Quantum mechanical treatment of light

1 Introduction

As we argued when discussing "Heisenberg's microscope", a consistent description of Nature requries that all of physics be treated in a fashion consistent with quantum mechanics so we must view Maxwell's equations as describing another classical system which must be replaced by a quantum formulation. As one learns from studying classical and then quantum field theory this can be done in an elegant way by mildly extending the canonical Hamiltonian formulation of mult-particle quantum mechanics to apply to a classical Hamiltonian description of a system of classical fields in which the dynamical variables, $\vec{E}(\vec{r},t)$ and $\vec{B}(\vec{r},t)$ are continuous functions of space.

Since you will see this treatment in a course in quantum field theory and the approach is somewhat formal, here we will take an easier and more immediate approach of specializing Maxwell's equations to a finite-volume system in which $\vec{E}(\vec{r},t)$ and $\vec{B}(\vec{r},t)$ are fields confined to a box of volume L^3 obeying periodic boundary conditions. If we plan on using perturbation theory to treat the interaction between light and matter then we naturally start with Maxwell's equations without ρ or $\vec{j}$. Such a system is readily expressed in terms of normal modes, each of which is a simple harmonic oscillator. The only complexity caused by the fact that $\vec{E}(\vec{r},t)$ and $\vec{B}(\vec{r},t)$ are continuous functions of $\vec{r}$ is that there are an infinite number of these modes carrying wave numbers that can be arbitrarily large. Since quantizing each of these simple harmonic oscillators is easy, this is the approach taken in this note. We are basically presenting the standard, organized treatment of these modes and their interpretation as "photons" carrying specific values of momentum and energy.

2 Maxwell's equations

We begin with the usual four esu Maxwell equations:

$$\vec{\nabla}\cdot\vec{E} = 4\pi\rho \qquad \vec{\nabla}\times\vec{B} = \frac{4\pi}{c}\vec{j} + \frac{1}{c}\frac{d\vec{E}}{dt} \tag{1}$$

$$\vec{\nabla}\cdot\vec{B} = 0 \qquad \vec{\nabla}\times\vec{E} = -\frac{1}{c}\frac{d\vec{B}}{dt} \tag{2}$$

We will satisfy the final two equaations in Eq. (2) by introducing the vector and scalar potential $\vec{A}(\vec{r},t)$ and $\phi(\vec{r},t)$ requiring:

$$\vec{B} = \vec{\nabla}\times\vec{A} \qquad -\vec{\nabla}\phi = \vec{E} + \frac{1}{c}\frac{d\vec{A}}{dt}. \tag{3}$$

When these equations are substituted into the two remaining equations in Eq. (1) these equations do not describe a predictive system since an arbitrary gauge transformation of one solution will result in a different solution with different values of

$\vec{A}(\vec{r},t)$ and $\phi(\vec{r},t)$ at a later time even though the initial conditions are the same if we consider gauge transformation which reduce to the identity transform at the time at which the initial conditions are specified. Thus, we must fix the gauge. The gauge can be fixed in many ways and the resulting quantum mechanics can look quite different depending on this choice. However, when done consistently, each system will predict the same observable consequences.

We will use "Coulomb" or "radiation" gauge and require $\vec{\nabla}\cdot\vec{A}=0$. The equations of motion for $\vec{A}(\vec{r},t)$ and $\phi(\vec{r},t)$ resulting from the two equations in Eq. (1) can be obtained by substituting Eq. (3) into Eq. (1):

$$-\vec{\nabla}^2\phi = 4\pi\rho \tag{4}$$

$$\vec{\nabla}\times\left(\vec{\nabla}\times\vec{A}\right) = -(\vec{\nabla}^{\,2})\vec{A} = \frac{4\pi}{c}\vec{j} - \frac{1}{c}\vec{\nabla}\frac{d\phi}{dt} - \frac{1}{c^2}\frac{d^2\vec{A}}{dt^2}. \tag{5}$$

Equation (4) is a simple constraint that determines ϕ directly in terms of ρ:

$$\phi(\vec{r},t) = \int d^3r'\,\frac{1}{|\vec{r}-\vec{r}\,'|}\rho(\vec{r}\,',t). \tag{6}$$

Although this instantaeous, "action-at-a-distance" relation between the separated points $\vec{r}$ and $\vec{r}\,'$ is superficially inconsistent with causality and special relativity, when combined with the other non-covariant parts of this Coulomb gauge treatment, the final result will be consistent with special relativity provided the treatment of the charged matter is consistent with relativity.

3 Normal modes of $\vec{\nabla}^{\,2}$

Equation (5) is the real dynamical equation that we must treat quantum mechanically. The classical normal modes of this system can be obtained by finding what in quantum mechanics would be called the eigenstates of the operator $\vec{\nabla}^{\,2}$. Therefore, we will write $\vec{A}(\vec{r},t)$ as the power series:

$$\vec{A}(\vec{r},t) = \sum_{\vec{k}_{\vec{n}}}\sum_{i=1,2}\hat{\mathcal{E}}_i(\vec{k}_{\vec{n}})q_i(\vec{k}_{\vec{n}})\frac{e^{i\vec{k}_{\vec{n}}\cdot\vec{r}}}{L^{\frac{3}{2}}} \tag{7}$$

where $\vec{n}=(n_1,n_2,n_3)$ is the usual vector of integers and $\vec{k}_{\vec{n}} = 2\pi(n_1,n_2,n_3)/L$. Here the two vectors $\left\{\mathcal{E}_i(\vec{k}_{\vec{n}})\right\}_{i=1,2}$ determine the two possible directions consistent with the Coulomb gauge condition $\vec{\nabla}\cdot\vec{A}=0$ and therefore must obey $\vec{k}_{\vec{n}}\cdot\hat{\mathcal{E}}_i(\vec{k}_{\vec{n}})=0$. To make later steps simpler, we will also require: $\mathcal{E}_i(\vec{k}_{\vec{n}}) = \mathcal{E}_i(-\vec{k}_{\vec{n}})$.

This expansion is somewhat odd to perform in classical mechanics since for convenience we have introduced complex basis functions $\frac{e^{i\vec{k}_{\vec{n}}\cdot\vec{r}}}{L^{\frac{3}{2}}}$ and complex coefficients $q_i(\vec{k}_{\vec{n}})$. Since the classical field $\vec{A}(\vec{r},t)$ is real we must require:

$$\sum_{\vec{k}_{\vec{n}}}\sum_{i=1,2}\hat{\mathcal{E}}_i(\vec{k}_{\vec{n}})q_i(\vec{k}_{\vec{n}})\frac{e^{i\vec{k}_{\vec{n}}\cdot\vec{r}}}{L^{\frac{3}{2}}} = \sum_{\vec{k}_{\vec{n}}}\sum_{i=1,2}\hat{\mathcal{E}}_i(\vec{k}_{\vec{n}})\left[q_i(\vec{k}_{\vec{n}})\right]^*\frac{e^{-i\vec{k}_{\vec{n}}\cdot\vec{r}}}{L^{\frac{3}{2}}} \tag{8}$$

where we have assumed that $\mathcal{E}_i(\vec{k}_{\vec{n}}) = \mathcal{E}_i(\vec{k}_{\vec{n}})^*$, an asumption that needs to be relaxed if we work with circular polarization. Since each of the modes in Eq. (8) are independent, that equation requires:

$$q_i(\vec{k}_{\vec{n}}) = \left[q_i(-\vec{k}_{\vec{n}})\right]^*. \tag{9}$$

Finally we can make this description look like orginary classical mechanics if we wrtie the complex variables $q_i(\vec{k}_{\vec{n}})$ in terms of their real and imaginary parts:

$$q_i(\vec{k}_{\vec{n}}) = q_i^{\mathrm{re}}(\vec{k}_{\vec{n}}) + i q_i^{\mathrm{im}}(\vec{k}_{\vec{n}}). \tag{10}$$

We are now working with the standard classical variables $q_i^{\mathrm{re}}(\vec{k}_{\vec{n}})$ and $q_i^{\mathrm{im}}(\vec{k}_{\vec{n}})$ except for the peculiar redundancy required by the reality condition in Eq. (9):

$$q_i^{\mathrm{re}}(\vec{k}_{\vec{n}}) = q_i^{\mathrm{re}}(-\vec{k}_{\vec{n}}) \qquad q_i^{\mathrm{im}}(\vec{k}_{\vec{n}}) = -q_i^{\mathrm{im}}(-\vec{k}_{\vec{n}}). \tag{11}$$

Thus, we must be careful to recognize that the same real dynamical variable has two names and will naturally appear in a given expression more than once because of these different names: The quantity $q_i^{\mathrm{re}}(\vec{k}_{\vec{n}})$ is an even fucntion of $\vec{k}$ while $q_i^{\mathrm{im}}(\vec{k}_{\vec{n}})$ is an odd function.

We can now use our these new variables $q_i^X(\vec{k}_{\vec{n}})$ where $X = \mathrm{re}$ or $X = \mathrm{im}$ to simplfy Eq. (5): by equating the coefficients of the independent normal modes on the left- and right-hand sides:

$$\vec{k}_{\vec{n}}^2\, q_i^X(\vec{k}_{\vec{n}}) = -\frac{1}{c^2}\frac{d^2 q_i^X(\vec{k}_{\vec{n}})}{dt^2}. \tag{12}$$

Thus, each of our real classical varilables $q_i^X(\vec{k}_{\vec{n}})$ obeys the classical equation of a simple harmonic oscillator with frequency squared $\omega(\vec{k}_{\vec{n}})^2 = \vec{k}_{\vec{n}}^2 c^2$. To treat light quantum mechanically, we need only quantize these simple harmonic oscillators.

4 Quantum mechanical treatment

Since our formulation of quantum mechanics is based on Hamiltonian mechanics, we begin with classical energy for the oscillators obeying the equations of motion given in Eq. (12). This can be easily constructed from the usual classical energy:

$$\begin{aligned}
E &= \frac{1}{8\pi}\int d^3r \left\{\vec{E}^2(\vec{r}) + \vec{B}^2(\vec{r})\right\} && (13)\\
&= \frac{1}{8\pi}\sum_{\vec{k}_{\vec{n}}}\sum_{i=1,2}\left\{\frac{1}{c^2}\dot{q}_i(\vec{k}_{\vec{n}})\dot{q}_i(-\vec{k}_{\vec{n}})\right. && (14)\\
&\qquad\qquad \left. + \left[\left(i\vec{k}_{\vec{n}}\times\mathcal{E}_i(\vec{k}_{\vec{n}})\right)\cdot\left(-i\vec{k}_{\vec{n}}\times\mathcal{E}_i(\vec{k}_{\vec{n}})\right)\right]\left(q_i(\vec{k}_{\vec{n}})q_i(-\vec{k}_{\vec{n}})\right)\right\} &&\\
&= \frac{1}{8\pi}\sum_{\vec{k}_{\vec{n}}}\sum_{i=1,2}\left\{\left|\frac{1}{c}\dot{q}_i(\vec{k}_{\vec{n}})\right|^2 + \vec{k}_{\vec{n}}^2\left|q_i(\vec{k}_{\vec{n}})\right|^2\right\} && (15)\\
&= \frac{1}{8\pi}\sum_{\vec{k}_{\vec{n}}}\sum_{i=1,2}\left\{\frac{1}{c^2}\left(\dot{q}_i^{\mathrm{re}}(\vec{k}_{\vec{n}})^2 + \dot{q}_i^{\mathrm{im}}(\vec{k}_{\vec{n}})^2\right) + \vec{k}_{\vec{n}}^2\left(q_i^{\mathrm{re}}(\vec{k}_{\vec{n}})^2 + q_i^{\mathrm{im}}(\vec{k}_{\vec{n}})^2\right)\right\}. && (16)
\end{aligned}$$

Equation (14) is obtained from Eq. (13) by inserting the sum over normal modes given in Eq. (7) and using the complex variables $q_i(\vec{k}_{\vec{n}})$, Eq. (15) results from using the reality condition in Eq. (9) and finally in Eq. (16) we introduce the real and imaginary parts of $q_i(\vec{k}_{\vec{n}})$ as defined in Eq. (10).

The final equation for the energy, Eq. (16) can be easilyused to determine the Hamiltonian and the canonical momentum $p_i^X(\vec{k}_{\vec{n}})$ conjugate to the variable $q_i^X(\vec{k}_{\vec{n}})$ where X = re or im:

$$p_i^X(\vec{k}_{\vec{n}}) = \widetilde{m}\dot{q}_i^X(\vec{k}_{\vec{n}}) \tag{17}$$

where the constant $\widetilde{m} = \frac{1}{2\pi c^2}$, twice the coefficient of $\dot{q}_i^{\rm re}(\vec{k}_{\vec{n}})^2$ in Eq. (16), now plays the role of a mass. Note, while $2 \times \frac{1}{8\pi}$ may appear to be the coefficient of $\dot{q}_i^X(\vec{k}_{\vec{n}})^2$, that coefficient is actually twice as large since the variable $\dot{q}_i^X(\vec{k}_{\vec{n}})$ appears twice, in both the $\vec{k}_{\vec{n}}$ and $-\vec{k}_{\vec{n}}$ terms. Thus, Eq. (16) can be rewritten as a sum of individual SHO Hamiltonians:

$$H = \sum_{\vec{k}_{\vec{n}}} \sum_{i=1,2} \Big\{ \frac{1}{2\widetilde{m}} \left(p_i^{\rm re}(\vec{k}_{\vec{n}}) + p_i^{\rm im}(\vec{k}_{\vec{n}})^2 \right) + \frac{1}{2}\widetilde{m}\,\omega(\vec{k}_{\vec{n}})^2 \left(q_i^{\rm re}(\vec{k}_{\vec{n}})^2 + q_i^{\rm im}(\vec{k}_{\vec{n}})^2 \right)^2 \Big\}. \tag{18}$$

where $\omega(\vec{k}_{\vec{n}})^2 = \vec{k}_{\vec{n}}^2 c^2$, as expected.

From this point on the quantum mechanics is straight-forward. We need only introduce the ground state $|0\rangle$ and the lowering operators for each of the many normal modes:

$$a_i^X(\vec{k}_{\vec{n}}) = \sqrt{\frac{\widetilde{m}\omega}{2\hbar}} q_i^X(\vec{k}_{\vec{n}}) + \sqrt{\frac{1}{2\widetilde{m}\omega\hbar}} q_i^X(\vec{k}_{\vec{n}}) \tag{19}$$

The appearance of the same variable twice for $\pm\vec{k}_{\vec{n}}$ is awkard as can be seen if we evaluate the commultation relations for the operators defined Eq. (19):

$$\begin{aligned} \left[a_i^{\rm re}(\vec{k}_{\vec{n}}), (a_j^{\rm re}(\vec{k}_{\vec{n}'}))^\dagger \right] &= \delta_{ij} \left\{ \delta_{\vec{n},\vec{n}'} + \delta_{\vec{n},-\vec{n}'} \right\} & (20) \\ \left[a_i^{\rm im}(\vec{k}_{\vec{n}}), (a_j^{\rm im}(\vec{k}_{\vec{n}'}))^\dagger \right] &= \delta_{ij} \left\{ \delta_{\vec{n},\vec{n}'} - \delta_{\vec{n},-\vec{n}'} \right\} & (21) \end{aligned}$$

Fortunately, since complex numbers are routine in quantum mechanics we can combine the two non-hermitian operators $a_i^{\rm re}(\vec{k}_{\vec{n}})$ and $a_i^{\rm re}(\vec{k}_{\vec{n}})$ to remove this redundancy and to define independent lowering operators:

$$a_i(\vec{k}_{\vec{n}}) = \frac{1}{\sqrt{2}} \left\{ a_i^{\rm re}(\vec{k}_{\vec{n}}) + i a_i^{\rm im}(\vec{k}_{\vec{n}}) \right\}. \tag{22}$$

The new operators obey the simple commutation relations:

$$\left[a_i(\vec{k}_{\vec{n}}), (a_j(\vec{k}_{\vec{n}'}))^\dagger \right] = \delta_{ij}\delta_{\vec{n},\vec{n}'}. \tag{23}$$

This commutation relation for the follows from Eqs. (20) and (21) if we recognize that $a_i^{\rm re}(\vec{k}_{\vec{n}})$ and $a_i^{\rm im}(\vec{k}_{\vec{n}})$ and their hermitian conjugates involve independent variables and therefore commute. With the differing signs of the $\pm\delta_{\vec{n},-\vec{n}'}$ terms in Eqs. (20) and (21), this unwanted term cancels and does not appear in Eq. (23). Following this derviation we can then forget about the $q^{\rm re}$ and $q^{\rm im}$ variables and simply work with the single set of lowering operators $a_i(\vec{k}_{\vec{n}})$ defined in Eq. (22) and their hermitian conjugates.

The important equation is the expression for $\vec{A}$ written in terms of the operators $a_i(\vec{k}_{\vec{n}})$ and $a_i(\vec{k}_{\vec{n}})^\dagger$. Starting with the familiar formula expressing the q's in terms of the a's and $a^\dagger$'s:

$$\begin{aligned}
\frac{1}{\sqrt{2}}\left\{q_i^{\mathrm{re}}(\vec{k}_{\vec{n}})+iq_i^{\mathrm{im}}\right\}(\vec{k}_{\vec{n}}) &= \frac{1}{\sqrt{2}}\sqrt{\frac{\hbar}{2\widetilde{m}\omega}}\Big\{\Big(a_i^{\mathrm{re}}(\vec{k}_{\vec{n}})+a_i^{\mathrm{re}}(\vec{k}_{\vec{n}})^\dagger\Big) \\
&\qquad +i\left(a_i^{\mathrm{im}}(\vec{k}_{\vec{n}})+a_i^{\mathrm{im}}(\vec{k}_{\vec{n}})^\dagger\right)\Big\} \qquad (24)\\
&= \frac{1}{\sqrt{2}}\sqrt{\frac{\hbar}{2\widetilde{m}\omega}}\Big\{a_i(\vec{k}_{\vec{n}})+a_i(-\vec{k}_{\vec{n}})^\dagger\Big\} \qquad (25)\\
&= \sqrt{\frac{2\pi\hbar c}{k}}\Big\{a_i(\vec{k}_{\vec{n}})+a_i(-\vec{k}_{\vec{n}})^\dagger\Big\} \qquad (26)
\end{aligned}$$

where we have used the relation $a_i^{\mathrm{im}}(\vec{k}_{\vec{n}})=-a_i^{\mathrm{im}}(-\vec{k}_{\vec{n}})$ to absorb the extra minus sign that results in from the 'i' in the $ia_i^{\mathrm{im}}(\vec{k}_{\vec{n}})$ term in Eq. (22) when we take its hermitian conjugate. Finally, Eq. (26) can be substitued into Eq. (7) and the sign of $\vec{k}$ changed in the second term to obtain:

$$\vec{A}(\vec{r})=\sum_{\vec{k}_{\vec{n}}}\sum_{i=1,2}\hat{\mathcal{E}}_i(\vec{k}_{\vec{n}})\sqrt{\frac{2\pi\hbar c}{Vk}}\Big(a_i(\vec{k}_{\vec{n}})e^{i\vec{k}_{\vec{n}}\cdot\vec{r}}+a_i(\vec{k}_{\vec{n}})^\dagger e^{-i\vec{k}_{\vec{n}}\cdot\vec{r}}\Big) \qquad (27)$$

for $V=L^3$.

The interpretation of the resulting quantum theory can be made directly if (following problem 81) we use the equations dervied above to substitute into the Hamiltonain given in Eq. (16) and the classical expression for the momentum carried by the electromagnetic field to write each in terms of $a_i(\vec{k}_{\vec{n}})$ and $a_i(\vec{k}_{\vec{n}})^\dagger$:

$$\begin{aligned}
H &= \frac{1}{8\pi}\int d^3r\Big\{\vec{E}^2(\vec{r})+\vec{B}^2(\vec{r})\Big\}=\sum_{j,n}\hbar\omega\Big\{a_j(\vec{k}_{\vec{n}})^\dagger a_j(\vec{k}_{\vec{n}})+\frac{1}{2}\Big\} \qquad (28)\\
\vec{P} &= \frac{1}{4\pi c}\int d^3r\ \vec{E}(\vec{r})\times\vec{B}(\vec{r})=\sum_{j,n}\hbar\vec{k}_{\vec{n}}\,a_j(\vec{k}_{\vec{n}})^\dagger a_j(\vec{k}_{\vec{n}}). \qquad (29)
\end{aligned}$$

Based on our previous multi-particle, second-quantizied formalism, we can recognize the quantum theory of light as that of a systems of bosons, each identified by only a momentum $\vec{k}_{\vec{n}}c$ and polarization vector $\hat{\mathcal{E}}_i(\vec{k}_{\vec{n}})$ and with energy $\omega(\vec{k}_{\vec{n}})$.

These results can be immediately used to describe the interaction of an atom with light if we use our familar kinetic energy term for an electron but replace the classical vector potential $\vec{A}(\vec{r})$ but the quantum operator given in Eq. (27) that creates and anihilates photons:

$$H=\sum_{j=1}^{Z}\left\{\frac{(\vec{p}_i+\frac{e}{c}\vec{A}(\vec{r}_i))^2}{2m}-\frac{Ze^2}{|\vec{r}_j|}\right\}+\frac{1}{2}\sum_{j\neq k;j,k=1}^{Z}\frac{1}{|\vec{r}_j-\vec{r}_k|}+\sum_{j,\vec{n}}\hbar\omega\Big\{a_j(\vec{k}_n)^\dagger a_j(\vec{k}_n)+\frac{1}{2}\Big\} \qquad (30)$$

is then the complete Hamiltonian for an atom with Z electrons interacting with the quantum electromagnetic field. Here $\vec{p}_i$, $\vec{r}_i$ and $\vec{A}$ are all operators.