Spontaneous emission and atomic line shift in causal perturbation theory

Karl-Peter Marzlin\textsuperscript{1, 2} and Bryce Fitzgerald\textsuperscript{1}

\textsuperscript{1}Department of Physics, St. Francis Xavier University, Antigonish, Nova Scotia, B2G 2W5, Canada
\textsuperscript{2}Department of Physics and Atmospheric Science, Dalhousie University, Halifax, Nova Scotia B3H 4R2, Canada

We derive spontaneous emission rate and line shift for two-level atoms coupled to the radiation field using causal perturbation theory. In this approach, employing the theory of distribution splitting prevents the occurrence of divergent integrals. Our method confirms the result for atomic decay rate but suggests that the cutoff frequency for the atomic line shift is determined by the atomic mass, rather than Bohr radius or electron mass.

I. INTRODUCTION

Quantum optics, the theory of the interaction between atoms and photons, has been developed since the early days of the laser and has led to many spectacular discoveries. It is an effective theory in which atoms are reduced to point objects with just a few internal energy eigenstates. Photons, described using the quantized electromagnetic field in Coulomb gauge, trigger transitions between these states by inducing an electric dipole moment. Hence, atoms are essentially modelled as a point dipole for each transition between internal states. When many-body effects are studied, each atomic energy level is represented by a non-relativistic quantum field \cite{1, 2}, so that quantum optics corresponds to a non-relativistic quantum field theory. As it happens often in quantum field theories, it is plagued by divergent integrals and needs to be regularized.

Like other effective theories, quantum optics possesses a natural smallest length scale: the size of an atom. This length scale is often used as a cutoff to regularize the theory, but this procedure is not universally applied. Many researchers often simply ignore diverging terms, stating that the line shift has been taken into account in the definition of resonance frequencies. Others apply methods that are similar to mass renormalization in high-energy physics. One may say that there is no general agreement on which method should be used.

Despite the fact that different methods lead to different predictions, this situation has not led to major problems. The reason is that very often one is only interested in the decay rate of isolated atoms, which does not depend on the regularization procedure. In dilute gases, the variation of the Lamb shift with the thermal properties of an atomic gas can also be safely ignored. However, this non-rigorous approach may fail in situations where the variation of line shifts with experimental parameters, such as temperature, density, or distance to a dielectric material, becomes more pronounced. Examples for such situations include spontaneous decay in photonic crystals \cite{3, 4} and absorbing dielectrics \cite{5–7}, corrections to the Lorentz-Lorenz formula in dense atomic gases \cite{8}, and plasmonics \cite{9}. It is therefore of importance to find a more systematic approach.

In this paper, we propose causal perturbation theory (CPT) as a tool to make finite predictions for effective theories such as quantum optics. We illustrate this approach by deriving spontaneous emission rate and line shift for two-level atoms coupled to the radiation field. After a short introduction to CPT in Sec. II, we describe in Sec. III our strategy to derive decay rate and line shift. Sec. IV contains the derivation of the main result. The physical implications of our results are discussed in Sec. V.

II. CAUSAL PERTURBATION THEORY

The Wightman axioms \cite{10} specify that field operators in quantum field theories correspond to operator-valued distributions. For relativistic quantum field theories, Epstein and Glaser \cite{11} have pointed out that ultraviolet divergencies appear because of the improper splitting of operator-valued distributions into retarded and advanced part. They demonstrated that perturbation theory will remain finite if one uses the proper theory of distribution splitting \cite{12} and causality. The details of how CPT is used in relativistic quantum electrodynamics have been explained in Ref. \cite{13}, for instance. We summarize the main aspects to explain the advantages of CPT for effective theories.

Time evolution in quantum field theories can be described using a perturbation expansion of the S-matrix,

\[ S = 1 + \sum_{n=1}^{\infty} \frac{1}{n!} \hat{T}_n \]

where \( \hat{T}_n = \int d^4 x_1 \cdots d^4 x_n T_n(x_1, \ldots, x_n) g(x_1) \cdots g(x_n) \),

with \( g(x) \) a test function that switches the interaction on and off. We will denote four vectors by italic letters \( x \) and three-dimensional vectors by \( \vec{x} \). The components of these vectors are specified using greek and latin indices, respectively. The first-order term takes the form

\[ T_1(x) = \frac{-i}{\hbar c} H_{\text{int}}(x), \]

where the interaction Hamiltonian \( H_{\text{int}}(x) \) is typically a product of field operators. In relativistic theories it is
often more convenient to express the first-order scattering amplitude in terms of the Lagrangian. However, the starting point of non-relativistic effective theories is usually the system Hamiltonian, so that we have adopted this approach here.

Since \( g(x) \) is a test function, integral \( \int d^4x T_1(x)g(x) \) is well defined. All higher order operators can be constructed recursively. For our purpose, it suffices to consider the second-order term

\[
T_2(x, y) = \begin{cases} 
A_2(x, y) = -T_1(x)T_1(y) & \text{for } x^0 > y^0 \\
R_2(x, y) = -T_1(y)T_1(x) & \text{for } x^0 < y^0
\end{cases}
\]

The usual procedure to separate \( T_2(x, y) \) into its retarded and advanced part would be to use step functions,

\[
T_s(x, y) = -\theta(x^0 - y^0)T_1(x)T_1(y) - \theta(y^0 - x^0)T_1(y)T_1(x).
\]

However, step functions are not test functions, so that this operation is not well defined. Even though step functions may appear benign, they are the cause for ultraviolet divergences. To illustrate this problem we consider the well-known one-dimensional distribution \( d(t) = (t \pm i\epsilon)^{-1} \), with \( \epsilon \to 0 \). The Sokhotski-Plemelj theorem implies that

\[
\int_{-\infty}^{\infty} dt \frac{g(t)}{t} = \mathcal{P} \int_{-\infty}^{\infty} dt \frac{g(t)}{t} = \mp i\pi g(0),
\]

where \( \mathcal{P} \) denotes the principal value. If we instead consider the “retarded” distribution \( d_R(t) = \theta(t)d(t) \), we obtain

\[
\int_{-\infty}^{\infty} dt d_R(t)g(t) = \lim_{\epsilon \to 0} \int_{\epsilon}^{\infty} dt \frac{g(t)}{t} = \mp i\pi g(0).
\]

This integral is only well defined for test functions which fulfill \( g(0) = 0 \). In general, multiplying \( d(t) \) with a step function will therefore lead to divergent results.

The key difference between CPT and ordinary perturbation theory is that Eq. (5) is avoided. Instead, retarded and advanced part of a distribution are defined using distribution splitting. Step function \( \theta(t) \) is replaced by a test function \( \theta_L(t) \), which depends on a parameter \( L \) and converges toward \( \theta(t) \) for \( L \to 0 \). For finite \( L \), the product \( \theta_L(t)d(t) \) is then a well-defined distribution. To avoid singularities in the limit \( L \to 0 \), \( g(t) \) is projected on the subspace of test functions for which the limit is well-defined. This can be achieved by a kind of projection operator \( P \). In the example above, \( P \) ensures that \( g(0) = 0 \). Distribution \( \theta_L(t)d(t)P \) is then well-defined even in the limit \( L \to 0 \).

It is important to note that \( P \) is generally not unique. Different choices for \( P \) can be characterized by a finite set of parameters, which play a similar role as renormalization parameters in conventional perturbation theory. In fact, it has been shown that CPT is equivalent to the theory of renormalization for relativistic field theories [13, 14]. One may say that CPT provides an alternative approach to perturbation theory, which is more rigorous with regard to distribution theory.

Since CPT is equivalent to renormalization theory and requires a deeper understanding of distribution theory, it has only been used by comparably few researchers to describe relativistic theories such as QED [15, 16] and self-interacting quantum fields [17–19]. It has also been applied to describe interacting quantum fields in curved space-time [20]. In non-relativistic models, distribution splitting has been used to describe singular potentials in the Schrödinger equation [21].

However, for effective theories CPT may help to clarify situations in which different methods produce competing results. Since the use of perturbation theory with retarded and advanced Green’s function is very common, the clear identification of the origin of singular terms in CPT can shed light on which methods are best suitable. The main limitation of CPT is that the underlying theory has to be causal. For effective theories with relativistic fields, or for solid state systems with finite propagation speed such as acoustic phonons, this is already the case. Other effective theories, including quantum optics, have to be modified to use CPT. In the next section we demonstrate this procedure at the example of spontaneous emission by a single two-level atom.

**III. CAUSAL PERTURBATION THEORY FOR SPONTANEOUS EMISSION**

A cold gas of atoms interacting with light is well described by representing the atoms by a set of non-relativistic field operators

\[
\Psi^{(nr)}_n(x) = (2\pi)^{-3/2} \int d^3k \hat{a}^{(n)}_k e^{i\vec{k}\cdot\vec{x}},
\]

which annihilate an atom with internal (electronic) energy \( E_n \) at position \( \vec{x} \) [1, 2]. Depending on the isotope, atoms either correspond to Bosons or Fermions. In this paper, we will consider bosonic atoms with two internal energy states, ground state and excited state \( (n = g, e) \).

CPT requires Einstein causality, so that we have to modify the above model slightly by replacing non-relativistic field operators by relativistic fields. In quantum optical models, the processes of exciting an atom, \( \hat{H}_{exc} \sim \int d^3x \Psi^{(nr)}_{e}\Psi^{(nr)*}_g(x) \), and de-exciting the atom, \( \hat{H}_{exc} = \hat{H}_{exc}^\dagger \), are usually described by two separate terms in the Hamiltonian. Since a real relativistic scalar field contains both annihilation and creation operators for particles, it cannot be used to accomplish such a construction. For this reason, we represent atoms with internal energy \( E_n \) by a complex
The Klein-Gordon field

\[ \hat{\psi}_n(x) = (2\pi)^{-3/2} \int \frac{d^3k}{\sqrt{\lambda_n|w_k|}} \left( a_k^{(n)} e^{-i\vec{k} \cdot \vec{x}} + b_k^{(n)*} e^{i\vec{k} \cdot \vec{x}} \right) , \]

(10)

where \( k \cdot x = k_\mu x^\mu \) with \( g_{\mu\nu} = \text{diag}(1,-1,-1,-1) \) and \( k_\mu = (w_k, -\vec{k}) \), where \( w_k = \sqrt{k^2 + \lambda_n^2} \). The factor \( \lambda_n = \hbar/(m_\alpha c) \) corresponds to the Compton wavelength of the atom divided by 2\( \pi \). The mass of an atom with internal energy \( E_n \) is given by \( m_n = m_0 + E_n/c^2 \). In our two-level model, the mass difference between ground and excited state is related to the atomic resonance frequency \( \omega_{eg} \) via \((m_e - m_g)c^2 = \hbar\omega_{eg} \). This is in line with the fact that the rest energy of a composite particle includes the binding energy.

Operators \( a_k^{(n)}, b_k^{(n)*} \) describe the annihilation of atoms and anti-atoms with internal energy \( E_n \) and center-of-mass momentum \( \hbar\vec{k} \). They obey the usual harmonic oscillator commutation relations \([a_k^{(n)}, a_{k'}^{(n)*}] = \delta(\vec{k} - \vec{k'})\delta_{nn'}\). We remark that for cold atoms we have \( \lambda_n w_k \approx 1 \), so that the particle-part (\( \propto a_k^{(n)} \)) of Eq. (10) reduces to the conventional field operator \( \hat{\psi}_n^{(\text{ab})}(\vec{x}) \) used in quantum optics.

The radiation field is described in Coulomb gauge, so that the electric field operator contains only contributions from transverse photons,

\[ \vec{E}(x) = \vec{E}^{(+)}(x) + \vec{E}^{(-)}(x) \]

(11)
\[ \vec{E}^{(+)}(x) = -i \int d^3k \sum_{\sigma=1}^2 \frac{\hbar\epsilon_{k\sigma}}{2(2\pi)^3\epsilon_0} a_{k\sigma}^{(-)} \vec{e}_{k\sigma} e^{-i\vec{k} \cdot \vec{x}} , \]

(12)

where \( \vec{E}^{(-)}(x) = \vec{E}^{(+)*}(x) \) as well as \( k_0 = |\vec{k}| \) and \([a_{k\sigma}^{(-)}, a_{k'\sigma'}^{(+)*}] = \delta(\vec{k} - \vec{k'})\delta_{\sigma\sigma'} \). The operator \( a_{k\sigma}^{(-)} \) annihilates a photon with momentum \( \hbar\vec{k} \) and polarization vector \( \vec{e}_{k\sigma} \). Coulomb gauge is almost universally used to describe the physics of atoms and light. It is not covariant, but it does not break causality since the commutator between transverse electric fields has support on the light cone.

We describe the interaction between atoms and light using electric dipole coupling in rotating-wave approximation [22],

\[ \hat{H}_\text{int}(x^0) = \int d^3x \hat{H}_\text{int}(x) \]

(13)
\[ \hat{H}_\text{int}(x) = -\vec{\hat{E}}(x) \cdot \left( \vec{d}_{eg} \hat{\Psi}_e^\dagger(x) \hat{\Psi}_e(x) + \text{h.c.} \right) , \]

(14)

where \( x = (x^0, \vec{x}) \). In this expression, \( \vec{d}_{eg} \) denotes matrix element \( \langle e|d|g \rangle \) of the atomic dipole moment operator \( \vec{d} \). In our model, it is a parameter that is fixed using experimental observations.

To describe spontaneous emission, we consider as initial state a single excited atom in a vacuum. Since \( \hat{H}_\text{int} \) can only excite or de-excite an atom, \( \hat{T}_2 \) describes a self-energy term as depicted in Fig. 1, in which an excited atom emits and re-absors a single photon. If we would use Eq. (5) to split \( T_2(x, y) \) of Eq. (4) in retarded and advanced part, we would have to to evaluate all orders of perturbation theory, which is not feasible. However, Knight and Allen [23] have shown that the Weisskopf-Wigner theory [24, 25] of spontaneous emission is equivalent to the ladder approximation in quantum field theory, which is depicted in Fig. 2. It is therefore sufficient to find \( T_2 \) in order to describe spontaneous emission.
IV. ATOMIC DECAY RATE AND LINE SHIFT

In CPT, it is essential that distributions are causal in the sense that their support lies on the light cone or is time-like. To verify that the support of an operator-valued distribution is causal, one expresses all operator products through normally ordered products by using

\[ E_i(x)E_j(y) =: E_i(x)E_j(y) : = -iD^{(+)}_{ij}(x - y), \quad i, j = 1, 2, 3 \]  

(18)

\[ \hat{\Psi}_i(x)\hat{\Psi}_j^\dagger(y) =: \hat{\Psi}_i(x)\hat{\Psi}_j^\dagger(y) : = -iD^{(+)}_{ij}(x - y), \quad i, j = e, g \]  

(19)

\[ \hat{\Psi}_i^\dagger(x)\hat{\Psi}_j(y) =: \hat{\Psi}_i^\dagger(x)\hat{\Psi}_j(y) : = -iD^{(+)}_{ij}(x - y), \quad i, j = e, g \]  

(20)

where

\[ D^{(+)}_{ij}(x) = \int \frac{d^3k}{(2\pi)^3} \frac{\hbar c k_0}{\omega} \left( \delta_{ij} - \frac{k_i k_j}{k^2} \right) e^{-ik\cdot x} \]  

(21)

\[ D^{(+)}_{ij}(x) = \frac{2i}{\pi^2} \int d^4k \delta(k \cdot k - \lambda_i^{-1} \theta(k_0)) e^{-ik\cdot x}. \]  

(22)

Applying this to \( D_2(x, y) \) of Eq. (17) yields

\[ D_2(x, y) = [ : \hat{\Psi}_i(x)\hat{\Psi}_e(y) : + : \hat{\Psi}_i^\dagger(y)\hat{\Psi}_e(x) : ]D_2(x - y) + \text{rest} \]  

(23)

\[ D_2(x) = \frac{d_{eg, d_{eg, i}}}{\hbar^2 c^2} (D^{(+)}_g)(x) D_{ij}(x) - D_g(x) D^{(-)}_{ij}(x), \]  

(24)

with \( D^{(-)}(x) = -D^{(+)}(-x) \). Distributions of the form \( D = D^{(+)} + D^{(-)} \) have causal support, so that \( D_2(x) \) is causal as well. The terms labelled “rest” either contain vacuum diagrams, which we may ignore, or are proportional to normally ordered products of ground state atoms or radiation field. Since our system initially does not contain photons or ground state atoms, these terms annihilate the initial state.

To split \( D_2(x) \) into retarded and advanced parts, we have to determine its singular order. This is best done in momentum space. We will denote the Fourier transformation of a function \( f(x) \) by \( \tilde{f}(p) \) [26]. We find

\[ \tilde{D}_2(p) = \frac{i(p \cdot p x_g^2 - 1)^3}{120 \epsilon_0 \hbar c (2\pi p \cdot p)^3 x_g^3} \Psi_\theta(p \cdot p x_g^2 - 1) \times \left( |\tilde{d}_{eg}^2|^2 (2p_0^2 - p \cdot p) - 2|p \cdot \tilde{d}_{eg}^2|^2 \right). \]  

(25)

For \( p \to \infty \), this distribution scales as \( p^\omega \) with singular order \( \omega = 2 \).

The general method to find the retarded part \( r(x) \) of a distribution \( d(x) \) includes two steps. First, \( d(x) \) is multiplied with a smooth smeared-out step function \( \theta_L(x \cdot v) \), which is monotonous and takes the value 0 when the argument is negative and 1 if it is larger than \( L \). \( v \) is an arbitrary time-like four vector. For fixed \( L \), the product \( \theta_L(x \cdot v)d(x) \) is a well-defined distribution, but it will diverge like \( L^{-4+\omega} \) for \( L \to 0 \). Despite of this, the linear functional

\[ \int d^4x \theta_L(x \cdot v)d(x)f(x) \]  

(26)

is well defined in the limit as long as all derivatives of test function \( f(x) \) up to order \( \omega \) vanish at \( x = 0 \). To warrant that this is the case for arbitrary test functions, \( \theta_L(x \cdot v)d(x) \) is modified so that contributions of these derivatives are removed. In momentum space, this is accomplished by evaluating the following integral (proposition 3.4 of Ref. [13]),

\[ \hat{r}(p_0) = \frac{i}{2\pi^{1/2}} \int dp_0 \frac{\tilde{d}(k_0)}{(k_0 - i0)^{\omega+1}}(p_0 - k_0 + i0). \]  

(27)

In this expression, the temporal axis for the \( k \)-integration has been rotated by an orthogonal transformation so that it is parallel to (time-like) four vector \( p_\mu \).

It will be explained below that for a very slow atom we can restrict our considerations to the case \( p = 0 \). To simplify the notation, we set \( p_0 = u x_g^{-1} \) with real dimensionless parameter \( u \). We then find the for retarded part \( \hat{R}_2(u) \) of \( D_2(p_0) \) the expression

\[ \hat{R}_2(u) = \frac{|d_{eg}^2|^2}{6(2\pi)^4 \hbar c x_g^3} \left\{ \frac{u^2 - 1}{2u^2} - \frac{5}{4} + \frac{11}{12} u^2 \right\}. \]  

(28)

To achieve our goal of deriving \( T_2(x, y) \), we need to subtract \( \hat{r}'(x, y) \) from \( R_2(x, y) \). Using the same methods as for \( D_2(x, y) \) we obtain

\[ R'_2(x, y) = [ : \hat{\Psi}_e^\dagger(x)\hat{\Psi}_e(y) : + : \hat{\Psi}_e^\dagger(y)\hat{\Psi}_e(x) : ]R_2(x - y) + \text{rest} \]  

(29)

\[ \hat{R}'_2(p) = \frac{(p \cdot p x_g^2 - 1)^3}{120 \epsilon_0 \hbar c (2\pi p \cdot p)^3 x_g^3} \Psi_\theta(p \cdot p x_g^2 - 1) \times \left( |d_{eg}^2|^2 (2p_0^2 - p \cdot p) - 2|p \cdot \tilde{d}_{eg}^2|^2 \right). \]  

(30)

The Fourier transform is given by

\[ \hat{R}'_2(p) \]  

(31)

We thus arrive at

\[ T_2(x, y) = [ : \hat{\Psi}_e^\dagger(x)\hat{\Psi}_e(y) : + : \hat{\Psi}_e^\dagger(y)\hat{\Psi}_e(x) : ]T_2(x - y) + \text{rest} \]  

(32)

\[ \tilde{T}_2(x) = R_2(x) - R'_2(x), \]  

(33)

with the Fourier transform of the right-hand side given by Eqs. (28) and (31). However, in appendix A we show that atoms are only affected by the combination

\[ \tilde{T}_2^{(s)}(p) = \frac{1}{2} (\tilde{T}_2(p) + \tilde{T}_2(-p)). \]  

(34)
For a resting atom with \( p_\mu = (u\hat{x}^{-1}, 0) \), this is given by

\[
\hat{T}_2(s) = \frac{|d_{eg}|^2}{12(2\pi)^4 \epsilon_0 \hbar \lambda_y^3} \left[ \frac{(u^2 - 1)^3}{2u^4} \left[ 2\pi i \theta(u^2 - 1) - \log((u^2 - 1)^2) \right] + \frac{1}{u^2} - \frac{5}{2} + \frac{11}{6} u^2 + C_0 + C_1 u + C_2 u^2 \right].
\] (35)

Here we have added a polynomial \( C_0 + C_1 u + C_2 u^2 \) because distribution splitting is not unique for singular orders \( \omega \geq 0 \). Different approaches to splitting a distribution may then differ by a polynomial in \( p \) of order \( \omega \) [13]. The coefficients \( C_i \), \( i = 0, 1, 2 \) play a similar role as renormalization parameters in the usual theory of renormalization. They are fixed by physical, rather than mathematical, considerations.

With this result, we are in the position to derive spontaneous emission rate and line shift for the model under consideration. It is shown in appendix A that operator \( \hat{T}_2 \) acts on a slow excited atom with center-of-mass wave-function \( \phi(\bar{x}) \) like \( \hat{T}_2(\phi) = Z(\phi) \), with complex factor

\[
Z \approx 2(2\pi)^2 c t_s \hat{T}_2(s)(u),
\] (36)

where \( u = \frac{\lambda_y}{\lambda} \) and \( t_s \) is the duration of the interaction between atom and radiation. Since \( m_e = m_e + \omega_{eg}/c^2 \), we have \( \lambda_y^{-1} = \lambda_y^{-1} + \omega_{eg}/c \). We therefore can set \( u = 1 + \delta u \), with positive \( \delta u = \omega_{eg}/(m_e c^2) \ll 1 \). The spontaneous emission rate \( \gamma = \operatorname{Im}(Z/t_s) \) is given by

\[
\gamma = \frac{\delta u^3 (2 + \delta u)^3 |d_{eg}|^2}{24\pi(1 + \delta u)^5 \epsilon_0 \hbar \lambda_y^3}.
\] (37)

To leading order in \( \delta u \) we find

\[
\gamma = \frac{|d_{eg}|^2 \omega_{eg}^3}{3\pi \epsilon_0 \hbar c^2}.
\] (38)

This result agrees with the standard result for the decay rate of a two-level atom [27] and does not depend on the distribution splitting scheme.

The atomic line shift \( \Delta = \operatorname{Re}(Z/t_s) \) is given by

\[
\Delta = \frac{|d_{eg}|^2}{144\pi^2 \epsilon_0 \hbar \lambda_y^3} \left[ -48 \delta u^3 \log(2\delta u) - 3(2C_0 + 15) \delta u^3 + 2 + 6(C_0 + C_1 + C_2) + (8 - 6C_0 + 6C_2) \delta u + 3(2C_0 + 7) \delta u^2 \right] + O(\delta u^4).
\] (39)

At this point, we have to specify the normalization parameters \( C_i \). For optical transitions, parameter \( \delta u \) is extremely small \((10^{-8} - 10^{-9})\), so that lower powers of \( \delta u \) will give larger contributions. We compare our result to a more precise expression for the Lamb shift for the 1s–2p transition in Hydrogen [28],

\[
\Delta_L \approx \frac{m_{e\text{lec}} c^2 \alpha^5}{\pi \hbar} \left( -25.25 + \frac{4}{3} \ln(\alpha^{-2}) \right),
\] (40)

where \( \alpha \approx 1/137 \) is the fine structure constant and \( m_{e\text{lec}} \) the mass of the electron. It is not hard to see that, in Eq. (39), terms of order \( \delta u^2 \) or lower will produce results that are many orders of magnitude larger than the observed line shift. Our distribution splitting scheme is therefore fixed by the condition that these terms disappear, which is achieved for \( C_0 = -\frac{7}{4}, C_1 = -\frac{29}{3}, \) and \( C_2 = 8 \). The final result for the line shift in a two-level model is then given by

\[
\Delta = -\frac{\gamma}{2\pi} \left[ 1 + 2 \log \left( \frac{2\hbar \omega_{eg}}{m_e c^2} \right) \right].
\] (41)

For a numerical comparison with a more precise prediction we use data for the 1s–2p transition in Hydrogen, i.e., \( \hbar \omega_{eg} = 0.75 \times 13.6 \) eV and \( d_{eg} = \sqrt{22'}3^{-5}e a_0 \), with \( a_0 \) the Bohr radius. We find

\[
\frac{\Delta}{\Delta_L} = -0.029 \log(\alpha) - 0.015 \log \left( \frac{m_e}{m_{e\text{lec}}} \right) - 0.0031 \approx 0.55.
\] (42)

The fact that the simple two-level model makes a prediction that is off by one order of magnitude is not surprising. The model only includes the line shift generated by a single transition, instead of a sum over all excited states. Furthermore, models like that of Eq. (9), where excitation and de-excitation are treated separately, can be considered as a consequence of the rotating-wave approximation, which neglects further off-resonant contributions to the line shift.

V. DISCUSSION

In the language of renormalization and regularization, one would identify the atomic Compton angular frequency \( m_e c^2/\hbar \), which appears in the logarithm in Eq. (41), with a cutoff frequency. CPT therefore predicts that in models where atoms are considered as point dipoles, it is the atomic mass that determines the cutoff. However, in most applications of atomic point-dipole models, the line shift is either ignored, or the cutoff is taken to be the Compton frequency of the electron, so that the results resemble the Lamb shift (40). It appears that, for atomic point-dipole models to be consistent with causality, the atomic Compton frequency should be used instead.

Another point of interest about our findings is connected to the choice (14) of the interaction between atoms and radiation. Electric dipole coupling \( \vec{d} \cdot \vec{E} \) is popular in quantum optics, but in relativistic theories minimal coupling \( \vec{p} \cdot \vec{A} \), with \( \vec{p} \) the momentum and \( \vec{A} \) the vector potential) is preferred. However, using minimal coupling would only be possible in Lorentz gauge, which is rarely used to describe non-relativistic atoms. In Coulomb gauge, the vector potential will be a transverse field, i.e.,
it obeys $\nabla \cdot \vec{A} = 0$. It can be shown that the commutator between transverse vector potentials has space-like support [29], so that CPT cannot be used for minimal coupling in Coulomb gauge.

One of the motivations for this work was the question whether separating a vector field operator $\vec{V}(\vec{r})$ into a transverse part $\vec{V}_t$ and a longitudinal part $\vec{V}_l$ would require the methods of distribution splitting. The transverse part is most easily constructed in momentum space by modifying its Fourier coefficients as $\vec{V}_{t,l}(k) = (\delta_{ij} - k_i k_j / k^2) \vec{V}_j(\vec{k})$, similar to Eq. (21). This modification does not change the singular order of a distribution, so that the separation into transverse and longitudinal parts can be performed in a conventional way.

Our last remark concerns the question whether the model discussed in this work is renormalizable, i.e., whether a finite set of normalization parameters is sufficient to describe the S-matrix to all orders in perturbation theory. While the answer to this question is beyond our present treatment, we should emphasize that a model with a cutoff frequency is determined by the atomic mass in contrast to the electron mass or Bohr radius. This difference arises because, in our effective theory, atoms are treated as point dipoles with inner structure, so that it is the mass of this object that is relevant.

The problem we considered serves as an illustration how CPT can be applied to effective models in quantum optics or condensed matter theory. Its rigorous approach may be useful to study models in which other methods produce contradicting results. We are planning to apply CPT to other systems in quantum optics, such as dense atomic gases or atoms in absorbing dielectrics.

**ACKNOWLEDGMENTS**

We are grateful to Mark Walton for discussions about this work, and to the Natural Sciences and Engineering Research Council of Canada (NSERC) for financial support. B. F. thanks Saint Francis Xavier University for a UCR summer research award.

**Appendix A: The action of $\hat{T}_2$ on a slow atom**

$\hat{T}_2$ has the general form

$$\hat{T}_2 = \int d^4x d^4y \, \mathcal{T}_2(x - y)g(x)g(y) \times \left( \hat{\Psi}_e^\dagger(x)\hat{\Psi}_e(y) : + : \hat{\Psi}_e^\dagger(y)\hat{\Psi}_e(x) : \right), \quad (A1)$$

and is applied to a state that describes a single non-relativistic excited atom,

$$\ket{\phi} = \int d^3k \, \tilde{\phi}(\vec{k}) a^{(e)c}_k \ket{0}. \quad (A2)$$

Using Eq. (10) we obtain

$$\hat{\Psi}_e(x)|\phi(\phi) = \phi(x)|\phi$$

$$\hat{\Phi}(x) = (2\pi)^{-2} \int \frac{d^3k}{\sqrt{\lambda_\omega w_k}} e^{-ik\cdot x} \tilde{\phi}(\vec{k}), \quad (A4)$$

where the time component of four vector $k$ is given by $w_k$ (this convention is throughout this appendix). This leads to

$$\hat{T}_2|\phi = \int d^4x d^4y \left( \mathcal{T}_2(x - y) + \mathcal{T}_2(y - x) \right) \times \phi(y)g(y)g(x)|\hat{\Psi}_e^\dagger(x)|\phi$$

$$\times \phi(x)g(x)\hat{\Psi}_e(x)|\phi \quad (A5)$$

In momentum space, this expression takes the form

$$\hat{T}_2|\phi = \int \frac{d^4p}{(2\pi)^2} (\bar{T}_2(p) + \bar{T}_2(-p)) \int d^4y \, e^{ip\cdot y} g(y)\phi(y) \times \int \frac{d^4x}{\sqrt{\lambda_\omega w_k}} e^{-i\vec{k}\cdot\vec{x}} \tilde{\phi}(\vec{k})\tilde{\phi}(p - k) \times \int \frac{d^4k}{\sqrt{\lambda_\omega w_k}} \tilde{\phi}(\vec{k} - p)a^{(e)c}_k \ket{0} \quad (A6)$$

$\hat{T}_2|\phi \approx \int \frac{d^4p}{(2\pi)^2} (\bar{T}_2(p) + \bar{T}_2(-p)) \int d^4y \, e^{ip\cdot y} g(y)\phi(y) \times \int \frac{d^4k}{\sqrt{\lambda_\omega w_k}} \tilde{\phi}(\vec{k} - p)a^{(e)c}_k \ket{0} \quad (A7)$

We now assume that $g(x)$ does not vary much over the spatial support of $\phi(x)$, so that we can set $\phi(x)g(x) \approx \phi(x)g(x^0)$. Using the convolution theorem, this relation implies that, in momentum space,

$$\tilde{\phi}(\vec{k})\tilde{\phi}(p - k) \approx (2\pi)^{3/2} \tilde{\phi}(\vec{k})\delta(\vec{p} - \vec{k})\tilde{\phi}(p_0 - k_0). \quad (A8)$$
Hence,

\[
\hat{T}_2(\phi) \approx \sqrt{2\pi} \int d^4p \left( \hat{T}_2(p) + \hat{T}_2(-p) \right) \frac{\tilde{\phi}(\tilde{p})}{\sqrt{\tilde{x}_p w_p}} \tilde{g}(p_0 - w_p) \\
\times \int \frac{d^3k'}{\sqrt{\tilde{x}_k w_k}} \tilde{g}(k' - p) a^{(c)\dagger}_k |0\rangle \\
\approx (2\pi)^2 \int d^4p \left( \hat{T}_2(p) + \hat{T}_2(-p) \right) \tilde{\phi}(\tilde{p}) \tilde{g}(p_0 - w_p) \\
\times \tilde{g}(w_p - p_0) a^{(c)\dagger}_p |0\rangle . 
\]

(A9)

If the width of wavepacket \( \tilde{\phi}(\tilde{k}) \) in momentum space is so narrow that dispersion can be neglected, we can set \( w_p = \sqrt{\tilde{p}^2 + \tilde{x}_p^2} \approx \tilde{x}_p \). This approximation is well justified for typical atomic gases with temperatures at or below room temperature. We furthermore assume that over the width of \( \tilde{\phi}(\tilde{k}) \) we can neglect the dependence of \( \hat{T}_2(p) \) on the spatial components of \( p \), so that \( \hat{T}_2(p) \approx \hat{T}_2(p_0) \). The action of \( T_2 \) then simplifies to

\[
\hat{T}_2(\phi) \approx (2\pi)^2 \int dp_0 \left( \hat{T}_2(p_0) + \hat{T}_2(-p_0) \right) \\
\times \tilde{g}(p_0 - \tilde{x}_p^{-1}) \tilde{g}(\tilde{x}_p^{-1} - p_0) \int d^3p \tilde{\phi}(\tilde{p}) a^{(c)\dagger}_p |0\rangle \\
= Z|\phi\rangle \\
Z = (2\pi)^2 \int dp_0 \left( \hat{T}_2(p_0) + \hat{T}_2(-p_0) \right) \\
\times \tilde{g}(p_0 - \tilde{x}_p^{-1}) \tilde{g}(\tilde{x}_p^{-1} - p_0). 
\]

(A11)

(A12)

(A13)

If \( \tilde{g}(p_0 - \tilde{x}_p^{-1}) \) is sufficiently narrow, we can set \( \hat{T}_2(p_0) \approx \hat{T}_2(\tilde{x}_p^{-1}) \). With this approximation we obtain

\[
Z \approx (2\pi)^2 (\hat{T}_2(\tilde{x}_p^{-1}) + \hat{T}_2(-\tilde{x}_p^{-1})) \\
\times \int_{-\infty}^{\infty} dp_0 \tilde{g}(p_0 - \tilde{x}_p^{-1}) \tilde{g}(\tilde{x}_p^{-1} - p_0) \\
= (2\pi)^2 (\hat{T}_2(\tilde{x}_p^{-1}) + \hat{T}_2(-\tilde{x}_p^{-1})) \int_{-\infty}^{\infty} dx^0 g^2(x^0). 
\]

(A14)

(A15)

If test function \( g(x^0) \) is close to unity during a time interval of width \( ct_g \) and drops quickly to zero outside the interval, the integral in this expression is approximately equal to \( ct_g \). Using this approximation results in Eq. (36).

[1] W. Zhang and D. F. Walls, Phys. Rev. A 49, 3799 (1994).
[2] G. Lenz, P. Meystre, and E. M. Wright, Phys. Rev. A 50, 1681 (1994).
[3] S. John and T. Quang, Phys. Rev. A 50, 1764 (1994).
[4] K.-P. Marzlin and W. Zhang, Phys. Rev. A 59, 2982 (1999).
[5] M. S. Yeung and T. K. Gustafson, Phys. Rev. A 54, 5227 (1996).
[6] M. Fleischhauer, Phys. Rev. A 60, 2534 (1999).
[7] H. T. Dung, S. Y. Buhmann, L. Knöll, D.-G. Welsch, S. Scheel, and J. Kastel, Phys. Rev. A 68, 043816 (2003).
[8] M. Fleischhauer and S. F. Yelin, Phys. Rev. A 59, 2427 (1999).
[9] J. L. Choquette and K.-P. Marzlin, Phys. Rev. A 85, 063841 (2012).
[10] R. F. Streater and A. S. Wightman, PCT, Spin and Statistics and All That (Princeton University Press, 1964).
[11] H. Epstein and V. Glaser, Ann. Poin. Phys. Theor. A19, 211 (1973).
[12] B. Malgrange, Sémin. Schwartz 4, Exp. No. 23 (1959-1960).
[13] G. Scharf, Finite Quantum Electrodynamics, 3rd ed. (Dover Publications, Mineola, 2014).
[14] G. Scharf, Gauge Field Theories: Spin One and Spin Two (Dover Publications, Mineola, 2016).
[15] J. M. G. Bondia, M. Carbalal, O. Miranda, and O. Rosas-Ortiz, AIP Conf. Proc. 809, 24 (2006).
[16] A. Aste, C. von Arx, and G. Scharf, Prog. Part. Nucl. Phys. 64, 61 (2010).
[17] D. R. Grigore, J. Phys. A 33, 8443 (2000).
[18] G. Pinter, Ann. Physik 10, 333 (2001).
[19] A. Aste and D. Trautmann, Can. J. Phys. 81, 1433 (2003).
[20] K. Frederenhagen and K. Rejzner, J. Math. Phys. 57, 031101 (2016).
[21] M. M. Calçada, J. T. Lunardi, L. A. Manzoni, and W. Monteiro, Front. Phys. 2, 23 (2014).
[22] M. Frasca, Ann. Phys. 306, 193 (2003).
[23] P. Knight and L. Allen, Phys. Lett. A 38, 99 (1972).
[24] V. Weisskopf and E. Wigner, Z. Phys. 63, 54 (1930).
[25] V. Weisskopf and E. Wigner, Z. Phys. 65, 18 (1930).
[26] Strictly speaking, we are using wavevector \( \tilde{k} \) instead of momentum \( \hbar k \) in all of our derivations.
[27] P. W. Milonni, The quantum vacuum (Academic Press, Boston, 1994).
[28] K. Pachucki, D. Leibfried, M. Weitz, A. Huber, W. König, and T. W. Hänsch, J. Phys. B 29, 177 (1996).
[29] O. Keller, Phys. Rev. A 58, 3407 (1998).

[30] M. Dütsch and K. Fredenhagen, Rev. Math. Phys. 16, 1291 (2004).