Polarization Vectors and Polarization Sums

1. Introduction

Consider a free electromagnetic field, which satisfies the source-free Maxwell equations:

$$\vec{\boldsymbol{\nabla}}\times\vec{\boldsymbol{E}}=-\frac{1}{c}\frac{\partial\vec{\boldsymbol{B}}}{\partial t}\,,\qquad\qquad \vec{\boldsymbol{\nabla}}\boldsymbol{\cdot}\vec{\boldsymbol{B}}=0\,, \tag{1}$$

$$\vec{\boldsymbol{\nabla}}\boldsymbol{\cdot}\vec{\boldsymbol{E}}=0\,,\qquad\qquad \vec{\boldsymbol{\nabla}}\times\vec{\boldsymbol{B}}=\frac{1}{c}\frac{\partial\vec{\boldsymbol{E}}}{\partial t}\,, \tag{2}$$

where we have employed gaussian units. In light of eq. (1), one can define a scalar potential ϕ and a vector potential $\vec{\boldsymbol{A}}$ such that,

$$\vec{\boldsymbol{E}}=-\vec{\boldsymbol{\nabla}}\phi-\frac{1}{c}\frac{\partial\vec{\boldsymbol{A}}}{\partial t}\,,\qquad\qquad \vec{\boldsymbol{B}}=\vec{\boldsymbol{\nabla}}\times\vec{\boldsymbol{A}}\,. \tag{3}$$

It is convenient to work in the Coulomb gauge where $\vec{\boldsymbol{\nabla}}\boldsymbol{\cdot}\vec{\boldsymbol{A}}=0$. In the presence of sources ρ and $\vec{\boldsymbol{J}}$, we showed in class that ϕ and $\vec{\boldsymbol{A}}$ satisfy

$$\phi(\vec{\boldsymbol{x}},t)=\int d^3x'\,\frac{\rho(\vec{\boldsymbol{x}}',t)}{|\vec{\boldsymbol{x}}-\vec{\boldsymbol{x}}'|}\,,\qquad\qquad \Box\vec{\boldsymbol{A}}=\frac{4\pi}{c}\vec{\boldsymbol{J}}_{tr}\,, \tag{4}$$

where $\vec{\boldsymbol{J}}_{tr}\equiv\vec{\boldsymbol{J}}-\vec{\boldsymbol{J}}_{long}$ and

$$\vec{\boldsymbol{J}}_{long}\equiv-\frac{1}{4\pi}\vec{\boldsymbol{\nabla}}\int d^3x'\,\frac{\vec{\boldsymbol{\nabla}}'\boldsymbol{\cdot}\vec{\boldsymbol{J}}(\vec{\boldsymbol{x}}',t)}{|\vec{\boldsymbol{x}}-\vec{\boldsymbol{x}}'|}\,. \tag{5}$$

Thus, when $\phi=\vec{\boldsymbol{J}}=0$, it follows that $\phi(\vec{\boldsymbol{x}},t)=0$ and $\vec{\boldsymbol{A}}$ satisfies the free wave equation,

$$\Box\vec{\boldsymbol{A}}(\vec{\boldsymbol{x}},t)=0\,. \tag{6}$$

Since $\vec{\boldsymbol{A}}(\vec{\boldsymbol{x}},t)$ is a *real* vector field, the most general solution of eq. (6) is

$$\vec{\boldsymbol{A}}(\vec{\boldsymbol{x}},t)=\sum_\lambda\int\frac{d^3k}{(2\pi)^3}\left[A_0(\vec{\boldsymbol{k}},\lambda)\,\hat{\boldsymbol{\epsilon}}_\lambda(\vec{\boldsymbol{k}})\,e^{i(\vec{\boldsymbol{k}}\boldsymbol{\cdot}\vec{\boldsymbol{x}}-\omega t)}+\text{c.c.}\right]\,, \tag{7}$$

where $A_0(\vec{\boldsymbol{k}},\lambda)$ is a complex amplitude, the abbreviation c.c. stands for "complex conjugate" of the preceding term, and $\omega\equiv kc$ (with $k\equiv|\vec{\boldsymbol{k}}|$). In light of eq. (3) with $\phi=0$, it follows that:

$$\vec{\boldsymbol{E}}(\vec{\boldsymbol{x}},t)=-\frac{1}{c}\frac{\partial\vec{\boldsymbol{A}}}{\partial t}=\sum_\lambda\int\frac{d^3k}{(2\pi)^3}\left[E_0(\vec{\boldsymbol{k}},\lambda)\,\hat{\boldsymbol{\epsilon}}_\lambda(\vec{\boldsymbol{k}})\,e^{i(\vec{\boldsymbol{k}}\boldsymbol{\cdot}\vec{\boldsymbol{x}}-\omega t)}+\text{c.c.}\right]\,, \tag{8}$$

$$\vec{\boldsymbol{B}}(\vec{\boldsymbol{x}},t)=\vec{\boldsymbol{\nabla}}\times\vec{\boldsymbol{A}}=\sum_\lambda\int\frac{d^3k}{(2\pi)^3}\left[E_0(\vec{\boldsymbol{k}},\lambda)\,\hat{\boldsymbol{k}}\times\hat{\boldsymbol{\epsilon}}_\lambda(\vec{\boldsymbol{k}})\,e^{i(\vec{\boldsymbol{k}}\boldsymbol{\cdot}\vec{\boldsymbol{x}}-\omega t)}+\text{c.c.}\right]\,, \tag{9}$$

where $E_0(\vec{\boldsymbol{k}},\lambda) \equiv ikA_0(\vec{\boldsymbol{k}},\lambda)$ is a complex amplitude, $\hat{\boldsymbol{\epsilon}}_\lambda(\vec{\boldsymbol{k}})$ is the *polarization vector* with two possible polarization states indicated by the subscript λ, and $\hat{\boldsymbol{k}} \equiv \vec{\boldsymbol{k}}/k$.

Note that $\vec{\boldsymbol{\nabla}}\boldsymbol{\cdot}\vec{\boldsymbol{B}} = 0$ is automatically satisfied since $\vec{\boldsymbol{k}}\boldsymbol{\cdot}\big[\hat{\boldsymbol{k}} \times \hat{\boldsymbol{\epsilon}}(\vec{\boldsymbol{k}})\big] = 0$. In contrast, the equation $\vec{\boldsymbol{\nabla}}\boldsymbol{\cdot}\vec{\boldsymbol{E}} = 0$ imposes a constraint on the polarization vectors,

$$\vec{\boldsymbol{k}}\boldsymbol{\cdot}\hat{\boldsymbol{\epsilon}}_\lambda(\vec{\boldsymbol{k}}) = 0\,, \tag{10}$$

which corresponds to the statement that electromagnetic waves are transverse. Focusing on a single mode $\vec{\boldsymbol{k}}$, it is convenient to establish a complex basis of unit vectors that span three dimensional space: $\{\hat{\boldsymbol{\epsilon}}_1(\vec{\boldsymbol{k}}), \hat{\boldsymbol{\epsilon}}_2(\vec{\boldsymbol{k}}), \hat{\boldsymbol{k}}\}$. Although $\hat{\boldsymbol{k}}$ is a real unit vector, $\hat{\boldsymbol{\epsilon}}_1(\vec{\boldsymbol{k}})$ and $\hat{\boldsymbol{\epsilon}}_2(\vec{\boldsymbol{k}})$ can be complex unit vectors depending on the choice of basis. The complex vectors that make up the basis set are mutually orthonormal. Hence, the basis vectors satisfy:

$$\hat{\boldsymbol{\epsilon}}_1^*(\vec{\boldsymbol{k}})\boldsymbol{\cdot}\hat{\boldsymbol{\epsilon}}_1(\vec{\boldsymbol{k}}) = \hat{\boldsymbol{\epsilon}}_2^*(\vec{\boldsymbol{k}})\boldsymbol{\cdot}\hat{\boldsymbol{\epsilon}}_2(\vec{\boldsymbol{k}}) = \hat{\boldsymbol{k}}\boldsymbol{\cdot}\hat{\boldsymbol{k}} = 0\,, \tag{11}$$

$$\hat{\boldsymbol{\epsilon}}_\lambda^*(\vec{\boldsymbol{k}})\boldsymbol{\cdot}\hat{\boldsymbol{\epsilon}}_\lambda(\vec{\boldsymbol{k}}) = \delta_{\lambda\lambda'}\,, \qquad \text{for } \lambda, \lambda' \in \{1,2\}, \tag{12}$$

$$\vec{\boldsymbol{k}}\boldsymbol{\cdot}\hat{\boldsymbol{\epsilon}}_\lambda(\vec{\boldsymbol{k}}) = \vec{\boldsymbol{k}}\boldsymbol{\cdot}\hat{\boldsymbol{\epsilon}}_\lambda^*(\vec{\boldsymbol{k}}) = 0\,, \qquad \text{for } \lambda, \lambda' \in \{1,2\}, \tag{13}$$

where

$$\delta_{\lambda\lambda'} = \begin{cases} 1\,, & \text{if } \lambda = \lambda', \\ 0\,, & \text{if } \lambda \neq \lambda'. \end{cases} \tag{14}$$

Suppose we have an electromagnetic wave moving in the z-direction, in which case $\hat{\boldsymbol{k}} = \hat{\boldsymbol{z}}$. If the wave is linearly polarized, then the polarization vectors are real: $\hat{\boldsymbol{\epsilon}}_1(k\hat{\boldsymbol{z}}) = \hat{\boldsymbol{x}}$ corresponds to an x-polarized wave and $\hat{\boldsymbol{\epsilon}}_2(k\hat{\boldsymbol{z}}) = \hat{\boldsymbol{y}}$ corresponds to a y-polarized wave. If the wave is circularly polarized, then the two *complex* polarization vectors are denoted by:[1]

$$\hat{\boldsymbol{\epsilon}}_\pm(k\hat{\boldsymbol{z}}) \equiv \mp\frac{1}{\sqrt{2}}\big(\hat{\boldsymbol{x}} \pm i\hat{\boldsymbol{y}}\big) = \frac{1}{\sqrt{2}}\big(\mp 1\,,\,-i\,,\,0\big)\,. \tag{15}$$

In the conventions employed in optics textbooks, $\hat{\boldsymbol{\epsilon}}_+(k\hat{\boldsymbol{z}})$ is the polarization vector of a left-circularly polarized wave (corresponding to an electric field vector that rotates in the counterclockwise direction when viewed by an observer facing the incoming wave) and $\hat{\boldsymbol{\epsilon}}_-(k\hat{\boldsymbol{z}})$ is the polarization vector of a right-circularly polarized wave (corresponding to an electric field vector that rotates in the clockwise direction when viewed by an observer facing the incoming wave).

When working with the circular polarization vectors, it is convenient to write

$$\hat{\boldsymbol{\epsilon}}_{\pm 1}(\vec{\boldsymbol{k}}) \equiv \hat{\boldsymbol{\epsilon}}_\pm(\vec{\boldsymbol{k}})\,, \qquad \hat{\boldsymbol{\epsilon}}_0(\vec{\boldsymbol{k}}) \equiv \hat{\boldsymbol{k}}\,. \tag{16}$$

This is the so-called *spherical basis* of unit vectors, which comprise the three complex unit

[1]The choice of the $\pm$ sign in eq. (15) is conventional and has been chosen to match the phase conventions employed in the definition of the spherical harmonics [see eq. (21)]. However, note that many authors, including Jackson [cf. eq. (7.22)], omit the $\pm$ in the definition of the circularly polarized polarization vectors.

vectors $\hat{\boldsymbol{\epsilon}}_m(\vec{\boldsymbol{k}})$ [where $m = -1, 0, +1$] that satisfy the following relations:

$$\hat{\boldsymbol{\epsilon}}^*_m(\vec{\boldsymbol{k}}) \boldsymbol{\cdot} \hat{\boldsymbol{\epsilon}}_{m'}(\vec{\boldsymbol{k}}) = \delta_{mm'}\,, \qquad \text{for } m \in \{-1, 0, +1\}, \tag{17}$$

$$\hat{\boldsymbol{\epsilon}}^*_m(\vec{\boldsymbol{k}}) = (-1)^m\, \hat{\boldsymbol{\epsilon}}_{-m}(\vec{\boldsymbol{k}})\,, \qquad \text{for } m \in \{-1, 0, +1\}, \tag{18}$$

$$\hat{\boldsymbol{\epsilon}}^*_m(\vec{\boldsymbol{k}}) \times \hat{\boldsymbol{\epsilon}}_{m'}(\vec{\boldsymbol{k}}) = im\hat{\boldsymbol{k}}\,\delta_{mm'}\,, \quad \text{for } m \in \{-1, +1\}, \tag{19}$$

$$i\hat{\boldsymbol{k}} \times \hat{\boldsymbol{\epsilon}}_m(\vec{\boldsymbol{k}}) = m\hat{\boldsymbol{\epsilon}}_m(\vec{\boldsymbol{k}})\,, \qquad \text{for } m \in \{-1, +1\}. \tag{20}$$

Moreover, note that for $\vec{\boldsymbol{k}} = k\hat{\boldsymbol{z}}$, the three unit vectors $\hat{\boldsymbol{\epsilon}}_{\pm 1}(k\hat{\boldsymbol{z}})$ and $\hat{\boldsymbol{\epsilon}}_0(k\hat{\boldsymbol{z}}) = \hat{\boldsymbol{z}}$ [cf. eq. (16)] are related to the spherical harmonics with $\ell = 1$ as follows:

$$rY_{1m}(\theta_r, \phi_r) = \sqrt{\frac{3}{4\pi}}\, \hat{\boldsymbol{\epsilon}}_m(k\hat{\boldsymbol{z}}) \boldsymbol{\cdot} \vec{\boldsymbol{x}}\,, \qquad \text{for } m \in \{-1, 0, +1\}, \tag{21}$$

where $\vec{\boldsymbol{x}} = r\bigl(\sin\theta_r \cos\phi_r, \sin\theta_r \sin\phi_r, \cos\theta_r\bigr)$ and $r \equiv |\vec{\boldsymbol{x}}|$ when expressed in spherical coordinates. Eq. (21) provides the motivation for the sign conventions chosen in eq. (15), as noted in footnote 1.

2. Polarization vectors of an electromagnetic wave moving in an arbitrary direction $\vec{\boldsymbol{k}}$

The explicit polarization vectors defined in eq. (15) correspond to an electromagnetic wave moving in the $\hat{\boldsymbol{k}} = \hat{\boldsymbol{z}}$ direction. To obtain an explicit form for $\hat{\boldsymbol{\epsilon}}_\pm(\vec{\boldsymbol{k}})$ where the direction of $\vec{\boldsymbol{k}}$ is arbitrary, we will perform an active transformation that rotates $k\hat{\boldsymbol{z}}$ into $\vec{\boldsymbol{k}}$. The latter expressed in spherical coordinates is $\vec{\boldsymbol{k}} = k\bigl(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta\bigr)$. If we represent the corresponding rotation by the 3×3 matrix $\mathcal{R}$ such that $\hat{\boldsymbol{k}} = \mathcal{R}\hat{\boldsymbol{z}}$, then we can parameterize $\mathcal{R}$ in terms of three Euler angles,

$$\mathcal{R}(\phi, \theta, \gamma) \equiv R(\hat{\boldsymbol{z}}, \phi) R(\hat{\boldsymbol{y}}, \theta) R(\hat{\boldsymbol{z}}, \gamma)\,, \tag{22}$$

where $R(\hat{\boldsymbol{n}}, \beta)$ is a 3×3 real orthogonal matrix with unit determinant that represents a rotation by an angle β about a fixed axis $\hat{\boldsymbol{n}}$. The matrix elements of $R(\hat{\boldsymbol{n}}, \beta)$ are given explicitly by Rodrigues' formula:

$$R_{ij}(\hat{\boldsymbol{n}}\,,\,\beta) = \exp(-i\beta\hat{\boldsymbol{n}} \boldsymbol{\cdot} \vec{\boldsymbol{\mathcal{S}}})_{ij} = n_i n_j + (\delta_{ij} - n_i n_j)\cos\beta - \sum_{k=1}^{3} \epsilon_{ijk} n_k \sin\beta\,, \tag{23}$$

where the $\vec{\boldsymbol{\mathcal{S}}} = (\mathcal{S}_1\,, \mathcal{S}_2\,, \mathcal{S}_3)$ are three 3×3 matrices whose matrix elements are given by $(\mathcal{S}_i)_{jk} = -i\epsilon_{ijk}$ and ϵ_{ijk} is the Levi-Civita symbol,

$$\epsilon_{ijk} = \begin{cases} +1\,, & \text{if } ijk \text{ is an even permutation of 123,} \\ +1\,, & \text{if } ijk \text{ is an odd permutation of 123,} \\ \;\;0\,, & \text{otherwise.} \end{cases} \tag{24}$$

Note that the angle γ in eq. (22) is arbitrary, since $R(\hat{\boldsymbol{z}}\,,\,\gamma)\hat{\boldsymbol{z}} = \hat{\boldsymbol{z}}$. That is,

$$\hat{\boldsymbol{k}} = \mathcal{R}\hat{\boldsymbol{z}} = (\sin\theta\cos\phi\,,\, \sin\theta\sin\phi\,,\, \cos\theta)\,, \tag{25}$$

independently of the choice of γ. Henceforth, we shall adopt a convention where $\gamma = 0$.

Two special cases are noteworthy. If $\theta = 0$ or $\theta = \pi$, then the unit vector $\hat{\boldsymbol{k}}$ is parallel to the z-axis and the azimuthal angle ϕ is undefined. Nevertheless, since $\vec{\boldsymbol{k}} \to -\vec{\boldsymbol{k}}$ corresponds in general to $\theta \to \pi - \theta$ and $\phi \to \phi + \pi$ (mod 2π), it is convenient to identify

$$\phi = \begin{cases} 0\,, & \text{for } \hat{\boldsymbol{k}} = \hat{\boldsymbol{z}}\,, \quad (\theta = 0)\,, \\ \pi\,, & \text{for } \hat{\boldsymbol{k}} = -\hat{\boldsymbol{z}}\,, \quad (\theta = \pi)\,. \end{cases} \tag{26}$$

In the convention with $\gamma = 0$ adopted above, the rotation matrix $\mathcal{R}$ defined in eq. (22) is given by

$$\mathcal{R}(\phi\,,\,\theta\,,\,0) = R(\hat{\boldsymbol{z}}\,,\,\phi)\,R(\hat{\boldsymbol{y}}\,,\,\theta) = \begin{pmatrix} \cos\theta\cos\phi & -\sin\phi & \sin\theta\cos\phi \\ \cos\theta\sin\phi & \cos\phi & \sin\theta\sin\phi \\ -\sin\theta & 0 & \cos\theta \end{pmatrix}\,, \tag{27}$$

after making use of eq. (23). We can now transform the polarization vectors $\hat{\boldsymbol{\epsilon}}_{\pm 1}(k\hat{\boldsymbol{z}})$ into $\hat{\boldsymbol{\epsilon}}_{\pm}(\vec{\boldsymbol{k}}, \lambda)$ by noting that under an active rotation, the components of a Cartesian tensor of rank one (i.e., vectors) transform as

$$\big[\hat{\boldsymbol{\epsilon}}_{\pm}(\vec{\boldsymbol{k}})\big]_i = \sum_{j=1}^{3} \mathcal{R}_{ij}(\phi, \theta, 0)\big[\hat{\boldsymbol{\epsilon}}_{\pm}(k\hat{\boldsymbol{z}})\big]_j\,. \tag{28}$$

Using eqs. (15) and (27), a simple computation yields

$$\hat{\boldsymbol{\epsilon}}_{\pm}(\vec{\boldsymbol{k}}) = \frac{1}{\sqrt{2}}\big(\mp\cos\theta\cos\phi + i\sin\phi\,,\, \mp\cos\theta\sin\phi - i\cos\phi\,,\, \pm\sin\theta\big)\,, \tag{29}$$

Note that the $\epsilon_{\pm}(\vec{\boldsymbol{k}})$ depend only on the direction of $\vec{\boldsymbol{k}}$ and not on its magnitude k. Moreover, for $\vec{\boldsymbol{k}} = k\hat{\boldsymbol{z}}$, eq. (29) reduces to eq. (15) when $\theta = \phi = 0$, as expected. One can also verify that eqs. (17)–(20) are satisfied by the $\hat{\boldsymbol{\epsilon}}_m(\vec{\boldsymbol{k}})$, where $\hat{\boldsymbol{\epsilon}}_0(\vec{\boldsymbol{k}}) = \hat{\boldsymbol{k}}$ and $\hat{\boldsymbol{\epsilon}}_{\pm 1}(\vec{\boldsymbol{k}}) = \hat{\boldsymbol{\epsilon}}_{\pm}(\vec{\boldsymbol{k}})$ are given by eq. (29).

Finally, it is instructive to examine the relation between $\epsilon_{\pm}(-\vec{\boldsymbol{k}})$ and $\epsilon_{\pm}(\vec{\boldsymbol{k}})$ As previously noted, if $\vec{\boldsymbol{k}}$ points in a direction with polar and azimuthal angles θ and ϕ, then the corresponding polar and azimuthal angles of the vector $-\vec{\boldsymbol{k}}$ are $\pi - \theta$ and $\phi + \pi$ (mod 2π). In particular, note the following identity:

$$\mathcal{R}(\phi + \pi, \pi - \theta, 0) = R(\phi, \theta, 0)D\,, \qquad \text{where } D \equiv R(\hat{\boldsymbol{x}}, \pi) = \text{diag}(+1, -1, -1). \tag{30}$$

Using eq. (15), one can check that $D\,\hat{\boldsymbol{\epsilon}}_{\pm}(k\hat{\boldsymbol{z}}) = -\hat{\boldsymbol{\epsilon}}_{\mp}(k\hat{\boldsymbol{z}})$. Hence, eqs. (28) and (30) yield:

$$\begin{aligned} \hat{\boldsymbol{\epsilon}}_{\pm}(-\vec{\boldsymbol{k}}) &= \mathcal{R}(\phi + \pi, \pi - \theta, 0)\hat{\boldsymbol{\epsilon}}_{\pm}(k\hat{\boldsymbol{z}}) = R(\phi, \theta, 0)D\,\hat{\boldsymbol{\epsilon}}_{\pm}(k\hat{\boldsymbol{z}}) = -R(\phi, \theta, 0)\hat{\boldsymbol{\epsilon}}_{\mp}(k\hat{\boldsymbol{z}}) \\ &= -\hat{\boldsymbol{\epsilon}}_{\mp}(\vec{\boldsymbol{k}})\,. \end{aligned} \tag{31}$$

Combining eqs. (18) and (31), we end up with

$$\hat{\boldsymbol{\epsilon}}_{\pm}(-\vec{\boldsymbol{k}}) = \hat{\boldsymbol{\epsilon}}^*_{\pm}(\vec{\boldsymbol{k}})\,, \tag{32}$$

where the upper and lower sign subscripts on the left and right hand sides of eqs. (31) and (32), respectively, are correlated as indicated.

3. Polarization sums

In computing the angular distribution of a scattered electromagnetic wave, if the final state polarization is not observed, then one must sum over the two possible polarizations of the scattered wave. If the initial state polarization is not observed, then one must *average* over the two possible polarizations of the incoming wave. To facilitate the computation of the sum over polarizations, I will derive a very useful formula.

Recall from your quantum mechanics class that given a complete orthonormal basis $\{|n\rangle\}$, the completeness relation states that

$$\sum_n |n\rangle\langle n| = \mathbb{1}\,, \tag{33}$$

where $\mathbb{1}$ is the identity operator. Applying this result to the complex basis $\{\hat{\boldsymbol{\epsilon}}_1, \hat{\boldsymbol{\epsilon}}_2, \hat{\boldsymbol{k}}\}$, it follows that

$$\hat{\boldsymbol{\epsilon}}_1^*\hat{\boldsymbol{\epsilon}}_1 + \hat{\boldsymbol{\epsilon}}_2^*\hat{\boldsymbol{\epsilon}}_2 + \hat{\boldsymbol{k}}\hat{\boldsymbol{k}} = \mathbb{1}\,. \tag{34}$$

If written in terms of the vector components, eq. (34) is equivalent to

$$(\hat{\boldsymbol{\epsilon}}_1^*)_i(\hat{\boldsymbol{\epsilon}}_1)_j + (\hat{\boldsymbol{\epsilon}}_2^*)_i(\hat{\boldsymbol{\epsilon}}_2)_j + \hat{\boldsymbol{k}}_i\hat{\boldsymbol{k}}_j = \delta_{ij}\,, \tag{35}$$

where the indices $i, j \in \{1, 2, 3\}$ label the x, y and z components of the corresponding unit vector. The unit vector $\hat{\boldsymbol{k}}$ indicates the direction of the wave propagation, which is also designated by the symbol $\hat{\boldsymbol{n}}$. Hence, it follows that

$$\boxed{\sum_{\lambda=1}^{2}(\hat{\boldsymbol{\epsilon}}_\lambda)_i(\hat{\boldsymbol{\epsilon}}_\lambda^*)_j = \delta_{ij} - \hat{\boldsymbol{n}}_i\hat{\boldsymbol{n}}_j\,, \qquad \text{where } \hat{\boldsymbol{n}} \equiv \frac{\vec{\boldsymbol{k}}}{k}}\,. \tag{36}$$

Eq. (36) is called the polarization sum formula. It can be applied to any basis of polarization vectors (linear, circular, or even elliptical polarization).

4. Illustrating the use of polarization sums

To illustrate the use of the polarization sum formula, consider the angular distribution of radiation obtained via the nonrelativistic Larmour formula,

$$\frac{dP}{d\Omega} = \frac{e^2}{4\pi c^3}|\hat{\boldsymbol{n}} \times (\hat{\boldsymbol{n}} \times \vec{\boldsymbol{a}})|^2\,, \tag{37}$$

where $\vec{\boldsymbol{a}}$ is the acceleration vector of a point particle of charge e. Consider the electric field of the emitted radiation. Using the completeness relation [eq. (34)],

$$\vec{\boldsymbol{E}} = \hat{\boldsymbol{\epsilon}}_1\big(\hat{\boldsymbol{\epsilon}}_1^*\boldsymbol{\cdot}\vec{\boldsymbol{E}}\big) + \hat{\boldsymbol{\epsilon}}_2\big(\hat{\boldsymbol{\epsilon}}_2^*\boldsymbol{\cdot}\vec{\boldsymbol{E}}\big)\,, \tag{38}$$

after noting that $\hat{\boldsymbol{k}} = \hat{\boldsymbol{n}}$ and $\hat{\boldsymbol{n}}\boldsymbol{\cdot}\vec{\boldsymbol{E}} = 0$ (for a transverse wave). Hence it follows that

$$|\vec{\boldsymbol{E}}|^2 = |\hat{\boldsymbol{\epsilon}}_1^*\boldsymbol{\cdot}\vec{\boldsymbol{E}}|^2 + |\hat{\boldsymbol{\epsilon}}_2^*\boldsymbol{\cdot}\vec{\boldsymbol{E}}|^2\,. \tag{39}$$

Since Larmour's formula was derived by evaluating $|\vec{\boldsymbol{E}}|^2$ for an accelerating point charge, it follows that if the polarization vector of the emitted radiation is $\hat{\boldsymbol{\epsilon}}$ then eq. (37) must be modified as follows:

$$\frac{dP}{d\Omega} = \frac{e^2}{4\pi c^3}\left|\hat{\boldsymbol{\epsilon}}^*\boldsymbol{\cdot}[\hat{\boldsymbol{n}}\times(\hat{\boldsymbol{n}}\times\vec{\boldsymbol{a}})]\right|^2 . \tag{40}$$

Using the relevant vector identities and noting that $\hat{\boldsymbol{\epsilon}}^*\boldsymbol{\cdot}\hat{\boldsymbol{n}} = 0$ [in light of eq. (13) after identifying $\hat{\boldsymbol{n}} = \hat{\boldsymbol{k}}$], it follows that

$$\hat{\boldsymbol{\epsilon}}^*\boldsymbol{\cdot}[\hat{\boldsymbol{n}}\times(\hat{\boldsymbol{n}}\times\vec{\boldsymbol{a}})] = (\hat{\boldsymbol{\epsilon}}^*\boldsymbol{\cdot}\hat{\boldsymbol{n}})(\vec{\boldsymbol{a}}\boldsymbol{\cdot}\hat{\boldsymbol{n}}) - (\hat{\boldsymbol{\epsilon}}^*\boldsymbol{\cdot}\vec{\boldsymbol{a}})(\hat{\boldsymbol{n}}\boldsymbol{\cdot}\hat{\boldsymbol{n}}) = -\hat{\boldsymbol{\epsilon}}^*\boldsymbol{\cdot}\vec{\boldsymbol{a}} . \tag{41}$$

Hence, Larmour's formula for the angular distribution for radiation observed with polarization vector $\hat{\boldsymbol{\epsilon}}$ (in the nonrelativistic limit) is given by

$$\frac{dP}{d\Omega} = \frac{e^2}{4\pi c^3}\left|\hat{\boldsymbol{\epsilon}}^*\boldsymbol{\cdot}\vec{\boldsymbol{a}}\right|^2 . \tag{42}$$

In class, I showed that by using eq. (42), one can derive the famous Thomson scattering cross section formula,

$$\frac{d\sigma}{d\Omega} = \left(\frac{e^2}{mc^2}\right)^2 \left|\hat{\boldsymbol{\epsilon}}_i\boldsymbol{\cdot}\hat{\boldsymbol{\epsilon}}_f^*\right|^2 , \tag{43}$$

for the scattering of an electron with an incoming electromagnetic wave with polarization vector $\hat{\boldsymbol{\epsilon}}_i$ which produces an emitted electromagnetic wave with polarization vector $\hat{\boldsymbol{\epsilon}}_f$. In order to compute the scattering cross section if neither the initial polarization nor the final polarization is observed, then one must average over initial state polarizations and sum over final state polarizations. That is,

$$\left(\frac{d\sigma}{d\Omega}\right)_{\rm unpolarized} = \frac{1}{2}\left(\frac{e^2}{mc^2}\right)^2 \sum_\lambda\sum_{\lambda'}\left|\hat{\boldsymbol{\epsilon}}_\lambda\boldsymbol{\cdot}\hat{\boldsymbol{\epsilon}}_{\lambda'}^*\right|^2 , \tag{44}$$

where we have simplified the notation by denoting the initial polarization by λ and the final polarization by λ' (thereby omitting the subscripts i and f). Assume that the incoming electromagnetic wave is propagating in the z direction and the outgoing electromagnetic wave is propagating in the $\hat{\boldsymbol{n}}$ direction. Then, eq. (36) implies that

$$\sum_{\lambda=1}^{2}(\hat{\boldsymbol{\epsilon}}_\lambda)_j(\hat{\boldsymbol{\epsilon}}_\lambda^*)_\ell = \delta_{j\ell} - \hat{\boldsymbol{z}}_j\hat{\boldsymbol{z}}_\ell , \tag{45}$$

$$\sum_{\lambda'=1}^{2}(\hat{\boldsymbol{\epsilon}}_{\lambda'})_j(\hat{\boldsymbol{\epsilon}}_{\lambda'}^*)_\ell = \delta_{j\ell} - \hat{\boldsymbol{n}}_j\hat{\boldsymbol{n}}_\ell . \tag{46}$$

Then eq. (44) yields

$$\frac{1}{2}\sum_\lambda\left|\hat{\boldsymbol{\epsilon}}_\lambda\boldsymbol{\cdot}\hat{\boldsymbol{\epsilon}}_{\lambda'}^*\right|^2 = \frac{1}{2}(\hat{\boldsymbol{\epsilon}}_{\lambda'})_j(\hat{\boldsymbol{\epsilon}}_{\lambda'}^*)_\ell\sum_\lambda(\hat{\boldsymbol{\epsilon}}_\lambda)_\ell(\hat{\boldsymbol{\epsilon}}_\lambda^*)_j = \frac{1}{2}(\hat{\boldsymbol{\epsilon}}_{\lambda'})_j(\hat{\boldsymbol{\epsilon}}_{\lambda'}^*)_\ell\big(\delta_{\ell j} - \hat{\boldsymbol{z}}_\ell\hat{\boldsymbol{z}}_j\big) = \frac{1}{2}\big(1 - |\hat{\boldsymbol{z}}\boldsymbol{\cdot}\hat{\boldsymbol{\epsilon}}_{\lambda'}|^2\big) , \tag{47}$$

where there is an implicit sum over twice-repeated indices following Einstein's summation convention. Hence it follows that

$$\sum_\lambda \sum_{\lambda'} \left|\hat{\boldsymbol{\epsilon}}_\lambda \boldsymbol{\cdot} \hat{\boldsymbol{\epsilon}}^*_{\lambda'}\right|^2 = \frac{1}{2}\sum_{\lambda'}\left(1 - |\hat{\boldsymbol{z}}\boldsymbol{\cdot}\hat{\boldsymbol{\epsilon}}_{\lambda'}|^2\right) = 1 - \frac{1}{2}\sum_{\lambda'} |\hat{\boldsymbol{z}}\boldsymbol{\cdot}\hat{\boldsymbol{\epsilon}}_{\lambda'}|^2 = 1 - \frac{1}{2}\hat{\boldsymbol{z}}_j\hat{\boldsymbol{z}}_\ell \sum_{\lambda'} (\hat{\boldsymbol{\epsilon}}_{\lambda'})_j(\hat{\boldsymbol{\epsilon}}^*_{\lambda'})_\ell$$
$$= 1 - \frac{1}{2}\hat{\boldsymbol{z}}_j\hat{\boldsymbol{z}}_\ell\left(\delta_{j\ell} - \hat{\boldsymbol{n}}_j\hat{\boldsymbol{n}}_\ell\right) = \frac{1}{2}\left[1 + (\hat{\boldsymbol{z}}\boldsymbol{\cdot}\hat{\boldsymbol{n}})^2\right] = \frac{1}{2}\left(1+\cos^2\theta\right), \tag{48}$$

after using $\hat{\boldsymbol{n}} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$. Hence, we conclude that the unpolarized Thomson cross section is given by,

$$\left(\frac{d\sigma}{d\Omega}\right)_{\text{unpolarized}} = \frac{1}{2}\left(\frac{e^2}{mc^2}\right)^2\left(1+\cos^2\theta\right). \tag{49}$$

Integrating over angles yields the famous result for the Thomson scattering cross section,

$$\begin{aligned}\sigma_T &= \frac{1}{2}\left(\frac{e^2}{mc^2}\right)^2\int\left(1+\cos^2\theta\right)d\Omega \\ &= \frac{1}{2}\left(\frac{e^2}{mc^2}\right)^2 2\pi\int_{-1}^{1}\left(1+\cos^2\theta\right)d\cos\theta \\ &= \frac{8\pi}{3}\left(\frac{e^2}{mc^2}\right)^2.\end{aligned} \tag{50}$$

We denote the *classical radius of the electron* by

$$r_c = \frac{e^2}{mc^2} \simeq 2.82\times 10^{-13}\ \text{cm}, \tag{51}$$

which would physically correspond to setting the rest energy of the electron to the electrostatic energy of a uniformly charged spherical surface of radius r_c. In terms of r_c, the Thomson scattering cross section is given by

$$\sigma_T = \frac{8\pi}{3}r_c^2. \tag{52}$$

Note that the concept of the classical radius of the electron does not make any sense from a quantum mechanics point of view, as quantum mechanical effects relevant for the properties of the electron cannot be ignored at distance scales shorter than the Compton wavelength of the electron, $\hbar/(mc) \simeq 2.426\times 10^{-10}$ cm, which is significantly larger than r_c.

In quantum mechanics, electromagnetic radiation can also be described as photons, in which case the Thomson scattering is interpreted as corresponding to the scattering process $\gamma + e^- \to \gamma + e^-$ (which is now called Compton scattering). In quantum electrodynamics, the total unpolarized cross section for Compton scattering can be derived in perturbation theory, and the result is given by the Klein-Nishima formula,

$$\left(\frac{d\sigma}{d\Omega}\right)_{\text{unpolarized}} = \frac{1}{2}\left(\frac{e^2}{mc^2}\right)^2 y^2\left(y + \frac{1}{y} - \sin^2\theta\right), \tag{53}$$

where

$$y \equiv \frac{1}{1+z(1-\cos\theta)} \quad \text{and} \quad z \equiv \frac{\hbar\omega}{mc^2}\,. \tag{54}$$

Integrating over angles to obtain the total cross section, we obtain

$$\sigma_{KN} = \frac{\pi}{z^3}\left(\frac{e^2}{mc^2}\right)^2 \left[\frac{2z(2+8z+9z^2+z^3)}{(1+2z)^2} - (2+2z-z^2)\ln(1+2z)\right]. \tag{55}$$

In the classical limit of long wavelengths (corresponding to $\hbar\omega \ll mc^2$), $z \to 0$ (and $y \to 1$) and the Klein-Nishima cross section tends to the Thomson cross section obtained in eqs. (44) and (50). One can check explicitly that eq. (55) reduces to eq. (50) in the limit as $z \to 0$ by expanding out the logarithm in eq. (55) for small z and showing that the term within the brackets in eq. (55) behaves as $8z^3/3$ as $z \to 0$. In case you are curious, Mathematica yields

$$\frac{2z(2+8z+9z^2+z^3)}{(1+2z)^2} - (2+2z-z^2)\ln(1+2z) = \frac{8z^3}{3}\big[1-2z+\mathcal{O}(z^2)\big]\,. \tag{56}$$