



**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. 2017. hal-01383508v2

**HAL Id: hal-01383508**

**<https://hal.science/hal-01383508v2>**

Preprint submitted on 6 Apr 2017

**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<sup>1,2</sup> and Christian Brouder<sup>1</sup>

<sup>1</sup>*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*

<sup>2</sup>*Synchrotron SOLEIL, L'Orme des Merisiers, Saint-Aubin, BP 48, 91192 Gif-sur-Yvette Cedex, France*

(Dated: April 6, 2017)

There is an incompatibility between gauge invariance and the semi-classical time-dependent perturbation theory commonly used to calculate light absorption and scattering cross-sections. There is an additional incompatibility between perturbation theory and the description of the electron dynamics by a semi-relativistic Hamiltonian.

In this paper, the gauge-dependence problem of exact perturbation theory is described, the proposed solutions are reviewed and it is concluded that none of them seems fully satisfactory. The problem is finally solved by using the fully relativistic absorption and scattering cross-sections given by quantum electrodynamics. Then, a new many-body Foldy-Wouthuysen transformation is presented to obtain correct semi-relativistic transition operators. This transformation considerably simplifies the calculation of relativistic corrections. 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.

We compare our result with the ones obtained by using several semi-relativistic time-dependent Hamiltonians. In the case of absorption, the final formula agrees with the result obtained from one of them. However, the correct scattering cross-section is not given by any of the semi-relativistic Hamiltonians.

PACS numbers: 78.70.Ck X-ray scattering, 78.70.Dm X-ray absorption spectra, 11.15.Bt Perturbation theory, applied to gauge field theories, 31.30.jx Nonrelativistic limits of Dirac-Fock calculations

## I. INTRODUCTION

There is a well-known conflict between time-dependent perturbation theory and gauge invariance (see section II for a non exhaustive list of references). Indeed,  $|\langle \phi_n | \psi(t) \rangle|^2$  gives the probability to find the system described by the state  $|\psi(t)\rangle$  in the eigenstate  $|\phi_n\rangle$  at time  $t$ , where  $|\psi(t)\rangle$  is a solution of the time-dependent Schrödinger equation for the Hamiltonian  $H(t) = H_0 + H_1(t)$ , while  $|\phi_n\rangle$  is an eigenstate of the time-independent Hamiltonian  $H_0$ . A time-dependent gauge transformation of  $H(t)$  will be assigned to  $H_1(t)$  but not to  $H_0$ , which must remain independent of time. The transition probability is then modified because the state  $|\psi(t)\rangle$  is gauge-transformed and not the state  $|\phi_n\rangle$  (see section II for a more detailed argument).

As a (not so well-known) consequence, the absorption and scattering cross-sections derived by semi-classical arguments are not gauge invariant. Since a basic principle of quantum physics states that an observable has to be gauge invariant to be physically meaningful, we meet a serious difficulty: “Until this problem is understood, therefore, it seems that no calculation can be trusted at all.”<sup>1</sup>

There is a further conflict between time-dependent perturbation theory and semi-relativistic physics.<sup>2</sup> The semi-relativistic approximation of  $|\phi_n\rangle$  is obtained by applying to it the time-independent Foldy-Wouthuysen transformation:  $|\phi_n^{\text{FW}}\rangle = U_{H_0}|\phi_n\rangle$ . The semi-relativistic approximation of  $|\psi(t)\rangle$  is derived from the time-dependent Foldy-Wouthuysen transformation  $|\psi^{\text{FW}}(t)\rangle =$

$U_{H(t)}|\psi(t)\rangle$ . Since  $U_{H(t)} \neq U_{H_0}$ , the transformed transition probability  $|\langle \phi_n^{\text{FW}} | \psi^{\text{FW}}(t) \rangle|^2$  is not equal to  $|\langle \phi_n | \psi(t) \rangle|^2$ , even if the Foldy-Wouthuysen transformations  $U_{H_0}$  and  $U_{H(t)}$  are known to all orders.

In this paper we discuss and solve these two conflicts. In a nutshell, the gauge problem is solved by deriving relativistic absorption and scattering cross-sections from quantum electrodynamics instead of the usual semi-classical argument where the incident light wave is described by a time-dependent potential. The semi-relativistic problem is solved by applying a new many-body Foldy-Wouthuysen transformation to the relativistic cross-sections instead of describing the dynamics of the system with a semi-relativistic Hamiltonian. The final result is a semi-relativistic absorption and scattering cross-section involving a new term that couples the spin and the position operators. In a companion paper, we show that this new term contributes significantly to the x-ray absorption of magnetic materials.<sup>3</sup>

We now describe the outline of this paper. Section 2 discusses the gauge transformation of transition probabilities and reviews the solutions to the gauge-dependence problem proposed in the literature. Since none of them was widely accepted, we turn to the quantum electrodynamics framework in section 3, where we derive the relativistic electric dipole, quadrupole and magnetic dipole relativistic transition operators. In section 4, we derive a new many-body Foldy-Wouthuysen transformation that we apply to the transition operators. They are used to obtain semi-relativistic absorption and scattering cross-sections in sections 5 and 6, where a new

spin-position term is derived. In section 7, the conflict between time-dependent perturbation theory and semi-relativistic methods is described and illustrated by the calculation of spin-position term using four different semi-relativistic Hamiltonians commonly used in the literature. The conclusion presents possible extensions of the present work.

## II. 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 and obtained different results.<sup>4</sup> This gave rise to a long series of papers up to this day.<sup>1,2,5-53</sup> In 1987, the same Lamb (then Nobel prize winner) still considered this as “one of the outstanding problems of modern quantum optics.”<sup>37</sup>

We quickly describe the meaning of gauge invariance and then consider its failure in semi-classical perturbation theory.

### 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  $\Phi$  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  $\Lambda$  is any smooth function of space and time. In classical electromagnetism, 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 (Dirac) 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  $\beta$  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  $\psi$  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, a 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*.<sup>54</sup>

The principle of gauge invariance has become a cornerstone of particle physics. Since general relativity can also be considered as a gauge theory<sup>55</sup>, 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  $\Phi$ . Indeed, under a gauge transformation  $\Phi$  becomes  $\Phi - \dot{\Lambda}$  and the term  $\dot{\Lambda}$  cannot be compensated for in the absence of a time derivative.

### B. Gauge dependence of transition probabilities

In time-dependent perturbation theory, a system is assumed to be in 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

$$P_{ng}(t) = |\langle\phi_n|\psi(t)\rangle|^2 = |\langle\phi_n|U_A(t, t_0)|\phi_g\rangle|^2. \quad (1)$$

Since we want to ensure that the breakdown of gauge invariance is not due to an approximation, we work with exact (i.e. to all orders) perturbation theory. If we carry out a gauge transformation of the perturbation from  $a$  to  $a' = a - \partial\Lambda$ , then the evolution operator becomes<sup>33</sup>

$$U_{A'}(t, t_0) = M_\Lambda(t)U_A(t, t_0)M_\Lambda^\dagger(t_0), \quad (2)$$

where  $A' = a_0 + a'$ .

Therefore  $\langle\phi_n|\psi'(t)\rangle = \langle\phi_n|U_{A'}(t, t_0)|\phi_g\rangle$  is

$$\int d\mathbf{r}d\mathbf{r}' \phi_n^*(\mathbf{r})e^{ie\Lambda(\mathbf{r}, t)}U_A(\mathbf{r}, t; \mathbf{r}', t_0)e^{-ie\Lambda(\mathbf{r}', t_0)}\phi_g(\mathbf{r}'),$$

which is generally different from  $\langle\phi_n|\psi(t)\rangle$  since  $\Lambda(\mathbf{r}, t)$  is an arbitrary function (take for example  $\Lambda(\mathbf{r}, t) = \mathbf{r} \cdot \mathbf{k}(t)$ , where  $\mathbf{k}(t)$  is an arbitrary function of time). Therefore, in general,  $P'_{ng}(t) \neq P_{ng}(t)$  and the transition probabilities calculated in the two gauges are different. Moreover, since the transition rate entering cross-sections is the derivative of the transition probability with respect

to time,<sup>56</sup> the arbitrariness of the transition rate is increased by the fact that an arbitrary function  $\Lambda(\mathbf{r}, t)$  enters the integrand. Indeed, several papers evaluate the discrepancy between the probability calculated with two different gauges, and they find that it is generally not small.<sup>4,39,51</sup> By properly choosing  $\Lambda$ , the discrepancy can even be made arbitrary large.<sup>49</sup>

The absence of gauge invariance is due to the fact that the operator is transformed but not the eigenstates of  $a_0$ . This is called a *hybrid transformation* in the literature.<sup>13</sup>

### C. Proposed solutions

The lack of gauge invariance of transition probabilities is an alarming problem to which several solutions have been proposed. Since no clear consensus appears to have emerged,<sup>46</sup> we present a critical review of these solutions.

The first one, called the *consistent procedure*, was proposed by Forney and coll. and Epstein.<sup>9,11</sup> It 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  $|\phi_g\rangle$  and  $|\phi_n\rangle$  are also transformed into time-dependent states  $|\phi'_g\rangle = M_\Lambda(t_0)|\phi_g\rangle$  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  $H_0 = H_{a_0}$ .

Since the reference states  $|\phi_n\rangle$  become time-dependent, we leave the standard framework of time-dependent perturbation theory where the initial Hamiltonian  $H_0$  does not depend on time. Moreover, it is not physically clear why the gauge transformation should be applied to  $H_0$  and not to the perturbation.

In spite of these difficulties, many authors proposed to use the consistent procedure. However, as noticed by Yang,<sup>2</sup> 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 in the perturbation to start the calculation). This gauge dependence would be a serious problem because we would have to select the “true” physical gauge for the perturbation.

A second solution appeared in a series of papers starting in 1976,<sup>2,7,15,16,21,23,24,29,33,35,40</sup> 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

$$H_0(t) = \frac{(\mathbf{p} - e\mathbf{a}_0 - e\mathbf{a}(t))^2}{2m} + eV, \quad (3)$$

where  $V$  describes the electron-electron and electron-nuclear interactions so that  $H = H_0 + e\phi$ : the perturbation is only the scalar potential  $\phi$ . 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 transitions are calculated between the time-dependent states  $|\phi_n(t)\rangle$ . The corresponding transition probabilities are indeed gauge invariant. This solution has been used up to this day,<sup>38,50,53</sup> although it was also strongly criticized.<sup>10,17,18,25–27,30–32,57,58</sup> 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. (3) 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,<sup>59</sup> Feuchtwang, Kazes and coll. proposed the following alternative solution.<sup>25,26,31,32,60</sup> 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.<sup>54</sup> Thus, two Lagrangians that differ by a total time derivative are equivalent.<sup>61</sup> Then, they remark that the addition of a total time derivative  $e\dot{\Lambda}$  to the Lagrangian induces a gauge transformation  $A \rightarrow A - \partial\Lambda$  of the Hamiltonian.<sup>47,54,60</sup> 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 at the first order of perturbation theory, 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.<sup>58</sup> 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

$$\langle\psi|H_A - H_{A'}|\psi'\rangle = e\langle\psi|\dot{\Lambda}|\psi'\rangle + ie\frac{E - E'}{\hbar}\langle\psi|\Lambda|\psi'\rangle. \quad (4)$$

If we consider the absorption cross-section of a photon of energy  $\hbar\omega$ , then energy conservation implies that  $E' = E + \hbar\omega$ . Thus, if  $\Lambda$  satisfies  $\dot{\Lambda} = -i\omega\Lambda$ , then  $\langle\psi|H_A - H_{A'}|\psi'\rangle = 0$ .<sup>40,58</sup> In other words, by restricting

the gauge transformations 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.<sup>8,40,51</sup>

Equation (4) 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 published semi-classical approaches where the photon is represented by an external potential, the transition probabilities are not gauge-invariant and no proposed solution has reached general acceptance. 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 initial and final states is the same and, in the Schrödinger picture, the gauge transformation is expressed in terms of time-independent operators instead of a time-dependent function  $\Lambda$ .<sup>62</sup> Equation (4) suggests that transition probabilities, which are now described through the so-called S-matrix, could be gauge invariant.

This is indeed the case, although a review of the literature on the gauge invariance of QED might look ambiguous because the kind of gauge transformation considered in different works can vary. In standard textbooks, “the S-matrix is gauge invariant by construction”<sup>63</sup> because only the so-called  $\xi$ -term is modified. In the most general gauge transformation, the space of states change from one gauge to the other.<sup>64</sup> 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 (built by acting on the vacuum with creation operators of the left, right, longitudinal and scalar photons) can have a negative norm. In the Lorenz gauge, the Lorenz condition cannot be satisfied as an operator equation,<sup>65</sup> it becomes a subsidiary condition used to determine a subspace of physical states with positive norm.

In other words, the state spaces of the Coulomb and

Lorenz gauges have a quite different nature and the relation between them is delicate. Haller managed to show that the usual gauges are equivalent by devising a common framework containing all of them.<sup>41</sup> Note also that the gauge-invariance can only be expected for the renormalized S-matrix.<sup>66–68</sup>

The gauge invariance under 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.<sup>69,70</sup> The case of finite BRST transformations is in progress.<sup>71,72</sup>

To summarize the discussion, the gauge invariance of the renormalized S-matrix is established for infinitesimal gauge transformations and for a reasonably large class of gauges.<sup>13,41,68,73–78</sup> 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.<sup>79</sup> Although it was proved that the S-matrix elements are often the same with and without the subsidiary condition,<sup>10,13,26,80,81</sup> this fails when the Hamiltonian is suddenly changed,<sup>82</sup> as in the sudden creation of a core hole in photoemission or x-ray absorption<sup>83,84</sup>. In that case, Gauss’ law has to be imposed in the Lorenz gauge and the Coulomb gauge result is recovered.<sup>82</sup>

We choose to use quantum electrodynamics in the Coulomb gauge because it is the most accurate gauge for low-energy many-body calculations.<sup>54,85</sup>

### III. RELATIVISTIC MATRIX ELEMENTS

Since we now have a gauge-invariant framework, we can calculate the relativistic matrix elements that will be used in x-ray scattering and absorption cross-sections.

#### A. The Hamiltonian

The quantum field Hamiltonian describing the interaction of light with matter in the Coulomb gauge is:<sup>13,54,86,87</sup>

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

where

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

where  $\phi$  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  $\psi$  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  $\beta_n$  act on the  $n$ th Dirac electron. It can be given a well-defined mathematical meaning if the electronic system is described with respect to the Dirac sea,<sup>88</sup> although the physical validity of the Dirac sea is sometimes disputed.<sup>89</sup>

The photon Hamiltonian is

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

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 in the Coulomb gauge. According to Bialynicki-Birula, the Hamiltonian  $H$  also describes the dynamics of gauge-invariant states in any gauge.<sup>87</sup> 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

$$S = \lim_{\epsilon \rightarrow 0} T(e^{-\frac{i}{\hbar} \int_{-\infty}^{\infty} H_\epsilon(t) dt}), \quad (5)$$

where  $H_\epsilon(t) = e^{-\epsilon|t|} e^{iH_0 t} H_{e\gamma} e^{-iH_0 t}$ . 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  $\epsilon \rightarrow 0$  can be shown to exist up to technical assumptions.<sup>90</sup> 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 wavefunctions. 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 ‘‘non-covariant’’ approach,<sup>91</sup> 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:

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

Up to second order,

$$\langle m|T|n\rangle = \langle m|H_{e\gamma}|n\rangle + \sum_p \frac{\langle m|H_{e\gamma}|p\rangle\langle p|H_{e\gamma}|n\rangle}{e_p - e_n + i\gamma}, \quad (6)$$

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<sup>92</sup> but they would bring us too far. From the physical point of view,  $\gamma$  describes the life-time of the state  $|p\rangle$ , which can decay by radiative or non-radiative relaxation. The sign of the damping term  $\gamma$  has been the object of some controversy.<sup>93-96</sup>

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.<sup>97</sup> A modern version of the Schrödinger picture of QFT is given by Hatfield.<sup>98</sup>

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:<sup>99</sup>

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

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

We denote  $|n\rangle = a_{\mathbf{k}, l}^\dagger |0\rangle |\Psi_n\rangle$  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  $|e\rangle = a_{\mathbf{k}', l'}^\dagger a_{\mathbf{k}, l}^\dagger |0\rangle |\Psi_m\rangle$  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:

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

and

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

where

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

### 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.<sup>65,100</sup>

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  $F_I(t) = e^{iH_0t/\hbar}F e^{-iH_0t/\hbar}$ . 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

$$\begin{aligned} [H_0, F] &= \frac{i\hbar}{e} \int \nabla \cdot \mathbf{j}(\mathbf{r})f(\mathbf{r})d\mathbf{r} = -\frac{i\hbar}{e} \int \mathbf{j}(\mathbf{r}) \cdot \nabla f(\mathbf{r})d\mathbf{r} \\ &= -i\hbar c \int \psi^\dagger(\mathbf{r})\boldsymbol{\alpha}\psi(\mathbf{r}) \cdot \nabla f(\mathbf{r})d\mathbf{r}. \end{aligned} \quad (7)$$

To find the electric dipole transition term we apply Eq. (7) 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  $e^{i\mathbf{k} \cdot \mathbf{r}} \simeq 1$

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

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. (7) 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

$$\langle a|H_{e\gamma}|n\rangle = \frac{e\Delta E}{i\hbar} \sqrt{\frac{\hbar}{2\epsilon_0 V \omega}} \langle \Psi_m | \int \psi^\dagger T \psi | \Psi_n \rangle, \quad (8)$$

where  $\Delta E = E_m - E_n$  and

$$T = \boldsymbol{\epsilon} \cdot \mathbf{r} + \frac{i}{2} \boldsymbol{\epsilon} \cdot \mathbf{r}\mathbf{k} \cdot \mathbf{r} - \frac{\hbar c}{2\Delta E} (\boldsymbol{\epsilon} \times \mathbf{k}) \cdot (\mathbf{r} \times \boldsymbol{\alpha}). \quad (9)$$

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 out to be the magnetic-dipole operator (see section IV D). Similarly,

$$\langle e|H_{e\gamma}|n\rangle = \frac{e\Delta E}{i\hbar} \sqrt{\frac{\hbar}{2\epsilon_0 V \omega'}} \langle \Psi_m | \int \psi^\dagger T' \psi | \Psi_n \rangle, \quad (10)$$

where

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

## IV. SEMI-RELATIVISTIC REPRESENTATION

In the previous sections, we have shown that gauge invariance is ensured by describing the interaction of light and matter with quantum electrodynamics, where photons are quantized and electrons are described by four-component Dirac spinor quantum fields.

However, in most solid-state calculations, we do not use Dirac spinors but two-component (Pauli) wavefunctions. Moreover, semi-relativistic expressions are often physically clearer. Therefore, we need to link the two representations by using a generalization of the Foldy-Wouthuysen transformation.

In this section, we first describe the Foldy-Wouthuysen transformation and its new many-body extension. Then, we use this framework to calculate the relativistic corrections to the dipole and quadrupole transitions. The calculations are considerably simpler than the usual approach, where the relativistic corrections are derived from a semi-relativistic Hamiltonian.

### A. The Foldy-Wouthuysen transformation

The idea of the Foldy-Wouthuysen transformation is the following. If  $H_D$  is a time-independent 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

$H_D|\psi_D\rangle = E|\psi_D\rangle$ , 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<sup>101</sup> 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 consists in successive transformations of the form  $U = e^{iS}$ .<sup>101,102</sup>

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,<sup>103</sup> and he showed that the correction is at an order beyond the one we consider in this paper.

## B. 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:

$$H_N^0 = \sum_{n=1}^N \beta_n m c^2,$$

where  $\beta_n$  is the matrix  $\beta$  acting on the  $n$ th electron (i.e.  $\beta_n = 1^{\otimes(n-1)} \otimes \beta \otimes 1^{\otimes(N-n)}$ ). This definition is valid because  $H_N^0$  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  $\eta$  (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  $H_N^0$  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  $\eta = \beta_1 \otimes \dots \otimes \beta_n$  is the natural many-body generalization of  $\beta$ . Indeed,  $\eta^\dagger = \eta$  and  $\eta^2 = 1$ . Moreover,  $\eta$  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,<sup>104–106</sup> 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  $U = U_1 \otimes \dots \otimes U_N$  is a tensor power of one-body Foldy-Wouthuysen operators, as proposed by Moshinsky and Nikitin.<sup>107</sup>

However, this tensor-power form does not hold at higher orders. Indeed, we show now that at the next order, the many-body Foldy-Wouthuysen Hamiltonian is the sum of one-body and two-body contributions. 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  $m^{-2}$ :

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

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

It rewrites

$$H_{\text{FW}} = \sum_{n=1}^N H_{\text{FW}}^n + H_{\text{FW}}^{MB}. \quad (11)$$

where each  $H_{\text{FW}}^n$  is the usual one-body Foldy-Wouthuysen Hamiltonian:

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

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 expanding to higher order.

The new two-body term  $H_{\text{FW}}^{MB}$  arises because  $V(\mathbf{r}_m - \mathbf{r}_n) = V(\mathbf{r}_{mn})$  is a two body operator:

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

By using:<sup>108</sup>

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

the derivatives in the last term can be rewritten

$$\begin{aligned} (\boldsymbol{\alpha}_n \cdot \nabla_n)(\boldsymbol{\alpha}_p \cdot \nabla_p)V(\mathbf{r}_{np}) &= \sum_{jk} \boldsymbol{\alpha}_n^j \boldsymbol{\alpha}_p^k \partial_j \partial_k V(\mathbf{r}_{np}) \\ &= \frac{e^2}{8\pi\epsilon_0} \left( -\frac{4\pi}{3} \boldsymbol{\alpha}_n \cdot \boldsymbol{\alpha}_p \delta(\mathbf{r}_{np}) - \frac{\boldsymbol{\alpha}_n \cdot \boldsymbol{\alpha}_p}{|\mathbf{r}_{np}|^3} \right. \\ &\quad \left. + 3 \frac{\boldsymbol{\alpha}_n \cdot \mathbf{r}_{np} \boldsymbol{\alpha}_p \cdot \mathbf{r}_{np}}{|\mathbf{r}_{np}|^5} \right). \end{aligned}$$

This expression looks superficially like some contributions to the Breit interaction as presented by Bethe and Salpeter.<sup>109</sup> 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 the 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.<sup>110</sup> At order  $m^{-2}$ , the transformation writes

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

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  $U$ .

### C. 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  $|\Psi\rangle = U^\dagger |\psi\rangle$ . Since  $U$  is written as a many-body operator, we translate the quantum field expression for  $D$  into the many-body formula  $D = \langle \Phi | \boldsymbol{\epsilon} \cdot \mathbf{R} | \Psi \rangle$ , where  $\mathbf{R} = \sum_{n=1}^N \mathbf{r}_n$ .<sup>111,112</sup> We calculate  $D = \langle \phi | U \boldsymbol{\epsilon} \cdot \mathbf{R} U^\dagger | \psi \rangle$ , where  $U = e^{iS}$  by using the Baker-Campbell-Hausdorff formula

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

where  $L(T) = [S, T]$  and  $L^n(T) = L(L^{n-1}(T))$ . If  $U = U_1 \otimes \dots \otimes U_N$ , 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

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

and

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

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

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

So that, for each particle, and up to  $O(m^{-2})$ ,

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

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

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

### D. Semi-relativistic multipole transitions

From Eq. (9), we write the multipole transitions

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

where

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

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-Hausdorff formula.

The term  $[S_n, \boldsymbol{\epsilon} \cdot \mathbf{r}_n \mathbf{k} \cdot \mathbf{r}_n]$  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

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

and

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

Note that  $\hbar \boldsymbol{\Sigma} = g \mathbf{S}$  with  $g = 2$  (because the spin operator is  $\mathbf{S} = \hbar \boldsymbol{\Sigma}/2$ ). Thus, we recover the fact that the Dirac equation gives a gyromagnetic factor  $g = 2$  to the electron. Moreover,  $\mathbf{L} + \hbar \boldsymbol{\Sigma} = \mathbf{L} + 2\mathbf{S}$  is 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. 113. The term  $M_2$  describes magnetic-dipole transitions. The multipole transitions are

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

### V. 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 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:<sup>114</sup>

$$w = \frac{2}{\hbar} \delta_{mn} \text{Im} \langle m | T | m \rangle + \frac{2\pi}{\hbar} \delta(e_n - e_m) |\langle n | T | m \rangle|^2. \quad (12)$$

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

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

where  $T_{\text{FW}}$  is:

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

with  $\alpha_0$  the fine structure constant and  $\Delta E = E_f - E_i = \hbar\omega$ .

It corresponds to the usual formula for the cross section<sup>115</sup> with two more terms: the third one and the last one.

The third term was already found by Christos Gougoussis in his PhD thesis,<sup>116</sup> but his final result was not in agreement with ours because of his use of the commutation relation, as described in section VII F. We rewrite it by using  $\boldsymbol{\pi} = (m/i\hbar)[\mathbf{r}, H_0^{\text{FW}}] + O(c^{-2})$ , where  $H_0^{\text{FW}}$  is the Foldy-Wouthuysen Hamiltonian, to get:

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

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 a companion paper.<sup>3</sup>

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.

## VI. 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  $\mathbf{k}_i, \epsilon_i$  and that in the final state the system is in state  $|F\rangle$  with a scattered photon  $\mathbf{k}_f, \epsilon_f$ . We do not consider the special case when  $\mathbf{k}_i, \epsilon_i = \mathbf{k}_f, \epsilon_f$ .

Eqs. (8), (10) and (12) yield:

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

where  $\gamma > 0$  and

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

The scattering cross-section is related to  $w$  by:<sup>117</sup>

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

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

$$\frac{d^2\sigma}{d\Omega d\omega_f} = (r_e m c^2)^2 \frac{\omega_f}{\omega_i} \sum_F \delta(E_f + \hbar\omega_f - E_i - \hbar\omega_i) \left| \sum_L \frac{\langle F | e_{-\mathbf{k}_f} \psi^\dagger \boldsymbol{\alpha} \cdot \boldsymbol{\epsilon}_f^* \psi | L \rangle \langle L | e_{\mathbf{k}_i} \psi^\dagger \boldsymbol{\alpha} \cdot \boldsymbol{\epsilon}_i \psi | I \rangle}{E_i - E_l + \hbar\omega_i + i\gamma} + \frac{\langle F | e_{\mathbf{k}_i} \psi^\dagger \boldsymbol{\alpha} \cdot \boldsymbol{\epsilon}_i \psi | L \rangle \langle L | e_{-\mathbf{k}_f} \psi^\dagger \boldsymbol{\alpha} \cdot \boldsymbol{\epsilon}_f^* \psi | I \rangle}{E_i - E_l - \hbar\omega_f} \right|^2.$$

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 standard one.<sup>118</sup> If we are interested in the resonant part

of the scattering cross section, then  $E_l > 0$  and

$$\frac{d^2\sigma}{d\Omega d\omega_f} = \left( \frac{r_e m}{\hbar^2} \right)^2 \frac{\omega_f}{\omega_i} \sum_f \delta(E_f + \hbar\omega_f - E_i - \hbar\omega_i) \left| \sum_{L>} (E_l - E_i)(E_f - E_l) \frac{\langle f | T_{\text{FW}}^{f'l}(\boldsymbol{\epsilon}_f) | l \rangle \langle l | T_{\text{FW}}^{li}(\boldsymbol{\epsilon}_i) | i \rangle}{E_i - E_l + \hbar\omega_i + i\gamma} \right|^2.$$

with

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

and

$$T_{\text{FW}}^{ij}(\boldsymbol{\epsilon}_f) = \sum_n \boldsymbol{\epsilon}_f^* \cdot \mathbf{r}_n - \frac{i}{2} \boldsymbol{\epsilon}_f^* \cdot \mathbf{r}_n \mathbf{k}_f \cdot \mathbf{r}_n - \frac{\hbar}{4m^2 c^2} \boldsymbol{\pi}_n \cdot (\boldsymbol{\epsilon}_f^* \times \boldsymbol{\Sigma}_n) + \frac{\hbar \beta_n}{2m \Delta E^{ij}} (\boldsymbol{\epsilon}_f^* \times \mathbf{k}_f) \cdot (\hbar \boldsymbol{\Sigma}_n + \boldsymbol{\Lambda}_n),$$

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.<sup>118</sup>

## VII. OTHER METHODS

In this section, we compare our semi-relativistic transition matrix elements with the ones obtained by using time-dependent perturbation theory where the time-evolution is described by several time-dependent 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). Before making this comparison, we first explain why using a time-dependent semi-relativistic Hamiltonian in a perturbation calculation can lead to incorrect results.

### A. Foldy-Wouthuysen subtelties

In this section, we assume that the exact time-dependent Foldy-Wouthuysen operator  $U$  is known. Thus, the following difficulties are not related to the use of an approximation, but to the interplay of the Foldy-Wouthuysen method with perturbation theory.

The first subtlety was noticed by Nieto.<sup>59,119</sup> 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. The time-dependent Dirac equation for  $|\Psi\rangle$  implies that  $|\psi\rangle$  is a solution of the time-dependent Schrödinger equation  $(i\hbar\partial_t - H')|\psi\rangle = 0$ , where  $H' = UHU^{-1} + i\hbar(\partial_t U)U^{-1}$  is the time-dependent Foldy-Wouthuysen Hamiltonian. In the following, an uppercase Greek letter ( $|\Phi\rangle$  or  $|\Psi\rangle$ ) refers to a solution of the Dirac equation and the corresponding lowercase letter ( $|\phi\rangle$  or  $|\psi\rangle$ ) to its Foldy-Wouthuysen transformation.

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 subtlety was observed by Yang,<sup>2</sup> and concerns the most straightforward way to use the Foldy-Wouthuysen Hamiltonian  $H'(t)$ , where the time dependence is now explicit, to compute transition probabilities. This Hamiltonian is split into a time-independent part  $H'_0$  and a time-dependent one  $H'_1(t)$ , so that  $H'(t) = H'_0 + H'_1(t)$ . The scalar product  $\langle\phi'_n|\psi(t)\rangle$ , where  $|\phi'_n\rangle$  is an eigenstate of  $H'_0$ , cannot be equal to the relativistic scalar product  $\langle\Phi_n|\Psi(t)\rangle$ . Indeed  $|\psi(t)\rangle = U(H_0 + H_1(t))|\Psi(t)\rangle$  but  $|\phi'_n\rangle \neq U(H_0 + H_1(t))|\Phi_n\rangle$  because  $|\phi'_n\rangle$  and  $|\Phi_n\rangle$  are independent of time whereas  $U(H_0 + H_1(t))$  depends on time. Since only the QED relativistic matrix elements were found to be gauge invariant,  $\langle\phi'_n|\psi(t)\rangle$  is generally not physically meaningful.

The two problems combine if first-order perturbation theory is naively applied with Foldy-Wouthuysen eigenstates and Hamiltonian. The Foldy-Wouthuysen interaction Hamiltonian  $H'_1(t) = H'(t) - H'_0 \neq U(H_0)(H(t) - H_0)U^\dagger(H_0)$ . As a consequence,  $\langle\phi'_n|H'_1(t)|\phi'_g\rangle$  is not equal to  $\langle\Phi_n|H_1(t)|\Phi_g\rangle$ .

To illustrate the variety of results that can be obtained by using first-order perturbation theory with semi-relativistic Hamiltonians, we now examine four Hamiltonians used in practice. To help comparing these Hamiltonians, we express them in a common one-particle framework.

## B. The Blume Hamiltonian

Blume discussed the interaction of light with magnetic matter by starting from the Hamiltonian:<sup>118,120</sup>

$$H^B = \frac{\boldsymbol{\pi}^2}{2m} + eV - \frac{e\hbar}{2m}\boldsymbol{\sigma} \cdot \mathbf{B} - \frac{e\hbar}{4m^2c^2}\boldsymbol{\sigma} \cdot (\mathbf{E} \times \boldsymbol{\pi}), \quad (13)$$

where  $\boldsymbol{\pi} = \mathbf{p} - e\mathbf{A}$ . This Hamiltonian is the sum of four terms: (i) the kinetic energy of the electron, (ii) an external potential, (iii) the Zeeman interaction between the electron and a 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  $H_\gamma$  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,<sup>118,120</sup> but this was corrected in the third one.<sup>121</sup> 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.

## C. Foldy-Wouthuysen Hamiltonian

We consider now the so-called ‘‘gauge-invariant’’ Foldy-Wouthuysen Hamiltonian for positive-energy states up to order  $1/(mc)^2$ .<sup>122</sup>

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

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 presently. A basic difference between  $H^B$  and  $H^{\text{FW}}$  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.<sup>99</sup> In the semi-classical treatment of light-matter interaction, the photons are represented by an external time-dependent potential and this term is present.

These Hamiltonians can be written  $H(\mathbf{A}, \Phi)$ , where the total vector potential  $\mathbf{A}$  and scalar potential  $\Phi$  are a sum  $\mathbf{A} = \mathbf{a}_0 + \mathbf{a}$ ,  $\Phi = \phi_0 + \phi$ , of static external potentials  $\mathbf{a}_0$  and  $\phi_0$  (representing the static internal and external fields) perturbed by dynamical potentials  $\mathbf{a}$  and  $\phi$  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  $H^{\text{FW}}$  lead to two different interactions:  $H_I^B = h_1 + h_2 + h_3 + h_4 + h_5 + h_6$  and

$H_I^{\text{FW}} = H_I^{\text{B}} + h_7$ , where

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

with  $\boldsymbol{\pi}_0 = \mathbf{p} - e\mathbf{a}_0$ . The *curl-term* in  $H^{\text{FW}}$  is the origin of the presence of  $h_7$  in  $H_I^{\text{FW}}$ . It originates from the term  $i\hbar(\partial_t U)U^{-1}$  in the time-dependent Foldy-Wouthuysen Hamiltonian. 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.<sup>120</sup> 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.<sup>3</sup>

#### D. Textbook Foldy-Wouthuysen Hamiltonian

Standard textbooks often derive a Foldy-Wouthuysen Hamiltonian  $H^{\text{TFW}}$  which is the same as  $H^{\text{FW}}$ , except for the fact that  $\boldsymbol{\pi}$  is replaced by  $\mathbf{p}$  in the spin-orbit term.<sup>123,124</sup> A mass-velocity term  $-(\mathbf{p} \cdot \mathbf{p})^2/8m^3c^2$  is often added<sup>123</sup> 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.

#### 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.<sup>125</sup> They wrote the most general gauge-invariant non-relativistic Lagrangian terms and fitted the coefficients of these terms to known QED processes.<sup>126</sup>

The corresponding NRQED Hamiltonian is the same as  $H^{\text{FW}}$  up to order  $c^{-2}$ , but its interpretation is different.<sup>126</sup> 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  $H^{\text{FW}}$ . 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 VII A.

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:<sup>127,128</sup>

$$\mathbf{p} = \frac{mi}{\hbar} [H_0, \mathbf{r}]. \quad (14)$$

However, if one considers the static Hamiltonian given by Blume (13), its commutator with  $\mathbf{r}$  is:

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

which is different from Eq. (14) 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  $H_0^{\text{TFW}}$  and  $H_0^{\text{FW}}$  are the same. In  $H_I^{\text{FW}}$  and  $H_I^{\text{B}}$ , 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  $\frac{e^2 \hbar}{4m^2 c^2} \boldsymbol{\sigma} \cdot [\nabla v_0 \times \mathbf{a}]$  in  $H_I^{\text{TFW}}$ .

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

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  $\Delta E = -\hbar\omega$  in the cross section so that  $h_5$  and  $h_6$  lead to the same contribution to the matrix element:

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

which corresponds to the *spin-position* interaction. Explicit calculations showed that this contribution can appear two times, one time or cancel completely, according to which Hamiltonian and which commutator was used. Starting from  $H^{\text{FW}}$ , the same absorption cross section as in our new approach can be derived. However, in the case of scattering, even with  $H^{\text{FW}}$ , there is a factor  $\Delta E/\hbar\omega$  which is not correct. The same kind of discrepancy was already observed in the literature.<sup>8,40</sup>

### VIII. CONCLUSION

This paper was written because of the gauge-dependence of transition probabilities in the semi-classical approach and because we observed, after other authors,<sup>127,129</sup> that different semi-relativistic Hamiltonians lead to different cross-sections.

Our solution makes essential use of quantum electrodynamics as the correct gauge-invariant framework to discuss the interaction of light with matter. It is well-known that the semi-classical and QED absorption cross-sections are identical in the Coulomb gauge.<sup>130</sup> This is compatible with our discussion because, to go from the Coulomb gauge to another gauge, the semi-classical approach only involves the operator  $M_\Lambda$ , while QED involves a redefinition of the space of states, including in an essential way non-physical polarizations and even ghost states in the BRST approach. This redefinition is able to maintain gauge invariance where the semi-classical  $M_\Lambda$  fails to do so.

In the present paper, the stationary states of the electronic system was taken to be eigenstates of  $H_e$ . The interaction Hamiltonian  $H_{e\gamma}$  can modify these states through various QED effects, for example the Breit interaction discussed by Bethe and Salpeter.<sup>109</sup> We expect these contributions to be small in x-ray spectroscopy.

The explicit calculation of the spin-position contribution at the K-edge of Fe, Co and Ni will be presented in a forthcoming publication.<sup>3</sup>

It was known since Heisenberg in 1928,<sup>131</sup> that the Thomson cross-section which is due to the  $A^2$  term in the non-relativistic approach, can be derived from the relativistic framework by using a sum over negative-energy states.<sup>99,132</sup> We intend to provide a more accurate discussion of the contribution of negative-energy states to the scattering cross-section by using our many-body Foldy-Wouthuysen approach.

#### Acknowledgments

We are very grateful to Uwe Gerstmann, Matteo Cailandra 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.

This work was supported by French state funds managed by the ANR within the Investissements d'Avenir programme under Reference No. ANR-11-IDEX-0004-02, and more specifically within the framework of the Cluster of Excellence MATISSE led by Sorbonne Universités.

### Appendix: General Foldy-Wouthuysen transformation

To derive a many-body Foldy-Wouthuysen transformation, we first notice that, in the one-body case,  $\beta$  endows the space of spinors with the structure of a Krein space, where  $\beta$  is then called a *fundamental symmetry*.<sup>133</sup> For quite a different purpose,<sup>134</sup> we investigated the tensor product of such spaces and showed that the fundamental symmetry of the  $N$ th 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  $\eta$  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.<sup>103,106,135</sup> It means that  $U$  is self-adjoint for the Krein-space structure.

Let us start with general considerations involving a self-adjoint operator  $\eta$  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  $U H_D U$  is an even operator. Any operator  $H$  can be written as the sum of its even part  $B_+ H B_+ + B_- H B_-$  and its odd part  $B_+ H B_- + B_- H B_+$ .

Our proof of the theorem is essentially a generalized and rigorous version of Eriksen's proof.<sup>135</sup> We use the fact that  $H_D$  is self-adjoint to define  $\lambda = \text{sign} H_D$  by functional calculus. The operator  $\lambda$  is called the *flat band Hamiltonian* in topological insulator theory.<sup>136</sup> 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  $\eta$  and  $\lambda$  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,<sup>137</sup> 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<sup>138</sup>

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

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:<sup>139,140</sup>

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

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  $\eta$  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  $\lambda$  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  $\Gamma_+$  projects onto the space of positive energy and  $\Gamma_-$  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  $\lambda$  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 state 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 handle. 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  $U^\dagger \eta U = \pm \text{sign}(H_D)$ .

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  $H_D$ :<sup>141</sup> if we multiply condition (ii) from the right by  $\eta U$  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  $P_t$  corresponding to the eigenstates of  $H_D$  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  $f(t) = \pm \text{sign} t$ , we need to add the 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  $f^2 = 1$  outside the origin. 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  $\text{sign} 0 = 1$  or  $\text{sign} 0 = -1$ . Both solutions are valid.

<sup>1</sup> D. H. Kobe and A. L. Smirl, Am. J. Phys. **46**, 624 (1978).

<sup>2</sup> K.-H. Yang, J. Phys. A: Math. Gen. **15**, 437 (1982).

<sup>3</sup> N. Bouldi, N. J. Vollmers, C. G. Delpy-Laplanche, Y. Joly, A. Juhin, P. Sainctavit, C. Brouder, M. Calandra,

- L. Paulatto, F. Mauri, and U. Gerstmann, submitted.
- <sup>4</sup> J. W. E. Lamb, Phys. Rev. **85**, 259 (1952).
  - <sup>5</sup> I. P. Grant, J. Phys. B: Atom. Molec. Phys. **7**, 1458 (1974).
  - <sup>6</sup> I. P. Grant and A. F. Starace, J. Phys. B: Atom. Molec. Phys. **8**, 1999 (1975).
  - <sup>7</sup> K.-H. Yang, Annals Phys. **101**, 6 (1976).
  - <sup>8</sup> R. Zeyher, H. Bilz, and M. Cardona, Solid State Commun. **19**, 57 (1976).
  - <sup>9</sup> J. J. Forney, A. Quattropani, and F. Bassani, Nuovo Cimento **37B**, 78 (1977).
  - <sup>10</sup> Y. Aharonov and C. K. Au, Phys. Rev. A **20**, 1553 (1979).
  - <sup>11</sup> S. T. Epstein, Chem. Phys. Lett. **65**, 417 (1979).
  - <sup>12</sup> G. Grynberg and E. Giacobino, J. Phys. B: Atom. Molec. Phys. **12**, L93 (1979).
  - <sup>13</sup> K. Haller and R. B. Sohn, Phys. Rev. A **20**, 1541 (1979).
  - <sup>14</sup> S. Olariu, I. Popescu, and C. B. Collins, Phys. Rev. D **20**, 3095 (1979).
  - <sup>15</sup> D. H. Kobe and K.-H. Yang, J. Phys. A: Math. Gen. **13**, 3171 (1980).
  - <sup>16</sup> D. H. Kobe and C.-T. Wen, Phys. Lett. **80A**, 121 (1980).
  - <sup>17</sup> C. Leubner and P. Zoller, J. Phys. B: Atom. Molec. Phys. **13**, 3613 (1980).
  - <sup>18</sup> Y. Aharonov and C. K. Au, Phys. Lett. **86A**, 269 (1981).
  - <sup>19</sup> M. I. Shirokov, Sov. Phys. JETP **54**, 645 (1981).
  - <sup>20</sup> D. H. Kobe, Int. J. Quant. Chem. **21**, 685 (1982).
  - <sup>21</sup> D. H. Kobe, E. C. T. Wen, and K.-H. Yang, Phys. Rev. D **26**, 1927 (1982).
  - <sup>22</sup> Y. Aharonov and C. K. Au, Phys. Lett. **88A**, 491 (1982).
  - <sup>23</sup> K.-H. Yang, Phys. Lett. **92A**, 71 (1982).
  - <sup>24</sup> K.-H. Yang, J. Phys. A: Math. Gen. **15**, 1201 (1982).
  - <sup>25</sup> T. E. Feuchtwang, E. Kazes, H. Grotch, and P. H. Cutler, Phys. Lett. **93A**, 4 (1982).
  - <sup>26</sup> E. Kazes, T. E. Feuchtwang, P. H. Cutler, and H. Grotch, Annals Phys. **142**, 80 (1982).
  - <sup>27</sup> Y. Aharonov and C. K. Au, Phys. Lett. **95A**, 412 (1983).
  - <sup>28</sup> D. Lee and A. C. Albrecht, J. Chem. Phys. **78**, 5373 (1983).
  - <sup>29</sup> K.-H. Yang, J. Phys. A: Math. Gen. **16**, 919 (1983).
  - <sup>30</sup> C. K. Au, J. Phys. B: Mol. Phys. **17**, L59 (1984).
  - <sup>31</sup> T. E. Feuchtwang, E. Kazes, and P. H. Cutler, J. Phys. A: Math. Gen. **17**, 1157 (1984).
  - <sup>32</sup> T. E. Feuchtwang, E. Kazes, P. H. Cutler, and H. Grotch, J. Phys. A: Math. Gen. **17**, 151 (1984).
  - <sup>33</sup> D. H. Kobe and K.-H. Yang, Phys. Rev. A **32**, 952 (1985).
  - <sup>34</sup> D. Adu-Gyamfi, J. Phys. A: Math. Gen. **19**, 3443 (1986).
  - <sup>35</sup> K.-H. Yang and D. H. Kobe, Annals Phys. **168**, 104 (1986).
  - <sup>36</sup> D. H. Kobe and S. M. Golsham, J. Phys. A: Math. Gen. **20**, 2813 (1987).
  - <sup>37</sup> J. W. E. Lamb, R. R. Schlicher, and M. O. Scully, Phys. Rev. A **36**, 2763 (1987).
  - <sup>38</sup> G. P. Arrighini, C. Guidotti, and N. Durante, Gazz. Chim. Ital. **118**, 703 (1988).
  - <sup>39</sup> N. Durante, G. P. Arrighini, and C. Guidotti, Z. Phys. D **8**, 63 (1988).
  - <sup>40</sup> K.-H. Yang, Ann. Phys. **186**, 209 (1988).
  - <sup>41</sup> K. Haller and E. Lim-Lombridas, Found. Phys. **24**, 217 (1994).
  - <sup>42</sup> R. G. Woolley, Molec. Phys. **94**, 409 (1998).
  - <sup>43</sup> R. G. Woolley, Inter. J. Quant. Chem. **74**, 531 (1999).
  - <sup>44</sup> R. G. Woolley, Proc. R. Soc. London A **456**, 1803 (2000).
  - <sup>45</sup> I. M. Savukov and W. R. Johnson, Phys. Rev. A **62**, 052506 (2000).
  - <sup>46</sup> A. M. Stewart, J. Phys. A: Math. Gen. **33**, 9165 (2000).
  - <sup>47</sup> S.-W. Qian and Z.-Y. Gu, Commun. Theor. Phys. **38**, 267 (2002).
  - <sup>48</sup> 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*.
  - <sup>49</sup> K. Rzazewski and R. W. Boyd, J. Modern Opt. **51**, 1137 (2004).
  - <sup>50</sup> S.-W. Qian and J.-S. Wang, Commun. Theor. Phys. **49**, 308 (2008).
  - <sup>51</sup> A. Stokes, J. Phys. B: At. Mol. Opt. Phys. **46**, 145505 (2013).
  - <sup>52</sup> A. D. Bandrauk, F. Fillion-Gourdeau, and E. Lorin, J. Phys. B: At. Mol. Opt. Phys. **46**, 153001 (2013).
  - <sup>53</sup> A. Mandal and K. L. C. Hunt, J. Chem. Phys. **144**, 044109 (2016).
  - <sup>54</sup> C. Cohen-Tannoudji, J. Dupont-Roc, and G. Grynberg, *Photons and Atoms. Introduction to Quantum Electrodynamics* (Wiley-VCH, New York, 1997).
  - <sup>55</sup> M. Blagojević and F. W. Hehl, *Gauge Theories of Gravitation* (Imperial College Press, London, 2013).
  - <sup>56</sup> A. Bohm, *Quantum Mechanics: Foundations and Applications* (Springer, New York, 1993), 3rd ed.
  - <sup>57</sup> T. E. Feuchtwang and E. Kazes, J. Phys. A: Math. Gen. **18**, 2859 (1985).
  - <sup>58</sup> T. E. Feuchtwang, E. Kazes, and P. H. Cutler, Found. Phys. **16**, 1263 (1986).
  - <sup>59</sup> T. Goldman, Phys. Rev. D **15**, 1063 (1977).
  - <sup>60</sup> E. Kazes, T. E. Feuchtwang, H. Grotch, and P. H. Cutler, Phys. Rev. D **27**, 1388 (1983).
  - <sup>61</sup> S. Weinberg, *The Quantum Theory of Fields. I Foundations* (Cambridge University Press, Cambridge, 1995).
  - <sup>62</sup> N. Nakanishi and I. Ojima, *Covariant Operator Formalism of Gauge Theories and Quantum Gravity* (World Scientific, Singapore, 1990).
  - <sup>63</sup> M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory* (Addison-Wesley, Reading, 1995).
  - <sup>64</sup> O. Steinmann, *Perturbative Quantum Electrodynamics and Axiomatic Field Theory* (Springer, Berlin, 2000).
  - <sup>65</sup> F. Strocchi, *An Introduction to Non-Perturbative Foundations of Quantum Field Theory* (Oxford Science Publications, Oxford, 2013).
  - <sup>66</sup> I. Bialynicki-Birula, Phys. Rev. D **2**, 2877 (1970).
  - <sup>67</sup> R. E. Kallosh and I. V. Tyutin, Soviet J. Nucl. Phys. **17**, 98 (1973).
  - <sup>68</sup> E. B. Manoukian, Int. J. Theor. Phys. **27**, 787 (1988).
  - <sup>69</sup> S. Weinberg, *The Quantum Theory of Fields II. Modern Applications* (Cambridge University Press, Cambridge, 1995).
  - <sup>70</sup> S. Hollands, Rev. Math. Phys. **20**, 1033 (2008).
  - <sup>71</sup> I. A. Batalin, P. M. Lavrov, and I. V. Tyutin, Int. J. Mod. Phys. **A29**, 1450166 (2014).
  - <sup>72</sup> P. Y. Moshin and A. A. Reshetnyak, Inter. J. Mod. Phys. A **30**, 1550021 (2015).
  - <sup>73</sup> H. Matsuda and R. Kubo, Prog. Theor. Phys. **63**, 275 (1980).
  - <sup>74</sup> B. L. Voronov, P. M. Lavrov, and I. V. Tyutin, Sov. J. Nucl. Phys. **36**, 292 (1982).
  - <sup>75</sup> F. Lenz, H. W. L. Naus, K. Ohta, and M. Thies, Annals Phys. **233**, 17 (1994).

- <sup>76</sup> T. Kashiwa and N. Tanimura, *Fortschr. Phys.* **45**, 381 (1997).
- <sup>77</sup> D. Grigore, *Ann. Phys. (Leipzig)* **10**, 439 (2001).
- <sup>78</sup> A. K. Das, R. R. Francisco, and J. Frenkel, *Phys. Rev. D* **88**, 085012 (2013).
- <sup>79</sup> K. Haller and L. F. Landovitz, *Phys. Rev.* **171**, 1749 (1968).
- <sup>80</sup> K. Haller and L. F. Landovitz, *Phys. Rev. D* **2**, 1498 (1970).
- <sup>81</sup> K. Haller, *Acta Phys. Aust.* **42**, 163 (1975).
- <sup>82</sup> K. Haller and L. F. Landovitz, *Phys. Rev. Lett.* **22**, 245 (1969).
- <sup>83</sup> K. Ohtaka and Y. Tanabe, *Rev. Mod. Phys.* **62**, 929 (1990).
- <sup>84</sup> E. Klevak, J. J. Kas, and J. J. Rehr, *Phys. Rev. B* **89**, 085123 (2014).
- <sup>85</sup> I. P. Grant, *Relativistic Quantum Theory of Atoms and Molecules* (Springer, New York, 2007).
- <sup>86</sup> W. Greiner and J. Reinhardt, *Field Quantization* (Springer, Berlin, 1996).
- <sup>87</sup> 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.
- <sup>88</sup> M. J. Esteban, M. Lewin, and E. Séré, *Bull. Amer. Math. Soc.* **45**, 535 (2008).
- <sup>89</sup> W. Kutzelnigg, *Chem. Phys.* **395**, 16 (2012).
- <sup>90</sup> C. Brouder, G. Panati, and G. Stoltz, *Ann. Henri Poincaré* **11**, 1285 (2010).
- <sup>91</sup> W. Heitler, *The Quantum Theory of Radiation* (Dover, New York, 1984), 3rd ed.
- <sup>92</sup> C. Brouder, G. Panati, and G. Stoltz, *Phys. Rev. Lett.* **103**, 230401 (2009).
- <sup>93</sup> B. A. Veklenko, *Sov. Phys. J.* **30**, 555 (1987).
- <sup>94</sup> S. Hassing and E. N. Svendsen, *J. Raman Spectr.* **35**, 87 (2004).
- <sup>95</sup> P. W. Milonni, R. Loudon, P. R. Berman, and S. M. Barnett, *Phys. Rev. A* **77**, 043835 (2008).
- <sup>96</sup> S. Mukamel, *Phys. Rev. A* **76**, 021803 (2007).
- <sup>97</sup> S. T. Epstein, *Phys. Rev.* **98**, 196 (1955).
- <sup>98</sup> B. Hatfield, *Quantum Field Theory of Point Particles and Strings* (Addison-Wesley, Reading, 1992).
- <sup>99</sup> P. Strange, *Relativistic Quantum Mechanics* (Cambridge University Press, Cambridge, 1998).
- <sup>100</sup> D. Baye, *Phys. Rev. C* **86**, 034306 (2012).
- <sup>101</sup> L. L. Foldy and S. A. Wouthuysen, *Phys. Rev.* **78**, 29 (1950).
- <sup>102</sup> W. Greiner, *Relativistic Quantum Mechanics* (Springer, Berlin, 2000), 3rd ed.
- <sup>103</sup> A. J. Silenko, *Phys. Rev. A* **93**, 022108 (2016).
- <sup>104</sup> Z. V. Chraplyvy, *Phys. Rev.* **91**, 388 (1953).
- <sup>105</sup> Z. V. Chraplyvy, *Phys. Rev.* **92**, 1310 (1953).
- <sup>106</sup> E. Eriksen, *Phys. Rev.* **111**, 1011 (1958).
- <sup>107</sup> M. Moshinsky and A. Nikitin, *Rev. Mex. Fis. Suppl.* **2** **50**, 66 (2004).
- <sup>108</sup> W. Weiglhofer, *Am. J. Phys.* **57**, 455 (1989).
- <sup>109</sup> H. A. Bethe and E. E. Salpeter, *Quantum Mechanics of One- and Two-Electron Atoms* (Plenum, New York, 1977).
- <sup>110</sup> J. D. Morrison and R. Moss, *Molec. Phys.* **41**, 491 (1980).
- <sup>111</sup> A. L. Fetter and J. D. Walecka, *Quantum Theory of Many-Particle Systems* (McGraw-Hill, Boston, 1971).
- <sup>112</sup> E. K. U. Gross, E. Runge, and O. Heinonen, *Many-Particle Theory* (Adam Hilger, Bristol, 1991).
- <sup>113</sup> C. Cohen-Tannoudji, B. Diu, and F. Laloë, *Mécanique Quantique* (Hermann, Paris, 1973).
- <sup>114</sup> J. D. Walecka, *Advanced Modern Physics. Theoretical Foundations* (World Scientific, Singapore, 2010).
- <sup>115</sup> C. Brouder, M. Alouani, and K. H. Bennemann, *Phys. Rev. B* **54**, 7334 (1996).
- <sup>116</sup> 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.
- <sup>117</sup> J. Als-Nielsen and D. McMorrow, *Elements of Modern X-Ray Physics* (Wiley, New York, 2000).
- <sup>118</sup> M. Blume, *J. Appl. Phys.* **51**, 3615 (1985).
- <sup>119</sup> M. M. Nieto, *Phys. Rev. Lett.* **38**, 1042 (1977).
- <sup>120</sup> M. Blume, SSRL Report **83/02**, 126 (1983).
- <sup>121</sup> 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.
- <sup>122</sup> L. L. Foldy, *Phys. Rev.* **87**, 688 (1952).
- <sup>123</sup> C. Itzykson and J.-B. Zuber, *Quantum Field Theory* (McGraw-Hill, New York, 1980).
- <sup>124</sup> J. D. Bjorken and S. D. Drell, *Relativistic Quantum Fields* (McGraw-Hill, New York, 1965).
- <sup>125</sup> W. E. Caswell and G. P. Lepage, *Phys. Lett.* **167B**, 437 (1986).
- <sup>126</sup> G. Paz, *Mod. Phys. Lett. A* **30**, 1550128 (2015).
- <sup>127</sup> M. Takahashi and N. Hiraoka, *Phys. Rev. B* **92**, 094441 (2015).
- <sup>128</sup> Y. Joly, S. DiMatteo, and O. Bunau, *Eur. Phys. J. Special Topics* **208**, 21 (2012).
- <sup>129</sup> S. D. Matteo (2016), private communication.
- <sup>130</sup> R. Loudon, *The Quantum Theory of Light* (Clarendon Press, Oxford, 2000), 3rd ed.
- <sup>131</sup> L. M. Brown, *Stud. Hist. Phil. Mod. Phys.* **33**, 211 (2002).
- <sup>132</sup> J. J. Sakurai, *Advanced Quantum Mechanics* (Addison-Wesley, Reading, 1967).
- <sup>133</sup> H. Baum, *Spin-Strukturen und Dirac-Operatoren über pseudoriemannschen Mannigfaltigkeiten* (Teubner, Leipzig, 1981).
- <sup>134</sup> C. Brouder, N. Bizi, and F. Besnard, arXiv:1504.03890.
- <sup>135</sup> E. Eriksen and M. Kolsrud, *Nuovo Cimento (Suppl.)* **18**, 1 (1960).
- <sup>136</sup> E. Prodan and H. Schulz-Baldes, *Bulk and Boundary Invariants for Complex Topological Insulators* (Springer, Berlin, 2016).
- <sup>137</sup> M. Pourahmadi, *Foundations of Time Series Analysis and Prediction Theory* (Wiley, New York, 2001).
- <sup>138</sup> S. Lang, *Linear Algebra*, Undergraduate Graduate Texts in Mathematics (Springer-Verlag, Berlin, 1987), 3rd ed.
- <sup>139</sup> L. J. Wallen, *Michigan J. Math.* **16**, 153 (1969).
- <sup>140</sup> A. A. Joye, in *XVIIth International Congress on Mathematical Physics*, edited by A. Jensen (World Scientific, Singapore, 2014), pp. 486–94.
- <sup>141</sup> F. A. Berezin and M. A. Shubin, *The Schrödinger Equation* (Kluwer, Dordrecht, 1991).