

Physics 221B
Spring 2011
Notes 38
Lagrangian and Hamiltonian Formulation
of the Classical Electromagnetic Field

In Notes 37, we derived the classical Hamiltonian description of the classical electromagnetic field by starting with the equations of motion and making a series of guesses for the Hamiltonian and the canonical variables. A more systematic, but more lengthy, procedure is to start with the classical Lagrangian for the field and proceed through the Legendre transformation. This procedure is outlined in these Notes, which are independent of Notes 37. I will not lecture on this material, and either set of Notes can be regarded as a starting point for the quantization carried out in Notes 39.

Quantization of the electromagnetic field is necessary for a consistent treatment of the interaction of radiation with matter, for if only part of nature is treated quantum mechanically while the rest is treated classically, then it becomes possible to violate the uncertainty principle in the quantum half by using the classical half to make measurements. Therefore if we believe the uncertainty principle is fundamental, then we must believe that all of nature is quantum mechanical. Unfortunately, the electromagnetic field is not the easiest field to quantize, and from a pedagogical standpoint it would be better to begin with some model field such as the one-dimensional, vibrating string. In these notes, however, we will proceed directly to the electromagnetic field, which is of the greatest physical interest.

It is not terribly difficult to get straight to the some of the most important features and ideas surrounding photons by working directly with the equations of motion for the mode amplitudes of the electromagnetic field (by noticing that they are harmonic oscillators, by transcribing the classical variables into quantum mechanical operators, etc.) In these notes, however, we will work instead with field Lagrangians and Hamiltonians, which is not only the more proper way to justify the various steps, but which is also important for the study of symmetries and invariants and other fundamental issues.

The quantization of the electromagnetic field raises an interesting question, namely, how in general do we go from a system whose classical mechanics is known to a proper quantum mechanical description of that system? The answer cannot be framed in terms of any completely deductive process, because quantum mechanics contains more information than classical mechanics. But if we had to make an outline of the various steps to follow, it would run something like this. First, we start with the classical equations of motion, which in the case of the electromagnetic field are Maxwell's equations. Next, we find a classical Lagrangian for these equations, $L(q, \dot{q})$, which is a function of the q 's and \dot{q} 's. The Lagrangian is associated with the action S , defined by

$$S = \int dt L. \tag{1}$$

The equations of motion should be derivable from *Hamilton's principle*, which says that $\delta S = 0$ along physically realizable motions. This is equivalent to the Euler-Lagrange equations,

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) = \frac{\partial L}{\partial q_i}. \quad (2)$$

The index i labels the degrees of freedom of the system. Next, we define the momenta by

$$p_i = \frac{\partial L}{\partial \dot{q}_i}, \quad (3)$$

and use them to construct the classical Hamiltonian,

$$H(q, p) = \sum_i p_i \dot{q}_i - L(q, \dot{q}), \quad (4)$$

where, as indicated, the Hamiltonian is regarded as a function of the q 's and p 's. Now the equations of motion are Hamilton's equations,

$$\begin{aligned} \dot{q}_i &= \frac{\partial H}{\partial p_i}, \\ \dot{p}_i &= -\frac{\partial H}{\partial q_i}. \end{aligned} \quad (5)$$

Finally, we transcribe the classical Hamiltonian into a quantum Hamiltonian by replacing the q 's and p 's by operators which satisfy the commutation relations,

$$[q_i, p_j] = i\hbar \delta_{ij}. \quad (6)$$

The last step is Dirac's quantization prescription, and it is the one which is nonunique, since the results vary depending on the ordering of operators and on the system of canonical variables used in the classical Hamiltonian. Ultimately, the correctness of the quantum Hamiltonian must be tested by experiment.

Fields in general have an infinite number of degrees of freedom, and the electromagnetic field in particular has (as we will see) two degrees of freedom for each spatial point \mathbf{r} . As in all field theories, the degrees of freedom are labelled, not by discrete indices i as in Eq. (2) above, but rather by continuous labels such as \mathbf{r} . These notes will assume that you are familiar with basic concepts of classical field theory, such as the labelling of continuous degrees of freedom by labels such as \mathbf{r} , the passage from discrete to a continuous number of degrees of freedom in model systems such as the vibrating string, etc. It is also assumed that you are familiar with the functional derivative and such notions as Lagrangian densities.

For reference, we list here the homogeneous Maxwell equations,

$$\begin{aligned} \nabla \times \mathbf{E} &= -\frac{1}{c} \frac{\partial \mathbf{B}}{\partial t}, \\ \nabla \cdot \mathbf{B} &= 0, \end{aligned} \quad (7)$$

and the inhomogeneous Maxwell equations,

$$\begin{aligned}\nabla \cdot \mathbf{E} &= 4\pi\rho, \\ \nabla \times \mathbf{B} &= \frac{4\pi}{c}\mathbf{J} + \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t}.\end{aligned}\tag{8}$$

The fields \mathbf{E} and \mathbf{B} are expressed in terms of the scalar and vector potential by

$$\mathbf{E} = -\nabla\phi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t},\tag{9a}$$

$$\mathbf{B} = \nabla \times \mathbf{A}.\tag{9b}$$

The potentials are not unique, but may be subjected to a gauge transformation, i.e., ϕ and \mathbf{A} may be replaced by ϕ' and \mathbf{A}' , where

$$\begin{aligned}\phi' &= \phi - \frac{1}{c}\frac{\partial \chi}{\partial t}, \\ \mathbf{A}' &= \mathbf{A} + \nabla\chi,\end{aligned}\tag{10}$$

for any scalar field χ . Expressions (9) cause the homogeneous Maxwell equations to be satisfied automatically, while the inhomogeneous Maxwell equations become

$$\nabla^2\phi = -4\pi\rho - \frac{1}{c}\frac{\partial}{\partial t}(\nabla \cdot \mathbf{A}),\tag{11a}$$

$$\nabla^2\mathbf{A} - \frac{1}{c^2}\frac{\partial^2\mathbf{A}}{\partial t^2} = -\frac{4\pi}{c}\mathbf{J} + \nabla\left(\frac{1}{c}\frac{\partial\phi}{\partial t} + \nabla \cdot \mathbf{A}\right).\tag{11b}$$

These are the equations for which we seek a Lagrangian description, in the first step in our quantization program.

The field Lagrangians which will be of interest to us in this course all have the form of a spatial integral of a Lagrangian density, where the latter is a function of the fields and their space and time derivatives. Writing $\psi = \psi(\mathbf{r}, t)$ for a generic field and $\dot{\psi} = \partial\psi/\partial t$ and $\nabla\psi$ for its time and space derivatives, respectively, we can express our Lagrangians in the form

$$L = \int d^3\mathbf{r} \mathcal{L}(\psi, \dot{\psi}, \nabla\psi),\tag{12}$$

where L is the Lagrangian and \mathcal{L} is the Lagrangian density. The action as usual is the time integral of the Lagrangian,

$$S = \int L dt = \int d^3\mathbf{r} dt \mathcal{L} = \int d^4x \mathcal{L},\tag{13}$$

where the last form is a 4-dimensional notation for the integral of the Lagrangian density over space and time. In order to get field equations which are relativistically covariant, it is necessary to have an action which is a Lorentz scalar (so that Hamilton's principle will be independent of Lorentz frame). But since the 4-volume element d^4x is a Lorentz scalar, the Lagrangian density \mathcal{L} must be also. This fact imposes severe constraints on the form of Lagrangian densities allowed in relativistic field theory.

When we make a variation in the field ψ , the action changes according to

$$\delta S = \int d^4x \left[\frac{\partial \mathcal{L}}{\partial \psi} \delta \psi + \frac{\partial \mathcal{L}}{\partial \dot{\psi}} \delta \dot{\psi} + \frac{\partial \mathcal{L}}{\partial (\nabla \psi)} \cdot \delta \nabla \psi \right]. \quad (14)$$

If we integrate this by parts and demand that $\delta S = 0$ for all variations $\delta \psi$, we obtain the Euler-Lagrange equations for the field ψ ,

$$\frac{\partial}{\partial t} \left(\frac{\partial \mathcal{L}}{\partial \dot{\psi}} \right) + \nabla \cdot \left(\frac{\partial \mathcal{L}}{\partial (\nabla \psi)} \right) = \frac{\partial \mathcal{L}}{\partial \psi}, \quad (15)$$

which are the field analogs of Eq. (2) in the discrete case. The Euler-Lagrange equations for the field are neater in 4-vector notation,

$$\frac{\partial}{\partial x^\mu} \left(\frac{\partial \mathcal{L}}{\partial \psi_{,\mu}} \right) = \frac{\partial \mathcal{L}}{\partial \psi}, \quad (16)$$

where we use the ‘‘comma notation’’ for derivatives,

$$\psi_{,\mu} = \frac{\partial \psi}{\partial x^\mu}. \quad (17)$$

However, in these notes we will not use covariant, 4-vector notation very much, since we will be interested at first in interactions between the electromagnetic field and nonrelativistic matter, and because our quantization scheme for the electromagnetic field will not be manifestly covariant anyway.

We now construct the Lagrangian density for the electromagnetic field. At first we assume that the charge and current densities ρ and \mathbf{J} appearing in Maxwell’s equations are simply given functions of (\mathbf{r}, t) , which of course are constrained to satisfy the continuity equation,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0. \quad (18)$$

The electromagnetic Lagrangian density will consist of a free-field part and a part involving the interaction with matter. The free-field part must be a Lorentz scalar, since Maxwell’s equations are Lorentz-covariant. There are only two Lorentz scalars which can be constructed out of the electromagnetic field, namely, $E^2 - B^2$ and $\mathbf{E} \cdot \mathbf{B}$. The latter is not invariant under parity, so we reject it, and take $(E^2 - B^2)/8\pi$ for the Lagrangian density for the free field. The factor $1/8\pi$ is necessary to make the Hamiltonian density equal to the energy density of the field, as we expect. As for the interaction Lagrangian, this must also be a Lorentz scalar, and we take it to be proportional to $j^\mu A_\mu$ in covariant notation. Combining these terms and adjusting the proportionality constant to make the equations of motion come out right, we have the Lagrangian density,

$$\mathcal{L} = \frac{E^2 - B^2}{8\pi} - \rho \phi + \frac{1}{c} \mathbf{J} \cdot \mathbf{A}, \quad (19)$$

where we revert to 3 + 1-notation.

It is of interest to check that the Euler-Lagrange equations for this Lagrangian do indeed reproduce Maxwell’s equations. To do this, we must first recognize that the fields \mathbf{E} and \mathbf{B} are not

the fundamental dynamical variables of the electromagnetic fields (the analog of the q 's in a discrete Lagrangian), for Maxwell's equations are first order in time and we expect to see something analogous to the \ddot{q} 's which occur in ordinary particle mechanics. But since we do see second time derivatives in Eqs. (11), we try to interpret the potentials (ϕ, \mathbf{A}) as the fundamental dynamical variables. That is, we take (ϕ, \mathbf{A}) as four fields like the generic ψ in Eq. (15) above. In this interpretation, the fields \mathbf{E} and \mathbf{B} seen in Eq. (19) are taken as merely convenient substitutions for the expressions in Eqs. (9). Working first with the scalar potential ϕ , we have

$$\frac{\partial \mathcal{L}}{\partial \phi} = 0, \quad \frac{\partial \mathcal{L}}{\partial(\nabla\phi)} = -\frac{\mathbf{E}}{4\pi}, \quad \frac{\partial \mathcal{L}}{\partial \dot{\phi}} = -\rho, \quad (20)$$

which, according to Eq. (15), give Maxwell's equation,

$$\nabla \cdot \mathbf{E} = 4\pi\rho. \quad (21)$$

As for the vector potential, we identify ψ in Eq. (15) with one component A_i of the vector potential, and find

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \dot{A}_i} &= -\frac{1}{4\pi c} E_i, \\ \frac{\partial \mathcal{L}}{\partial A_{i,j}} &= -\frac{1}{4\pi} (A_{i,j} - A_{j,i}) = \frac{1}{4\pi} \epsilon_{ijk} B_k, \\ \frac{\partial \mathcal{L}}{\partial A_i} &= \frac{1}{c} J_i, \end{aligned} \quad (22)$$

where again we use comma notation,

$$A_{i,j} = \frac{\partial A_i}{\partial x_j}, \quad (23)$$

and where the second calculation in Eq. (22) is assisted by noticing that

$$B^2 = (\nabla \times \mathbf{A})^2 = \epsilon_{ijk} A_{k,j} \epsilon_{ilm} A_{m,\ell} = A_{m,\ell} A_{m,\ell} - A_{m,\ell} A_{\ell,m}. \quad (24)$$

Finally, the Euler-Lagrange equation for A_i is seen to be the Maxwell equation,

$$\nabla \times \mathbf{B} = \frac{4\pi}{c} \mathbf{J} + \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t}. \quad (25)$$

Thus, Eq. (19) is indeed the correct classical electromagnetic field Lagrangian. To follow our quantization prescription, our next step would be to find the classical field Hamiltonian. The standard program for passing from a field Lagrangian to a field Hamiltonian runs as follows, where we revert to the generic notation used in Eq. (12)–(17). First, we define the field $\pi(\mathbf{r})$ conjugate to $\psi(\mathbf{r})$ by

$$\pi(\mathbf{r}) = \frac{\delta L}{\delta \dot{\psi}(\mathbf{r})} = \frac{\partial \mathcal{L}}{\partial \dot{\psi}}, \quad (26)$$

which is the field analog of Eq. (3). Next, we define the field Hamiltonian by

$$H = \int d^3\mathbf{r} \pi(\mathbf{r}) \dot{\psi}(\mathbf{r}) - L, \quad (27)$$

which is the field analog of Eq. (4). In cases when the Lagrangian can be written as the spatial integral of a Lagrangian density, the Hamiltonian can be written as the integral a Hamiltonian density,

$$H = \int d^3\mathbf{r} \mathcal{H}, \quad (28)$$

where \mathcal{H} is a function of the fields ψ and π ,

$$\mathcal{H} = \mathcal{H}(\psi, \pi) = \pi\dot{\psi} - \mathcal{L}. \quad (29)$$

Finally, Hamilton's equations are

$$\begin{aligned} \dot{\psi}(\mathbf{r}) &= \frac{\delta H}{\delta \pi(\mathbf{r})} = \frac{\partial \mathcal{H}}{\partial \pi}, \\ \dot{\pi}(\mathbf{r}) &= -\frac{\delta H}{\delta \psi(\mathbf{r})} = -\frac{\partial \mathcal{H}}{\partial \psi}, \end{aligned} \quad (30)$$

which are the field analogs of Eqs. (5). This standard program works fine on simple fields such as the vibrating string.

However, when applied to the electromagnetic field Lagrangian (19), the standard program runs into certain difficulties which have no analog in the usual Lagrangians encountered in nonrelativistic particle mechanics (of the kinetic-minus-potential type). We must deal with these difficulties before proceeding to the classical field Hamiltonian.

There are actually two difficulties. The first is that the variable ϕ is not an independent dynamical variable, for the ‘‘equation of evolution’’ of ϕ , seen in Eq. (11a), does not involve any time derivatives. That is, if we imagine that $\mathbf{A}(\mathbf{r}, t)$ were known, then Eq. (11a) would determine ϕ as a function of (\mathbf{r}, t) simply by inverting the Laplacian. Thus, $\phi(\mathbf{r}, t)$, regarded as a dynamical variable, is a function of the other dynamical variables $\mathbf{A}(\mathbf{r}, t)$. This same difficulty makes its appearance in another form as soon as we attempt to compute the momentum conjugate to the scalar potential ϕ according to Eq. (26), for we find that this momentum vanishes identically. Therefore there is no evolution equation for this momentum in the usual sense. We fix this difficulty by inverting the Laplacian in Eq. (11a), to solve for ϕ in terms of the other variables, and then using the result to eliminate ϕ from the Lagrangian.

The second difficulty is that Eqs. (11) do not have a unique solution for given initial conditions, due to the gauge degree of freedom in the potentials. For if (ϕ, \mathbf{A}) are solutions of Eqs. (11), then so are any fields (ϕ', \mathbf{A}') related to (ϕ, \mathbf{A}) by the gauge transformation (10). Thus, the general solution of Eqs. (11) involves an arbitrary gauge transformation which develops in the course of time. To see this difficulty in another way, we notice that the gauge transformation (10) only affects the longitudinal part \mathbf{A}_{\parallel} of the vector potential \mathbf{A} , and not the transverse part, because the vector field $\nabla\chi$, when transformed to \mathbf{k} -space, is parallel (longitudinal) to \mathbf{k} . And if we attempt to project out the longitudinal part of Eq. (11b) to find an equation for $\ddot{\mathbf{A}}_{\parallel}$, we find that no such equation exists. That is, Eq. (11b) places no constraint on the value of \mathbf{A}_{\parallel} or its time derivatives, which is logical

since \mathbf{A}_{\parallel} can be changed into any longitudinal field we want by means of a gauge transformation. Therefore \mathbf{A}_{\parallel} is only a gauge degree of freedom, not a true dynamical degree of freedom. The way to fix this difficulty is to adopt a definite gauge convention, which will fix the value of \mathbf{A}_{\parallel} . In these notes, we will adopt Coulomb gauge,

$$\nabla \cdot \mathbf{A} = 0, \quad (31)$$

which is equivalent to $\mathbf{A}_{\parallel} = 0$.

After these two fixes, two of the four field variables inherent in the potentials (ϕ, \mathbf{A}) have been eliminated, and only two remain (the transverse components of the vector potential). The Lagrangian at this stage depends only on the transverse fields \mathbf{A}_{\perp} and $\dot{\mathbf{A}}_{\perp}$, and the two field variables inherent in \mathbf{A}_{\perp} are the true dynamical variables of the electromagnetic field. Thus, we can say that the electromagnetic field possesses two degrees of freedom at each spatial point \mathbf{r} .

We now make a digression into the subject of transverse and longitudinal vector fields, which is important in understanding the quantum mechanics of the electromagnetic field, especially in the Coulomb gauge. Let $\mathbf{F}(\mathbf{r})$, $\mathbf{G}(\mathbf{r})$, etc., be vector fields in \mathbf{r} -space, where we suppress any time-dependence. We will normally take these vector fields to be real. We define the Fourier transforms of such vector fields by

$$\begin{aligned} \tilde{\mathbf{F}}(\mathbf{k}) &= \int \frac{d^3\mathbf{r}}{(2\pi)^{3/2}} e^{-i\mathbf{k}\cdot\mathbf{r}} \mathbf{F}(\mathbf{r}), \\ \mathbf{F}(\mathbf{r}) &= \int \frac{d^3\mathbf{k}}{(2\pi)^{3/2}} e^{i\mathbf{k}\cdot\mathbf{r}} \tilde{\mathbf{F}}(\mathbf{k}), \end{aligned} \quad (32)$$

where we use tildes to denote quantities in \mathbf{k} -space. The Fourier transformed field $\tilde{\mathbf{F}}$ is complex, but it satisfies the identity,

$$\tilde{\mathbf{F}}(\mathbf{k}) = \tilde{\mathbf{F}}(-\mathbf{k})^*, \quad (33)$$

on account of the reality of \mathbf{F} in \mathbf{r} -space. Of course, the operator ∇ in \mathbf{r} -space is equivalent to the operator $i\mathbf{k}$ in \mathbf{k} -space, and $\partial/\partial\mathbf{k} = \nabla_{\mathbf{k}}$ in \mathbf{k} -space is equivalent to the operator $-i\mathbf{r}$ in \mathbf{r} -space. Finally, we note the Parseval identity,

$$\int d^3\mathbf{r} \mathbf{F}(\mathbf{r}) \cdot \mathbf{G}(\mathbf{r}) = \int d^3\mathbf{k} \tilde{\mathbf{F}}(\mathbf{k}) \cdot \tilde{\mathbf{G}}(\mathbf{k})^*. \quad (34)$$

An arbitrary vector field in \mathbf{k} -space can be decomposed into its longitudinal and transverse parts, which are respectively parallel and perpendicular to \mathbf{k} . That is, we write

$$\tilde{\mathbf{F}}(\mathbf{k}) = \tilde{\mathbf{F}}_{\perp}(\mathbf{k}) + \tilde{\mathbf{F}}_{\parallel}(\mathbf{k}), \quad (35)$$

where

$$\tilde{\mathbf{F}}_{\parallel}(\mathbf{k}) = \frac{1}{k^2} \mathbf{k}[\mathbf{k} \cdot \tilde{\mathbf{F}}(\mathbf{k})], \quad (36)$$

so that

$$\mathbf{k} \cdot \tilde{\mathbf{F}}_{\perp}(\mathbf{k}) = 0, \quad \mathbf{k} \times \tilde{\mathbf{F}}_{\parallel}(\mathbf{k}) = 0. \quad (37)$$

On transforming Eq. (35) back to \mathbf{r} -space, we obtain the longitudinal and transverse components of \mathbf{F} in that space,

$$\mathbf{F}(\mathbf{r}) = \mathbf{F}_\perp(\mathbf{r}) + \mathbf{F}_\parallel(\mathbf{r}) \quad (38)$$

where \mathbf{F}_\perp and \mathbf{F}_\parallel are respectively the inverse Fourier transforms of $\tilde{\mathbf{F}}_\perp$ and $\tilde{\mathbf{F}}_\parallel$. Therefore, by Eq. (37), the transverse and longitudinal fields in \mathbf{r} -space satisfy

$$\nabla \cdot \mathbf{F}_\perp(\mathbf{r}) = 0, \quad \nabla \times \mathbf{F}_\parallel(\mathbf{r}) = 0. \quad (39)$$

The process of projecting out the transverse and longitudinal parts of a vector field is a local operation in \mathbf{k} -space, but it is nonlocal in \mathbf{r} -space. That is, the value of \mathbf{F}_\perp at some point \mathbf{r} depends on the values of \mathbf{F} at all other points of \mathbf{r} -space. In fact, the relation between $\mathbf{F}(\mathbf{r})$ and $\mathbf{F}_\perp(\mathbf{r})$ can be written in the form,

$$F_{\perp i}(\mathbf{r}) = \int d^3\mathbf{r}' \Delta_{ij}^\perp(\mathbf{r} - \mathbf{r}') F_j(\mathbf{r}'), \quad (40)$$

where the kernel of the integral transform, Δ_{ij}^\perp , is called the *transverse delta function*. Notice that it is really a tensor in the indices i, j . We will now derive two useful expressions for the transverse delta function.

One way to project out the transverse part of \mathbf{F} is first to transform \mathbf{F} over to \mathbf{k} -space, then to project out the transverse part in \mathbf{k} -space, then to transform back to \mathbf{r} -space. If we do this, we find

$$\mathbf{F}_\perp(\mathbf{r}) = \int \frac{d^3\mathbf{k}}{(2\pi)^{3/2}} e^{i\mathbf{k}\cdot\mathbf{r}} \left(\mathbf{l} - \frac{\mathbf{k}\mathbf{k}}{k^2} \right) \int \frac{d^3\mathbf{r}'}{(2\pi)^{3/2}} e^{-i\mathbf{k}\cdot\mathbf{r}'} \mathbf{F}(\mathbf{r}'), \quad (41)$$

where \mathbf{l} is the identity tensor and where we use dyadic notation in the term $\mathbf{k}\mathbf{k}/k^2$. Comparing this to Eq. (40), we can easily read off the transverse delta function, which is

$$\Delta_{ij}^\perp(\mathbf{r} - \mathbf{r}') = \int \frac{d^3\mathbf{k}}{(2\pi)^3} e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')} \left(\delta_{ij} - \frac{k_i k_j}{k^2} \right). \quad (42)$$

Thus, the transverse delta function is just the Fourier transform of the transverse projection operator in \mathbf{k} -space.

A second expression for the transverse delta function is obtained by working directly in \mathbf{r} -space. In the decomposition (38), we note that since \mathbf{F}_\parallel is curl-free, it can be represented as the gradient of a scalar, say,

$$\mathbf{F}_\parallel = \nabla f. \quad (43)$$

We can solve for f by taking the divergence of Eq. (38), and noting $\nabla \cdot \mathbf{F}_\perp = 0$. Thus, we have

$$\nabla^2 f = \nabla \cdot \mathbf{F}. \quad (44)$$

We solve this by inverting the Laplacian, to find

$$f(\mathbf{r}) = -\frac{1}{4\pi} \int d^3\mathbf{r}' \frac{\nabla' \cdot \mathbf{F}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} = \frac{1}{4\pi} \int d^3\mathbf{r}' \nabla' \left(\frac{1}{|\mathbf{r} - \mathbf{r}'|} \right) \cdot \mathbf{F}(\mathbf{r}'), \quad (45)$$

where we are careful to write ∇' for derivatives with respect to \mathbf{r}' , and where we have integrated by parts and thrown away boundary terms to get the second integral. We note that in the second integral, we could replace ∇' by $-\nabla$, since it acts only on a function of $\mathbf{r} - \mathbf{r}'$. Finally, by using Eq. (43) to obtain the longitudinal part of \mathbf{F} and subtracting this from \mathbf{F} itself, we obtain the transverse part. We find

$$\mathbf{F}_\perp(\mathbf{r}) = \mathbf{F}(\mathbf{r}) + \frac{1}{4\pi} \nabla \int d^3\mathbf{r}' \nabla \left(\frac{1}{|\mathbf{r} - \mathbf{r}'|} \right) \cdot \mathbf{F}(\mathbf{r}'), \quad (46)$$

which by comparison with Eq. (40) gives

$$\Delta_{ij}^\perp(\mathbf{r} - \mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}') \delta_{ij} + \frac{1}{4\pi} \nabla_i \nabla_j \left(\frac{1}{|\mathbf{r} - \mathbf{r}'|} \right). \quad (47)$$

This can be cast into somewhat more symmetrical form by reexpressing the first term:

$$\Delta_{ij}^\perp(\mathbf{r} - \mathbf{r}') = -\frac{1}{4\pi} \left[\nabla^2 \left(\frac{1}{|\mathbf{r} - \mathbf{r}'|} \right) \delta_{ij} - \nabla_i \nabla_j \left(\frac{1}{|\mathbf{r} - \mathbf{r}'|} \right) \right]. \quad (48)$$

We note one final identity involving transverse and longitudinal vector fields. It is not in general true that the transverse and longitudinal parts of a vector field in ordinary space are perpendicular at a given point, i.e., $\mathbf{F}_\perp(\mathbf{r}) \cdot \mathbf{G}_\parallel(\mathbf{r})$ does not in general vanish. This is because the projection process is nonlocal in \mathbf{r} -space. On the other hand, we obviously have $\tilde{\mathbf{F}}_\perp \cdot \tilde{\mathbf{G}}_\parallel = 0$ at each point of \mathbf{k} -space. Therefore by the Parseval identity (34), we do find an orthogonality of sorts in \mathbf{r} -space, namely

$$\int d^3\mathbf{r} \mathbf{F}_\perp(\mathbf{r}) \cdot \mathbf{G}_\parallel(\mathbf{r}) = 0. \quad (49)$$

In all these operations involving transverse and longitudinal vector fields, we have assumed that the fields in question are sufficiently well behaved to legitimize the various steps. In particular, we assume that all fields fall off rapidly enough at infinity to allow the neglect of boundary terms in the integrations by parts.

Let us now return to the equations of motion (11), impose Coulomb gauge, and eliminate ϕ as discussed above. Since we will be using Coulomb gauge, \mathbf{A} will be purely transverse, $\mathbf{A} = \mathbf{A}_\perp$. We will generally omit the \perp subscript on \mathbf{A} . In Coulomb gauge, Eq. (11a) becomes

$$\nabla^2 \phi(\mathbf{r}, t) = -4\pi \rho(\mathbf{r}, t), \quad (50)$$

which of course is familiar from electrostatics. The solution is

$$\phi(\mathbf{r}, t) = \int d^3\mathbf{r}' \frac{\rho(\mathbf{r}', t)}{|\mathbf{r} - \mathbf{r}'|}, \quad (51)$$

which shows quite clearly that ϕ is not a dynamical variable, since it is determined by the supposedly given charge density ρ . Notice, however, that we are not doing electrostatics here, but rather electrodynamics, and that Eq. (51) gives ϕ at one spatial point and time in terms of ρ at all other spatial points at the same time. In other words, there is no retardation in Eq. (51), and it would

seem that this equation violates causality. As it turns out, however, there is no problem, because ϕ itself is not measurable. What is measurable is \mathbf{E} , which is given in terms of the potentials by Eq. (9a). Notice that in Coulomb gauge, the two terms of Eq. (9a) are longitudinal and transverse, so that

$$\mathbf{E}_{\parallel} = -\nabla\phi, \quad \mathbf{E}_{\perp} = -\frac{1}{c}\frac{\partial\mathbf{A}}{\partial t}. \quad (52)$$

Because of Eq. (51), the longitudinal electric field \mathbf{E}_{\parallel} does involve instantaneous interactions. This fact does not violate causality, however, because it turns out that the transverse electric field \mathbf{E}_{\perp} also has an instantaneous (non-retarded) part, and the two cancel one another out when the total electric field \mathbf{E} is computed. Thus, the total field \mathbf{E} , which is what is physically measurable, only involves retarded interactions.

Coulomb gauge also simplifies Eq. (11b) somewhat. In fact, since the left-hand side is purely transverse, the right-hand side must be also, and we suspect that the longitudinal part of $-4\pi\mathbf{J}/c$ must be cancelled by $(1/c)\nabla(\partial\phi/\partial t)$, which of course is purely longitudinal. Indeed, we have

$$\begin{aligned} \nabla\left(\frac{\partial\phi}{\partial t}\right) &= \nabla\int d^3\mathbf{r}'\frac{1}{|\mathbf{r}-\mathbf{r}'|}\frac{\partial\rho(\mathbf{r}',t)}{\partial t} = -\nabla\int d^3\mathbf{r}'\frac{1}{|\mathbf{r}-\mathbf{r}'|}\nabla'\cdot\mathbf{J}(\mathbf{r}',t) \\ &= -\nabla\int d^3\mathbf{r}'\nabla\left(\frac{1}{|\mathbf{r}-\mathbf{r}'|}\right)\cdot\mathbf{J}(\mathbf{r}',t), \end{aligned} \quad (53)$$

where we use the continuity equation and integration by parts. But the final expression can be recognized as 4π times $\mathbf{J}_{\parallel}(\mathbf{r},t)$. Altogether, Eq. (11b) becomes

$$\nabla^2\mathbf{A} - \frac{1}{c^2}\frac{\partial^2\mathbf{A}}{\partial t^2} = -\frac{4\pi}{c}\mathbf{J}_{\perp}. \quad (54)$$

Since the process of projecting out the transverse part of the current is nonlocal, the vector potential and hence the transverse electric field contain an instantaneous component, as mentioned above. Equation (54) is the one equation of motion for the true dynamical variables ($\mathbf{A} = \mathbf{A}_{\perp}$) of the electromagnetic field. Expressed in terms of electric and magnetic fields, it is equivalent to

$$\nabla\times\mathbf{B} = \frac{4\pi}{c}\mathbf{J}_{\perp} + \frac{1}{c}\frac{\partial\mathbf{E}_{\perp}}{\partial t}. \quad (55)$$

Now for some general comments on Coulomb gauge. Coulomb gauge has certain advantages and disadvantages. The main disadvantage is that it is not relativistically covariant. This will not bother us too much, because our first applications will involve interactions of the field with nonrelativistic matter. But if we were also to treat the matter relativistically, then the loss of manifest covariance would be more annoying. Of course, even though the choice of gauge is not covariant, all physical results must be properly covariant in a relativistic theory, so no physical principles are violated by an unfortunate choice of gauge. But clearly, a proper relativistic treatment of the electromagnetic field should employ a covariant choice of gauge, which would normally be Lorentz gauge. The general development of relativistic quantum electrodynamics is normally carried out in Lorentz gauge. Furthermore, a proper covariant treatment requires one to face up to the pseudo-degrees of

freedom in the electromagnetic field (the scalar potential and the longitudinal vector potential) in a more careful manner than we have done. This subject involves some new formalism which is outside the scope of this course.

Another disadvantage of the Coulomb gauge is that retardation effects are somewhat obscured, since the potentials ϕ and \mathbf{A} both contain an instantaneous (nonretarded) part. On the other hand, retardation effects are small for low velocity systems in which the light transit time is small in comparison to typical times scales of the mechanical motion of the system. In such cases, the dominant electromagnetic interactions of the particles of the system are just the instantaneous Coulomb interactions, which are captured by a nonretarded potential as in Eq. (51). For such systems, the Coulomb gauge is an advantage, because of the simple way in which the Coulomb potential emerges. Of course, we are accustomed to using such potentials in the nonrelativistic Schrödinger equation for atoms, molecules and nuclei. In such systems, retardation effects can be regarded as relativistic corrections (for example, in helium they are of the same order as the fine structure effects).

Another advantage of Coulomb gauge is pedagogical, in that it allows one to quantize the electromagnetic field in the most painless manner, without invoking the extra formalism needed for covariant quantization schemes. For this reason Coulomb gauge is usually used in introductory treatments.

We have now eliminated ϕ and \mathbf{A}_{\parallel} from our equations of motion, resulting in Eq. (55). Let us now eliminate them from the Lagrangian,

$$L = \int d^3\mathbf{r} \left(\frac{E^2 - B^2}{8\pi} - \rho\phi + \frac{1}{c} \mathbf{J} \cdot \mathbf{A} \right). \quad (56)$$

First we work on the term involving E^2 . We write $\mathbf{E} = \mathbf{E}_{\parallel} + \mathbf{E}_{\perp}$ according to Eq. (52), and notice that when we integrate E^2 over all space the cross terms cancel according to Eq. (49). As for the integral of E_{\parallel}^2 , this becomes

$$\frac{1}{8\pi} \int d^3\mathbf{r} E_{\parallel}^2 = \frac{1}{8\pi} \int d^3\mathbf{r} |\nabla\phi|^2 = -\frac{1}{8\pi} \int d^3\mathbf{r} \phi \nabla^2 \phi = \frac{1}{2} \int d^3\mathbf{r} \rho\phi, \quad (57)$$

and we see that it cancels one half of the $-\rho\phi$ term in the Lagrangian. The $\rho\phi$ term which remains can in turn be expressed purely in terms of the charge density,

$$-\frac{1}{2} \int d^3\mathbf{r} \rho(\mathbf{r}, t) \phi(\mathbf{r}, t) = -\frac{1}{2} \int d^3\mathbf{r} d^3\mathbf{r}' \frac{\rho(\mathbf{r}, t) \rho(\mathbf{r}', t)}{|\mathbf{r} - \mathbf{r}'|}. \quad (58)$$

Insofar as the dynamics of the field are concerned, this term is actually just a constant, and could be dropped from the Lagrangian with no effect on the equations of motion. We will keep it, however, because it takes on a dynamical significance when we introduce the degrees of freedom corresponding to the matter. Next, as for the integral of E_{\perp}^2 and B^2 , we leave these as they are because \mathbf{E}_{\perp} and \mathbf{B} can be expressed purely in terms of $\mathbf{A} = \mathbf{A}_{\perp}$. Finally, in the term involving $\mathbf{J} \cdot \mathbf{A}$, we note that

\mathbf{J} could be replaced by \mathbf{J}_\perp , since the integral of $\mathbf{J}_\parallel \cdot \mathbf{A}$ would vanish in accordance with Eq. (49). Altogether, the Lagrangian becomes

$$L = \int d^3\mathbf{r} \left(\frac{E_\perp^2 - B^2}{8\pi} + \frac{1}{c} \mathbf{J}_\perp \cdot \mathbf{A} \right) - \frac{1}{2} \int d^3\mathbf{r} d^3\mathbf{r}' \frac{\rho(\mathbf{r}, t)\rho(\mathbf{r}', t)}{|\mathbf{r} - \mathbf{r}'|}, \quad (59)$$

where \mathbf{J}_\perp could be replaced by \mathbf{J} .

Next we introduce the matter degrees of freedom. We suppose we have n nonrelativistic particles, with masses m_α , charges q_α , and positions \mathbf{x}_α , $\alpha = 1, \dots, n$. We use the symbol \mathbf{x} for the particle positions, partly to avoid confusion with the variable \mathbf{r} which labels the point at which the various fields are evaluated. Of course, the \mathbf{x}_α are dynamical variables, whereas \mathbf{r} is not (it merely labels the dynamical variables of the field). The charge and current density produced by the matter are

$$\begin{aligned} \rho(\mathbf{r}, t) &= \sum_\alpha q_\alpha \delta(\mathbf{r} - \mathbf{x}_\alpha(t)), \\ \mathbf{J}(\mathbf{r}, t) &= \sum_\alpha q_\alpha \dot{\mathbf{x}}_\alpha \delta(\mathbf{r} - \mathbf{x}_\alpha(t)), \end{aligned} \quad (60)$$

which satisfy the continuity equation (18). When we substitute these into the Lagrangian (59), we find

$$\int d^3\mathbf{r} \mathbf{J} \cdot \mathbf{A} = \sum_\alpha q_\alpha \dot{\mathbf{x}}_\alpha \cdot \mathbf{A}(\mathbf{x}_\alpha), \quad (61)$$

and

$$\frac{1}{2} \int d^3\mathbf{r} d^3\mathbf{r}' \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} = \frac{1}{2} \sum_{\alpha\beta} \frac{q_\alpha q_\beta}{|\mathbf{x}_\alpha - \mathbf{x}_\beta|}. \quad (62)$$

The final expression is, of course, the instantaneous Coulomb energy of the assemblage of particles, but it has the embarrassing feature of including the infinite self-energies. To obtain a result which is finite, we simply throw the diagonal terms ($\alpha = \beta$) away.

Finally, we introduce the nonrelativistic kinetic energy of the particles into the Lagrangian. The Lagrangian can then be expressed as a sum of three terms, one for the free field, one for the matter, and one for the interaction,

$$L = L_{\text{em}} + L_{\text{matter}} + L_{\text{int}}, \quad (63)$$

where

$$L_{\text{em}} = \int d^3\mathbf{r} \left(\frac{E_\perp^2 - B^2}{8\pi} \right), \quad (64a)$$

$$L_{\text{matter}} = \frac{1}{2} \sum_\alpha m_\alpha |\dot{\mathbf{x}}_\alpha|^2 - \sum_{\alpha < \beta} \frac{q_\alpha q_\beta}{|\mathbf{x}_\alpha - \mathbf{x}_\beta|}, \quad (64b)$$

$$L_{\text{int}} = \sum_\alpha \frac{q_\alpha}{c} \dot{\mathbf{x}}_\alpha \cdot \mathbf{A}(\mathbf{x}_\alpha). \quad (64c)$$

The nonrelativistic expression for the kinetic energy of the particles is the only term in this Lagrangian which destroys the relativistic covariance of the equations of motion. (Of course, the gauge

is not covariant.) We could create a relativistically covariant theory by using the relativistic expression for the kinetic energy, but the resulting theory would still not be correct physically, because it would have no mechanism for the creation or annihilation of particles. To incorporate the latter effects it is necessary to introduce fields for the matter as well as the radiation. As it stands, however, the Lagrangian (64) gives an adequate description of the interaction of matter with radiation in many nonrelativistic systems.

We now wish to go over to a Hamiltonian description. According to Eq. (26), the first step should be the computation of the momentum field $\boldsymbol{\pi}(\mathbf{r})$ which is conjugate to $\mathbf{A}(\mathbf{r})$, which in the present case would appear to be

$$\boldsymbol{\pi}(\mathbf{r}) = \frac{\delta L}{\delta \dot{\mathbf{A}}(\mathbf{r})} = \frac{1}{4\pi c^2} \dot{\mathbf{A}}(\mathbf{r}) = -\frac{1}{4\pi c} \mathbf{E}_\perp. \quad (65)$$

This is in fact correct, but it is not as obvious as it seems because the field \mathbf{A} is purely transverse and in computing the functional derivative in Eq. (26) we should employ only transverse variations $\delta \mathbf{A}$. Rather than worrying about a “transverse functional derivative,” however, we will take another approach, which is to go over to \mathbf{k} -space where the transverse nature of the fields is easier to deal with. Another advantage of \mathbf{k} -space is that the free field, at least, is decoupled in \mathbf{k} -space, so that it is easy to identify the normal modes. The normal modes are defined as the variables which evolve at a definite frequency. In quantum mechanics, they represent the oscillators whose excitations are identified with photons.

We will also switch over to periodic boundary conditions in a box of volume $V = L^3$. This is, of course, a pedagogical crutch. Periodic boundary conditions mean that all spatial integrals introduced above are to be reinterpreted as integrals only over the volume V , all quantities in \mathbf{r} -space are periodic and can be represented by discrete Fourier series in \mathbf{k} -space, and all Fourier series are taken over a lattice in \mathbf{k} -space with a fundamental cell size of $(\Delta k)^3 = (2\pi/L)^3$. The use of periodic boundary conditions means that our fields will be represented by a discrete set of variables in \mathbf{k} -space, so ordinary partial derivatives can be used instead of functional derivatives. Of course, proper physical results are recovered by taking the limit $V \rightarrow \infty$.

We will use the following formalism for periodic boundary conditions. If $\mathbf{F}(\mathbf{r})$ is a typical real field in \mathbf{r} -space, presumed periodic, we will expand it according to

$$\mathbf{F}(\mathbf{r}) = \frac{1}{\sqrt{V}} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}} \mathbf{F}_{\mathbf{k}}, \quad (66)$$

which serves to define the quantities $\mathbf{F}_{\mathbf{k}}$. We omit the tilde used previously for Fourier transformed quantities, since the \mathbf{k} -subscript indicates the space in question. The factor $1/\sqrt{V}$ is introduced for convenience. The inverse Fourier transform is

$$\mathbf{F}_{\mathbf{k}} = \frac{1}{\sqrt{V}} \int_V d^3\mathbf{r} e^{-i\mathbf{k}\cdot\mathbf{r}} \mathbf{F}(\mathbf{r}), \quad (67)$$

where the integral is taken over the volume V . These conventions cause the Parseval identity [see

Eq. (34)] to take on a particularly convenient form,

$$\int_V d^3\mathbf{r} \mathbf{F}(\mathbf{r}) \cdot \mathbf{G}(\mathbf{r}) = \sum_{\mathbf{k}} \mathbf{F}_{\mathbf{k}} \cdot \mathbf{G}_{\mathbf{k}}^*. \quad (68)$$

Finally, when we let $V \rightarrow \infty$ and go over to the continuous case, we make the transcriptions,

$$\sum_{\mathbf{k}} \rightarrow \frac{V}{(2\pi)^3} \int d^3\mathbf{k}, \quad (69)$$

$$\mathbf{F}_{\mathbf{k}} \rightarrow \frac{(2\pi)^{3/2}}{\sqrt{V}} \tilde{\mathbf{F}}(\mathbf{k}), \quad (70)$$

and

$$\delta_{\mathbf{k},\mathbf{k}'} \rightarrow \frac{(2\pi)^3}{V} \delta(\mathbf{k} - \mathbf{k}'). \quad (71)$$

Let us now return to our fields of interest and expand them in Fourier series, as appropriate in the case of periodic boundary conditions. For the transverse vector potential, we have

$$\mathbf{A}(\mathbf{r}) = \frac{1}{\sqrt{V}} \sum_{\mathbf{k}} \mathbf{A}_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}}. \quad (72)$$

We note that since $\mathbf{A}(\mathbf{r})$ is real, we have

$$\mathbf{A}_{\mathbf{k}} = \mathbf{A}_{-\mathbf{k}}^*. \quad (73)$$

Similarly, the transverse electric field and magnetic field are given in terms of the coefficients $\mathbf{A}_{\mathbf{k}}$:

$$\begin{aligned} \mathbf{E}_{\perp}(\mathbf{r}) &= -\frac{1}{c} \dot{\mathbf{A}}(\mathbf{r}) = -\frac{1}{c} \frac{1}{\sqrt{V}} \sum_{\mathbf{k}} \dot{\mathbf{A}}_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}}, \\ \mathbf{B}(\mathbf{r}) &= \nabla \times \mathbf{A}(\mathbf{r}) = \frac{1}{\sqrt{V}} \sum_{\mathbf{k}} i(\mathbf{k} \times \mathbf{A}_{\mathbf{k}}) e^{i\mathbf{k}\cdot\mathbf{r}}. \end{aligned} \quad (74)$$

Next, we introduce polarization vectors, which are transverse unit vectors which span the plane perpendicular to \mathbf{k} . These vectors are real for linear polarization and complex for circular or elliptic polarizations, in the usual way in optics. We write $\epsilon_{\mathbf{k}\mu}$, for $\mu = 1, 2$, for the two polarization vectors, which we think of as attached to each lattice point of \mathbf{k} -space. These vectors satisfy the orthonormality and completeness relations,

$$\epsilon_{\mathbf{k}\mu}^* \cdot \epsilon_{\mathbf{k}\mu'} = \delta_{\mu\mu'}, \quad (75a)$$

$$\hat{\mathbf{k}} \cdot \epsilon_{\mathbf{k}\mu} = 0, \quad (75b)$$

$$\sum_{\mu=1}^2 \epsilon_{\mathbf{k}\mu} \epsilon_{\mathbf{k}\mu}^* + \hat{\mathbf{k}} \hat{\mathbf{k}} = \mathbf{I}, \quad (75c)$$

so that if \mathbf{X} is an arbitrary vector, we have

$$\mathbf{X} = \sum_{\mu=1}^2 X_{\mu} \epsilon_{\mathbf{k}\mu} + X_k \hat{\mathbf{k}}, \quad (76)$$

where

$$X_\mu = \boldsymbol{\epsilon}_{\mathbf{k}\mu}^* \cdot \mathbf{X}, \quad X_k = \hat{\mathbf{k}} \cdot \mathbf{X}. \quad (77)$$

In particular, since $\mathbf{A}_{\mathbf{k}}$ is transverse, it can be expanded in terms of the two polarization vectors,

$$\mathbf{A}_{\mathbf{k}} = \sum_{\mu=1}^2 A_{\mathbf{k}\mu} \boldsymbol{\epsilon}_{\mathbf{k}\mu}, \quad (78)$$

where the expansion coefficients are given by

$$A_{\mathbf{k}\mu} = \boldsymbol{\epsilon}_{\mathbf{k}\mu}^* \cdot \mathbf{A}_{\mathbf{k}}. \quad (79)$$

Thus, we can now write the vector potential in the form,

$$\mathbf{A}(\mathbf{r}) = \frac{1}{\sqrt{V}} \sum_{\lambda} \boldsymbol{\epsilon}_{\lambda} A_{\lambda} e^{i\mathbf{k} \cdot \mathbf{r}}, \quad (80)$$

where we let λ be an abbreviation for the compound index (\mathbf{k}, μ) . As we will see, the index λ labels the modes of the electromagnetic field (a photon wave number and polarization).

It is now easy to express the Lagrangian (64) in terms of the quantities $A_{\mathbf{k}\mu} = A_{\lambda}$. For example, with the assistance of the Parseval identity (68), the term involving the spatial integral of E_{\perp}^2 becomes

$$\frac{1}{8\pi} \int_V d^3\mathbf{r} E_{\perp}^2 = \frac{1}{8\pi c^2} \sum_{\mathbf{k}} \dot{\mathbf{A}}_{\mathbf{k}} \cdot \dot{\mathbf{A}}_{\mathbf{k}}^* = \frac{1}{8\pi c^2} \sum_{\lambda} |\dot{A}_{\lambda}|^2. \quad (81)$$

Similarly, for the magnetic term we have

$$\begin{aligned} \frac{1}{8\pi} \int_V d^3\mathbf{r} B^2 &= \frac{1}{8\pi} \sum_{\mathbf{k}} (\mathbf{k} \times \mathbf{A}_{\mathbf{k}}) \cdot (\mathbf{k} \times \mathbf{A}_{\mathbf{k}}^*) = \frac{1}{8\pi} \sum_{\mathbf{k}} k^2 |\mathbf{A}_{\mathbf{k}}|^2 - |\mathbf{k} \cdot \mathbf{A}_{\mathbf{k}}|^2 \\ &= \frac{1}{8\pi c^2} \sum_{\lambda} \omega_k^2 |A_{\lambda}|^2, \end{aligned} \quad (82)$$

where we use $\mathbf{k} \cdot \mathbf{A}_{\mathbf{k}} = 0$ (since \mathbf{A} is transverse), and where we define

$$\omega_k = ck. \quad (83)$$

Thus, the free-field Lagrangian becomes

$$L_{\text{em}} = \frac{1}{8\pi c^2} \sum_{\lambda} (|\dot{A}_{\lambda}|^2 - \omega_k^2 |A_{\lambda}|^2). \quad (84)$$

We see that it has the form of a sum of uncoupled, complex harmonic oscillators.

We now pause to count the number of degrees of freedom of our Lagrangian. The quantities $\mathbf{A}_{\mathbf{k}}$ or A_{λ} contain the same information as the original field $\mathbf{A}(\mathbf{r})$. Thus, there are two complex numbers $A_{\lambda} = A_{\mathbf{k}\mu}$ for each lattice point in \mathbf{k} -space, or four real numbers. But in view of Eq. (73), the quantities $A_{-\mathbf{k}\mu}$ at the opposite lattice point $-\mathbf{k}$ are not independent of the quantities $A_{\mathbf{k}\mu}$, so we have an average of one complex number or two real numbers per lattice point in \mathbf{k} -space. In

the limit $V \rightarrow \infty$, we get two real numbers per point in \mathbf{k} -space, which is the count of the degrees of freedom of the electromagnetic field. But since functions in \mathbf{r} -space and \mathbf{k} -space are invertible Fourier transforms of each other, it is appropriate to say that we also have two (real) degrees of freedom per point of \mathbf{r} -space.

It is important not to overcount the degrees of freedom of the electromagnetic field. The \mathbf{k} -sums such as those in Eq. (84) count the degrees of freedom twice, in the sense that the same four (real) degrees of freedom occur in the term involving $A_{\mathbf{k}\mu}$ and $A_{-\mathbf{k}\mu}$. In order to obtain sums which count the degrees of freedom only once, we divide \mathbf{k} -space into two halves, and write

$$\sum_{\mathbf{k}(\frac{1}{2})} = \text{sum over } \frac{1}{2} \text{ of } \mathbf{k}\text{-space.} \quad (85)$$

Similarly, we define

$$\sum_{\lambda(\frac{1}{2})} = \sum_{\mathbf{k}(\frac{1}{2})} \sum_{\mu}. \quad (86)$$

We also require our polarization vectors to satisfy,[†]

$$\epsilon_{\mathbf{k}\mu} = \epsilon_{-\mathbf{k}\mu}^*, \quad (87)$$

which is possible because both \mathbf{k} and $-\mathbf{k}$ have the same transverse plane. This allows us to write,

$$A_{\mathbf{k}\mu} = A_{-\mathbf{k}\mu}^*. \quad (88)$$

Finally, if $\lambda = (\mathbf{k}, \mu)$, then we define

$$-\lambda = (-\mathbf{k}, \mu), \quad (89)$$

so that

$$A_{\lambda} = A_{-\lambda}^*. \quad (90)$$

With these notational changes, let us reexpress the Lagrangian in terms of sums which run only over half of \mathbf{k} -space. Consider, for example, the first term of Eq. (84). We have

$$\sum_{\lambda} |\dot{A}_{\lambda}|^2 = \sum_{\lambda(\frac{1}{2})} (\dot{A}_{\lambda} \dot{A}_{\lambda}^* + \dot{A}_{-\lambda} \dot{A}_{-\lambda}^*). \quad (91)$$

But in view of Eq. (88), the two terms in the sum on the right hand side are equal, and we have

$$\sum_{\lambda} |\dot{A}_{\lambda}|^2 = 2 \sum_{\lambda(\frac{1}{2})} |\dot{A}_{\lambda}|^2. \quad (92)$$

A similar argument applies to the second term in Eq. (85). Altogether, we have

$$L_{\text{em}} = \frac{1}{4\pi c^2} \sum_{\lambda(\frac{1}{2})} (|\dot{A}_{\lambda}|^2 - \omega_k^2 |A_{\lambda}|^2). \quad (93)$$

[†] This convention is convenient for the next several steps of the argument. Later we relax this requirement, and allow the polarization vectors at \mathbf{k} and $-\mathbf{k}$ to be independently chosen.

As for the vector potential itself, it can also be written as a sum over half of \mathbf{k} -space,

$$\mathbf{A}(\mathbf{r}) = \frac{1}{\sqrt{V}} \sum_{\lambda(\frac{1}{2})} \left(\epsilon_{\lambda} A_{\lambda} e^{i\mathbf{k}\cdot\mathbf{r}} + \epsilon_{\lambda}^* A_{\lambda}^* e^{-i\mathbf{k}\cdot\mathbf{r}} \right). \quad (94)$$

At this point it would be straightforward to work out the Euler-Lagrange equations, working with the four real degrees of freedom per \mathbf{k} -point in the half-space. These are the real and imaginary parts of $A_{\mathbf{k}\mu}$ for $\mu = 1, 2$, which are the q 's or generalized coordinates for the electromagnetic field. However, it is somewhat more economical to work directly with the complex numbers $A_{\mathbf{k}\mu}$. For this we require a formalism for complex coordinates in classical mechanics.

Consider a Lagrangian $L(x_1, x_2, \dot{x}_1, \dot{x}_2)$ which is a function of two real coordinates x_1, x_2 and their velocities, and define complex coordinates by

$$\begin{aligned} X &= \frac{x_1 + ix_2}{\sqrt{2}}, & x_1 &= \frac{X + X^*}{\sqrt{2}}, \\ X^* &= \frac{x_1 - ix_2}{\sqrt{2}}, & x_2 &= \frac{X - X^*}{i\sqrt{2}}. \end{aligned} \quad (95)$$

Then it is not hard to show that if we transform the Lagrangian to the complex coordinates, so that $L = L(X, X^*, \dot{X}, \dot{X}^*)$, then the Euler-Lagrange equations become

$$\begin{aligned} \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{X}} \right) &= \frac{\partial L}{\partial X}, \\ \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{X}^*} \right) &= \frac{\partial L}{\partial X^*}. \end{aligned} \quad (96)$$

These equations are somewhat more economical, because the second is the complex conjugate of the first.

As for the canonical momenta, it is convenient to define

$$\begin{aligned} P &= \frac{\partial L}{\partial \dot{X}} = \frac{p_1 + ip_2}{\sqrt{2}}, \\ P^* &= \frac{\partial L}{\partial \dot{X}^*} = \frac{p_1 - ip_2}{\sqrt{2}}, \end{aligned} \quad (97)$$

where p_1 and p_2 are the usual (real) momenta conjugate to x_1 and x_2 . In this way, p_1 and p_2 are related to the real and imaginary parts of P just as x_1 and x_2 are related to the real and imaginary parts of X . Equation (97) follows from the relations,

$$\begin{aligned} \frac{\partial}{\partial X} &= \frac{1}{\sqrt{2}} \left(\frac{\partial}{\partial x_1} - i \frac{\partial}{\partial x_2} \right), \\ \frac{\partial}{\partial X^*} &= \frac{1}{\sqrt{2}} \left(\frac{\partial}{\partial x_1} + i \frac{\partial}{\partial x_2} \right). \end{aligned} \quad (98)$$

When the Hamiltonian is expressed in terms of the complex coordinates and momenta, we find

$$H = p_1 \dot{x}_1 + p_2 \dot{x}_2 - L = P \dot{X}^* + P^* \dot{X} - L. \quad (99)$$

Finally, Hamilton's equations become

$$\dot{X} = \frac{\partial H}{\partial P^*}, \quad \dot{P} = -\frac{\partial H}{\partial X^*}. \quad (100)$$

Let us now return to our Lagrangian in the forms indicated by Eqs. (93), (94) and (64), and let us compute the canonical momenta. As for the momenta corresponding to the complex coordinates A_λ , we have

$$\pi_\lambda = \frac{\partial L}{\partial \dot{A}_\lambda^*} = \frac{1}{4\pi c^2} \dot{A}_\lambda, \quad (101)$$

which initially is defined only over the first half of \mathbf{k} -space. But then we extend the definition to the other half by demanding

$$\pi_{-\lambda} = \pi_\lambda^*, \quad (102)$$

which makes $\pi_\lambda = \dot{A}_\lambda/4\pi c^2$ true for all λ . We now define the field $\boldsymbol{\pi}(\mathbf{r})$ in real space by

$$\boldsymbol{\pi}(\mathbf{r}) = \frac{1}{\sqrt{V}} \sum_\lambda \boldsymbol{\epsilon}_\lambda \pi_\lambda e^{i\mathbf{k}\cdot\mathbf{r}} = \frac{1}{4\pi c^2} \dot{\mathbf{A}}(\mathbf{r}) = -\frac{1}{4\pi c} \mathbf{E}_\perp(\mathbf{r}). \quad (103)$$

which is the best we can do in writing down the field conjugate to $\mathbf{A}(\mathbf{r})$, given that $\mathbf{A}(\mathbf{r})$ is transverse. This is the same definition we guessed earlier in Eq. (65). Note that $\boldsymbol{\pi}(\mathbf{r})$ is also transverse. As for the momentum conjugate to the (real) particle positions \mathbf{x}_α , we have

$$\mathbf{p}_\alpha = \frac{\partial L}{\partial \dot{\mathbf{x}}_\alpha} = m_\alpha \dot{\mathbf{x}}_\alpha + \frac{q_\alpha}{c} \mathbf{A}(\mathbf{x}_\alpha). \quad (104)$$

It is now straightforward to compute the Euler-Lagrange equations in the complex coordinates A_λ , and to show that they give Maxwell's equation in the form (54). Similarly, the equations of motion for the particles turn into the Newton-Lorentz equations,

$$m_\alpha \ddot{\mathbf{x}} = q_\alpha \left(\mathbf{E} + \frac{\dot{\mathbf{x}}_\alpha}{c} \times \mathbf{B} \right). \quad (105)$$

A demonstration of these facts will be left as an exercise.

Instead, we will move on to the Hamiltonian, which according to Eq. (99) must be given by

$$H = \sum_{\lambda(\frac{1}{2})} (\pi_\lambda \dot{A}_\lambda^* + \pi_\lambda^* \dot{A}_\lambda) + \sum_\alpha \mathbf{p}_\alpha \cdot \dot{\mathbf{x}}_\alpha - L. \quad (106)$$

But by Eqs. (88) and (102), we can write the first sum as an unrestricted sum over all of \mathbf{k} -space,

$$\sum_{\lambda(\frac{1}{2})} (\pi_\lambda \dot{A}_\lambda^* + \pi_\lambda^* \dot{A}_\lambda) = \sum_\lambda \pi_\lambda \dot{A}_\lambda^* = \frac{1}{4\pi c^2} \sum_\lambda |\dot{A}_\lambda|^2. \quad (107)$$

As for the second sum in Eq. (106), by Eq. (104) it becomes

$$\sum_\alpha \mathbf{p}_\alpha \cdot \dot{\mathbf{x}}_\alpha = \sum_\alpha m_\alpha |\dot{\mathbf{x}}_\alpha|^2 + \sum_\alpha \frac{q_\alpha}{c} \dot{\mathbf{x}}_\alpha \cdot \mathbf{A}(\mathbf{x}_\alpha), \quad (108)$$

of which the first term is twice the kinetic energy and the second is the same as the interaction Lagrangian (64c). Combining these and subtracting L , Eq. (106) becomes

$$H = H_{\text{em}} + H_{\text{matter}}, \quad (109)$$

where

$$H_{\text{em}} = \frac{1}{8\pi c^2} \sum_{\lambda} (|\dot{A}_{\lambda}|^2 + \omega_k^2 |A_{\lambda}|^2) = \int d^3\mathbf{r} \left(\frac{E_{\perp}^2 + B^2}{8\pi} \right), \quad (110a)$$

$$H_{\text{matter}} = \frac{1}{2} \sum_{\alpha} m_{\alpha} |\dot{\mathbf{x}}_{\alpha}|^2 + \sum_{\alpha < \beta} \frac{q_{\alpha} q_{\beta}}{|\mathbf{x}_{\alpha} - \mathbf{x}_{\beta}|}. \quad (110b)$$

In this form it is apparent that the Hamiltonian consists of the total energy of the system, including the energy of the transverse electromagnetic field (both electric and magnetic) in H_{em} , the kinetic energy of the particles in the first term of H_{matter} , and the longitudinal electric field energy in the second term of H_{matter} , which is otherwise the electrostatic potential energy of interaction of the particles. There is no “interaction” energy.

On the other hand, it is customary in classical Hamiltonian mechanics to express the Hamiltonian as a function of the q 's and p 's, not the q 's and \dot{q} 's. If we transform over to the momentum variables π_{λ} and \mathbf{p}_{α} , we have

$$H_{\text{em}} = \frac{1}{2} \sum_{\lambda} \left(4\pi c^2 |\pi_{\lambda}|^2 + \frac{\omega_k^2}{4\pi c^2} |A_{\lambda}|^2 \right), \quad (111a)$$

$$H_{\text{matter}} = \sum_{\alpha} \frac{1}{2m_{\alpha}} \left[\mathbf{p}_{\alpha} - \frac{q_{\alpha}}{c} \mathbf{A}(\mathbf{x}_{\alpha}) \right]^2 + \sum_{\alpha < \beta} \frac{q_{\alpha} q_{\beta}}{|\mathbf{x}_{\alpha} - \mathbf{x}_{\beta}|}. \quad (111b)$$

The factors of $4\pi c^2$ are somewhat unsymmetrical in the first term (the transverse field energy), which is why field theorists (see Sakurai, *Advanced Quantum Mechanics*, p. 12) generally prefer Heaviside-Lorentz units (which get rid of the 4π 's in the Lagrangian and Hamiltonian). We will also get rid of these factors, by performing a canonical transformation,

$$\pi'_{\lambda} = \sqrt{4\pi c^2} \pi_{\lambda}, \quad A'_{\lambda} = \frac{1}{\sqrt{4\pi c^2}} A_{\lambda}, \quad (112)$$

which saves some writing. The field Hamiltonian now becomes

$$H_{\text{em}} = \frac{1}{2} \sum_{\lambda} \left(|\pi'_{\lambda}|^2 + \omega_k^2 |A'_{\lambda}|^2 \right) = \sum_{\lambda(\frac{1}{2})} \left(|\pi'_{\lambda}|^2 + \omega_k^2 |A'_{\lambda}|^2 \right), \quad (113)$$

where the second sum is taken only over half of \mathbf{k} -space.

Let us now compute the equations of evolution of the free field, by applying Hamilton's equations in the form (100) to H_{em} . The half-sum version of H_{em} is more convenient for this purpose. We find

$$\begin{aligned} \dot{A}'_{\lambda} &= \frac{\partial H_{\text{em}}}{\partial \pi'_{\lambda}} = \pi'_{\lambda}, \\ \dot{\pi}'_{\lambda} &= -\frac{\partial H_{\text{em}}}{\partial A'_{\lambda}} = -\omega_k^2 A'_{\lambda}, \end{aligned} \quad (114)$$

which are, of course, the equations of a (complex) harmonic oscillator. These equations have the solution,

$$\begin{aligned} A'_\lambda(t) &= C_1 e^{-i\omega_k t} + C_2 e^{+i\omega_k t}, \\ \pi'_\lambda(t) &= -i\omega_k C_1 e^{-i\omega_k t} + i\omega_k C_2 e^{+i\omega_k t}, \end{aligned} \quad (115)$$

where C_1 and C_2 are complex constants. We regard a term which goes as $e^{-i\omega t}$ ($e^{+i\omega t}$) as a positive (negative) frequency term.

At this point it is convenient to introduce normal mode variables, or simply “normal variables,” which not only evolve at a single frequency, but which also disentangle the coupling of the \mathbf{k} and $-\mathbf{k}$ lattice points inherent in the coefficients A_λ and π_λ . To motivate these definitions, we first note that A'_λ has both positive and negative frequencies, as shown by the free field solutions (115), but the linear combination $A'_\lambda + i\pi'_\lambda/\omega_k$ evolves only at the positive frequency ω_k . This motivates the definition of the normal variable a_λ below, and then we obtain three more definitions by taking complex conjugations and making the replacement $\lambda \rightarrow -\lambda$. Altogether, we make the definitions,

$$a_\lambda = N \left(A'_\lambda + \frac{i}{\omega_k} \pi'_\lambda \right), \quad (116a)$$

$$a_\lambda^* = N \left(A_\lambda'^* - \frac{i}{\omega_k} \pi_\lambda'^* \right), \quad (116b)$$

$$a_{-\lambda} = N \left(A_\lambda'^* + \frac{i}{\omega_k} \pi_\lambda'^* \right), \quad (116c)$$

$$a_{-\lambda}^* = N \left(A'_\lambda - \frac{i}{\omega_k} \pi'_\lambda \right), \quad (116d)$$

where N is a normalization constant to be determined. We notice that a_λ looks like $q + ip$ for a harmonic oscillator, which is essentially an annihilation operator in quantum mechanics. Indeed, we will see that the normal variables a_λ and a_λ^* become annihilation and creation operators upon quantization. Equations (116) specify an invertible transformation between the variables $(A'_\lambda, A_\lambda'^*, \pi'_\lambda, \pi_\lambda'^*)$ and $(a_\lambda, a_\lambda^*, a_{-\lambda}, a_{-\lambda}^*)$. The inverse transformation is

$$\begin{aligned} A'_\lambda &= \frac{1}{2N} (a_\lambda + a_{-\lambda}^*), \\ \pi'_\lambda &= \frac{\omega_k}{2iN} (a_\lambda - a_{-\lambda}^*), \end{aligned} \quad (117)$$

which is augmented by $A'_\lambda = A_{-\lambda}'^*$, $\pi'_\lambda = \pi_{-\lambda}'^*$.

The free-field equation of motion for a_λ is

$$\dot{a}_\lambda = -i\omega_k a_\lambda, \quad (118)$$

and the solution is

$$a_\lambda(t) = a_\lambda(0) e^{-i\omega_k t}, \quad (119)$$

so a_λ does evolve at the positive frequency ω_k . The variable $a_{-\lambda}$ also evolves at the positive frequency ω_k , while a_λ^* and $a_{-\lambda}^*$ evolve at the negative frequency $-\omega_k$.

To interpret the normal variables a_λ , let us suppose that only one of the a_λ 's is non-zero, say the one for $\lambda = \lambda_0 = (\mathbf{k}_0, \mu_0)$, and let us ask what the corresponding electromagnetic fields are. Of course, if $a_{\lambda_0} \neq 0$, then $a_{\lambda_0}^* \neq 0$ also. For simplicity, we work with the free-field case. First, by Eq. (117), the given time evolution of a_{λ_0} gives

$$A'_{\lambda_0}(t) = \frac{C}{2N} e^{-i\omega_0 t}, \quad A'_{-\lambda_0}(t) = \frac{C^*}{2N} e^{+i\omega_0 t}, \quad (120)$$

where $C = a_{\lambda_0}(0)$ is a complex constant. Thus, two of the A_λ 's are nonzero, those for $\lambda = \pm\lambda_0$, and there are two terms in the sum (80). These give

$$\mathbf{A}(\mathbf{r}, t) = \frac{1}{2N} \sqrt{\frac{4\pi c^2}{V}} [C \boldsymbol{\epsilon}_{\lambda_0} e^{i(\mathbf{k}_0 \cdot \mathbf{r} - \omega_0 t)} + C^* \boldsymbol{\epsilon}_{\lambda_0}^* e^{-i(\mathbf{k}_0 \cdot \mathbf{r} - \omega_0 t)}], \quad (121)$$

which is a light wave of polarization $\boldsymbol{\epsilon}_{\lambda_0}$ propagating in the \mathbf{k}_0 direction. The constant $C = a_{\lambda_0}(0)$ determines the amplitude and overall phase of the light wave. Thus, the normal variable a_λ can be thought of as a mode amplitude for mode λ . If we had allowed only the normal variable a_λ for $\lambda = -\lambda_0$ to be nonzero, we would have found a light wave propagating in the $-\mathbf{k}_0$ direction. It is in this sense that the transformation (116) disentangles the coupling between modes \mathbf{k} and $-\mathbf{k}$ which is inherent in the quantities A_λ . We note that unlike the relation $A_\lambda = A_{-\lambda}^*$, the mode variables a_λ and $a_{-\lambda}$ are independent, i.e.,

$$a_\lambda \neq a_{-\lambda}^*. \quad (122)$$

Let us now transform the electromagnetic field Hamiltonian over to the normal variables. For the normalization it is convenient to take

$$N = \sqrt{\frac{\omega_k}{2\hbar}}. \quad (123)$$

It may seem strange to introduce \hbar into a classical calculation, but we do it anyway because this choice for N makes a_λ and a_λ^* dimensionless, as required for creation and annihilation operators. Next, we note that

$$a_\lambda^* a_\lambda + a_{-\lambda}^* a_{-\lambda} = \frac{2N^2}{\omega_k^2} (|\pi'_\lambda|^2 + \omega_k^2 |A'_\lambda|^2), \quad (124)$$

so that the Hamiltonian becomes,

$$H_{\text{em}} = \sum_\lambda \hbar \omega_k a_\lambda^* a_\lambda. \quad (125)$$

The a_λ 's are complex phase space coordinates, and are not themselves canonical q 's and p 's in the usual sense in classical mechanics. But it is easy to introduce such q 's and p 's; we simply write

$$\begin{aligned} a_\lambda &= \frac{Q_\lambda + iP_\lambda}{\sqrt{2\hbar}}, \\ a_\lambda^* &= \frac{Q_\lambda - iP_\lambda}{\sqrt{2\hbar}}, \end{aligned} \quad (126)$$

so that the Q_λ 's and P_λ 's are real q 's and p 's associated with each mode of the electromagnetic field. In terms of these variables, the field Hamiltonian becomes

$$H_{\text{em}} = \frac{1}{2} \sum_{\lambda} \omega_k (P_\lambda^2 + Q_\lambda^2). \quad (127)$$

This takes care of the field Hamiltonian. The matter Hamiltonian (111b) involves the vector potential, which we should also express in terms of normal variables. Using Eqs. (94), (112) and (117), we find

$$\mathbf{A}(\mathbf{r}) = \sqrt{\frac{2\pi\hbar c^2}{V}} \sum_{\lambda} \frac{1}{\sqrt{\omega_k}} (\boldsymbol{\epsilon}_\lambda a_\lambda e^{i\mathbf{k}\cdot\mathbf{r}} + \boldsymbol{\epsilon}_\lambda^* a_\lambda^* e^{-i\mathbf{k}\cdot\mathbf{r}}). \quad (128)$$

We remark that although the definitions (116) were motivated by the free field solutions, we use those definitions even in the case of fields interacting with matter. We can also express the fields \mathbf{E}_\perp and \mathbf{B} in terms of normal variables. As for \mathbf{E}_\perp , we use Eq. (103) to express \mathbf{E}_\perp in terms of $\boldsymbol{\pi}$ and then in terms of the quantities π_λ , and we then use Eqs. (112) and (117) to express the π_λ in terms of the normal variables. This is easier than differentiating \mathbf{A} with respect to t , because the equations for \dot{a}_λ are not as simple as Eq. (118) in the case of the interacting field. The result is

$$\mathbf{E}_\perp(\mathbf{r}) = \frac{1}{c} \sqrt{\frac{2\pi\hbar c^2}{V}} \sum_{\lambda} \sqrt{\omega_k} (i\boldsymbol{\epsilon}_\lambda a_\lambda e^{i\mathbf{k}\cdot\mathbf{r}} - i\boldsymbol{\epsilon}_\lambda^* a_\lambda^* e^{-i\mathbf{k}\cdot\mathbf{r}}). \quad (129)$$

As for the magnetic field, it is obtained simply from $\mathbf{B} = \nabla \times \mathbf{A}$:

$$\mathbf{B}(\mathbf{r}) = \sqrt{\frac{2\pi\hbar c^2}{V}} \sum_{\lambda} \frac{1}{\sqrt{\omega_k}} [i(\mathbf{k} \times \boldsymbol{\epsilon}_\lambda) a_\lambda e^{i\mathbf{k}\cdot\mathbf{r}} - i(\mathbf{k} \times \boldsymbol{\epsilon}_\lambda^*) a_\lambda^* e^{-i\mathbf{k}\cdot\mathbf{r}}]. \quad (130)$$

There is one more topic involving the classical matter-field system which will be of interest to us, namely, the constants of motion. Quite generally, constants of motion are associated with symmetries of a Lagrangian or Hamiltonian; in the case of Lagrangians, the relation between continuous symmetries and constants of motion is called *Noether's theorem*. The theorem is easy. Suppose we have a Lagrangian $L = L(q, \dot{q})$, and let $q_i \rightarrow q_i + \delta q_i$ be an infinitesimal symmetry. The increment δq_i is in general a function of the q 's, so we write $\delta q_i = \epsilon F_i(q)$ for it, where the ϵ is a reminder that the increment is small. To say that the replacement $q_i \rightarrow q_i + \delta q_i$ is a symmetry means that the Lagrangian is invariant under such a replacement.

For example, a translation of particle positions \mathbf{x}_α is specified by $\mathbf{x}_\alpha \rightarrow \mathbf{x}_\alpha + \mathbf{a}$, where \mathbf{a} is the displacement vector (the same for all particles α , since the system as a whole is to be displaced). We obtain an infinitesimal translation by replacing \mathbf{a} by $\epsilon \mathbf{a}$, where ϵ is a small scale factor, so that $\delta \mathbf{x}_\alpha = \epsilon \mathbf{a}$. Similarly, a rotation of a particle system is specified by $\mathbf{x}_\alpha \rightarrow \mathbf{R} \mathbf{x}_\alpha$, where $\mathbf{R} = \exp(\theta \hat{\mathbf{n}} \cdot \mathbf{J})$ is a rotation (in the notation of Notes 11). The rotation becomes infinitesimal if θ is infinitesimal, so that $\mathbf{R} = \mathbf{I} + \theta \hat{\mathbf{n}} \cdot \mathbf{J}$. Then the infinitesimal symmetry is $\mathbf{x}_\alpha \rightarrow \mathbf{x}_\alpha + \theta \hat{\mathbf{n}} \times \mathbf{x}_\alpha$, or $\delta \mathbf{x}_\alpha = \theta \hat{\mathbf{n}} \times \mathbf{x}_\alpha$.

To return to the general case, suppose the Lagrangian is invariant under the replacement $q_i \rightarrow q_i + \epsilon F_i(q)$. Then we have

$$\delta L = \epsilon \sum_i \left(\frac{\partial L}{\partial q_i} F_i + \frac{\partial L}{\partial \dot{q}_i} \frac{dF_i}{dt} \right) = 0. \quad (131)$$

But the second term can be rewritten,

$$\frac{\partial L}{\partial \dot{q}_i} \frac{dF_i}{dt} = \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} F_i \right) - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) F_i, \quad (132)$$

so that when the Euler-Lagrange equations (2) are used, two terms cancel, and we are left with

$$\frac{d}{dt} \sum_i \frac{\partial L}{\partial \dot{q}_i} F_i = 0. \quad (133)$$

In other words, the quantity

$$\sum_i p_i F_i(q) \quad (134)$$

is a constant of motion.

Let us apply Noether's theorem to the Lagrangian (64). First we deal with displacements. As noted above, the particles transform according to $\mathbf{x}_\alpha \rightarrow \mathbf{x}_\alpha + \mathbf{a}$. How do the fields transform? Last semester in our discussion of translation operators on wave functions, we argued that a wave function would transform under translations according to $\psi(\mathbf{r}) \rightarrow \psi(\mathbf{r} - \mathbf{a})$, because this causes concentrations of probability to move forward by \mathbf{a} (as we expect under an active displacement). The same argument applies to classical fields; if we want a wave packet of light waves, for example, to move forward by \mathbf{a} under a translation, then we must have $\mathbf{A}(\mathbf{r}) \rightarrow \mathbf{A}(\mathbf{r} - \mathbf{a})$ (and similarly for \mathbf{E} and \mathbf{B}).

Let us check that the Lagrangian (64) is invariant under translations, so defined. First, the term L_{em} in Eq. (64a) is certainly invariant, because after the displacement we can write $\mathbf{r}' = \mathbf{r} - \mathbf{a}$ to change dummy variables of integration. Next, L_{matter} in Eq. (64b) is also invariant, because the displacement vector \mathbf{a} is constant, so $\dot{\mathbf{x}}_\alpha$ does not change under the displacement, and because the vector difference $\mathbf{x}_\alpha - \mathbf{x}_\beta$ in the potential energy is also invariant under the displacement. Finally, the term L_{int} in Eq. (64c) is also invariant, because the displacement shifts the argument of \mathbf{A} backward by \mathbf{a} under the transformation of the field, but then the fact that the argument itself is \mathbf{x}_α , which gets shifted forward by \mathbf{a} , cancels out the shift. If this is not clear, we can go back to an expression for L_{int} obtained from Eqs. (60) and (61),

$$L_{\text{int}} = \int d^3\mathbf{r} \sum_\alpha q_\alpha \delta(\mathbf{r} - \mathbf{x}_\alpha) \dot{\mathbf{x}}_\alpha \cdot \mathbf{A}(\mathbf{r}), \quad (135)$$

which shows that after $\mathbf{x}_\alpha \rightarrow \mathbf{x}_\alpha + \mathbf{a}$ and $\mathbf{A}(\mathbf{r}) \rightarrow \mathbf{A}(\mathbf{r} - \mathbf{a})$ we can change dummy variables of integration, $\mathbf{r}' = \mathbf{r} - \mathbf{a}$, to recover the original form for L_{int} .

Thus, the whole Lagrangian is invariant under translations; and in particular, it is invariant under infinitesimal translations, which we obtain by replacing \mathbf{a} by $\epsilon\mathbf{a}$. Under an infinitesimal translation, we have $\delta\mathbf{x}_\alpha = \epsilon\mathbf{a}$, and

$$\mathbf{A}(\mathbf{r} - \epsilon\mathbf{a}) = \mathbf{A}(\mathbf{r}) - \epsilon\mathbf{a} \cdot \nabla\mathbf{A}, \quad (136)$$

so that $\delta\mathbf{A} = -\epsilon\mathbf{a} \cdot \nabla\mathbf{A}$. Therefore by Noether's theorem (134) we have the conserved quantity,

$$\sum_\alpha \mathbf{a} \cdot \mathbf{p}_\alpha - \int d^3\mathbf{r} \mathbf{a} \cdot \nabla\mathbf{A} \cdot \boldsymbol{\pi}. \quad (137)$$

[In the final expression, $\mathbf{a} \cdot \nabla \mathbf{A} \cdot \boldsymbol{\pi}$ means $(\mathbf{a} \cdot \nabla \mathbf{A}) \cdot \boldsymbol{\pi} = a_i A_{j,i} \pi_j$; the rule is that ∇ only acts on the operand immediately following, and the \cdot only connects adjacent operands, unless overruled by parentheses.]

Since the vector \mathbf{a} is arbitrary, we actually obtain a vector of conserved quantities,

$$\mathbf{P} = \sum_{\alpha} \mathbf{p}_{\alpha} - \int d^3\mathbf{r} \nabla \mathbf{A} \cdot \boldsymbol{\pi}. \quad (138)$$

We define \mathbf{P} to be the momentum of the matter-field system, because it is the generator of translations. If we use Eqs. (103) and (104) for \mathbf{p}_{α} and $\boldsymbol{\pi}$, we obtain another form for the momentum,

$$\mathbf{P} = \sum_{\alpha} \left[m_{\alpha} \dot{\mathbf{x}}_{\alpha} + \frac{q_{\alpha}}{c} \mathbf{A}(\mathbf{x}_{\alpha}) \right] + \frac{1}{4\pi c} \int d^3\mathbf{r} \nabla \mathbf{A} \cdot \mathbf{E}_{\perp}. \quad (139)$$

We will call the second term in this equation $\mathbf{P}_{\text{trans}}$, the momentum associated with the transverse electric field, since it can be written in terms of the integral of $\mathbf{E}_{\perp} \times \mathbf{B}$:

$$\mathbf{P}_{\text{trans}} = \frac{1}{4\pi c} \int d^3\mathbf{r} \nabla \mathbf{A} \cdot \mathbf{E}_{\perp} = \frac{1}{4\pi c} \int d^3\mathbf{r} \mathbf{E}_{\perp} \times \mathbf{B}. \quad (140)$$

To prove the second equality, we first expand the cross product,

$$\mathbf{E}_{\perp} \times \mathbf{B} = \mathbf{E}_{\perp} \times (\nabla \times \mathbf{A}) = \nabla \mathbf{A} \cdot \mathbf{E}_{\perp} - \mathbf{E}_{\perp} \cdot \nabla \mathbf{A}, \quad (141)$$

where the first term gives us the first integral in Eq. (140). As for the second term in Eq. (141), it vanishes on integration, for if we write out the i -th component of this term, we have

$$(-\mathbf{E}_{\perp} \cdot \nabla \mathbf{A})_i = -E_{\perp j} A_{i,j} = -(E_{\perp j} A_i)_{,j} + E_{\perp j,j} A_i. \quad (142)$$

Of these terms, the first is an exact derivative which vanishes on integration, and the second vanishes because $E_{\perp j,j} = \nabla \cdot \mathbf{E}_{\perp} = 0$ (since \mathbf{E}_{\perp} is transverse).

The second term in the sum in Eq. (139) can also be reexpressed. We call this term \mathbf{P}_{long} , the momentum associated with the longitudinal electric field, because it can be written in the form,

$$\mathbf{P}_{\text{long}} = \sum_{\alpha} \frac{q_{\alpha}}{c} \mathbf{A}(\mathbf{x}_{\alpha}) = \frac{1}{4\pi c} \int d^3\mathbf{r} \mathbf{E}_{\parallel} \times \mathbf{B}. \quad (143)$$

To prove the second equality, we first rewrite the sum,

$$\begin{aligned} \sum_{\alpha} \frac{q_{\alpha}}{c} \mathbf{A}(\mathbf{x}_{\alpha}) &= \frac{1}{c} \int d^3\mathbf{r} \mathbf{A}(\mathbf{r}) \sum_{\alpha} q_{\alpha} \delta(\mathbf{r} - \mathbf{x}_{\alpha}) \\ &= \frac{1}{c} \int d^3\mathbf{r} \rho(\mathbf{r}) \mathbf{A}(\mathbf{r}) = -\frac{1}{4\pi c} \int d^3\mathbf{r} \nabla^2 \phi \mathbf{A}. \end{aligned} \quad (144)$$

Next, we rewrite the integral in Eq. (143),

$$\frac{1}{4\pi c} \int d^3\mathbf{r} \mathbf{E}_{\parallel} \times \mathbf{B} = -\frac{1}{4\pi c} \int d^3\mathbf{r} \nabla \phi \times (\nabla \times \mathbf{A}) = -\frac{1}{4\pi c} \int d^3\mathbf{r} (\nabla \mathbf{A} \cdot \nabla \phi - \nabla \phi \cdot \nabla \mathbf{A}). \quad (145)$$

In the final integral, the first term integrates to zero, for we have

$$(\nabla \mathbf{A} \cdot \nabla \phi)_i = A_{j,i} \phi_{,j} = (A_{j,i} \phi)_{,j} - A_{j,ij} \phi, \quad (146)$$

in which the first term is an exact derivative which integrates to zero, and the second term involves $A_{j,ij} = (\nabla \cdot \mathbf{A})_{,i}$ which vanishes because \mathbf{A} is transverse. As for the second term in the final integral in Eq. (145), we transform this according to

$$(\nabla \phi \cdot \nabla \mathbf{A})_i = \phi_{,j} A_{i,j} = (\phi_j A_i)_{,j} - \phi_{,jj} A_i, \quad (147)$$

of which the first term in the final expression integrates to zero, while the second gives us the final integral in Eq. (144). This proves Eq. (143).

Adding \mathbf{P}_{long} and $\mathbf{P}_{\text{trans}}$, we obtain the momentum \mathbf{P}_{em} of the electromagnetic field,

$$\mathbf{P}_{\text{em}} = \mathbf{P}_{\text{long}} + \mathbf{P}_{\text{trans}} = \frac{1}{4\pi c} \int d^3\mathbf{r} \mathbf{E} \times \mathbf{B}, \quad (148)$$

which is the usual expression from electromagnetic theory. Finally, the total momentum of the matter-field system is the kinetic momentum of the particles plus the momentum of the field,

$$\mathbf{P} = \sum_{\alpha} m_{\alpha} \dot{\mathbf{x}}_{\alpha} + \frac{1}{4\pi c} \int d^3\mathbf{r} \mathbf{E} \times \mathbf{B}. \quad (149)$$

Let us now apply Noether's theorem to find the conserved quantity associated with rotations of the matter-field system, which will be the total angular momentum of the system. Under a rotation specified by a proper rotation matrix \mathbf{R} , the particle positions transform according to $\mathbf{x}_{\alpha} \rightarrow \mathbf{R}\mathbf{x}_{\alpha}$. How do the fields transform? First let us recall that quantum wave functions transform under rotations according to $\psi(\mathbf{r}) \rightarrow \psi(\mathbf{R}^{-1}\mathbf{r})$ [see Eq. (15.13)], which as discussed in Notes 15 is the necessary transformation law to make concentrations of probability move forward under the active rotation. As for classical electromagnetic fields, we expect wave packets of light waves also to be rotated in a forward manner under an active rotation, so we expect the point of application \mathbf{r} of fields such as $\mathbf{A}(\mathbf{r})$ or $\mathbf{E}(\mathbf{r})$ to be replaced by $\mathbf{R}^{-1}\mathbf{r}$ under a rotation. But these fields are vector fields; how do the vectors themselves transform? A simple geometrical picture will convince you that the direction of these fields should be rotated by \mathbf{R} (not \mathbf{R}^{-1}) under an active transformation. Thus, we will require a vector field such as \mathbf{A} to transform according to

$$\mathbf{A}(\mathbf{r}) \rightarrow \mathbf{R}\mathbf{A}(\mathbf{R}^{-1}\mathbf{r}), \quad (150)$$

and similarly for \mathbf{E} , \mathbf{B} , etc.

Let us now check to see that the matter-field Lagrangian (64) is invariant under rotations. Certainly the term L_{em} is invariant, because, for example, the term E_{\perp}^2 in the integrand transforms under rotations into

$$E_{\perp}^2 = [\mathbf{E}_{\perp}(\mathbf{r})]^2 \rightarrow [\mathbf{R}\mathbf{E}(\mathbf{R}^{-1}\mathbf{r})]^2 = [\mathbf{E}_{\perp}(\mathbf{R}^{-1}\mathbf{r})]^2, \quad (151)$$

after which a change of variable of integration $\mathbf{r}' = \mathbf{R}^{-1}\mathbf{r}$ restores the original form of L_{em} . Likewise the term L_{matter} is easily seen to be invariant under $\mathbf{x}_{\alpha} \rightarrow \mathbf{R}\mathbf{x}_{\alpha}$, since it only involves dot products of

vectors such as \mathbf{x}_α , $\dot{\mathbf{x}}_\alpha$. Finally, L_{int} is also invariant under rotations, which is perhaps most easily seen in the form (135): we rotate \mathbf{x}_α and $\mathbf{A}(\mathbf{r})$, and then make the change of variable $\mathbf{r}' = \mathbf{R}^{-1}\mathbf{r}$ in the integral, which restores the original form of L_{int} .

Since the Lagrangian is invariant under all rotations, it is invariant in particular under infinitesimal rotations. Let $\mathbf{R} = \mathbf{1} + \theta \hat{\mathbf{n}} \cdot \mathbf{J}$ be an infinitesimal rotation, so that under $\mathbf{x}_\alpha \rightarrow \mathbf{R}\mathbf{x}_\alpha$ we have $\delta\mathbf{x}_\alpha = \theta \hat{\mathbf{n}} \times \mathbf{x}_\alpha$, and under $\mathbf{A}(\mathbf{r}) \rightarrow \mathbf{R}\mathbf{A}(\mathbf{R}^{-1}\mathbf{r})$ we have

$$\mathbf{R}\mathbf{A}(\mathbf{R}^{-1}\mathbf{r}) = (\mathbf{1} + \theta \hat{\mathbf{n}} \cdot \mathbf{J})\mathbf{A}(\mathbf{r} - \theta \hat{\mathbf{n}} \times \mathbf{r}) = \mathbf{A}(\mathbf{r}) + \delta\mathbf{A}(\mathbf{r}), \quad (152)$$

or,

$$\delta\mathbf{A} = \theta [\hat{\mathbf{n}} \times \mathbf{A} - (\hat{\mathbf{n}} \times \mathbf{r}) \cdot \nabla \mathbf{A}]. \quad (153)$$

Then by applying Noether's theorem we obtain the conserved quantity,

$$\sum_{\alpha} \mathbf{p}_{\alpha} \cdot (\hat{\mathbf{n}} \times \mathbf{x}_{\alpha}) + \int d^3\mathbf{r} [(\hat{\mathbf{n}} \times \mathbf{A}) \cdot \boldsymbol{\pi} - (\hat{\mathbf{n}} \times \mathbf{r}) \cdot \nabla \mathbf{A} \cdot \boldsymbol{\pi}]. \quad (154)$$

But $\hat{\mathbf{n}}$ is an arbitrary unit vector, so we actually have a vector of conserved quantities,

$$\begin{aligned} \mathbf{J} &= \sum_{\alpha} \mathbf{x}_{\alpha} \times \mathbf{p}_{\alpha} + \int d^3\mathbf{r} [\mathbf{A} \times \boldsymbol{\pi} - \mathbf{r} \times (\nabla \mathbf{A} \cdot \boldsymbol{\pi})] \\ &= \sum_{\alpha} \left[m_{\alpha} \mathbf{x}_{\alpha} \times \dot{\mathbf{x}}_{\alpha} + \frac{q_{\alpha}}{c} \mathbf{x}_{\alpha} \times \mathbf{A}(\mathbf{x}_{\alpha}) \right] \\ &\quad + \frac{1}{4\pi c} \int d^3\mathbf{r} [\mathbf{r} \times (\nabla \mathbf{A} \cdot \mathbf{E}_{\perp}) - \mathbf{A} \times \mathbf{E}_{\perp}], \end{aligned} \quad (155)$$

where we use Eqs. (103) and (104) for \mathbf{p}_{α} and $\boldsymbol{\pi}$. We define \mathbf{J} to be the total angular momentum of the matter-field system,[†] because it is the generator of rotations.

We will call the two terms in the final integral in Eq. (155) the *orbital* and *spin* angular momentum of the classical electromagnetic field, respectively, and we will write

$$\mathbf{L} = \frac{1}{4\pi c} \int d^3\mathbf{r} \mathbf{r} \times (\nabla \mathbf{A} \cdot \mathbf{E}_{\perp}), \quad (156)$$

$$\mathbf{S} = -\frac{1}{4\pi c} \int d^3\mathbf{r} \mathbf{A} \times \mathbf{E}_{\perp}. \quad (157)$$

The justification for this terminology will be given later; for now we merely remark that the orbital angular momentum is associated with the rotation of the point of application \mathbf{r} of the field, whereas the spin is associated with the rotation of the direction of the field itself. We also remark that the decomposition of the angular momentum of the classical electromagnetic field, or of a photon in the quantum theory, is a tricky subject. For example, as we will see, the obvious candidates for orbital and spin angular momentum operators are not even defined on the state space of a single photon; only the total angular momentum is defined.

[†] Obviously not to be confused with the current density.

In any case, we now define $\mathbf{J}_{\text{trans}}$ as the sum of \mathbf{L} and \mathbf{S} , and assert that it can be written in terms of the transverse electric field,

$$\mathbf{J}_{\text{trans}} = \mathbf{L} + \mathbf{S} = \frac{1}{4\pi c} \int d^3\mathbf{r} \mathbf{r} \times (\mathbf{E}_{\perp} \times \mathbf{B}). \quad (158)$$

Similarly, we define \mathbf{J}_{long} as the second term in the sum in Eq. (155), and assert that it can be written in terms of the longitudinal electric field,

$$\mathbf{J}_{\text{long}} = \sum_{\alpha} \frac{q_{\alpha}}{c} \mathbf{x}_{\alpha} \times \mathbf{A}(\mathbf{x}_{\alpha}) = \frac{1}{4\pi c} \int d^3\mathbf{r} \mathbf{r} \times (\mathbf{E}_{\parallel} \times \mathbf{B}). \quad (159)$$

The proof of the second equalities in Eqs. (158) and (159) is left as an exercise.

Accepting these results, we can see that the sum of \mathbf{J}_{long} and $\mathbf{J}_{\text{trans}}$ is the angular momentum \mathbf{J}_{em} of the electromagnetic field,

$$\mathbf{J}_{\text{em}} = \mathbf{J}_{\text{long}} + \mathbf{J}_{\text{trans}} = \frac{1}{4\pi c} \int d^3\mathbf{r} \mathbf{r} \times (\mathbf{E} \times \mathbf{B}), \quad (160)$$

and that the total angular momentum of the system is the kinetic angular momentum of the matter plus the field angular momentum,

$$\mathbf{J} = \sum_{\alpha} m_{\alpha} \mathbf{x}_{\alpha} \times \dot{\mathbf{x}}_{\alpha} + \frac{1}{4\pi c} \int d^3\mathbf{r} \mathbf{r} \times (\mathbf{E} \times \mathbf{B}). \quad (161)$$