infty} H_\epsilon(t)dt}), \tag{7}$$

where $H_\epsilon(t) = e^{-\epsilon|t|}e^{iH_0t}H_{e\gamma}e^{-iH_0t}$. The adiabatic switching factor $e^{-\epsilon|t|}$ enables us to describe physical processes as matrix elements of S between eigenstates of $H_0 = H_e + H_\gamma$. The limit can be shown to exist up to technical assumptions.[104] Note that H_0 is not quadratic because of the Coulomb interaction term in H_e. The eigenstates of H_e are correlated multi-electronic wave-functions. As a consequence, we are not in the textbook framework, the time-dependence of $H_\epsilon(t)$ cannot be calculated explicitly and the Feynman diagram technique is no longer available to describe electrons. We can bypass this problem with the so-called "old-fashioned" (i.e. non-covariant) approach,[105] using matrix elements of $H_\epsilon(t)$ between eigenstates of H_0. Then, cross-sections are expressed in terms of the S-matrix and T-matrix elements related by:

$$\langle m|S|n\rangle \;=\; \delta_{mn} - 2i\pi\delta(e_m - e_n)\langle m|T|n\rangle,$$

where, up to second order,

$$\langle m|T|n\rangle = \langle m|H_{e\gamma}|n\rangle + \sum_p \frac{\langle m|H_{e\gamma}|p\rangle\langle p|H_{e\gamma}|n\rangle}{e_p - e_n + i\gamma}, \tag{8}$$

where $|m\rangle$, $|p\rangle$ and $|n\rangle$ are eigenstates of H_0 with energy e_m, e_p and e_n, respectively. The term $i\gamma$ was added as a heuristic way to avoid divergence at resonance (i.e. when the states $|n\rangle$ and $|p\rangle$ are degenerate). More sophisticated methods exist to deal with such degeneracies[106] but they would bring us too far. From the physical point of view, γ describes the life-time of the state $|p\rangle$, which can decay by photon or Auger emission. The sign of the damping term γ has been the object of some controversy.[107–110]

Let us stress again that, since H_e is not quadratic, we essentially work in the Schrödinger picture, where the operators are independent of time, instead of the standard interaction picture which is used in most textbooks. Both approaches are equivalent.[111] A modern version of the Schrödinger picture of QFT is given by Hatfield.[112]

Our purpose is now to calculate the matrix elements $\langle m|H_{e\gamma}|n\rangle$, where $H_{e\gamma}$ is independent of time. The second quantized expression for the photon field in the Schrödinger picture is:[9]

$$\mathbf{A}(\mathbf{r}) = \sum_{\mathbf{k},l}\sqrt{\frac{\hbar}{2\epsilon_0 V\omega_{\mathbf{k}}}}\left(\boldsymbol{\epsilon}_{\mathbf{k},l}a_{\mathbf{k},l}\mathrm{e}^{i\mathbf{k}\cdot\mathbf{r}} + \boldsymbol{\epsilon}^\star_{\mathbf{k},l}a^\dagger_{\mathbf{k},l}\mathrm{e}^{-i\mathbf{k}\cdot\mathbf{r}}\right).$$

Note that we do not assume the polarization vectors $\boldsymbol{\epsilon}_{\mathbf{k},l}$ to be real.

We denote $|n\rangle = a^\dagger_{\mathbf{k},l}|0\rangle|\Psi_n\rangle$ an eigenstate of H_0 where one photon is present in mode $\mathbf{k}, l$ and the electrons are in state $|\Psi_n\rangle$ with energy E_n. The energy of $|n\rangle$ is $e_n = \hbar\omega_{\mathbf{k},l} + E_n$. The interaction Hamiltonian $H_{e\gamma}$ is linear in $\mathbf{A}$ which is linear in photon creation and annihilation operators so that only one-photon transitions are possible. The state $|n\rangle$ can make transitions towards $|a\rangle = |0\rangle|\Psi_m\rangle$ by absorption and $|e\rangle = a^\dagger_{\mathbf{k},l}a^\dagger_{\mathbf{k}',l'}|0\rangle|\Psi_m\rangle$ by emission. From now on, we denote $\omega = \omega_{\mathbf{k},l}$, $\boldsymbol{\epsilon} = \boldsymbol{\epsilon}_{\mathbf{k},l}$, $\omega' = \omega_{\mathbf{k}',l'}$ and $\boldsymbol{\epsilon}' = \boldsymbol{\epsilon}_{\mathbf{k}',l'}$. The corresponding matrix elements are:

$$\langle a|H_{e\gamma}|n\rangle = -ec\sqrt{\frac{\hbar}{2\epsilon_0 V\omega}}\boldsymbol{\epsilon}\cdot\langle\Psi_m|\int\psi^\dagger\boldsymbol{\alpha}\psi\mathrm{e}^{i\mathbf{k}\cdot\mathbf{r}}|\Psi_n\rangle,$$

and

$$\langle e|H_{e\gamma}|n\rangle = -ec\sqrt{\frac{\hbar}{2\epsilon_0 V\omega'}}\boldsymbol{\epsilon}'^\star\cdot\langle\Psi_m|\int\psi^\dagger\boldsymbol{\alpha}\psi\mathrm{e}^{-i\mathbf{k}'\cdot\mathbf{r}}|\Psi_n\rangle,$$

where

$$\int\psi^\dagger\boldsymbol{\alpha}\psi\mathrm{e}^{\pm i\mathbf{k}\cdot\mathbf{r}} \;=\; \int\psi^\dagger(\mathbf{r})\boldsymbol{\alpha}\psi(\mathbf{r})\mathrm{e}^{\pm i\mathbf{k}\cdot\mathbf{r}}d\mathbf{r}.$$

C. Electric dipole and multipole transitions

To carry out a multipole expansion of the previous matrix elements, we shall continue working with quantum fields instead of the usual many-body expressions. In that framework, the expressions are simpler because there is no electron index and we can use the following well-known trick.[78,113]

Let $F = \int\psi^\dagger(\mathbf{r})f(\mathbf{r})\psi(\mathbf{r})d\mathbf{r}$, where f is some function of $\mathbf{r}$. To calculate the commutator of F with some Hamiltonian H_0, we go to the interaction picture and define $F_I(t) = e^{iH_0t/\hbar}Fe^{-iH_0t/\hbar}$. Then, the time-derivative $\dot{F}_I$ of F_I is given by $-i\hbar\dot{F}_I(t) = [H_0, F_I(t)]$. Now, we notice that F is related to the density operator $\rho(\mathbf{r}) = \psi^\dagger(\mathbf{r})\psi(\mathbf{r})$ by $F = \int\rho(\mathbf{r})f(\mathbf{r})d\mathbf{r}$. Thus, $-i\hbar\dot{F}_I(t) = -i\hbar\int\dot{\rho}(\mathbf{r},t)f(\mathbf{r})d\mathbf{r} = [H_0, F_I(t)]$. If H_0 conserves the electric charge, the continuity equation $e\dot{\rho}(\mathbf{r}) = -\nabla\cdot\mathbf{j}$ holds, where $\mathbf{j}$ is the electric current operator. By taking $t = 0$ to recover the operators in the Schrödinger picture, we obtain

$$\begin{aligned}[H_0, F] \;&=\; \frac{i\hbar}{e}\int\nabla\cdot\mathbf{j}(\mathbf{r})f(\mathbf{r})d\mathbf{r} = -\frac{i\hbar}{e}\int\mathbf{j}(\mathbf{r})\cdot\nabla f(\mathbf{r})d\mathbf{r}\\ &=\; -i\hbar c\int\psi^\dagger(\mathbf{r})\boldsymbol{\alpha}\psi(\mathbf{r})\cdot\nabla f(\mathbf{r})d\mathbf{r}.\end{aligned} \tag{9}$$

To find the electric dipole transition term we apply Eq. (9) with $f(\mathbf{r}) = \boldsymbol{\epsilon}\cdot\mathbf{r}$ and $H_0 = H_e$:

$$[H_e, \int\psi^\dagger(\mathbf{r})\boldsymbol{\epsilon}\cdot\mathbf{r}\psi(\mathbf{r})d\mathbf{r}] = -i\hbar c\int\psi^\dagger(\mathbf{r})\boldsymbol{\alpha}\psi(\mathbf{r})\cdot\boldsymbol{\epsilon}d\mathbf{r},$$

and we obtain in the dipole approximation $\mathrm{e}^{i\mathbf{k}\cdot\mathbf{r}} \simeq 1$

$$\langle a|H_{e\gamma}|n\rangle = \frac{e(E_m - E_n)}{i\hbar}\sqrt{\frac{\hbar}{2\epsilon_0 V\omega}}\langle\Psi_m|\int\psi^\dagger\boldsymbol{\epsilon}\cdot\mathbf{r}\psi|\Psi_n\rangle.$$

To deal with electric quadrupole and magnetic dipole transitions, we expand to the first order: $e^{i\mathbf{k}\cdot\mathbf{r}} \simeq 1 + i\mathbf{k}\cdot\mathbf{r}$. We apply Eq. (9) with $f(\mathbf{r}) = \boldsymbol{\epsilon}\cdot\mathbf{r}\mathbf{k}\cdot\mathbf{r}$ and $H_0 = H_e$:

$$[H_e, \psi^\dagger \boldsymbol{\epsilon}\cdot\mathbf{r}\mathbf{k}\cdot\mathbf{r}\psi] = -i\hbar c\psi^\dagger\boldsymbol{\alpha}\psi\cdot(\boldsymbol{\epsilon}\mathbf{k}\cdot\mathbf{r} + \mathbf{k}\boldsymbol{\epsilon}\cdot\mathbf{r}),$$

where we removed the integral sign for notational convenience. Thus,

$$\psi^\dagger\boldsymbol{\epsilon}\cdot\boldsymbol{\alpha}\mathbf{k}\cdot\mathbf{r}\psi \;=\; \frac{i}{\hbar c}[H_e, \psi^\dagger\boldsymbol{\epsilon}\cdot\mathbf{r}\mathbf{k}\cdot\mathbf{r}\psi] - \psi^\dagger\boldsymbol{\epsilon}\cdot\mathbf{r}\mathbf{k}\cdot\boldsymbol{\alpha}\psi.$$

If we add $\psi^\dagger\boldsymbol{\epsilon}\cdot\boldsymbol{\alpha}\mathbf{k}\cdot\mathbf{r}\psi$ to both terms we obtain

$$\begin{aligned} 2\psi^\dagger\boldsymbol{\epsilon}\cdot\boldsymbol{\alpha}\mathbf{k}\cdot\mathbf{r}\psi \;&=\; \frac{i}{\hbar c}[H_e, \psi^\dagger\boldsymbol{\epsilon}\cdot\mathbf{r}\mathbf{k}\cdot\mathbf{r}\psi] \\ &\quad -\psi^\dagger(\boldsymbol{\epsilon}\times\mathbf{k})\cdot(\mathbf{r}\times\boldsymbol{\alpha})\psi. \end{aligned}$$

Finally, up to electric quadrupole transitions

$$\langle a|H_{e\gamma}|n\rangle = \frac{e\Delta E}{i\hbar}\sqrt{\frac{\hbar}{2\epsilon_0 V\omega}}\langle\Psi_m|\int\psi^\dagger T\psi|\Psi_n\rangle, \quad (10)$$

where $\Delta E = E_m - E_n$ and

$$T = \boldsymbol{\epsilon}\cdot\mathbf{r} + \frac{i}{2}\boldsymbol{\epsilon}\cdot\mathbf{r}\mathbf{k}\cdot\mathbf{r} - \frac{\hbar c}{2\Delta E}(\boldsymbol{\epsilon}\times\mathbf{k})\cdot(\mathbf{r}\times\boldsymbol{\alpha}). \quad (11)$$

The first term of T is the usual electric-dipole operator, the second one is the electric-quadrupole operator and the third one will turn to be the magnetic-dipole operator (see section V C). Similarly,

$$\langle e|H_{e\gamma}|n\rangle = \frac{e\Delta E}{i\hbar}\sqrt{\frac{\hbar}{2\epsilon_0 V\omega'}}\langle\Psi_m|\int\psi^\dagger T'\psi|\Psi_n\rangle, \quad (12)$$

where

$$T' = \boldsymbol{\epsilon}'^\star\cdot\mathbf{r} - \frac{i}{2}\boldsymbol{\epsilon}'^\star\cdot\mathbf{r}\mathbf{k}'\cdot\mathbf{r} + \frac{\hbar c}{2\Delta E}(\boldsymbol{\epsilon}'^\star\times\mathbf{k}')\cdot(\mathbf{r}\times\boldsymbol{\alpha}).$$

V. SEMI-RELATIVISTIC REPRESENTATION

In the previous sections, we have shown that, to ensure gauge invariance, it was safe to describe the interaction of light and matter with quantum electrodynamics, where electrons are described by four-component Dirac spinors. However, in most solid-state calculations, we do not use Dirac spinors but two-component (Pauli) wavefunctions. Therefore, we need to link the two representations by using a generalization of Foldy-Wouthuysen transformation.

The idea of the Foldy-Wouthuysen transformation is the following. If H_D is a time-dependent relativistic Hamiltonian, it has the form

$$H_D \;=\; H^0 + \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix},$$

where $H^0 = mc^2\beta$ and each H_{ij} is a 2x2 matrix. We write H_D as the sum of even and odd parts $H_D = H^0 + \mathcal{E} + \mathcal{O}$, where

$$\mathcal{E} = \begin{pmatrix} H_{11} & 0 \\ 0 & H_{22} \end{pmatrix}, \quad \mathcal{O} = \begin{pmatrix} 0 & H_{12} \\ H_{21} & 0 \end{pmatrix},$$

satisfy $\beta\mathcal{E}\beta = \mathcal{E}$ and $\beta\mathcal{O}\beta = -\mathcal{O}$. Note that H^0 is also even. If $|\psi_D\rangle$ is a solution of the Dirac equation $i\hbar\partial_t|\psi_D\rangle = H_D|\psi_D\rangle$, where H_D is the Dirac Hamiltonian, then the upper two components of $|\psi_D\rangle$ are called the large components and the lower two the small components. The Dirac equation couples the large and small components of $|\psi_D\rangle$ through the odd terms of H_D. Foldy and Wouthuysen[114] looked for a unitary operator U that decouples the large and small components of $|\psi\rangle = U|\psi_D\rangle$. In other words, $H = UH_DU^\dagger$ has only even components: $H = \beta H\beta$. The method proposed by Foldy and Wouthuysen consist in successive transformations of the form $U = e^{iS}$. [114,115].

This transformation does not satisfy Eriksen's condition $U = \beta U^\dagger\beta$ discussed in the Appendix. This is because the product $U = e^{iS^{(2)}}e^{iS^{(1)}}$ does not satisfy this equation even if $e^{iS^{(1)}}$ and $e^{iS^{(2)}}$ do. Silenko recently derived the correction that must be applied to go from Foldy-Wouthuysen to Eriksen transformations,[116] and he showed that the correction is at an order beyond the one considered.

A. Many-body Foldy-Wouthuysen transformation

To generalize the Foldy-Wouthuysen approach to the many-body Dirac Hamiltonian we face the following problem. The generalization of H^0 is imposed by the many-body Dirac Hamiltonian:

$$H^0_N \;=\; \sum_{n=1}^{N}\beta_n mc^2,$$

where β_n is the matrix β acting on the nth electron (i.e. $\beta_n = 1^{\otimes(n-1)}\otimes\beta\otimes 1^{\otimes(N-n)}$). This definition is valid because H^0_N commutes with the projector P_N onto the space of antisymmetric N-body states.

We show in the Appendix that a Foldy-Wouthuysen transformation can be defined whenever we have a self-adjoint operator η (with $\eta^2 = 1$) to define parity. In the one-body case, $\beta^2 = 1$ and $\eta = \beta$ defines parity. But in the many-body case the operator $\sum_{n=1}^{N}\beta_n$ suggested by H^0_N cannot be used for that purpose because its square is not proportional to the identity (it contains products $\beta_n\beta_m$). It turns out that $\eta = \beta_1\otimes\cdots\otimes\beta_n$ is the natural many-body generalization of β. Indeed, $\eta^\dagger = \eta$, $\eta^2 = 1$. Moreover, η commutes with P_N, which allows us to work with tensor products instead of antisymmetric tensor products.

In the literature, the Foldy-Wouthuysen transformation was studied for two-body Hamiltonians,[117–119], but

the results were rather complicated and not easy to extend to the many-body case.

The even and odd parts of H_N are then $H_N^0 + \mathcal{E}$ and $\mathcal{O}$, respectively:

$$\begin{aligned}\mathcal{E} &= e\sum_{n=1}^{N}\phi_0(\mathbf{r}_n) + e\sum_{m\neq n} V(\mathbf{r}_m - \mathbf{r}_n),\\ \mathcal{O} &= \sum_{n=1}^{N} c\boldsymbol{\alpha}_n\cdot\boldsymbol{\pi}_n = \sum_{n=1}^{N}\mathcal{O}_n,\end{aligned}$$

where $V(\mathbf{r}) = \frac{e}{8\pi\epsilon_0|\mathbf{r}|}$ is the Coulomb potential and $\boldsymbol{\pi}_n = -i\hbar\nabla_n - e\mathbf{a}_0(\mathbf{r}_n)$.

At first order in c^{-1}, the Foldy-Wouthuysen operator is $U = e^{iS^{(1)}}$ where

$$S^{(1)} = -\frac{i}{2mc^2}\sum_n \beta_n\mathcal{O}_n.$$

Indeed, it can be checked that $i[S^{(1)}, H_N^0] = -\mathcal{O}$ removes the odd term of H_D. At this order $U = U_1\otimes\cdots\otimes U_N$ is a tensor power of one-body Foldy-Wouthuysen operators, as proposed by Moshinksy and Nikitin.[120]

From that point, the usual formal Foldy-Wouthuysen transformation $U = e^{iS^{(1)}}e^{iS^{(2)}}$ can be carried out almost unchanged and we find, with m as expansion parameter, at order m^{-2}:

$$\begin{aligned}H_{\mathrm{FW}} &= H_N^0 + \mathcal{E} + \frac{1}{2mc^2}\sum_{n=1}^{N}\beta_n\mathcal{O}_n^2\\ &\quad -\frac{1}{8m^2c^4}\sum_{n=1}^{N}[\mathcal{O}_n,[\mathcal{O}_n, e\varphi_n + eV]]\\ &\quad +\frac{1}{8m^2c^4}\sum_{p\neq n}\beta_p\beta_n[\mathcal{O}_p[\mathcal{O}_n, V]].\end{aligned}$$

This Hamiltonian obeys $\eta H_{\mathrm{FW}}\eta = H_{\mathrm{FW}}$ which makes it a Foldy-Wouthuysen Hamiltonian.

It rewrites

$$H_{\mathrm{FW}} = \sum_{n=1}^{N} H_{\mathrm{FW}}^n + H_{\mathrm{FW}}^{MB}. \tag{13}$$

where H_{FW}^n are the usual one body Foldy-Wouthuysen Hamiltonians:

$$\begin{aligned}H_{\mathrm{FW}}^n &= \beta_n mc^2 + e\phi_0(\mathbf{r}_n) + \sum_{p\neq n} eV(\mathbf{r}_n - \mathbf{r}_p)\\ &\quad + \frac{1}{2m}\beta_n\boldsymbol{\pi}_n^2 - e\hbar\boldsymbol{\Sigma}_n\cdot\mathbf{b}_0(\mathbf{r}_n) - \frac{\hbar^2 e}{8m^2c^2}\nabla\cdot\mathbf{E}_n\\ &\quad + \frac{\hbar e}{8m^2c^2}\boldsymbol{\Sigma}\cdot(\boldsymbol{\pi}_n\times\mathbf{E}_n - \mathbf{E}_n\times\boldsymbol{\pi}_n)\end{aligned}$$

where

$$\mathbf{b}_0(\mathbf{r}_n) = \nabla\times\mathbf{a}_0(\mathbf{r}_n)$$

and

$$\mathbf{E}_n = -\nabla\phi_n(\mathbf{r}_p) - \sum_{p\neq n}\nabla V(\mathbf{r}_n - \mathbf{r}_p).$$

The mass velocity term $\frac{\beta_n}{8m^3c^2}(\mathbf{p}_n\cdot\mathbf{p}_n)^2$ would be obtained by pushing the development one order beyond. The other term H_{FW}^{MB} arises because $V(\mathbf{r}_m - \mathbf{r}_n) = V(\mathbf{r}_{mn})$ is a two body operator:

$$\begin{aligned}H_{\mathrm{FW}}^{n,p} &= \frac{\hbar e}{8m^2c^2}\sum_{p\neq n}^{N}\Big(\hbar\Delta V(\mathbf{r}_{np})\\ &\quad - \boldsymbol{\Sigma}\cdot(\boldsymbol{\pi}_n\times\nabla V(\mathbf{r}_{np}) - \nabla V(\mathbf{r}_{np})\times\boldsymbol{\pi}_n)\\ &\quad - \hbar\beta_n\beta_p(\boldsymbol{\alpha}_n\cdot\nabla_n)(\boldsymbol{\alpha}_p\cdot\nabla_p)V(\mathbf{r}_{np})\Big).\end{aligned}$$

By using:[121]

$$\partial_j\partial_k V(\mathbf{r}) = \frac{e^2}{8\pi\epsilon_0}\Big(-\delta_{jk}\frac{4\pi}{3}\delta(\mathbf{r}) - \delta_{jk}\frac{1}{r^3} + \frac{3r^jr^k}{r^5}\Big),$$

the derivatives in the last term can be rewritten

$$\begin{aligned}&(\boldsymbol{\alpha}_p\cdot\nabla_p)(\boldsymbol{\alpha}_q\cdot\nabla_q)V(\mathbf{r}_p - \mathbf{r}_q) = \sum_{jk}\boldsymbol{\alpha}_p^j\boldsymbol{\alpha}_q^k\partial_j\partial_k V(\mathbf{r}_{pq})\\ &= \frac{e^2}{8\pi\epsilon_0}\Big(-\frac{4\pi}{3}\boldsymbol{\alpha}_p\cdot\boldsymbol{\alpha}_q\delta(\mathbf{r}_{pq}) - \frac{\boldsymbol{\alpha}_p\cdot\boldsymbol{\alpha}_q}{|\mathbf{r}_{pq}|^3}\\ &\qquad + 3\frac{\boldsymbol{\alpha}_p\cdot\mathbf{r}_{pq}\boldsymbol{\alpha}_q\cdot\mathbf{r}_{pq}}{|\mathbf{r}_{pq}|^5}\Big).\end{aligned}$$

This expression looks superficially like some contributions to the Breit interaction as presented by Bethe and Salpeter.[122] However, they are different since the Breit interaction is due to the exchange of a photon and not to a semi-relativistic effect. Note that the last two terms are singular. It is known that expansion of the Foldy-Wouthuysen transformation as a power serie in $1/c^2$ becomes more and more singular because of the presence of the Coulomb potential.[123] At order m^{-2}, the transformation writes

$$\begin{aligned}U = 1 &+ \frac{1}{2mc^2}\sum_n\beta_n\mathcal{O}_n - \frac{1}{8m^2c^4}\Big(\sum_n\beta_n\mathcal{O}_n\Big)^2\\ &+ \frac{1}{4m^2c^4}\sum_n\beta_n\Big[\sum_m\beta_m\mathcal{O}_m, \mathcal{E}\Big]\end{aligned}$$

and it obeys $U = \eta U^\dagger\eta$. We also checked that U^2 is odd in H_D after paying attention to the discontinuity at zero discussed in the Appendix. Thus, the positive (negative) energy eigenstate of H_D are transformed into even (odd) states by the action of U.

B. Semi-relativistic dipole transitions

Matrix elements such as $D = \langle\Phi|\int\psi^\dagger\boldsymbol{\epsilon}\cdot\mathbf{r}\psi|\Psi\rangle$ are now evaluated by expressing the positive energy Dirac wavefunctions $|\Phi\rangle$ and $|\Psi\rangle$ in terms of the Foldy-Wouthuysen

ones $|\phi\rangle$ and $|\psi\rangle$: $|\Phi\rangle = U^\dagger|\phi\rangle$ and $|\Psi\rangle = U^\dagger|\psi\rangle$. We use the many-body expression for $D = \langle\Phi|\boldsymbol{\epsilon}\cdot\mathbf{R}|\Psi\rangle$, where $\mathbf{R} = \sum_{n=1}^N \mathbf{r}_n$, since U is written as a many-body operator. We calculate $D = \langle\phi|U\epsilon\cdot\mathbf{R}U^\dagger|\psi\rangle$, where $U = e^{iS}$ by using the Baker-Campbell-Hausdorff formula

$$e^{iS}Te^{-iS} = T + i[S,T] + \sum_{n=2}^{\infty} i^n \frac{L^n(T)}{n!},$$

where $L(T) = [S,T]$ and $L^n(T) = L(L^{n-1}(T))$. If $U = U_1 \otimes \cdots \otimes U_N$, where $U_i = e^{iS_i}$, we can calculate the action of U on each variable independently. Removing temporarily the constant $-i/2mc^2$, we take the one-body operator $S = \beta\mathcal{O}$ and compute

$$\begin{aligned} L(\hat{\epsilon}\cdot\mathbf{r}) &= c[\beta\boldsymbol{\alpha}\cdot(\mathbf{p}-e\mathbf{a}_0), \hat{\epsilon}\cdot\mathbf{r}] = c\sum_{ij}\beta\alpha^i\epsilon^j[p_i, r_j] \\ &= -i\hbar c\sum_{ij}\beta\alpha^i\epsilon^j\delta_{ij} = -i\hbar c\beta\boldsymbol{\alpha}\cdot\hat{\epsilon}, \end{aligned}$$

and

$$\begin{aligned} L^2(\hat{\epsilon}\cdot\mathbf{r}) &= -i\hbar c^2[\beta\boldsymbol{\alpha}\cdot(\mathbf{p}-e\mathbf{a}_0), \beta\boldsymbol{\alpha}\cdot\hat{\epsilon}] \\ &= -i\hbar c^2\sum_{ij}(p_i - ea_{0i})\epsilon_j[\beta\alpha^i, \beta\alpha^j] \\ &= i\hbar c^2\sum_{ij}(p_i - ea_{0i})\epsilon_j[\alpha^i, \alpha^j], \end{aligned}$$

where we used $\beta\alpha_i = -\alpha_i\beta$ and $\beta^2 = 1$. We compute

$$[\alpha^i, \alpha^j] = 2i\sum_k \epsilon_{ijk}\begin{pmatrix}\sigma^k & 0\\ 0 & \sigma^k\end{pmatrix} = 2i\sum_k\epsilon_{ijk}\Sigma^k,$$

which defines Σ^k the components of $\boldsymbol{\Sigma}$. Therefore,

$$L^2(\hat{\epsilon}\cdot\mathbf{r}) = -2\hbar c^2(\mathbf{p}-e\mathbf{a}_0)\cdot(\hat{\epsilon}\times\boldsymbol{\Sigma}).$$

So that, for each particle, and up to $O(c^{-3})$,

$$\begin{aligned} U_n\boldsymbol{\epsilon}\cdot\mathbf{r}_nU_n^\dagger &= \boldsymbol{\epsilon}\cdot\mathbf{r}_n - i\frac{\hbar}{2mc}\beta_n\boldsymbol{\alpha}_n\cdot\boldsymbol{\epsilon} \\ &\quad -\frac{\hbar}{4m^2c^2}\boldsymbol{\pi}_n\cdot(\boldsymbol{\epsilon}\times\boldsymbol{\Sigma}_n). \end{aligned}$$

The many-body version is obtained by summing the right-hand side over n.

In the matrix elements $D = \langle\phi|U\boldsymbol{\epsilon}\cdot\mathbf{R}U^\dagger|\psi\rangle$, recall that $|\psi\rangle = \eta|\psi\rangle$ and $|\phi\rangle = \eta|\phi\rangle$ because $|\Psi\rangle$ and $|\Phi\rangle$ are positive energy states, as shown in the Appendix. Therefore, $\langle\phi|U\boldsymbol{\epsilon}\cdot\mathbf{R}U^\dagger|\psi\rangle = \langle\phi|\eta U\boldsymbol{\epsilon}\cdot\mathbf{R}U^\dagger\eta|\psi\rangle$ and all the terms that are odd in $U\boldsymbol{\epsilon}\cdot\mathbf{R}U^\dagger$ are eliminated by the matrix elements. This eliminates the term proportional to $\beta_n\boldsymbol{\alpha}_n$ and we are left with

$$D = \sum_{n=1}^N\langle\phi|\boldsymbol{\epsilon}\cdot\mathbf{r}_n - \frac{\hbar}{4m^2c^2}\boldsymbol{\pi}_n\cdot(\boldsymbol{\epsilon}\times\boldsymbol{\Sigma}_n)|\psi\rangle.$$

C. Semi-relativistic multipole transitions

From Eq. (11), we write the multipole transitions

$$M = \frac{i}{2}M_1 - \frac{\hbar c}{2\Delta E}M_2,$$

where

$$\begin{aligned} M_1 &= \sum_n\langle\phi|U\boldsymbol{\epsilon}\cdot\mathbf{r}_n\mathbf{k}\cdot\mathbf{r}_nU^\dagger|\psi\rangle, \\ M_2 &= \sum_n\langle\phi|U(\boldsymbol{\epsilon}\times\mathbf{k})\cdot(\mathbf{r}_n\times\boldsymbol{\alpha}_n)U^\dagger|\psi\rangle, \end{aligned}$$

correspond to the electric quadrupole and magnetic dipole transitions, respectively. Since multipole transitions are smaller than dipole ones, it is enough to use the first two terms of the Baker-Campbell-Hausdorf formula.

The term $[S_n, \boldsymbol{\epsilon}\cdot\mathbf{r}_n\mathbf{k}_n\cdot\mathbf{r}_n]$ is odd and disappears in the matrix element. Thus, at the order we consider,

$$M_1 = \sum_n\langle\phi|\boldsymbol{\epsilon}\cdot\mathbf{r}_n\mathbf{k}\cdot\mathbf{r}_n|\psi\rangle.$$

Let $T_2 = (\boldsymbol{\epsilon}\times\mathbf{k})\cdot(\mathbf{r}\times\boldsymbol{\alpha})$. We write

$$\begin{aligned} [\beta\mathcal{O}, T_2] &= c[\beta\boldsymbol{\alpha}\cdot\mathbf{p}, T_2] - ec[\beta\boldsymbol{\alpha}\cdot\mathbf{a}_0, T_2] \\ &= c\beta(\{\boldsymbol{\alpha}\cdot\mathbf{p}, T_2\} - e\{\boldsymbol{\alpha}\cdot\mathbf{a}_0, T_2\}). \end{aligned}$$

The anticommutators are

$$\begin{aligned} \{\boldsymbol{\alpha}\cdot\mathbf{p}, T_2\} &= \sum_{ijkl}\epsilon_{jkl}(\boldsymbol{\epsilon}\times\mathbf{k})_j(\alpha_i\alpha_lp_ir_k + r_kp_i\alpha_l\alpha_i) \\ &= 2(\boldsymbol{\epsilon}\times\mathbf{k})\cdot(\hbar\boldsymbol{\Sigma}+\mathbf{L}), \end{aligned}$$

and

$$\begin{aligned} \{\boldsymbol{\alpha}\cdot\mathbf{a}_0, T_2\} &= \sum_{ijkl}\epsilon_{ikl}(\boldsymbol{\epsilon}\times\mathbf{k})_ia_jr_k\{\alpha_j, \alpha_l\} \\ &= 2(\boldsymbol{\epsilon}\times\mathbf{k})\cdot(\mathbf{r}\times\mathbf{a}_0). \end{aligned}$$

Note that $\hbar\boldsymbol{\Sigma} = g\mathbf{S}$ with $g = 2$ (because $\mathbf{S} = \hbar\boldsymbol{\Sigma}/2$). Thus, we recover the fact that the Dirac equation gives a gyromagnetic factor $g = 2$ to the electron. Moreover, $\mathbf{L} + \hbar\boldsymbol{\Sigma} = \mathbf{L} + 2\mathbf{S}$, the total magnetic moment of the electron.

Finally, since $\mathbf{r}_n\times\boldsymbol{\alpha}_n$ is odd,

$$M_2 = \sum_n\frac{\beta_n}{mc}\langle\phi|(\boldsymbol{\epsilon}\times\mathbf{k})\cdot(\hbar\boldsymbol{\Sigma}_n + \boldsymbol{\Lambda}_n)|\psi\rangle,$$

where $\boldsymbol{\Lambda}_n = \mathbf{L}_n - e\mathbf{r}_n\times\mathbf{a}_0(\mathbf{r}_n)$ is the moment of the mechanical momentum as defined in Ref. 124. The term M_2 describes magnetic-dipole transitions. The multipole transitions are

$$\begin{aligned} M &= \sum_n\langle\phi|\frac{i}{2}\boldsymbol{\epsilon}\cdot\mathbf{r}_n\mathbf{k}\cdot\mathbf{r}_n \\ &\quad -\frac{\hbar\beta_n}{2m\Delta E}(\boldsymbol{\epsilon}\times\mathbf{k})\cdot(\hbar\boldsymbol{\Sigma}_n + \boldsymbol{\Lambda}_n)|\psi\rangle. \end{aligned}$$

VI. ABSORPTION CROSS-SECTION

The absorption cross section is calculated by assuming that initially the system of electrons is in state $|I\rangle$ that can be transformed into a Foldy-Wouthuysen eigenstate $|i\rangle$, with energy E_i, and that a photon $\mathbf{k}, \boldsymbol{\epsilon}$ is present. In the final state there is no photon and the system is in state $|F\rangle$ ($|f\rangle$ after transformation).

The transition probability per unit time from state m to state n is related to the T-matrix elements by:[125]

$$w = \frac{2}{\hbar}\delta_{mn}\,\mathrm{Im}\langle m|T|m\rangle + \frac{2\pi}{\hbar}\delta(e_n - e_m)|\langle n|T|m\rangle|^2. \quad (14)$$

and must be divided by c/v (rate at which the photon crosses a unit of surface) to obtain the cross section. Since we consider real transitions (i.e. $m \neq n$), only the second term is present.

From (10) and using the result of transformation derived in the previous section:

$$\sigma \;=\; 4\pi^2\alpha_0\hbar\omega\sum_f |\langle f|T_{\mathrm{FW}}|i\rangle|^2\delta(E_f - E_i - \hbar\omega),$$

where T_{FW} is:

$$\begin{aligned} T_{\mathrm{FW}} \;=\; & \sum_n \boldsymbol{\epsilon}\cdot\mathbf{r}_n + \frac{i}{2}\boldsymbol{\epsilon}\cdot\mathbf{r}_n\mathbf{k}\cdot\mathbf{r}_n - \frac{\hbar}{4m^2c^2}\boldsymbol{\pi}_n\cdot(\boldsymbol{\epsilon}\times\boldsymbol{\Sigma}_n) \\ & -\frac{\beta_n}{2m\omega}(\boldsymbol{\epsilon}\times\mathbf{k})\cdot(\hbar\boldsymbol{\Sigma}_n + \boldsymbol{\Lambda}_n), \end{aligned}$$

with α_0 the fine structure constant and $\Delta E = E_f - E_i = \hbar\omega$.

It corresponds to the usual formula for the cross section[126] with two more terms: the third one and the last one.

The third term was already present in the PhD thesis of Christos Gougoussis.[127] We rewrite it using $\mathbf{p} - e\mathbf{a}_0 = (m/i\hbar)[\mathbf{r}, H_0^{\mathrm{FW}}] + O(c^{-2})$, where H_0^{FW} is the Foldy-Wouthuysen Hamiltonian, to get:

$$\begin{aligned} -\frac{\hbar}{4m^2c^2}\langle f|(\mathbf{p} - e\mathbf{a})\cdot(\boldsymbol{\epsilon}\times\boldsymbol{\Sigma})|i\rangle & \\ = \frac{i}{4mc^2}(E_i - E_f)\langle f|\mathbf{r}\cdot(\boldsymbol{\epsilon}\times\boldsymbol{\Sigma})|i\rangle & \\ = \frac{i\hbar\omega}{4mc^2}\langle f|(\boldsymbol{\epsilon}\times\mathbf{r})\cdot\boldsymbol{\Sigma}|i\rangle. & \end{aligned}$$

We call *spin-position operator* the operator $(\boldsymbol{\epsilon}\times\mathbf{r})\cdot\boldsymbol{\Sigma}$. Its evaluation at the K-edge of materials will be presented in our companion paper.[10]

The amplitude of the last term depends on the choice of the space origin in the Coulomb gauge for $\mathbf{a}_0$. It does not make the cross section gauge dependent because the states are changed accordingly when choosing the origin of the gauge. If the origin of the gauge is chosen at the atom position, fields larger than 10^6 T are required for this term to be significant. Such fields are way beyond laboratory accessible values.

VII. SCATTERING CROSS-SECTION

The scattering cross section is calculated by assuming that initially the system of electrons is in state $|I\rangle$ with a photon $\mathbf{k}_i, \boldsymbol{\epsilon}_i$ and that in the final state the system is in state $|F\rangle$ with a scattered photon $\mathbf{k}_f, \boldsymbol{\epsilon}_f$. We do not consider the special case when $\mathbf{k}_i, \boldsymbol{\epsilon}_i = \mathbf{k}_f, \boldsymbol{\epsilon}_f$.

Eqs. (10), (12) and (14) yield:

$$\begin{aligned} w = \frac{2\pi}{\hbar}\sum_F \delta(E_f + \hbar\omega_f - E_i - \hbar\omega_i)\Big|\sum_L \frac{e^2c^2\hbar}{2\epsilon_0 V}\frac{1}{\sqrt{\omega_i\omega_f}} & \\ \frac{\langle F|e_{-\mathbf{k}_f}\psi^\dagger\boldsymbol{\alpha}\cdot\boldsymbol{\epsilon}_f^\star\psi|L\rangle\langle L|e_{\mathbf{k}_i}\psi^\dagger\boldsymbol{\alpha}\cdot\boldsymbol{\epsilon}_i\psi|I\rangle}{E_i - E_l + \hbar\omega_i + i\gamma} & \\ + \frac{\langle F|e_{\mathbf{k}_i}\psi^\dagger\boldsymbol{\alpha}\cdot\boldsymbol{\epsilon}_i\psi|L\rangle\langle L|e_{-\mathbf{k}_f}\psi^\dagger\boldsymbol{\alpha}\cdot\boldsymbol{\epsilon}_f^\star\psi|I\rangle}{E_i - E_l - \hbar\omega_f}\Big|^2, & \end{aligned}$$

where $\gamma > 0$ and

$$e_{\mathbf{k}}\psi^\dagger\boldsymbol{\alpha}\cdot\boldsymbol{\epsilon}\psi \;=\; \sum_{j=1}^{3}\int e^{i\mathbf{k}\cdot\mathbf{r}}\psi^\dagger(\mathbf{r})\alpha^j\psi(\mathbf{r})\epsilon^j d\mathbf{r}.$$

The scattering cross-section is related to w by:[128]

$$\frac{d^2\sigma}{d\Omega d\omega_f} \;=\; \frac{V^2}{(2\pi)^3}\omega_f^2\frac{1}{\hbar c^4}w.$$

Since the electric charge is related to the classical electron radius r_e by $e^2 = 4\pi\epsilon_0 r_e mc^2$, we obtain the relativistic Kramers-Heisenberg scattering cross-section:

$$\begin{aligned} \frac{d^2\sigma}{d\Omega d\omega_f} \;=\; & (r_e mc^2)^2\frac{\omega_f}{\omega_i}\sum_F \delta(E_f + \hbar\omega_f - E_i - \hbar\omega_i) \\ & \Big|\sum_L \frac{\langle F|e_{-\mathbf{k}_f}\psi^\dagger\boldsymbol{\alpha}\cdot\boldsymbol{\epsilon}_f^\star\psi|L\rangle\langle L|e_{\mathbf{k}_i}\psi^\dagger\boldsymbol{\alpha}\cdot\boldsymbol{\epsilon}_i\psi|I\rangle}{E_i - E_l + \hbar\omega_i + i\gamma} \\ & + \frac{\langle F|e_{\mathbf{k}_i}\psi^\dagger\boldsymbol{\alpha}\cdot\boldsymbol{\epsilon}_i\psi|L\rangle\langle L|e_{-\mathbf{k}_f}\psi^\dagger\boldsymbol{\alpha}\cdot\boldsymbol{\epsilon}_f^\star\psi|I\rangle}{E_i - E_l - \hbar\omega_f}\Big|^2. \end{aligned}$$

In this expression, the sum over $|L\rangle$ involves a complete set of states, with positive and negative energies. Since E_i is usually the positive energy of the ground state including the electron rest energy, we have $E_i = mc^2 + E_i' > 0$, where E_i' is the usual (negative) ground state energy. If $|L\rangle$ is a positive energy state, we have $E_l = mc^2 + E_l'$ with $E_l' > E_i'$ and the first term is resonant at $\hbar\omega_i = E_l' - E_i'$. If $|L\rangle$ is a negative energy state, then $E_l = -mc^2 - E_l'$ and $E_i - E_l - \hbar\omega_f = 2mc^2 + E_i' - E_l' - \hbar\omega_f$ cannot be resonant in standard experimental conditions.

We show that the resonant scattering term has a semi-relativistic expansion close to but different from the one given by Blume.

If we are interested in the resonant part of the scatter-

ing cross section, then $E_l > 0$ and

$$\frac{d^2\sigma}{d\Omega d\omega_f} = (\frac{r_e m}{\hbar^2})^2 \frac{\omega_f}{\omega_i} \sum_f \delta(E_f + \hbar\omega_f - E_i - \hbar\omega_i) \Big| \sum_{L^>} (E_l - E_i)(E_f - E_l) \frac{\langle f|T'^{fl}_{\rm FW}(\boldsymbol{\epsilon}_f)|l\rangle\langle l|T^{li}_{\rm FW}(\boldsymbol{\epsilon}_i)|i\rangle}{E_i - E_l + \hbar\omega_i + i\gamma}\Big|.$$

with

$$T^{ij}_{\rm FW}(\boldsymbol{\epsilon}_i) = \sum_n \boldsymbol{\epsilon}_i \cdot \mathbf{r}_n + \frac{i}{2}\boldsymbol{\epsilon}_i \cdot \mathbf{r}_n \mathbf{k}_i \cdot \mathbf{r}_n - \frac{\hbar}{4m^2c^2}\boldsymbol{\pi}_n \cdot (\boldsymbol{\epsilon}_i \times \boldsymbol{\Sigma}_n) - \frac{1}{2m\Delta E^{ij}}(\boldsymbol{\epsilon}_i \times \mathbf{k}_i) \cdot (\hbar\boldsymbol{\Sigma}_n + \boldsymbol{\Lambda}_n),$$

and

$$T'^{ij}_{\rm FW}(\boldsymbol{\epsilon}_f) = \sum_n \boldsymbol{\epsilon}^\star_f \cdot \mathbf{r}_n - \frac{i}{2}\boldsymbol{\epsilon}^\star_f \cdot \mathbf{r}_n \mathbf{k}_f \cdot \mathbf{r}_n - \frac{\hbar}{4m^2c^2}\boldsymbol{\pi}_n \cdot (\boldsymbol{\epsilon}^\star_f \times \boldsymbol{\Sigma}_n) + \frac{1}{2m\Delta E^{ij}}(\boldsymbol{\epsilon}^\star_f \times \mathbf{k}_f) \cdot (\hbar\boldsymbol{\Sigma}_n + \boldsymbol{\Lambda}_n),$$

where $\Delta E^{ij} = E_i - E_j$.

As in the absorption case, the spin-position term in the transition operator is not present in the usual formula.[4]

VIII. CONCLUSION

We saw that four different semi-relativistic Hamiltonians can be used to describe the interaction of light with matter. We concluded that gauge invariance cannot be convincingly enforced in the semi-classical approach and we had recourse to quantum electrodynamics to obtain *relativistic gauge-invariant* absorption and scattering cross-sections. We developed a time-independent many-body Foldy-Wouthuysen transformation that gave us the proper semi-relativistic expression for these cross-sections in the Coulomb gauge. We can now compare our result to those obtained with the four Hamiltonians.

For the absorption cross-section, $H^{\rm B}$, $H^{\rm FW}$ and the NRQED Hamiltonians give the same result as our calculation. However if the Foldy-Wouthuysen transformation is also carried out for electric quadrupole transitions, then only $H^{\rm B}$ is still correct. Unfortunately, even if Blume started from the correct Hamiltonian, he subsequently neglected the term giving rise to the spin-position interaction.

For the scattering cross-section, a term similar to the spin-position operator arises from h_5 but multiplied by $\hbar\omega/\Delta E$. Since there is no energy conservation for transition with the intermediate states no Hamiltonian is valid for the scattering cross-section. We showed that the standard expression given by Blume overlooks a term which is not negligible in x-ray absorption spectroscopy. It will be interesting to evaluate the size of this term in the scattering cross-section.

Finally, we have shown that the absorption and scattering cross sections are indeed gauge invariant if they are written in the framework of quantum electrodynamics. This is reassuring since only gauge-invariant quantities are physically meaningful. The main conclusion of this work is that, to calculate a semi-relativistic cross-section, it is better to start from a fully relativistic cross-section than from a semi-relativistic (semi-classical) Hamiltonian.

A pending question is how Thomson scattering, which is due to the A^2 term in the non-relativistic approach, can be derived form in the relativistic framework, which has no A^2 term. We intend to address this question in a forthcoming publication.

Acknowledgments

We are very grateful to Uwe Gerstmann, Matteo Calandra and Nora Jenny Vollmers for encouraging us to work on the problem of the relativistic effects in x-ray absorption spectroscopy. We thank Alexander Silenko for this help concerning the Foldy-Wouthuysen transformations. Discussions with Amélie Juhin, Sergio Di Matteo, Yves Joly and Philippe Sainctavit are gratefully acknowledged. We are very grateful to Maria Esteban for her guidance through the mathematical literature on the many-body Dirac equation.

Appendix: General Foldy-Wouthuysen transformation

To derive a many-body Foldy-Wouthuysen transformation, we first notice that, in the one-body case, β endows the space of spinors with the structure of a Krein space, where β is called a *fundamental symmetry.*[129] For quite a different purpose,[130] we developed the tensor product of such spaces and showed that the fundamentaly symmetry of the Nth tensor power is essentially $\eta = \beta^{\otimes N}$. The abstract Krein-space framework leads us naturally to the following theorem:

Assume that H_D and η are self-adjoint operators and $\eta^2 = 1$. Then, there is a unitary operator U such that $U = \eta U^\dagger \eta$ and $\eta U H_D U^\dagger \eta = U H_D U^\dagger$. Moreover, if $|\psi_D\rangle$ is an eigenstate of H_D with positive (resp. negative) eigenvalue, then $|\psi\rangle = U|\psi_D\rangle$ satisfies $|\psi\rangle = \eta|\psi\rangle$ (resp. $|\psi\rangle = -\eta|\psi\rangle$).

The condition $U = \eta U^\dagger \eta$ does not appear in Foldy and Wouthuysen works. It was added by Eriksen.[116,117,131] It means that U is self-adjoint for the Krein-space structure.

Let us start with general considerations involving a self-adjoint operator η such that $\eta^2 = 1$. It can be used

to define projectors $B_\pm = (1 \pm \eta)/2$. It is clear that $B_+ + B_- = 1$, $B_\pm^2 = B_\pm$, $B_\pm^\dagger = B_\pm$ and $B_+B_- = B_-B_+ = 0$. A vector $|\psi\rangle$ is said to be even (odd) if $\eta|\psi\rangle = |\psi\rangle$ ($\eta|\psi\rangle = -|\psi\rangle$). Then, any vector $|\psi\rangle$ can be written as the sum of its even part $B_+|\psi\rangle$ and its odd part $B_-|\psi\rangle$. An operator H is said to be even (odd) if it transforms an even state into an even (odd) state and an odd state into an odd (even) state. An operator H is even (odd) if and only if $\eta H \eta = H$ ($\eta H \eta = -H$). Thus, the theorem states that UH_DU is an even operator. Any operator H can be written as the sum of its even part $B_+HB_+ + B_-HB_-$ and its odd part $B_+HB_- + B_-HB_+$.

Our proof of the theorem is essentially a generalized and rigorous version of Eriksen's proof.[131] We use the fact that H_D is self-adjoint to define $\lambda = \mathrm{sign} H_D$ by functional calculus. The operator λ is called the *flat band Hamiltonian* in topological insulator theory.[132] In physical terms, let $|\psi_D\rangle$ be an eigenstate of H_D for the energy E, then $\lambda|\psi_D\rangle = |\psi_D\rangle$ if $E \geq 0$ and $\lambda|\psi_D\rangle = -|\psi_D\rangle$ if $E < 0$. Since η and λ are self-adjoint and $\eta^2 = \lambda^2 = 1$, they are bounded and $\eta\lambda$ is unitary: $\eta\lambda(\eta\lambda)^\dagger = \eta\lambda\lambda\eta = \eta^2 = 1$ and $(\eta\lambda)^\dagger\eta\lambda = 1$. By the spectral theorem for unitary operators,[133] there is a unique family of orthogonal projections P_t such that

$$\eta\lambda \;=\; \int_{-\pi}^{\pi} e^{it} P_t dt.$$

In the finite dimensional case we could write this[134]

$$\eta\lambda \;=\; \sum_n e^{it_n} |\phi_n\rangle\langle\phi_n|.$$

Thus,

$$\lambda\eta \;=\; (\eta\lambda)^\dagger = \int_{-\pi}^{\pi} e^{-it} P_t dt = \int_{-\pi}^{\pi} e^{it} P_{-t} dt,$$

and, by unicity of P_t, $\eta\lambda = \eta(\eta\lambda)^\dagger\eta$ implies $P_t = \eta P_{-t}\eta$. We can now define a unitary square root U of $\eta\lambda$ by functional calculus:[135,136]

$$U \;=\; \sqrt{\eta\lambda} = \int_{-\pi}^{\pi} e^{it/2} P_t dt,$$

which satisfies

$$\eta U^\dagger \eta \;=\; \int_{-\pi}^{\pi} e^{-it/2} \eta P_t \eta dt = \int_{-\pi}^{\pi} e^{-it/2} P_{-t} dt = U.$$

We now show that this U satisfies the intertwining relation $\eta U = U\lambda$. Indeed, the relation $U^2 = \eta\lambda$ implies $U = U^\dagger\eta\lambda$. By multiplying from the left with η and using $\eta U^\dagger \eta = U$ we find $\eta U = U\lambda$. This important relation implies that $H = \eta H \eta$ and that $|\psi\rangle = U|\psi_D\rangle$ is even if $|\psi_D\rangle$ is a positive energy state and odd if $|\psi_D\rangle$ is a negative energy state.

The first property is easy to show:

$$\eta H \eta \;=\; \eta U H_D U^\dagger \eta = U\lambda H_D \lambda U^\dagger = U H_D \lambda^2 U^\dagger = H,$$

because λ commutes with H_D since it is a function of H_D.

To show the second property, let $\Gamma_\pm = (1 \pm \lambda)/2$, so that Γ_+ projects onto the space of positive energy and Γ_- of negative energy, and recall that $B_\pm = (1 \pm \eta)/2$. For a one-body system, $B_\pm$ projects onto the large/small components. Then, $U\Gamma_\pm = U/2 \pm U\lambda/2 = U/2 \pm \eta U/2 = B_\pm U$, which can be used to show that the Foldy-Wouthuysen wavefunctions $|\psi\rangle = U|\psi_D\rangle$ corresponding to positive energy have only even components. Indeed, let $|\psi_D\rangle$ be an eigenstate of H_D corresponding to a positive energy. By definition of λ we have $\Gamma_+|\psi_D\rangle = |\psi_D\rangle$ and $\Gamma_-|\psi_D\rangle = 0$. Thus, $U\Gamma_+|\psi_D\rangle = U|\psi_D\rangle = |\psi\rangle$ and $U\Gamma_+ = B_+U$ implies $|\psi\rangle = B_+U|\psi_D\rangle = B_+|\psi\rangle$. Thus $\eta|\psi\rangle = \eta B_+|\psi\rangle = B_+|\psi\rangle = |\psi\rangle$ and $|\psi\rangle$ is even. Similarly $0 = B_-|\psi\rangle$, so that the odd part of $|\psi\rangle$ is zero.

For a one-body system, even components and large components are identical. Indeed a Dirac one-body wavefunction can be written

$$|\psi_D\rangle \;=\; \begin{pmatrix} \phi \\ \psi \end{pmatrix},$$

If $\eta = \beta$, then the even part and the odd parts of $|\psi_D\rangle$ are, respectively,

$$\begin{pmatrix} \phi \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 \\ \psi \end{pmatrix}.$$

so that the small components of $|\psi\rangle$ are zero for a positive-energy state. This is not true for many body systems. For example, if we neglect antisymmetrization for notational convenience, a two body wavefunction can be obtained as the tensor product of one-body wavefunctions:

$$|\psi_D\rangle \;=\; \begin{pmatrix} \phi_1 \\ \psi_1 \end{pmatrix} \otimes \begin{pmatrix} \phi_2 \\ \psi_2 \end{pmatrix}.$$

Then, the even part of $|\psi_D\rangle$ is

$$\begin{pmatrix} \phi_1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} \phi_2 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \psi_1 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ \psi_2 \end{pmatrix},$$

while its odd part is

$$\begin{pmatrix} \phi_1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ \psi_2 \end{pmatrix} + \begin{pmatrix} 0 \\ \psi_1 \end{pmatrix} \otimes \begin{pmatrix} \phi_2 \\ 0 \end{pmatrix}.$$

The characterization of U as the square root of $\eta\lambda$ is not easy to handel. We give now a much simpler characterization:

Let U be a unitary operator continuously defined (outside zero) in terms of H_D such that: (i) $U = \eta U^\dagger \eta$; (ii) $\eta U H_D U^\dagger \eta = U H_D U^\dagger$; (iii) $U^2(-H_D) = -U^2(H_D)$. Then $U^\dagger \eta U = \pm \mathrm{sign}(H_D)$.

To prove this, define $Z = U^\dagger \eta U$. Clearly, $Z^\dagger = Z$. Moreover, Z is defined in terms H_D since U is. However, for Z to be a function of H_D in the sense of functional calculus, Z needs to commute with H_D:[137]if we multiply

condition (*ii*) from the right by ηU we find $\eta U H_D = U H_D U^\dagger \eta U = U H_D Z$. Hence,

$$Z H_D \;=\; U^\dagger \eta U H_D = U^\dagger U H_D Z = H_D Z.$$

Thus, there is a real function $f(t)$ and a family of orthogonal projections P_t such that

$$Z \;=\; \int_{-\infty}^{\infty} f(t) dP_t.$$

Moreover, $Z^2 = 1$ because $Z^2 = U^\dagger \eta U U^\dagger \eta U = U^\dagger \eta^2 U = U^\dagger U = 1$. Therefore, $f^2(t) = 1$ for every t. Finally, observe that $Z = \eta^2 U^\dagger \eta U = \eta U^2$, and condition (*iii*) implies that Z is an odd function of H_D: $f(-t) = -f(t)$. To conclude that $f(t) = \pm \mathrm{sign} t$, we need to add a condition of continuity on f outside zero. Indeed, functional calculus is valid for measurable functions and we could build a non-continuous odd function f such that $f^2 = 1$. In practice this does not take place because U is smoothly defined in terms of H_D, except at zero. No odd continuous function can satisfy $f^2 = 1$ over $\mathbb{R}$. It has to be discontinuous at zero. Since it is crucial that $f^2 = 1$ everywhere, we can choose either $\mathrm{sign} 0 = 1$ or $\mathrm{sign} 0 = -1$. Both solutions are valid.

[1] D. Ceresoli, U. Gerstmann, A. P. Seitsonen, and F. Mauri, Phys. Rev. B **81**, 060409(R) (2010).
[2] C. Pickard and F. Mauri, Phys. Rev. B **63**, 245101 (2001).
[3] M. Blume, SSRL Report **83/02**, 126 (1983).
[4] M. Blume, J. Appl. Phys. **51**, 3615 (1985).
[5] M. Blume, in *Resonant Anomalous X-Ray Scattering*, edited by G. Materlik, C. J. Sparks, and K. Fischer (North-Holland, Amsterdam, 1994), pp. 495–512.
[6] M. Takahashi and N. Hiraoka, Phys. Rev. B **92**, 094441 (2015).
[7] S. D. Matteo (2016), private communication.
[8] L. L. Foldy, Phys. Rev. **87**, 688 (1952).
[9] P. Strange, *Relativistic Quantum Mechanics* (Cambridge University Press, Cambridge, 1998).
[10] N. Bouldi, N. J. Vollmers, C. Gougoussis, A. Juhin, P. Sainctavit, C. Brouder, M. Calandra, F. Mauri, and U. Gerstmann, (2017).
[11] C. Itzykson and J.-B. Zuber, *Quantum Field Theory* (McGraw-Hill, New York, 1980).
[12] J. D. Bjorken and S. D. Drell, *Relativistic Quantum Fields* (McGraw-Hill, New York, 1965).
[13] M. M. Nieto, Phys. Rev. Lett. **38**, 1042 (1977).
[14] T. Goldman, Phys. Rev. D **15**, 1063 (1977).
[15] K.-H. Yang, J. Phys. A: Math. Gen. **15**, 437 (1982).
[16] W. E. Caswell and G. P. Lepage, Phys. Lett. **167B**, 437 (1986).
[17] G. Paz, Mod. Phys. Lett. A **30**, 1550128 (2015).
[18] Y. Joly, S. DiMatteo, and O. Bunau, Eur. Phys. J. Special Topics **208**, 21 (2012).
[19] J. W. E. Lamb, Phys. Rev. **85**, 259 (1952).
[20] D. H. Kobe, Int. J. Quant. Chem. **21**, 685 (1982).
[21] S. T. Epstein, Chem. Phys. Lett. **65**, 417 (1979).
[22] Y. Aharonov and C. K. Au, Phys. Rev. A **20**, 1553 (1979).
[23] I. P. Grant, J. Phys. B: Atom. Molec. Phys. **7**, 1458 (1974).
[24] K.-H. Yang, Annals Phys. **101**, 6 (1976).
[25] R. Zeyher, H. Bilz, and M. Cardona, Solid State Commun. **19**, 57 (1976).
[26] J. J. Forney, A. Quattropani, and F. Bassani, Nuovo Cimento **37B**, 78 (1977).
[27] D. H. Kobe and A. L. Smirl, Am. J. Phys. **46**, 624 (1978).
[28] G. Grynberg and E. Giacobino, J. Phys. B: Atom. Molec. Phys. **12**, L93 (1979).
[29] K. Haller and R. B. Sohn, Phys. Rev. A **20**, 1541 (1979).
[30] S. Olariu, I. Popescu, and C. B. Collins, Phys. Rev. D **20**, 3095 (1979).
[31] D. H. Kobe and K.-H. Yang, J. Phys. A: Math. Gen. **13**, 3171 (1980).
[32] D. H. Kobe and C.-T. Wen, Phys. Lett. **80A**, 121 (1980).
[33] C. Leubner and P. Zoller, J. Phys. B: Atom. Molec. Phys. **13**, 3613 (1980).
[34] Y. Aharonov and C. K. Au, Phys. Lett. **86A**, 269 (1981).
[35] M. I. Shirokov, Sov. Phys. JETP **54**, 645 (1981).
[36] D. H. Kobe, E. C. T. Wen, and K.-H. Yang, Phys. Rev. D **26**, 1927 (1982).
[37] Y. Aharonov and C. K. Au, Phys. Lett. **88A**, 491 (1982).
[38] K.-H. Yang, Phys. Lett. **92A**, 71 (1982).
[39] K.-H. Yang, J. Phys. A: Math. Gen. **15**, 1201 (1982).
[40] T. E. Feuchtwang, E. Kazes, H. Grotch, and P. H. Cutler, Phys. Lett. **93A**, 4 (1982).
[41] E. Kazes, T. E. Feuchtwang, P. H. Cutler, and H. Grotch, Annals Phys. **142**, 80 (1982).
[42] Y. Aharonov and C. K. Au, Phys. Lett. **95A**, 412 (1983).
[43] D. Lee and A. C. Albrecht, J. Chem. Phys. **78**, 5373 (1983).
[44] K.-H. Yang, J. Phys. A: Math. Gen. **16**, 919 (1983).
[45] C. K. Au, J. Phys. B: Mol. Phys. **17**, L59 (1984).
[46] T. E. Feuchtwang, E. Kazes, and P. H. Cutler, J. Phys. A: Math. Gen. **17**, 1157 (1984).
[47] T. E. Feuchtwang, E. Kazes, P. H. Cutler, and H. Grotch, J. Phys. A: Math. Gen. **17**, 151 (1984).
[48] D. H. Kobe and K.-H. Yang, Phys. Rev. A **20**, 2813 (1987).
[49] D. Adu-Gyamfi, J. Phys. A: Math. Gen. **19**, 3443 (1986).
[50] K.-H. Yang and D. H. Kobe, Annals Phys. **168**, 104 (1986).
[51] D. H. Kobe and S. M. Golsham, J. Phys. A: Math. Gen. **20**, 2813 (1987).
[52] J. W. E. Lamb, R. R. Schlicher, and M. O. Scully, Phys. Rev. A **36**, 2763 (1987).
[53] G. P. Arrighini, C. Guidotti, and N. Durante, Gazz. Chim. Ital. **118**, 703 (1988).
[54] N. Durante, G. P. Arrighini, and C. Guidotti, Z. Phys. D **8**, 63 (1988).
[55] K.-H. Yang, Annals Phys. **186**, 209 (1988).
[56] K. Haller and E. Lim-Lombridas, Found. Phys. **24**, 217 (1994).
[57] R. G. Woolley, Molec. Phys. **94**, 409 (1998).
[58] R. G. Woolley, Inter. J. Quant. Chem. **74**, 531 (1999).
[59] R. G. Woolley, Proc. R. Soc. London A **456**, 1803 (2000).

[60] I. M. Savukov and W. R. Johnson, Phys. Rev. A **62**, 052506 (2000).
[61] A. M. Stewart, J. Phys. A: Math. Gen. **33**, 9165 (2000).
[62] S.-W. Qian and Z.-Y. Gu, Commun. Theor. Phys. **38**, 267 (2002).
[63] R. G. Woolley, in *New Trends in Quantum Systems in Chemistry and Physics. Volume 2. Advanced Problems and Complex Systems. Paris, France 1999*, edited by J. Maruani, C. Minot, R. McWeeny, Y. Smeyers, and S. Wilson (Kluwer Academic Publishers, Dordrecht, 2002), vol. 7 of *Progress in Theoretical Chemistry and Physics*.
[64] K. Rzazewski and R. W. Boyd, J. Modern Opt. **51**, 1137 (2004).
[65] A. Stokes, J. Phys. B: At. Mol. Opt. Phys. **46**, 145505 (2013).
[66] A. D. Bandrauk, F. Fillion-Gourdeau, and E. Lorin, J. Phys. B: At. Mol. Opt. Phys. **46**, 153001 (2013).
[67] A. Mandal and K. L. C. Hunt, J. Chem. Phys. **144**, 044109 (2016).
[68] I. P. Grant and A. F. Starace, J. Phys. B: Atom. Molec. Phys. **8**, 1999 (1975).
[69] S.-W. Qian and J.-S. Wang, Commun. Theor. Phys. **49**, 308 (2008).
[70] C. Cohen-Tannoudji, J. Dupont-Roc, and G. Grynberg, *Photons and Atoms. Introduction to Quantum Electrodynamics* (Wiley-VCH, New York, 1997).
[71] M. Blagojević and F. W. Hehl, *Gauge Theories of Gravitation* (Imperial College Press, London, 2013).
[72] T. E. Feuchtwang and E. Kazes, J. Phys. A: Math. Gen. **18**, 2859 (1985).
[73] T. E. Feuchtwang, E. Kazes, and P. H. Cutler, Found. Phys. **16**, 1263 (1986).
[74] E. Kazes, T. E. Feuchtwang, H. Grotch, and P. H. Cutler, Phys. Rev. D **27**, 1388 (1983).
[75] S. Weinberg, *The Quantum Theory of Fields. I Foundations* (Cambridge University Press, Cambridge, 1995).
[76] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory* (Addison-Wesley, Reading, 1995).
[77] O. Steinmann, *Perturbative Quantum Electrodynamics and Axiomatic Field Theory* (Springer, Berlin, 2000).
[78] F. Strocchi, *An Introduction to Non-Perturbative Foundations of Quantum Field Theory* (Oxford Science Publications, Oxford, 2013).
[79] R. E. Kallosh and I. V. Tyutin, Soviet J. Nucl. Phys. **17**, 98 (1973).
[80] E. B. Manoukian, Int. J. Theor. Phys. **27**, 787 (1988).
[81] I. Bialynicki-Birula, Phys. Rev. D **2**, 2877 (1970).
[82] S. Weinberg, *The Quantum Theory of Fields II. Modern Applications* (Cambridge University Press, Cambridge, 1995).
[83] S. Hollands, Rev. Math. Phys. **20**, 1033 (2008).
[84] I. A. Batalin, P. M. Lavrov, and I. V. Tyutin, Int. J. Mod. Phys. **A29**, 1450166 (2014).
[85] P. Y. Moshin and A. A. Reshetnyak, Inter. J. Mod. Phys. A **30**, 1550021 (2015).
[86] H. Matsuda and R. Kubo, Prog. Theor. Phys. **63**, 275 (1980).
[87] T. Kashiwa and N. Tanimura, Fortschr. Phys. **45**, 381 (1997).
[88] D. Grigore, Ann. Phys. (Leipzig) **10**, 439 (2001).
[89] F. Lenz, H. W. L. Naus, K. Ohta, and M. Thies, Annals Phys. **233**, 17 (1994).
[90] B. L. Voronov, P. M. Lavrov, and I. V. Tyutin, Sov. J. Nucl. Phys. **36**, 292 (1982).
[91] A. K. Das, R. R. Francisco, and J. Frenkel, Phys. Rev. D **88**, 085012 (2013).
[92] K. Haller and L. F. Landovitz, Phys. Rev. **171**, 1749 (1968).
[93] K. Haller and L. F. Landovitz, Phys. Rev. D **2**, 1498 (1970).
[94] K. Haller, Acta Phys. Aust. **42**, 163 (1975).
[95] K. Haller and L. F. Landovitz, Phys. Rev. Lett. **22**, 245 (1969).
[96] K. Ohtaka and Y. Tanabe, Rev. Mod. Phys. **62**, 929 (1990).
[97] E. Klevak, J. J. Kas, and J. J. Rehr, Phys. Rev. B **89**, 085123 (2014).
[98] W. Greiner and J. Reinhardt, *Field Quantization* (Springer, Berlin, 1996).
[99] I. Bialynicki-Birula, in *Quantum Electrodynamics and Quantum Optics*, edited by A. O. Barut (Plenum Press, New York, 1984), vol. 110 of *NATO ASI Series B*, pp. 41–61.
[100] M. J. Esteban, M. Lewin, and E. Séré, Bull. Amer. Math. Soc. **45**, 535 (2008).
[101] W. Kutzelnigg, Chem. Phys. **395**, 16 (2012).
[102] C. Rovelli, Phys. Rev. Lett. **80**, 4613 (1998).
[103] N. P. Landsman, Phys. Rev. Lett. **83**, 1070 (1999).
[104] C. Brouder, G. Panati, and G. Stoltz, Ann. Henri Poincaré **11**, 1285 (2010).
[105] W. Heitler, *The Quantum Theory of Radiation* (Dover, New York, 1984), 3rd ed.
[106] C. Brouder, G. Panati, and G. Stoltz, Phys. Rev. Lett. **103**, 230401 (2009).
[107] B. A. Veklenko, Sov. Phys. J. **30**, 555 (1987).
[108] S. Hassing and E. N. Svendsen, J. Raman Spectr. **35**, 87 (2004).
[109] P. W. Milonni, R. Loudon, P. R. Berman, and S. M. Barnett, Phys. Rev. A **77**, 043835 (2008).
[110] S. Mukamel, Phys. Rev. A **76**, 021803 (2007).
[111] S. T. Epstein, Phys. Rev. **98**, 196 (1955).
[112] B. Hatfield, *Quantum Field Theory of Point Particles and Strings* (Addison-Wesley, Reading, 1992).
[113] D. Baye, Phys. Rev. C **86**, 034306 (2012).
[114] L. L. Foldy and S. A. Wouthuysen, Phys. Rev. **78**, 29 (1950).
[115] W. Greiner, *Relativistic Quantum Mechanics* (Springer, Berlin, 2000), 3rd ed.
[116] A. J. Silenko, Phys. Rev. A **93**, 022108 (2016).
[117] E. Eriksen, Phys. Rev. **111**, 1011 (1958).
[118] Z. V. Chraplyvy, Phys. Rev. **91**, 388 (1953).
[119] Z. V. Chraplyvy, Phys. Rev. **92**, 1310 (1953).
[120] M. Moshinsky and A. Nikitin, Rev. Mex. Fis. Suppl. 2 **50**, 66 (2004).
[121] W. Weiglhofer, Am. J. Phys. **57**, 455 (1989).
[122] H. A. Bethe and E. E. Salpeter, *Quantum Mechanics of One- and Two-Electron Atoms* (Plenum, New York, 1977).
[123] J. D. Morrison and R. Moss, Molec. Phys. **41**, 491 (1980).
[124] C. Cohen-Tannoudji, B. Diu, and F. Laloë, *Mécanique Quantique* (Hermann, Paris, 1973).
[125] J. D. Walecka, *Advanced Modern Physics. Theoretical Foundations* (World Scientific, Singapore, 2010).
[126] C. Brouder, M. Alouani, and K. H. Bennemann, Phys. Rev. B **54**, 7334 (1996).
[127] C. Gougoussis, *Excitations électroniques et magnétisme des matériaux: calcul ab initio de l'absorption X et du*

dichroïsme circulaire magnétique au seuil K (Université Pierre et Marie Curie-Paris 6, 2009), ph.D. thesis.

[128] J. Als-Nielsen and D. McMorrow, *Elements of Modern X-Ray Physics* (Wiley, New York, 2000).

[129] H. Baum, *Spin-Strukturen und Dirac-Operatoren über pseudoriemannschen Mannigfaltigkeiten* (Teubner, Lepzig, 1981).

[130] C. Brouder, N. Bizi, and F. Besnard, , (), arXiv:1504.03890.

[131] E. Eriksen and M. Kolsrud, Nuovo Cimento (Suppl.) **18**, 1 (1960).

[132] E. Prodan and H. Schulz-Baldes, *Bulk and Boundary Invariants for Complex Topological Insulators* (Springer, Berlin, 2016).

[133] M. Pourahmadi, *Foundations of Time Series Analysis and Prediction Theory* (Wiley, New York, 2001).

[134] S. Lang, *Linear Algebra*, Undergraduate Graduate Texts in Mathematics (Springer-Verlag, Berlin, 1987), 3rd ed.

[135] L. J. Wallen, Michigan J. Math. **16**, 153 (1969).

[136] A. A. Joye, in *XVIIth International Congress on Mathematical Physics*, edited by A. Jensen (World Scientific, Singapore, 2014), pp. 486–94.

[137] F. A. Berezin and M. A. Shubin, *The Schrödinger Equation* (Kluwer, Dordrecht, 1991).