Obtaining the Hamiltonian of the electromagnetic field from the Lagrangian and finding photon solutions

Mads Juul Damgaard

January 24, 2017

Abstract

The path integral in quantum mechanics is introduced for discretized lattices obeying Schrödinger's equation, and the formalism is extended to include first order time derivatives in the Lagrangian. Then a prescription for converting the Lagrangian of a quantum lattice to the corresponding Hamiltonian is derived. This result will be used in an attempt to obtain a Hamiltonian of the free electromagnetic field. After introducing a Coulomb-like gauge restriction for the free field, a theory for photons will be developed, which for each polarization and wave vector describes the photons as decoupled two-dimensional harmonic oscillators. Lastly, a new phase operator for the photons will be derived.

1 Introduction

The semiclassical approach to quantum electrodynamics (QED) is a very powerful approach and has provided us with many great results. One example of this is the interaction between light and matter, where an electromagnetic wave is treated as a time-dependent potential in the Hamiltonian. Another great example is when calculating the energies and orbitals of atoms, which can be approximated very accurately simply by having them interact via a Coulomb potential. We know, however, that a full QED requires the fields to be quantum mechanical as well, so that the electrons are able to excite photons in the field. Of course, making a quantum system of fields rather than a quantum system of particles takes some care. After all, a field configuration contains much more information than the position vectors of a configuration of particles. The positions of N particles in three dimensions can be described by $3\times N$ real numbers, and the corresponding quantum system can be described by a function of $3\times N$ coordinates. But a field is described by a real number in every point in space, and therefore one might fear the a quantum field would have to be described by a function on an infinite-dimensional space. Hopefully, there is a solution to this problem.

The conventional way of quantizing field is fittingly called Quantum Field Theory (QFT). In this thesis, I will take a different approach to the problem, though. In conventional QFT, the aim is to be able to obtain a perturbation series for every possible scattering, providing the scattering amplitudes, and in that formalism, it does not really matter what the underlying Hamiltonian of the system is. In this thesis we will, however, deal with the latter subject and try to find a Hamiltonian of a quantum field, starting from its Lagrangian. The emphasis will be on the electromagnetic field, and the formalism developed in this theory will provide us with a very useful description of photons.

Originally, I wanted the thesis to include electrons in the theory as well as the electromagnetic field, but there is not room for it. I will continue the work after this is handed in, however, and hope to propose a full QED theory. This is why it is mentioned throughout the thesis that I might make a longer rendition of this paper in a near future.

2 Phonons in a lattice

To understand how to quantize a field, let us first look at a discrete lattice in quantum mechanics. For simplicity we well consider the one-dimensional case with a chain of N atoms. Each atom is allowed to move in one direction, and we will let x_j denote this displacement. The wave function for the field will

now be a function of all these x_j's so that $|\psi\rangle = \psi : \mathbb{R}^N \to \mathbb{C}$. We then define operators $\hat{x}_j$ and $\hat{p}_j$ by

$$\hat{x}_j |\psi\rangle = (x_1, \ldots, x_j, \ldots, x_N) \mapsto x_j \psi(x_1, \ldots, x_j, \ldots, x_N), \tag{1}$$

$$\hat{p}_j |\psi\rangle = (x_1, \ldots, x_j, \ldots, x_N) \mapsto -i\hbar \frac{\partial}{\partial x_j} \psi(x_1, \ldots, x_j, \ldots, x_N). \tag{2}$$

We will keep $\hbar$ in this section, but in later sections, we will use natural units instead. Now let the Hamiltonian of the system be given by

$$H = \sum_{j=1}^{N} \frac{\hat{p}_j^2}{2m} + \frac{1}{2} K (\hat{x}_{j+1} - \hat{x}_j)^2 + \frac{1}{2} m \Omega^2 \hat{x}_j^2, \tag{3}$$

where K, m and Ω are non-negative real numbers, and where $\hat{x}_{N+1} = \hat{x}_1$ so that we have periodic boundary conditions. The Hamiltonian describes a system of N harmonic oscillators (with mass equal to 0 and $\omega_0^2 = \Omega^2$) where each pair of neighbor atoms is coupled with a spring. We will now analyze the solutions to this Hamiltonian following the derivation by Lancaster and Blundell [1]. First we introduce Fourier transforms of $\hat{x}_j$ and $\hat{p}_j$, namely

$$\tilde{x}_k = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} \hat{x}_j e^{-ikja}, \tag{4}$$

$$\tilde{p}_k = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} \hat{p}_j e^{-ikja}. \tag{5}$$

Here we have introduced a spacing between the atoms given by a. The inverse transforms are given by

$$\hat{x}_j = \frac{1}{\sqrt{N}} \sum_{k} \tilde{x}_k e^{ikja}, \tag{6}$$

$$\hat{p}_j = \frac{1}{\sqrt{N}} \sum_{k} \tilde{p}_k e^{ikja}, \tag{7}$$

where k takes values of $2\pi n/(Na)$ with $n \in \mathbb{Z}$, $-\frac{N}{2} < n \leq \frac{N}{2}$. Note the identities

$$\sum_j e^{ikja} = \sum_j e^{i\frac{2\pi n}{N} j} = N\delta_{k,l}, \tag{8}$$

$$\sum_k e^{ikja} = \sum_n e^{i\frac{2\pi n}{N} j} = N\delta_{j,l}, \tag{9}$$

with $l \in \{Nz | z \in \mathbb{Z}\}$. With these, it is easy to plug in and see that relations (6) and (7) hold. Now we can start to express H in terms of the transformed operators. We have

$$\begin{aligned} \sum_j \hat{x}_j^2 &= \sum_j \Big(\frac{1}{\sqrt{N}} \sum_k \tilde{x}_k e^{ikja} \Big) \Big(\frac{1}{\sqrt{N}} \sum_{k'} \tilde{x}_{k'} e^{ik'ja} \Big) \\ &= \frac{1}{N} \sum_j \sum_k \sum_{k'} \tilde{x}_k \tilde{x}_{k'} e^{i(k+k')ja} \\ &= \sum_k \tilde{x}_k \tilde{x}_{-k}. \end{aligned} \tag{10}$$

A similar calculation for $\hat{p}_j$ yields

$$\sum_j \hat{p}_j^2 = \sum_k \tilde{p}_k \tilde{p}_{-k}. \tag{11}$$

Finally, the coupling term is given by

$$\begin{aligned} \sum_j (\hat{x}_{j+1} - \hat{x}_j)^2 &= \frac{1}{N} \sum_j \sum_k \sum_{k'} \tilde{x}_k \tilde{x}_{k'} e^{i(k+k')ja} (e^{ika} - 1)(e^{ik'a} - 1) \\ &= \sum_k \tilde{x}_k \tilde{x}_{-k} (2 - 2\cos ka) \\ &= \sum_k \tilde{x}_k \tilde{x}_{-k} \big(4 \sin^2 \frac{ka}{2} \big). \end{aligned} \tag{12}$$

We can now express H as

$$H = \sum_k \frac{1}{2m}\tilde{p}_k\tilde{p}_{-k} + \frac{1}{2}m\omega_k^2\tilde{x}_k\tilde{x}_{-k}, \tag{13}$$

where $\omega_k^2 = \frac{4K}{m}\sin^2\frac{ka}{2} + \Omega^2$. This can be simplified even further by introducing the creation and annihilation operators,

$$\hat{a}_k = \sqrt{\frac{m\omega_k}{2\hbar}}\big(\tilde{x}_k + \frac{i}{m\omega_k}\tilde{p}_k\big), \tag{14}$$

$$\hat{a}_k^\dagger = \sqrt{\frac{m\omega_k}{2\hbar}}\big(\tilde{x}_{-k} - \frac{i}{m\omega_k}\tilde{p}_{-k}\big), \tag{15}$$

with inversions

$$\tilde{x}_k = \sqrt{\frac{\hbar}{2m\omega_k}}\big(\hat{a}_k + \hat{a}_{-k}^\dagger\big), \tag{16}$$

$$\tilde{p}_k = -i\sqrt{\frac{\hbar m\omega_k}{2}}\big(\hat{a}_k - \hat{a}_{-k}^\dagger\big). \tag{17}$$

The Hamiltonian now becomes

$$H = \sum_k \frac{\hbar\omega_k}{2}\big(\hat{a}_k\hat{a}_k^\dagger + \hat{a}_{-k}^\dagger\hat{a}_{-k}\big) = \sum_k \frac{\hbar\omega_k}{2}\big(\hat{a}_k\hat{a}_k^\dagger + \hat{a}_k^\dagger\hat{a}_k\big), \tag{18}$$

where, to get the last equality, we have used that $\omega_k = \omega_{-k}$ and changed indices in the last term. If we calculate the commutation relation between $\tilde{x}_k$ and $\tilde{p}_{k'}$, we will get $[\tilde{x}_k, \tilde{p}_{k'}] = \delta_{k,-k'}$, and it follows that $[\hat{a}_k, \hat{a}_{k'}^\dagger] = \delta_{k,k'}$. Now the Hamiltonian can finally be written as

$$H = \sum_k \hbar\omega_k\left(\hat{a}_k^\dagger\hat{a}_k + \frac{1}{2}\right). \tag{19}$$

This Hamiltonian looks almost exactly like the Hamiltonian for a set of decoupled harmonic oscillators where every oscillator represents a mode of the field. Since these are quantized modes in a lattice, let us call them phonons. It would be nice to be able to show that the state vector can expanded in terms of eigenstates $|n_1, n_2, \ldots, n_N\rangle$, where each n_i is a non-negative integer representing the occupation number of phonons with distinct wave vector k_i. Actually, to prove this, we only need to show that there exists a state with a nonvanishing norm, call it $|0\rangle$, where $\hat{a}_k|0\rangle = 0$ for all allowed values of k. It would be nice to see a proof of this for all N and for all parameters, but unfortunately I have not been able to find any. We will assume, however, in the following that such a state can always be found. In this case, we can take $|0\rangle$ to be normalized, i.e. $\langle 0\,|\,0\rangle = 1$, and define

$$|n_1, n_2, \ldots, n_N\rangle = \frac{1}{\sqrt{n_1!n_2!\cdots n_N!}}(\hat{a}_{k_1}^\dagger)^{n_1}(\hat{a}_{k_2}^\dagger)^{n_2}\cdots(\hat{a}_{k_N}^\dagger)^{n_N}|0\rangle\,. \tag{20}$$

It immediately follows for any k_i that

$$\hat{a}_{k_i}^\dagger|n_1, \ldots, n_i, \ldots, n_N\rangle = \sqrt{n_i + 1}\,|n_1, \ldots, n_i + 1, \ldots, n_N\rangle\,. \tag{21}$$

Using the commutation relations, $[\hat{a}_k, \hat{a}_{k'}^\dagger] = \delta_{k,k'}$, it is easy to show from eq. (20) that

$$\hat{a}_{k_i}|n_1, \ldots, n_i, \ldots, n_N\rangle = \sqrt{n_i}\,|n_1, \ldots, n_i - 1, \ldots, n_N\rangle\,, \tag{22}$$

and finally, by repeated use of eq. (22), we obtain

$$\langle n_1, n_2, \ldots, n_N\,|\,n_1', n_2', \ldots, n_N'\rangle = \delta_{n_1,n_1'}\delta_{n_2,n_2'}\cdots\delta_{n_N,n_N'}. \tag{23}$$

We have now found an orthogonal set of eigenvectors of H. We have not shown whether or not this set is a complete basis of the Hilbert space. However, whenever a given perturbation of H, call it H', can be expressed in terms of $\hat{a}_k$ and $\hat{a}_k^\dagger$, or equivalently in terms of $\hat{x}_j$ and $\hat{p}_j$, then the $|n_1, n_2, \ldots, n_N\rangle$-states will only interfere among themselves when evolving in time.

So far, we have only been dealing with a one-dimensional lattice. The extension to more than one dimension is easy, though. Just replace j and k with vectors $\boldsymbol{j}$ and $\boldsymbol{k}$, and replace the normalizing factor $1/\sqrt{N}$ with $1/\sqrt{N^D}$ where D is the number of dimensions. This means that e.g. relation (4) becomes

$$\tilde{x}_k = \frac{1}{\sqrt{N^D}} \sum_{\boldsymbol{j}} \hat{x}_{\boldsymbol{j}} e^{-i\boldsymbol{k}\cdot\boldsymbol{j}a}, \tag{24}$$

and ω_k^2 becomes $\omega_{\boldsymbol{k}}^2 = 4K \sin^2 \frac{a\boldsymbol{k}}{2} + M^2$. Also replace $|n_1, n_2, \dots, n_N\rangle$ by $|n_1, n_2, \dots, n_{N^D}\rangle$ as there are now N^D possible $\boldsymbol{k}$'s.

We have now seen an example, though still only under the assumption that $|0\rangle$ exists, on how the solutions to the quantum lattice can be expanded in phonon modes with a finite number of discrete $\boldsymbol{k}$-values. This is a promising result when we want to use the quantization of discretized fields as an approach to developing a QED theory. Instead of having a quantum system where the number of particle coordinates, on which the wave functions are defined, grows to infinity when $a \to 0$, it now seems we might be able to reinterpret the system as a Hilbert space of bosons, where the effect of letting a tend to zero only makes the momentum space resolution more dense. If this is true for a given theory, it would certainly seem probable that one would be able to obtain a well-defined continuum limit of the theory.

The Hilbert space of bosons as discussed in this section, where any number of bosons is allowed, is called a (bosonic) Fock space. This kind of Hilbert space is very important to QFT, and indeed any theory where the number of particles might vary. A Fock space is simply a Hilbert space of vectors $(\psi_0, \psi_1, \psi_2, \dots)$ where each ψ_n is a vector in a (smaller) Hilbert space describing n particles. The norm of a vector,

$$|(\psi_0, \psi_1, \psi_2, \dots)| = \sqrt{|\psi_0|^2 + |\psi_2|^2 + |\psi_2|^2 + \dots}, \tag{25}$$

should be finite and 1 when normalized. The probability of finding n particles in the system when making a measurement is then simply given by $|\psi_n|$. The internal probabilities are of course contained in ψ_n in the usual way.

As a final remark of this section, note that $\omega_{\boldsymbol{k}}^2 = \frac{4K}{m} \sin^2 \frac{a\boldsymbol{k}}{2} + \Omega^2$ goes as $\frac{K}{m} a^2 \boldsymbol{k}^2 + \Omega^2$ for small a. If we let $\frac{K}{m} = a^{-2}$ as $a \to 0$, we obtain the relativistic dispersion relation for a particle with mass Ω. This is not entirely a coincidence as it has to do with the Lagrangian of the system being Lorentz invariant in that limit. This will be discussed further in a later section.

3 The path integral

We will now introduce the path integral i quantum mechanics. The outline of the derivation follows Srednicky [2]. To keep in line with most textbooks we will first rename x_j to q_j instead and let $\boldsymbol{q} = (q_1, \dots, q_N)$. The Hilbert space is hence the space of all square-integrable functions over $\boldsymbol{q}$. We will keep the $\hat{p}_j$'s. Let the Hamiltonian of a system be given by

$$H = \frac{1}{2m}\hat{P}^2 + \hat{V} = \frac{1}{2m}\hat{P}^2 + V(\boldsymbol{q}), \tag{26}$$

where $\hat{P}^2 = \sum_{j=1}^N \hat{p}_j^2$. This is just the normal Schrödinger equation describing either a particle in N dimensions, N/D particles in D dimensions or a lattice of N/D lattice atoms free to move in D dimensions. Let $|\boldsymbol{q}\rangle$ and $|\boldsymbol{p}\rangle$ denote the position and momentum eigenstates respectively, normalized by $\langle \boldsymbol{q} \,|\, \boldsymbol{q}' \rangle = \delta^N(\boldsymbol{q}-\boldsymbol{q}')$ and $\langle \boldsymbol{p} \,|\, \boldsymbol{p}' \rangle = \delta^N(\boldsymbol{p}-\boldsymbol{p}')$. Their inner product is given by $\langle \boldsymbol{q} \,|\, \boldsymbol{p} \rangle = (2\pi)^{-N/2} \exp(i\boldsymbol{p}\cdot\boldsymbol{q})$. Let us now look at the matrix element of the time propagation operator for a time interval δt:

$$\left\langle \boldsymbol{q}_2 \,\middle|\, e^{-iH\delta t} \,\middle|\, \boldsymbol{q}_1 \right\rangle = \langle \boldsymbol{q}_2 |\, e^{-\frac{i}{2m}\hat{P}^2\delta t - i\hat{V}\delta t} \,|\boldsymbol{q}_1\rangle = \langle \boldsymbol{q}_2 |\, e^{-\frac{i}{2m}\hat{P}^2\delta t} e^{-i\hat{V}\delta t} e^{O(\delta t^2)} \,|\boldsymbol{q}_1\rangle \,. \tag{27}$$

To get the last equality, we have used the fact that $e^{A+B} = e^A e^B e^{-1/2[A,B]+\cdots}$ for any operators A and B. If δt becomes small enough, we can set $e^{O(\delta t^2)} = 1$, so let us do that for now. Inserting the identity

operator $\int d^N\boldsymbol{p}_1\,|\boldsymbol{p}_1\rangle\,\langle\boldsymbol{p}_1| = \hat{I}$ as well, yields

$$\begin{aligned}\langle\boldsymbol{q}_2\,|\,e^{-iH\delta t}\,|\,\boldsymbol{q}_1\rangle &= \int d^N\boldsymbol{p}_1\,\langle\boldsymbol{q}_2|\,e^{-\frac{i}{2m}\hat{P}^2\delta t}\,|\boldsymbol{p}_1\rangle\,\langle\boldsymbol{p}_1|\,e^{-i\hat{V}\delta t}\,|\boldsymbol{q}_1\rangle \\ &= \int d^N\boldsymbol{p}_1 e^{-i\frac{\boldsymbol{p}_1^2}{2m}\delta t - iV(\boldsymbol{q}_1)\delta t}\,\langle\boldsymbol{q}_2\,|\,\boldsymbol{p}_1\rangle\,\langle\boldsymbol{p}_1\,|\,\boldsymbol{q}_1\rangle \\ &= \int \frac{d^N\boldsymbol{p}_1}{(2\pi)^N} e^{-i\frac{\boldsymbol{p}_1^2}{2m}\delta t - iV(\boldsymbol{q}_1)\delta t} e^{i\boldsymbol{p}_1\cdot(\boldsymbol{q}_2-\boldsymbol{q}_1)} \\ &= \int \frac{d^N\boldsymbol{p}_1}{(2\pi)^N} e^{-iH(\boldsymbol{p}_1,\boldsymbol{q}_1)\delta t} e^{i\boldsymbol{p}_1\cdot(\boldsymbol{q}_2-\boldsymbol{q}_1)}.\end{aligned} \tag{28}$$

Here we have used $\langle\boldsymbol{q}\,|\,\boldsymbol{p}\rangle = (2\pi)^{-N/2}\exp(i\boldsymbol{p}\cdot\boldsymbol{q})$ to get the third line and then introduced $H(\boldsymbol{p},\boldsymbol{q}) = \boldsymbol{p}^2 + V(\boldsymbol{q})$. We can use this equation to approximate $\langle\boldsymbol{q}''\,|\,e^{-iHt}\,|\,\boldsymbol{q}'\rangle$ for any t by

$$\begin{aligned}\langle\boldsymbol{q}''\,|\,e^{-iHt}\,|\,\boldsymbol{q}'\rangle &= \int\prod_{i=1}^{M} d^N\boldsymbol{q}_i\;\langle\boldsymbol{q}''|\,e^{-iH\delta t}\,|\boldsymbol{q}_M\rangle\,\langle\boldsymbol{q}_M|\,e^{-iH\delta t}\,|\boldsymbol{q}_{M-1}\rangle\ldots\langle\boldsymbol{q}_1|\,e^{-iH\delta t}\,|\boldsymbol{q}'\rangle \\ &= \int\prod_{i=1}^{M}(d^N\boldsymbol{q}_i)\prod_{j=0}^{M}\Big(\frac{d^N\boldsymbol{p}_j}{(2\pi)^N}\Big)\,e^{i\sum_{k=0}^{M}\left(\boldsymbol{p}_k\cdot\dot{\boldsymbol{q}}_k - H(\boldsymbol{p}_k,\boldsymbol{q}_k)\right)\delta t},\end{aligned} \tag{29}$$

where $\delta t = t/(M+1)$ and $\dot{\boldsymbol{q}}_k = (\boldsymbol{q}_{k+1}-\boldsymbol{q}_k)/\delta t$, $\boldsymbol{q}_0 = \boldsymbol{q}'$. If we simplify the notation by replacing $\prod_{i=1}^{M}(d^N\boldsymbol{q}_i)\prod_{j=0}^{M}\big(\frac{d^N\boldsymbol{p}_j}{(2\pi)^N}\big)$ with $\mathcal{D}\boldsymbol{q}\,\mathcal{D}\boldsymbol{p}$ and take the limit $\delta t\to 0$, this becomes

$$\langle\boldsymbol{q}''\,|\,e^{-iHt}\,|\,\boldsymbol{q}'\rangle = \int\mathcal{D}\boldsymbol{q}\,\mathcal{D}\boldsymbol{p}\,e^{i\int_{t'}^{t''}dt\,\boldsymbol{p}\cdot\dot{\boldsymbol{q}} - H(\boldsymbol{p},\boldsymbol{q})}. \tag{30}$$

If we had kept the factor of $\exp(O(\delta t^2))$, we would have ended up with a factor of $(1+O(\delta t))$ on the RHS of eq. (30). This factor tends toward 1 for $M\to\infty$. It is interesting to note, however, before we move on, that we can still only justify removing the factor if the integral in eq. (28) still remains convergent after the integrand is approximated. Luckily, this is true in our case, but if we for instance had replaced $\exp(-iH\delta t)$ with $(1-iH\delta t)(1+O(\delta t^2))$ instead, the integral would not have remained convergent. We could therefore not have done that.

By eq. (30), we now have a way of finding any matrix element of the propagator just by doing an integral. In our case with the Schrödinger Hamiltonian, we can simplify the integral even further by evaluating the Gaussian integral over the $\boldsymbol{p}_i$'s. The multidimensional Gaussian integral is of much use in quantum field theory. The useful identity for evaluating it is

$$\int d^N\boldsymbol{x}\,e^{-\frac{i}{2}\boldsymbol{x}^TA\boldsymbol{x}+\boldsymbol{b}^T\boldsymbol{x}} = \left(\frac{(2\pi)^N}{\det A}\right)^{\frac{1}{2}} e^{\frac{i}{2}\boldsymbol{b}^TA^{-1}\boldsymbol{b}}, \tag{31}$$

where $\boldsymbol{b}$ is a constant vector and A is any real, positive definite matrix. With $H(\boldsymbol{p},\boldsymbol{q}) = \boldsymbol{p}^2 + V(\boldsymbol{q})$, the RHS of eq. (30) therefore becomes

$$\begin{aligned}\int\mathcal{D}\boldsymbol{q}\,\mathcal{D}\boldsymbol{p}\,e^{i\int_{t'}^{t''}dt\,\boldsymbol{p}\cdot\dot{\boldsymbol{q}} - \frac{1}{2m}\boldsymbol{p}^2 - V(\boldsymbol{q})} &= \Big(\frac{m}{2\pi}\Big)^{\frac{1}{2}N(M+1)}\int\mathcal{D}\boldsymbol{q}\,e^{i\int_{t'}^{t''}dt\,\frac{1}{2}m\dot{\boldsymbol{q}}^2 - V(\boldsymbol{q})} \\ &= \Big(\frac{m}{2\pi}\Big)^{\frac{1}{2}N(M+1)}\int\mathcal{D}\boldsymbol{q}\,e^{i\int_{t'}^{t''}dt\,L(\dot{\boldsymbol{q}},\boldsymbol{q})}.\end{aligned} \tag{32}$$

Here $L(\dot{\boldsymbol{q}},\boldsymbol{q}) = \frac{1}{2}m\dot{\boldsymbol{q}}^2 - V(\boldsymbol{q})$ is simply the classical Lagrangian for the system of particles moving in a potential $V(\boldsymbol{q})$, and so $\int_{t'}^{t''}dt\,L(\dot{\boldsymbol{q}},\boldsymbol{q})$ is just the classical action. The whole integral is known as a path integral, and it is clear to see why. If we are looking at a system of particles, each set of $\boldsymbol{q}_i$'s simply corresponds to a particular path through space. Every possible path contributes with an amplitude of $\exp(i\int_{t'}^{t''}dt\,L(\dot{\boldsymbol{q}},\boldsymbol{q}))$. Note that the normalization factor, $(m/(2\pi))^{N(M+1)/2}$, is dependent on the discretization of the paths. If we are looking at a lattice, each set of $\boldsymbol{q}_i$'s corresponds to particular evolution of the lattice configuration.

The path integral, besides being practical, also gives a neat understanding of quantum mechanics when governed by the Schrödinger equation. In classical mechanics the particles follow a path given by the equations of motion (EOM). This path is characterized by being a stationary point of the action. In quantum mechanics the particles take every possible path and each of these contributes with an amplitude with modulus 1, but in the neighborhood of the stationary path, the phases of the amplitudes interfere positively. It therefore makes sense that the dynamics of the quantum system would somewhat resemble the classical dynamics given by the EOM.

The potential in a classical field Lagrangian typically consists of local operators. A Lagrangian for a field described by a function of space, $\varphi(\boldsymbol{x}, t)$, could e.g. have terms like φ^2 or $\nabla^2\varphi$ in the potential. This enables us to also write $V(\boldsymbol{q})$ as a sum of local operators, as well as the kinetic term. In the formal limit $N \to \infty$, where the spacing between the lattice atoms goes to zero, we can write L as an integral over space of a *Lagrangian density*, $\mathcal{L}(\dot{\varphi}(x), \varphi(x), \nabla^2\varphi(x), \dots)$. Here $x = (t, \boldsymbol{x})$ is the space-time coordinates. We can therefore formally write

$$\int \mathcal{D}\boldsymbol{q}\, e^{i\int_{t'}^{t''} dt\, L(\dot{\boldsymbol{q}},\boldsymbol{q})} = \int \mathcal{D}\boldsymbol{q}\, e^{i\iint_{t'}^{t''} \mathcal{L}\, dt\, d^3\boldsymbol{x}}. \tag{33}$$

In this way the path integral can be written very compactly, but when we actually have to do an integral, we still need to choose a lattice spacing and a bounded volume of space. The choice of the spacing affects the system and the theory, but the hope is that the dynamics of the field bosons converge in the continuum limit, as mentioned above.

Let us now look at the Lagrangian of electromagnetism. In natural units, which will be used from now on, it is given by

$$\mathcal{L}_{EM} = \frac{1}{2}\big(\boldsymbol{E}^2 - \boldsymbol{B}^2\big) = \frac{1}{2}\big((-\nabla V - \frac{\partial}{\partial t}\boldsymbol{A})^2 - (\nabla \times \boldsymbol{A})^2\big). \tag{34}$$

When the action is varied with respect to the evolution of the four-potential $A^\mu = (V, \boldsymbol{A})$, the Euler-Lagrange equations yield the Gauss-Ampère law[1] (the Gauss-Faraday law is implicit already). A natural thought is now to try to work our way backwards from eq. (33) to the field Hamiltonian. If we succeed in this, the dynamics of the corresponding quantum system should resemble classical electromagnetism. However, there are two problems we need to deal with first. One is that $\mathcal{L}_{EM}$ is missing a kinetic term for V, i.e. $(\partial V/\partial t)^2$. Furthermore, there is also the term with the first-order time derivative of $\boldsymbol{A}$, with which we do not know deal yet. We will deal with the second problem first.

4 A more general path integral

Let us go back to eq. (28) and introduce a term $\sum_{j=1}^N C_j(\boldsymbol{q})\,\hat{p}_j$ in the Hamiltonian so that

$$H = \frac{1}{2m}\hat{P}^2 + V(\boldsymbol{q}) + \sum_{j=1}^N C_j(\boldsymbol{q})\,\hat{p}_j, \tag{35}$$

where the C_j's are real functions of the position vector $\boldsymbol{q} = (q_1, \dots, q_N)$. For all $j \in U$, where U is a subset of $\{1, \dots, N\}$, let $C_j(\boldsymbol{q}) = 0$ everywhere. Let also $\partial C_j/\partial q_i = 0$ everywhere for all $j, i \in \overline{U}$, where $\overline{U} = \{1, \dots, N\} \setminus U$ is the complement of U. We define the operators $\hat{C}_j$ by $\hat{C}_j\,|\boldsymbol{q}\rangle = C_j(\boldsymbol{q})\,|\boldsymbol{q}\rangle$. Note that because of our limitations on the C_j's, we have $[\hat{C}_j, \hat{p}_i] = 0$ for all $j, i \in \overline{U}$, and so all the non-vanishing terms in $\sum_{j=1}^N C_j(\boldsymbol{q})\,\hat{p}_j$ now commute. Finally define $\boldsymbol{C}(\boldsymbol{q}) = \big(C_1(\boldsymbol{q}), \dots, C_N(\boldsymbol{q})\big)$. Instead of eq. (28), we now obtain

$$\begin{aligned}
\big\langle \boldsymbol{q}_2 \,\big|\, e^{-iH\delta t} \,\big|\, \boldsymbol{q}_1 \big\rangle &= \int d^N\boldsymbol{p}_1\, \langle \boldsymbol{q}_2|\, e^{-\frac{i}{2m}\hat{P}^2\delta t}\, |\boldsymbol{p}_1\rangle\, \langle \boldsymbol{p}_1|\, e^{-i\sum_j \hat{C}_j\hat{p}_j\delta t} e^{-i\hat{V}\delta t}\, |\boldsymbol{q}_1\rangle \\
&= \int d^N\boldsymbol{p}_1 e^{-i\frac{\boldsymbol{p}_1^2}{2m}\delta t - iV(\boldsymbol{q}_1)\delta t - i\boldsymbol{C}(\boldsymbol{q}_1)\cdot\boldsymbol{p}_1\delta t}\, \langle \boldsymbol{q}_2 \,|\, \boldsymbol{p}_1\rangle\, \langle \boldsymbol{p}_1 \,|\, \boldsymbol{q}_1\rangle \\
&= \int \frac{d^N\boldsymbol{p}_1}{(2\pi)^N} e^{-i\frac{\boldsymbol{p}_1^2}{2m}\delta t - iV(\boldsymbol{q}_1)\delta t} e^{i\boldsymbol{p}_1\cdot(\boldsymbol{q}_2 - \boldsymbol{q}_1 - \boldsymbol{C}(\boldsymbol{q}_1)\delta t)},
\end{aligned} \tag{36}$$

[1] This is not true if the action is varied w.r.t. e.g. $\boldsymbol{E}$ and $\boldsymbol{B}$ instead.

where we have used the fact that $\hat{C}_j$ and $\hat{p}_j$ commute for all j, and that all the terms in $\sum_{j=1}^{N} C_j(\boldsymbol{q})\,\hat{p}_j$ commute, to get the second equality. We see that eq. (29) now becomes

$$\langle \boldsymbol{q}'' \,|\, e^{-iHt} \,|\, \boldsymbol{q}' \rangle = \int \prod_{i=1}^{M} d^N \boldsymbol{q}_i \prod_{j=0}^{M} \frac{d^N \boldsymbol{p}_j}{(2\pi)^N}\, e^{i\sum_{k=0}^{M} \left(\boldsymbol{p}_k \cdot (\dot{\boldsymbol{q}}_k - \boldsymbol{C}(\boldsymbol{q}_k)) - \frac{\boldsymbol{p}_k^2}{2m} - V(\boldsymbol{q}_k)\right)\delta t}. \tag{37}$$

The Gaussian integration can now be evaluated the same way by using eq. (31), and the end result is

$$\langle \boldsymbol{q}'' \,|\, e^{-iHt} \,|\, \boldsymbol{q}' \rangle = \left(\frac{m}{2\pi}\right)^{\frac{1}{2}N(M+1)} \int \mathcal{D}\boldsymbol{q}\, e^{i\int_{t'}^{t''} dt\, L(\dot{\boldsymbol{q}},\boldsymbol{q})}. \tag{38}$$

This is just as before, but now with the more general $L(\dot{\boldsymbol{q}}, \boldsymbol{q}) = \frac{1}{2}m(\dot{\boldsymbol{q}} - \boldsymbol{C}(\boldsymbol{q}))^2 - V(\boldsymbol{q})$ instead of just $L(\dot{\boldsymbol{q}}, \boldsymbol{q}) = \frac{1}{2}m\dot{\boldsymbol{q}}^2 - V(\boldsymbol{q})$. This greatly increases the variety of Lagrangians we can use to describe a quantum theory. Working backwards to the Hamiltonian is easy. It is given by $H = \hat{P}^2/(2m) + V(\boldsymbol{q}) + \sum_j \hat{C}_j \hat{p}_j$ and can therefore just be read off the Lagrangian. Note that because $[\hat{C}_j, \hat{p}_j] = 1$, the Hamiltonian is hermitian as required.

So far we have only looked at cases with $A = m^{-1}$ in eq. (31). We can make it a bit more general than that with an arbitrary A, as long as A is still real and positive definite. In general we have

$$H = \frac{1}{2}\hat{\boldsymbol{p}}^T A\, \hat{\boldsymbol{p}} + V(\boldsymbol{q}) + \boldsymbol{C}(\boldsymbol{q})^T \hat{\boldsymbol{p}} \longleftrightarrow L(\dot{\boldsymbol{q}}, \boldsymbol{q}) = \frac{1}{2}(\dot{\boldsymbol{q}} - \boldsymbol{C}(\boldsymbol{q}))^T A^{-1} (\dot{\boldsymbol{q}} - \boldsymbol{C}(\boldsymbol{q})) - V(\boldsymbol{q}), \tag{39}$$

where $\hat{\boldsymbol{p}} = (\hat{p}_1, \ldots, \hat{p}_N)$, and we can now easily convert between the two. We have left out the normalization factor in relation (39). This is given by $-1/2(M+1)(N\ln(2\pi) + \ln(\det A))$ and depend on the lattice spacing and on δt. Using the wrong normalization factor in the path integral results in an exponentially increasing or decreasing overall norm of the quantum state. The state can theoretically be normalized, however, and one can still compare individual probabilities and get the right result.

We can also loosen the restrictions on $\boldsymbol{C}$ a bit, and we will need this in the following. We could have chosen a $\boldsymbol{C}$ on the form $\boldsymbol{C}(\boldsymbol{q}) = \boldsymbol{C}_1(\boldsymbol{q}) + \boldsymbol{C}_2(\boldsymbol{q}) + \ldots + \boldsymbol{C}_K(\boldsymbol{q})$, $K \in \mathbb{N}$ instead, where every $\boldsymbol{C}_k$, $k \in \{1, 2, \ldots, K\}$, is a vector (not a vector element) that fulfills the restrictions from before. However, K has to be independent of the lattice spacing, and accordingly on N. The consequence is that not all the $\hat{C}_j$'s and $\hat{p}_i$'s commute with each other, but we can separate $\exp(-i\sum_j \hat{C}_j \hat{p}_j)$ in K exponential factors and only gain K factors on the form $\exp(O(\delta t^2))$ in the second expression in eq. (36). This means that only a finite number of factors, independent on N, has to be approximated as 1. To get the second equality in eq. (36) from there, we will again (only) get K factors of $\exp(O(\delta t^2))$ when we have to permute the K factors of $\exp(-i\sum_j \hat{C}_{k,j}\hat{p}_j \delta t)$ step by step to let the operators hit their eigenstates. This can be done by moving the front factor of $\exp(-i\sum_j \hat{C}_{k,j}\hat{p}_j\delta t)$ to the back to let the $\hat{p}_j$'s act on $\langle \boldsymbol{p}_1|$ and repeating this K times. The restrictions on $\boldsymbol{C}(\boldsymbol{q})$ will be recapped in the end of this section.

It is interesting to note that this result is exactly the same as in classical mechanics, just with $\hat{\boldsymbol{p}} \to \boldsymbol{p}$ in the Hamiltonian. In classical mechanics, the Hamiltonian is found by solving

$$\boldsymbol{p} = \nabla_{\dot{\boldsymbol{q}}} L(\dot{\boldsymbol{q}}, \boldsymbol{q}) \tag{40}$$

for $\dot{\boldsymbol{q}}$ and then substituting it in[2]

$$H = \boldsymbol{p}^T \dot{\boldsymbol{q}} - L(\dot{\boldsymbol{q}}, \boldsymbol{q}). \tag{41}$$

This is known as a *Legendre transform*. For the Lagrangian in relation (39), eq. (40) yields

$$\boldsymbol{p} = A^{-1}(\dot{\boldsymbol{q}} - \boldsymbol{C}(\boldsymbol{q})) \Leftrightarrow \dot{\boldsymbol{q}} = A\,\boldsymbol{p} + \boldsymbol{C}(\boldsymbol{q}). \tag{42}$$

Substituting this $\dot{\boldsymbol{q}}$ in eq. (41) gives us

$$\begin{aligned} H &= \boldsymbol{p}^T A\, \boldsymbol{p} + \boldsymbol{p}^T \boldsymbol{C}(\boldsymbol{q}) - \frac{1}{2}(A\,\boldsymbol{p})^T A^{-1} A \boldsymbol{p} + V(\boldsymbol{q}) \\ &= \boldsymbol{p}^T A\, \boldsymbol{p} - \frac{1}{2}\boldsymbol{p}^T A^T \boldsymbol{p} + V(\boldsymbol{q}) + \boldsymbol{p}^T \boldsymbol{C}(\boldsymbol{q}) \\ &= \frac{1}{2}\boldsymbol{p}^T A\, \boldsymbol{p} + V(\boldsymbol{q}) + \boldsymbol{C}(\boldsymbol{q})^T \boldsymbol{p}. \end{aligned} \tag{43}$$

[2]This applies to all Lagrangians that are strictly convex functions of $\dot{\boldsymbol{q}}$ (see Durhuus and Solovej [3]).

This is exactly the Hamiltonian we wanted. It is a remarkable result that the correspondence between the Hamiltonian and the Lagrangian in quantum mechanics is so similar to the classical correspondence, at least for a great variety of Hamiltonians. It also makes the conversion between them easy to remember. The positive definiteness of A is in fact also a requirement[3] for the Legendre transform to work in the classical picture, see for example Durhuus and Solovej [3]. We still need to take into account the restrictions on $\boldsymbol{C}(\boldsymbol{q})$, however. These can be rephrased this way: $\boldsymbol{C}$ has to be on the form $\boldsymbol{C}(\boldsymbol{q}) = \sum_{k=1}^{K} \boldsymbol{C}_k(\boldsymbol{q})$, where for each of the vectors $\boldsymbol{C}_k$, the following statement is true: If vector $\boldsymbol{C}_k$ is dependent on some of the elements in $\boldsymbol{q}$, then the corresponding components of $\boldsymbol{C}_k$ has to be zero everywhere. $K \in \mathbb{N}$ should be independent on the lattice spacing.

5 Lorentz invariance of the Lagrangian

Before we can convert the electromagnetic Lagrangian of eq. (34) via relation (39), we still need a kinetic term for V. The simple solution to this, which we are going to try in the next section, is to add a term

$$\mathcal{L}_{gauge} = \frac{\xi}{2}(\partial_\mu A^\mu)^2 = \frac{\xi}{2}\big(\frac{\partial}{\partial t}V - \nabla \cdot \boldsymbol{A}\big)^2 \tag{44}$$

to $\mathcal{L}_{EM}$, where ξ is a real non-negative parameter that can be varied. The reason for choosing this particular term is that it does not break the Lorentz invariance of $\mathcal{L}_{EM}$. In fact, it only removes the gauge freedom in the classical EOM and therefore does not change the dynamics of $\boldsymbol{E}$ and $\boldsymbol{B}$. There is a good reason why we do not want to break the Lorentz invariance. The reason is that one can argue, as is generally accepted, that a Lorentz invariant Lagrangian also leads to a Lorentz invariant quantum theory. We will now use this opportunity, before we move on, to talk about why this is so.

If we look at the path integral of a Lagrangian density,

$$\int \mathcal{D}\boldsymbol{q}\, e^{i \iint_{t'}^{t''} \mathcal{L}\, dt\, d^3\boldsymbol{x}}, \tag{45}$$

the first thing to notice is that the integration is over a space-time volume. Under a Lorentz transformation $x^\mu = (t, \boldsymbol{x})$ transforms as

$$(x')^\mu = \Lambda^\mu_\nu x^\nu, \quad \Lambda^\mu_\nu = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{46}$$

where $(x')^\mu = (t', \boldsymbol{x}')$ is the transformed coordinate, and where we have taken the boost to be in the x^1-direction. Since the Jacobian of this transformation is 1, we can transform the integration variables in the path integral without gaining any extra factors in the integrand. In other words, the integration measure is unchanged when we integrate (or sum) the Lagrangian density. Therefore, if $\mathcal{L}$ is Lorentz invariant, meaning that $\mathcal{L}'(x') = \mathcal{L}(x)$, then the path integral over the same space-time volume is invariant as well. If the fields in the Lagrangian also change under the transformation, we need the Jacobian of this transformation to be 1 as well because of the integration over the field configurations in the path integral. If A^μ is the field on which $\mathcal{L}$ depends, then the Jacobian is indeed 1, since A^μ also transforms by

$$(A')^\mu = \Lambda^\mu_\nu A^\nu. \tag{47}$$

This seems promising. The question is now, if this means that the corresponding quantum theory is Lorentz invariant as well. For this to be true, there must for every quantum state exist a corresponding state in the transformed system which, when time-evolved by the Hamiltonian and then Lorentz transformed back, will produce the same end state, as when the original state is time-evolved in the first inertial frame. In accordance with the special theory relativity, a translation might also be required in this procedure to obtain the same end state. This Lorentz transformation of the state should of course be defined independently of time, and any measurement of a particles position at x in the original system should yield the same outcome probabilities as a measurement in x' in the transformed system.

[3] If we are not counting the possible allowance for a negative definite matrix.

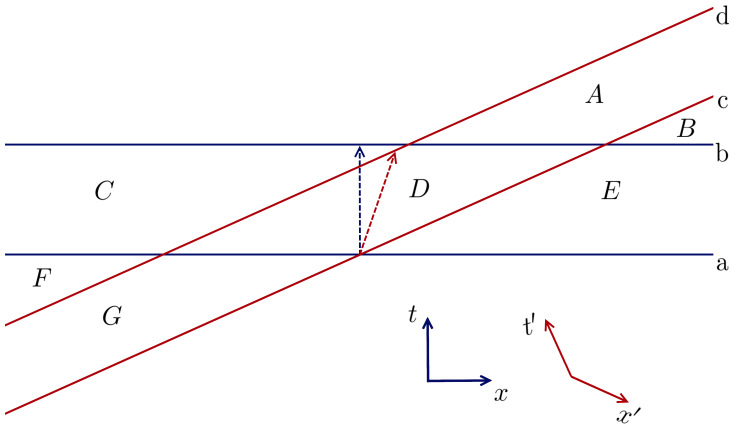

Figure 1: A volume of space-time, over which a path integral is performed in two different inertial frames. Only the temporal dimension and one spacial dimension, denoted by x, is drawn in the figure. The blue, horizontal lines, a and b, are two hyperplanes with constant t, and the red, tilted lines, c and d, are two hyperplanes with constant t', where t and t' are the different time coordinates of the two inertial frames, whose relative velocity lies in the x-direction. Letters A through G represent different space-time volumes bounded by the hyperplanes a, b, c and d.

Figure 1 represents a volume of space-time where the two blue, horizontal lines, labeled a and b, are volumes of space with constant t, and the two red, tilted lines, c and d, are volumes of space with constant t'. Only one spacial dimension is drawn in the figure, so one must imagine the two others. The capital letters label the volumes of space-time bounded by the volumes a, b, c and d. Let us start by considering a path integral from a quantum field in a, not into b, but only over the space-time volume E such that the resulting state is both partly in a, in c and in b. To make this work, we have to choose δt and the spacing, however small, in such a way that if x_1 is a space-time coordinate where a and c intersect, then $x_1 + (\delta t, \boldsymbol{a})$ lies in c, where $\boldsymbol{a}$ is the lattice vector pointing in the direction of the boost. This means δt has to be small compared to $\boldsymbol{a}$, when the Lorentz boost is small. Let us call this particular operation $\hat{K}_E$ where the index tells us what space-time volume we are doing the path integral over. We could then further propagate this state by $\hat{K}_D$ as well. The point is now that since we have to integrate over all intermediate states when we connect the two path integrals, the operations $\hat{K}_D\hat{K}_E$ and $\hat{K}_{D\cup E}$ are identical.

Before we move on with the argument, let us address the matter of the infinite path integral in $\hat{K}_E$. As we see in fig. 1, E is unbounded to the right and extends to infinity. Since we cannot do infinitely many integrals at once, we need a way to cut E for $\hat{K}_E$ to make sense. The quick way to deal with this is to assume that the fields vanishes after a certain large distance, call it R, but in reality, quantum fields never vanishes: even the vacuum state has a certain continuous probability distribution of the displacement. In this rendition of the paper, however, the purpose of this section is just to give an idea of the advantage of working with Lagrangians when wanting to prove Lorentz invariance of a theory. Therefore, the problem will be dealt with the easy way here, and we will assume that the fields vanishes outside a certain R.

Moving on, we now need to show that $\hat{K}_E$ is unitary. Having assumed the fields to vanish after a certain large distance, we only have to integrate over a finite space-time volume. Let us now extend the path integral over an additional space-time volume, so that the hyperplane containing the end state no longer lies in c but once again has $t = const.$. We also change $H(t, \boldsymbol{x})$ such that the interaction between two points, both lying beyond c in the newly added space-time volume, disappears. The path integral therefore yields the same result as before when we simply ended the integration in c, but we can now argue that since the path integral now simply propagates the state by $\exp(-iH(t, \boldsymbol{x}))$ with the new $H(t, \boldsymbol{x})$, and since $H(t, \boldsymbol{x})$ is still hermitian for all t, at least for all the Hamiltonians we have dealt with above, $\hat{K}_E = \exp(-iH(t, \boldsymbol{x}))$ is unitary. This is an important result because we can now define the boost operator about a point in the intersection of a and c, call it $\hat{S}_{ac}$, by

$$\hat{S}_{ac} = \hat{K}_G^\dagger \hat{K}_F^\dagger \hat{K}_B \hat{K}_E, \tag{48}$$

and the resulting state will be normalized. Because of the Lorentz invariance of the path integral,

propagating the state from c to d in the boosted system is simply done by $\hat{K}_A\hat{K}_D\hat{K}_G$. Therefore, boosting a state from a to c, then propagating it from c to d and then boosting it back from d to b via $\hat{S}^\dagger_{bd}$ is done by the operation:

$$\hat{S}^\dagger_{bd}\hat{K}_A\hat{K}_D\hat{K}_G\hat{S}_{ac} = \hat{K}_C\hat{K}_F\hat{K}^\dagger_B\hat{K}^\dagger_A\hat{K}_A\hat{K}_D\hat{K}_G\hat{K}^\dagger_G\hat{K}^\dagger_F\hat{K}_B\hat{K}_E = \hat{K}_C\hat{K}_D\hat{K}_E, \tag{49}$$

where we have used $[\hat{K}^\dagger_B, \hat{K}_D] = [\hat{K}_F, \hat{K}_D] = [\hat{K}^\dagger_B, \hat{K}_F] = 0$. The result is just the operator propagating the state in a into b in the original inertial frame, and we can therefore conclude that, under the conditions, the theory is Lorentz invariant.

6 The Hamiltonian with a Lorenz gauge term

Let us now see what happens if we try to convert the Lagrangian density

$$\mathcal{L} = \mathcal{L}_{EM} + \mathcal{L}_{gauge} = \frac{1}{2}\big(-\nabla V - \frac{\partial}{\partial t}\boldsymbol{A}\big)^2 - \frac{1}{2}\big(\nabla\times\boldsymbol{A}\big)^2 + \frac{\xi}{2}(\frac{\partial}{\partial t}V - \nabla\cdot\boldsymbol{A})^2 \tag{50}$$

to a Hamiltonian density. Reading off from relation (39), forgetting the normalization factor for now, we immediately obtain

$$\mathcal{H} = \frac{1}{2\xi}\hat{\Pi}^2_V + \frac{1}{2}\hat{\Pi}^2_{\boldsymbol{A}} + (\nabla\cdot\boldsymbol{A})\hat{\Pi}_V - (\nabla V)\hat{\Pi}_{\boldsymbol{A}} + \frac{1}{2}\big(\nabla\times\boldsymbol{A}\big)^2. \tag{51}$$

Here $\hat{\Pi}_V$ and $\hat{\Pi}_{\boldsymbol{A}}$ are the dense analogues of what we would have called $\hat{P}_V$ and $\hat{P}_{\boldsymbol{A}}$ in the discrete version. The first thing one might notice about this Hamiltonian density is that the only term which could have a binding effect on the eigensolutions is the $(\nabla\times\boldsymbol{A})^2$ term, and it can only bind two out of the four degrees of freedom of the field. V and the divergence part of $\boldsymbol{A}$ are, though coupled to each other, essentially free.

If we analyze the effect of letting ξ tend to zero, we see that this only has an effect analogous to letting the "mass" of the V-field tend to zero. This fact might help to understand the singularity when not including the gauge term, but it does not help to bind V. But what happens if we let ξ tend toward infinity instead? We have already argued that the quantum system should resemble the classical path, which is the stationary point of the action integral. When we crank up ξ, the deviation from the classical path will have to be smaller before the amplitudes start to interfere negatively. One might therefore predict that the dynamics of the quantum system would be closer to the classical path. The stationary point of the gauge term alone is simply $\partial V/\partial t = \nabla\cdot\boldsymbol{A}$, which is just the Lorenz gauge condition. Therefore, we might expect a solution to the quantum system to obey this equation when ξ goes to infinity. This actually seems to be the case if we look at eq. (51). Letting $\xi\to\infty$ makes the first term vanish, which only leaves one term to generate the movement of V, namely $(\nabla\cdot\boldsymbol{A})\hat{\Pi}_V$. Since $\hat{\Pi}_V$ is just the generator of, not translations, but displacements, this term has exactly this effect, and the solution therefore does obey $\partial V/\partial t = \nabla\cdot\boldsymbol{A}$. This is a remarkable result; we have managed to obtain a gauge restriction just by introducing a term in the Lagrangian and without restricting the integration of the path integral in any way. The Lorenz gauge condition is the only gauge condition, to my knowledge, where this is possible.

One of the problems with the Hamiltonian of eq. (51) emerges when we want to include electrons in the theory. The obvious and natural way to couple the electrons to V is by including a term proportional to $-V(\boldsymbol{q})$ in the Hamiltonian, where $\boldsymbol{q}$ are the positions of the electrons, but this results in a paradoxical situation: If V are to resemble its classical version, it has to decrease like $-1/r$, but this means an infinite energy because of the $-V(\boldsymbol{q})$ term. In fact, since we expect the quantum field to seek toward a configuration of lowest energy, one would expect V to increase as $r\to 0$ more likely. It is also a problem in it self that for any configuration of V, there exists a V with a lower energy, and so the energies of the quantum system are not bounded from below when containing electrons. This is not a problem in classical electrodynamics since here, V are restricted by the EOM, namely the Gauss-Ampère law. We cannot assume the EOM for the quantum field, however.

There is another reason why we would want to get rid of the gauge freedom. As we have seen, A^μ has two unbounded degrees of freedom, and therefore, we are not able to obtain any bounded and normalizable eigensolutions. If we can find a way to restrict the A^μ field in such a way that these unbounded degrees of freedom disappears, it would have yet another desirable effect. From experiments we know that photons

interfere in a similar way to bosons having two spinor indices. This is analogous to the classical case where the electromagnetic waves can have two polarizations. It is easy to see that the degrees of freedom of a field are connected to the number of spinor indices of the corresponding field bosons. Consider for example a field, A, with only two, uncoupled components. A boson excited in one field component of A will only interfere with other bosons in that same component, and they can therefore be thought of as having two spinor indices. Photons are a bit more complicated than this, though. If photons are incident on a screen with different angles, the interference becomes less significant as the angle between them increases. This can be deduced classically. The photons, despite having only two spinor indices, must therefore still interfere in a three-dimensional manner. After the next section, where we will show how to Fourier transform the path integral and the Lagrangian, we will proceed to put restrictions $V = 0$ and $\nabla \cdot \boldsymbol{A} = 0$ on A^μ. This will actually result in a simple and beautiful description of free photons, which has all these properties. These restrictions are not allowed when we introduce electrons to the theory, however, but we will not look further into that problem in this rendition of the paper.

7 Fourier transforming the Lagrangian and Hamiltonian

We can change the basis of $\mathcal{H}$ by making a Fourier transformation of the field. To show how this is done, let us look at the discretized field again. Consider a one-dimensional lattice of N lattice atoms with one degree of freedom. Assume N to be even, and let a denote the lattice spacing. We will label the individual displacements by q_j and the corresponding moments by p_j, where j run from 1 to N in integer steps. Since the field is real-valued, we will be careful to ensure it stays that way. We will therefore make the Fourier transforms in terms of sine and cosine functions instead of the complex exponential functions. Let therefore $\tilde{q}_{k,s}$ and $\tilde{p}_{k,s}$ be given by

$$\tilde{q}_{k,s} = \frac{(-i)^{\frac{1}{2}-\frac{s}{2}}}{\sqrt{2N}} \sum_j q_j \big(e^{ikja} + s\, e^{-ikja}\big), \tag{52}$$

$$\tilde{p}_{k,s} = \frac{(-i)^{\frac{1}{2}-\frac{s}{2}}}{\sqrt{2N}} \sum_j p_j \big(e^{ikja} + s\, e^{-ikja}\big), \tag{53}$$

where s can be either -1 or $+1$, and where k takes values of $2\pi n/(Na)$ with n running from $1/2$ to $N - 1/2$ in integer steps. Note that we can also express $\tilde{q}_{k,s}$, in terms of sines and cosines, as

$$\tilde{q}_{k,+1} = \frac{\sqrt{2}}{\sqrt{N}} \sum_j q_j \cos(kja), \quad \tilde{q}_{k,-1} = \frac{\sqrt{2}}{\sqrt{N}} \sum_j q_j \sin(kja). \tag{54}$$

We will now show that the transformation between $\tilde{\boldsymbol{q}}$ and $\boldsymbol{q}$ is unitary, where $\tilde{\boldsymbol{q}}$ is a vector containing all the $\tilde{q}_{k,s}$'s and $\boldsymbol{q}$ is a vector containing all the q_j's. Let us express $\tilde{q}^2$ in terms of the $q_{k,s}$'s. We see that

$$\tilde{\boldsymbol{q}}^2 = \sum_k \sum_s \tilde{q}_{k,s}^2 = \sum_k \sum_s \frac{(-i)^{1-s}}{2N} \sum_j \sum_{j'} q_j q_{j'} \big(e^{ikja} + s\, e^{-ikja}\big)\big(e^{ikj'a} + s\, e^{-ikj'a}\big). \tag{55}$$

Noting that $(-i)^{1-s} = s$, and that only even powers of s will survive the summation of s, we get

$$\begin{aligned} \tilde{\boldsymbol{q}}^2 =& \frac{1}{2N} \sum_k \sum_s \sum_j \sum_{j'} q_j q_{j'} s^2 \big(e^{ik(j-j')a} + e^{-ik(j-j')a}\big) \\ =& \frac{2}{N} \sum_k \sum_j \sum_{j'} q_j q_{j'} \cos(k(j-j')a). \end{aligned} \tag{56}$$

We now use the fact that since $k = 2\pi n/(Na)$, where n runs from $1/2$ to $N - 1/2$ in integer steps, $\sum_k \cos(k(j-j')a) = (N/2)\delta_{j,j'}$. Therefore eq. (56) becomes

$$\tilde{\boldsymbol{q}}^2 = \sum_j q_j^2 = \boldsymbol{q}^2. \tag{57}$$

This proves that the norm is always preserved under the transformation between $\boldsymbol{q}$ and $\tilde{\boldsymbol{q}}^2$, and it follows that the transformation is unitary. By the same calculation, the similar result can be proven for the

transformation between $\boldsymbol{p}$ and $\tilde{\boldsymbol{p}}^2$. This means that the Jacobian of the transformations are 1 and we can therefore change integration variables from $\boldsymbol{q}$ to $\tilde{\boldsymbol{q}}^2$ or from $\boldsymbol{p}$ to $\tilde{\boldsymbol{p}}^2$ without gaining any extra factor in the integrand. From this fact, we immediately obtain

$$\hat{I} = \int d^N \boldsymbol{q} \, |\boldsymbol{q}\rangle \, \langle \boldsymbol{q}| = \int d^N \tilde{\boldsymbol{q}} \, |\tilde{\boldsymbol{q}}\rangle \, \langle \tilde{\boldsymbol{q}}| \,, \tag{58}$$

$$\hat{I} = \int d^N \boldsymbol{p} \, |\boldsymbol{p}\rangle \, \langle \boldsymbol{p}| = \int d^N \tilde{\boldsymbol{p}} \, |\tilde{\boldsymbol{p}}\rangle \, \langle \tilde{\boldsymbol{p}}| \,. \tag{59}$$

Here, the state $|\tilde{\boldsymbol{q}}\rangle$ is actually equal to the corresponding state $|\boldsymbol{q}\rangle$, they are only labeled differently. Equation (58) shows, however, that if we change the basis such that $\boldsymbol{q} \to \tilde{\boldsymbol{q}}$, then this would send $\hat{I}$ into $\hat{I}$, and it follows that this new basis is also orthonormal. Therefore, the transformation from the $\boldsymbol{q}$-basis to the $\tilde{\boldsymbol{q}}$-basis is unitary, and the same also applies for $\boldsymbol{p}$ and $\tilde{\boldsymbol{p}}$. Due to the unitarity in both cases, we also know that both $\boldsymbol{p} \cdot \boldsymbol{q}$ and $\langle \tilde{\boldsymbol{q}} \, | \, \tilde{\boldsymbol{p}} \rangle$ are preserved under the Fourier transformation, which tells us that

$$\langle \tilde{\boldsymbol{q}} \, | \, \tilde{\boldsymbol{p}} \rangle = \langle \boldsymbol{q} \, | \, \boldsymbol{p} \rangle = (2\pi)^{-\frac{N}{2}} \exp(i \boldsymbol{p} \cdot \boldsymbol{q}) = (2\pi)^{-\frac{N}{2}} \exp(i \tilde{\boldsymbol{p}} \cdot \tilde{\boldsymbol{q}}). \tag{60}$$

Before we can transform the path integral of eq. (36), we need to define Fourier transforms of $\hat{\boldsymbol{p}}$, $\boldsymbol{C}$ and V as well. Since we already have used $\tilde{p}_{k,s}$ to denote the Fourier coefficients labeling the states, we will let $\check{\boldsymbol{p}}$ denote the Fourier transformed $\hat{\boldsymbol{p}}$. We define

$$\check{p}_{k,s} = \frac{(-i)^{\frac{1}{2}-\frac{s}{2}}}{\sqrt{2N}} \sum_j \hat{p}_j \big(e^{ikja} + s \, e^{-ikja} \big), \tag{61}$$

$$\tilde{C}_{k,s}(\tilde{\boldsymbol{q}}) = \frac{(-i)^{\frac{1}{2}-\frac{s}{2}}}{\sqrt{2N}} \sum_j C_j(\boldsymbol{q}) \big(e^{ikja} + s \, e^{-ikja} \big), \tag{62}$$

and define $\tilde{V}$ simply by $\tilde{V}(\tilde{\boldsymbol{q}}) = V(\boldsymbol{q})$. With exactly the same calculations, as when we showed $\sum_{k,s} \tilde{q}_{k,s}^2 = \sum_j q_j^2$, one can show

$$\sum_{k,s} \check{p}_{k,s}^2 = \sum_j \hat{p}_j^2, \tag{63}$$

and

$$\sum_{k,s} \tilde{C}_{k,s}(\tilde{\boldsymbol{q}}) \tilde{p}_{k,s} = \sum_j C_j(\boldsymbol{q}) p_j. \tag{64}$$

Note also that $\check{p}_{k,s} \, |\tilde{\boldsymbol{p}}\rangle = \tilde{p}_{k,s} \, |\tilde{\boldsymbol{p}}\rangle$. This follows directly from eq. (53) and (62) using the fact that $\hat{p}_j \, |\tilde{\boldsymbol{p}}\rangle = p_j \, |\tilde{\boldsymbol{p}}\rangle$.

We can now transform the path integral. Using eq. (36), we have

$$\big\langle \tilde{\boldsymbol{q}}_2 \, \big| \, e^{-iH\delta t} \, \big| \, \tilde{\boldsymbol{q}}_1 \big\rangle = \big\langle \boldsymbol{q}_2 \, \big| \, e^{-iH\delta t} \, \big| \, \boldsymbol{q}_1 \big\rangle = \int \frac{d^N \boldsymbol{p}_1}{(2\pi)^N} e^{-i\frac{\boldsymbol{p}_1^2}{2m}\delta t - iV(\boldsymbol{q}_1)\delta t} e^{i\boldsymbol{p}_1 \cdot (\boldsymbol{q}_2 - \boldsymbol{q}_1 - \boldsymbol{C}(q_1)\delta t)}. \tag{65}$$

Changing the integration variables and using eq. (63), eq. (64), $\tilde{V}(\tilde{\boldsymbol{q}}) = V(\boldsymbol{q})$ and $\boldsymbol{p} \cdot \boldsymbol{q} = \tilde{\boldsymbol{p}} \cdot \tilde{\boldsymbol{q}}$, we obtain

$$\big\langle \tilde{\boldsymbol{q}}_2 \, \big| \, e^{-iH\delta t} \, \big| \, \tilde{\boldsymbol{q}}_1 \big\rangle = \int \frac{d^N \tilde{\boldsymbol{p}}_1}{(2\pi)^N} e^{-i\frac{\tilde{\boldsymbol{p}}_1^2}{2m}\delta t - i\tilde{V}(\tilde{\boldsymbol{q}}_1)\delta t} e^{i\tilde{\boldsymbol{p}}_1 \cdot (\tilde{\boldsymbol{q}}_2 - \tilde{\boldsymbol{q}}_1 - \tilde{\boldsymbol{C}}(\tilde{\boldsymbol{q}}_1)\delta t)}. \tag{66}$$

Equations (58) and (59) further allow us to continue as in eq. (37), and we arrive at

$$\big\langle \tilde{\boldsymbol{q}}'' \, \big| \, e^{-iHt} \, \big| \, \tilde{\boldsymbol{q}}' \big\rangle = \Big(\frac{m}{2\pi} \Big)^{\frac{1}{2}N(M+1)} \int \mathcal{D}\tilde{\boldsymbol{q}} \, e^{i \int_{t'}^{t''} dt \, \tilde{L}(\dot{\tilde{\boldsymbol{q}}}, \tilde{\boldsymbol{q}})}, \tag{67}$$

with $\tilde{L}(\dot{\tilde{\boldsymbol{q}}}, \tilde{\boldsymbol{q}}) = \frac{1}{2} m (\dot{\tilde{\boldsymbol{q}}} - \tilde{\boldsymbol{C}}(\tilde{\boldsymbol{q}}))^2 - \tilde{V}(\tilde{\boldsymbol{q}})$. We now know that Fourier transforming the path integral has this very straightforward result. All this can easily be extended to three dimensions by changing j and k into three-dimensional vectors $\boldsymbol{j}$ and $\boldsymbol{k}$, with the components of $\boldsymbol{j}$ running from 1 to N. For $\boldsymbol{k}$, which

is now given by $2\pi\boldsymbol{n}/(Na)$, $\boldsymbol{n} = (n_1, n_2, n_3)$, n_1 should run from $1/2$ to $N - 1/2$ like before, but n_2 and n_3 should both run all the way from $-N + 1/2$ to $N - 1/2$. In three dimensions, $\tilde{q}_{k,s}$ then becomes

$$\tilde{q}_{\boldsymbol{k},s} = \frac{(-i)^{\frac{1}{2}-\frac{s}{2}}}{\sqrt{2N^3}} \sum_{\boldsymbol{j}} q_{\boldsymbol{j}} \big(e^{i\boldsymbol{k}\cdot\boldsymbol{j}a} + s\, e^{-i\boldsymbol{k}\cdot\boldsymbol{j}a}\big), \tag{68}$$

and the rest of the previous variables and operators will be defined in the same manner. When we want to look at a field with four degrees of freedom, we can just quadruple the number of $q_{\boldsymbol{j}}$'s.

We will now see what happens if we make a Fourier transformation of our $\mathcal{L}$ from before:

$$\mathcal{L} = \frac{1}{2}\big(- \nabla V - \frac{\partial}{\partial t}\boldsymbol{A}\big)^2 - \frac{1}{2}\big(\nabla \times \boldsymbol{A}\big)^2 + \frac{\xi}{2}(\frac{\partial}{\partial t}V - \nabla \cdot \boldsymbol{A})^2. \tag{69}$$

Let $(\tilde{V}_{\boldsymbol{k},s}, \tilde{\boldsymbol{A}}_{\boldsymbol{k},s}) = (\tilde{V}_{\boldsymbol{k},s}, \tilde{A}_{1,\boldsymbol{k},s}, \tilde{A}_{2,\boldsymbol{k},s}, \tilde{A}_{3,\boldsymbol{k},s})$ be the variable coefficients of the Fourier transformed A^μ, resolved individually in sine and cosine functions like above. Keeping in line with previous notation, bold letters will not be used when denoting the elements of a vector. Now, for every $\boldsymbol{k}$ and s, let us define $\tilde{A}_{\perp_1\boldsymbol{k},s}$, $\tilde{A}_{\perp_2\boldsymbol{k},s}$ and $\tilde{A}_{\parallel\boldsymbol{k},s}$ by

$$\begin{pmatrix} \tilde{A}_{\perp_1\boldsymbol{k},s} \\ \tilde{A}_{\perp_2\boldsymbol{k},s} \\ \tilde{A}_{\parallel\boldsymbol{k},s} \end{pmatrix} = U\tilde{\boldsymbol{A}}_{\boldsymbol{k},s} = \begin{pmatrix} u_{1,1} & u_{1,2} & u_{1,3} \\ u_{2,1} & u_{2,2} & u_{2,3} \\ u_{3,1} & u_{3,2} & u_{3,3} \end{pmatrix} \begin{pmatrix} \tilde{A}_{1,\boldsymbol{k},s} \\ \tilde{A}_{2,\boldsymbol{k},s} \\ \tilde{A}_{3,\boldsymbol{k},s} \end{pmatrix}, \tag{70}$$

where U is a unitary, real matrix. We choose $\tilde{A}_{\perp_1\boldsymbol{k},s}$ and $\tilde{A}_{\perp_2\boldsymbol{k},s}$ to be perpendicular to $\boldsymbol{k}$ and each other, and $\tilde{A}_{\parallel\boldsymbol{k},s}$ to be parallel to $\boldsymbol{k}$, in the sense that we choose $\boldsymbol{u}_1 \cdot \boldsymbol{k} = \boldsymbol{u}_2 \cdot \boldsymbol{k} = \boldsymbol{u}_2 \cdot \boldsymbol{u}_1 = 0$ and $\boldsymbol{u}_1 \cdot \boldsymbol{k} = k$, where for $i \in \{1, 2, 3\}$, $\boldsymbol{u}_i = (u_{i,1}, u_{i,2}, u_{i,3})$, and where $k = \sqrt{\boldsymbol{k}^2}$. If we take the lattice spacing to be small enough, we can treat the differential operators in eq. (69) to act on continuous sine and cosine functions. Of course, this does not apply for all $\boldsymbol{k}$ at once. But we can always get the range of $\boldsymbol{k}$ we want to by making the lattice spacing small enough. We see that the Fourier transformed Lagrangian then becomes

$$\tilde{L} = \frac{1}{2}\sum_{\boldsymbol{k},s} \big(\big(- (-s\tilde{V}_{\boldsymbol{k},-s}\boldsymbol{k}) - \frac{\partial}{\partial t}\tilde{\boldsymbol{A}}_{\boldsymbol{k},s}\big)^2 - \big(\boldsymbol{k} \times (-s\tilde{\boldsymbol{A}}_{\boldsymbol{k},-s})\big)^2 + \xi\big(\frac{\partial}{\partial t}\tilde{V}_{\boldsymbol{k},s} - \boldsymbol{k} \cdot (-s\tilde{\boldsymbol{A}}_{\boldsymbol{k},-s})\big)^2\big), \tag{71}$$

where the factor $(-s)$ and the change of sign in the index in e.g. $(-s\tilde{V}_{\boldsymbol{k},-s}\boldsymbol{k})$ comes from differentiating the sine and cosine functions. Since $\tilde{\boldsymbol{A}}_{\boldsymbol{k},s} = \boldsymbol{u}_1\tilde{A}_{\perp_1\boldsymbol{k},s} + \boldsymbol{u}_2\tilde{A}_{\perp_2\boldsymbol{k},s} + \boldsymbol{u}_3\tilde{A}_{\parallel\boldsymbol{k},s}$, we can use how $\boldsymbol{u}_1$, $\boldsymbol{u}_2$ and $\boldsymbol{u}_3$ relates to $\boldsymbol{k}$ to obtain

$$\begin{aligned} \tilde{L} =& \frac{1}{2}\sum_{\boldsymbol{k},s} \big(\big(\frac{\partial}{\partial t}\tilde{A}_{\perp_1\boldsymbol{k},s}\big)^2 + \big(\frac{\partial}{\partial t}\tilde{A}_{\perp_2\boldsymbol{k},s}\big)^2 + \big(ks\tilde{V}_{\boldsymbol{k},-s} - \frac{\partial}{\partial t}\tilde{A}_{\parallel\boldsymbol{k},s}\big)^2 - k^2\tilde{A}^2_{\perp_1\boldsymbol{k},s} - k^2\tilde{A}^2_{\perp_2\boldsymbol{k},s} \dots \\ & \dots + \xi\big(\frac{\partial}{\partial t}\tilde{V}_{\boldsymbol{k},s} + ks\tilde{A}_{\parallel\boldsymbol{k},-s}\big)^2\big). \end{aligned} \tag{72}$$

The result turned out exactly as one might have expected. It is now clear to see how the to parameters describing the rotational part of $\boldsymbol{A}$ becomes decoupled in simple harmonic oscillators.

We can now use eq. (39) once again to Fourier transform the Hamiltonian of eq. (51). We have already seen via eq. (58) how we are allowed to go to the $\tilde{\boldsymbol{q}}$-basis. We can use the same argument to go to the basis of $(\tilde{A}_{\perp_1\boldsymbol{k},s}, \tilde{A}_{\perp_2\boldsymbol{k},s}, \tilde{A}_{\parallel\boldsymbol{k},s})$ instead of the $\tilde{\boldsymbol{A}}_{\boldsymbol{k},s}$-basis. The result is

$$\begin{aligned} \tilde{H} =& \sum_{\boldsymbol{k},s}\Big(\frac{1}{2\xi}\hat{p}^2_{\tilde{V}_{\boldsymbol{k},s}} + \frac{1}{2}\sum_i \big(\hat{p}^2_{\tilde{A}_{i,\boldsymbol{k},s}}\big) + \frac{1}{2}k^2\tilde{A}^2_{\perp_1\boldsymbol{k},s} + \frac{1}{2}k^2\tilde{A}^2_{\perp_2\boldsymbol{k},s} - (ks\tilde{A}_{\parallel\boldsymbol{k},-s})\hat{\Pi}_{\tilde{V}_{\boldsymbol{k},s}} \dots \\ & \dots + (ks\tilde{V}_{\boldsymbol{k},-s})\hat{\Pi}_{\tilde{A}_{\parallel\boldsymbol{k},s}}\Big). \end{aligned} \tag{73}$$

Eigensolutions to this Hamiltonian can be found, but we will not analyze this further here, at least not in this rendition of the paper. Note that the Fourier transform does not change the fact that V is essentially free, which is a problem when we want to introduce electrons into the theory as the energies will not be bounded from below.

8 Free photons described as two-dimensional harmonic oscillators

In this section, we will put restrictions $V = 0$ and $\nabla \cdot \boldsymbol{A} = 0$ on the A^μ field. In classical electromagnetism, this is just the Coulomb gauge condition when applied to a electromagnetic field in the absence of any charges. Clearly, these restrictions will not work when we introduce electrons in the theory, since we need those parts of the field to give us the Coulomb repulsion between the electrons. In this rendition of the paper, it will not be discussed how to construct a theory that contains electrons, but in a later rendition, I will propose one possible way of doing this. That solution will have the same description of photons, as will be introduced in the following. So even though we "cheat" here by assuming those restrictions, it is still my belief that it is possible to use the following description of photons in a full QED theory including electrons as well.

Setting $V = 0$ and $\nabla \cdot \boldsymbol{A} = 0$ in our quantum system means collapsing the Hilbert space into the subspace where $V = 0$ and $\tilde{A}_\parallel = 0$ and also restricting the path integral accordingly. Actually, $\nabla \cdot \boldsymbol{A} = 0$ is a more loose restriction than $\tilde{A}_\parallel = 0$, since the former do not put any restrictions on the freedom to add a constant vector $\boldsymbol{A}_0$. But setting $\tilde{A}_\parallel = 0$ certainly causes the divergence of $\boldsymbol{A}$ to vanish, and this is what we will do. This changes the Lagrangian of eq. (72) to

$$\tilde{L} = \frac{1}{2} \sum_{\boldsymbol{k},s} \big(\big(\frac{\partial}{\partial t} \tilde{A}_{\perp_1 \boldsymbol{k},s} \big)^2 + \big(\frac{\partial}{\partial t} \tilde{A}_{\perp_2 \boldsymbol{k},s} \big)^2 - k^2 \tilde{A}^2_{\perp_1 \boldsymbol{k},s} - k^2 \tilde{A}^2_{\perp_2 \boldsymbol{k},s} \big). \tag{74}$$

The Hamiltonian, which is now defined on a Hilbert space of only $\tilde{A}_{\perp_1}$ and $\tilde{A}_{\perp_2}$, therefore becomes

$$\tilde{H} = \sum_{\boldsymbol{k},s} \big(\frac{1}{2} \hat{p}^2_{\tilde{A}_{\perp_1 \boldsymbol{k},s}} + \frac{1}{2} \hat{p}^2_{\tilde{A}_{\perp_2 \boldsymbol{k},s}} + k^2 \tilde{A}^2_{\perp_1 \boldsymbol{k},s} + k^2 \tilde{A}^2_{\perp_2 \boldsymbol{k},s} \big). \tag{75}$$

This is just a Hamiltonian describing a system of $2N^3$ decoupled harmonic oscillators. Of course, our derivation does not hold for all $\boldsymbol{k}$ at once, as mentioned in the last section, but we can still get any desired range of $\boldsymbol{k}$ by choosing the lattice spacing a to be small enough.

Since $\tilde{H}$ is separable in all coordinates, let us look at the solutions for the oscillators of $\tilde{A}_{\perp_1 \boldsymbol{k},+1}$ and $\tilde{A}_{\perp_1 \boldsymbol{k},-1}$ for a particular $\boldsymbol{k}$. The oscillators describe a standing cosine wave and a sine wave respectively, both with wave vector $\boldsymbol{k}$. For simplicity, let us rename $\tilde{A}_{\perp_1 \boldsymbol{k},+1}$ and $\tilde{A}_{\perp_1 \boldsymbol{k},-1}$ as A_{cos} and A_{sin} in the following. We will also rename $\hat{p}_{\tilde{A}_{\perp_1 \boldsymbol{k},+1}}$ and $\hat{p}_{\tilde{A}_{\perp_1 \boldsymbol{k},-1}}$ as $\hat{p}_{cos}$ and $\hat{p}_{sin}$. Our hope is now that the solutions to this two-dimensional oscillator can be interpreted as photons moving with wave vectors either $\boldsymbol{k}$ or $-\boldsymbol{k}$. Since $\omega_0 = k$ for the oscillators, the energies certainly fit the dispersion relation for massless particles, which is promising.

If we take A_{cos} to lie on the x-axis and take A_{sin} to lie on the y-axis, we see that a movement of the wave function in the counterclockwise direction relates to a translation of the field in the $\boldsymbol{k}$-direction. This means the generator of rotations in the (A_{cos}, A_{sin})-plane,

$$\hat{L} = \hat{A}_{cos} \hat{p}_{sin} - \hat{A}_{sin} \hat{p}_{cos}, \tag{76}$$

should be proportional to the moment operator in the direction of $\boldsymbol{k}$ for the photons with wave vector $\boldsymbol{k}$ or $-\boldsymbol{k}$. In terms of the normal ladder operators,

$$\hat{a}_{cos} = \sqrt{\frac{k}{2}} \big(\hat{A}_{cos} + \frac{i}{k} \hat{p}_{cos} \big), \tag{77}$$

$$\hat{a}_{sin} = \sqrt{\frac{k}{2}} \big(\hat{A}_{sin} + \frac{i}{k} \hat{p}_{sin} \big), \tag{78}$$

we can express $\hat{L}$ and $\tilde{H}$ as

$$\hat{L} = i \big(\hat{a}_{cos} \hat{a}^\dagger_{sin} - \hat{a}^\dagger_{cos} \hat{a}_{sin} \big) \tag{79}$$

and

$$\tilde{H} = k \big(\hat{a}^\dagger_{cos} \hat{a}_{cos} + \hat{a}^\dagger_{sin} \hat{a}_{sin} + 1 \big). \tag{80}$$

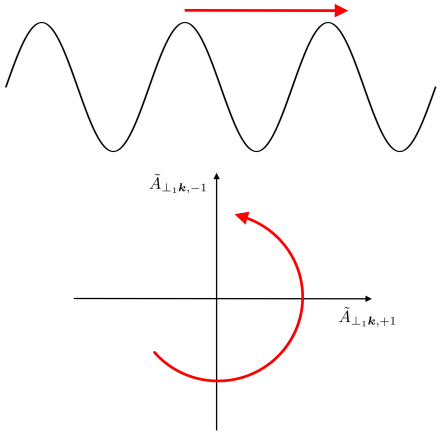

Figure 2: A two-dimensional harmonic oscillator. The two red arrows illustrates that a counterclockwise movement of a wave function in the $(\tilde{A}_{\perp_1 \boldsymbol{k},+1}, \tilde{A}_{\perp_1 \boldsymbol{k},-1})$-plane, also denoted as the (A_{cos}, A_{sin})-plane in this section, corresponds to a forward movement (in the $\boldsymbol{k}$-direction) of a wave with wave vector $\boldsymbol{k}$ in the $\tilde{A}_{\perp_1}$ field.

From the rotational symmetry of $\tilde{H}$, we know that $\tilde{H}$ and $\hat{L}$ commute (feel free to check this). Therefore, we can find a set of simultaneous eigenvectors. To do this, we will follow the procedure of Cohen-Tannoudji et al. [4]. First, we define

$$\hat{a}_d = \frac{1}{\sqrt{2}}\big(\hat{a}_{cos} - i\hat{a}_{sin}\big), \tag{81}$$

$$\hat{a}_g = \frac{1}{\sqrt{2}}\big(\hat{a}_{cos} + i\hat{a}_{sin}\big). \tag{82}$$

It can easily be shown, that the commutation relations between $\hat{a}_d$, $\hat{a}_d^\dagger$, $\hat{a}_g$ and $\hat{a}_g^\dagger$ are similar to those of $\hat{a}_{cos}$, $\hat{a}_{cos}^\dagger$, $\hat{a}_{sin}$ and $\hat{a}_{sin}^\dagger$, meaning that

$$[\hat{a}_d, \hat{a}_d^\dagger] = [\hat{a}_g, \hat{a}_g^\dagger] = 1 \tag{83}$$

are the only non-vanishing commutators between the four. We can now define

$$|n_d, n_g\rangle = \frac{1}{\sqrt{n_d! n_q!}} (\hat{a}_d^\dagger)^{n_d} (\hat{a}_g^\dagger)^{n_g} \, |0,0\rangle \,, \tag{84}$$

where $|0,0\rangle$ is the ground state. With similar arguments as when we derived eq. (23) in section 2, we can show that these states form an orthonormal basis, i.e. $\big\langle n_d, n_g \,\big|\, n'_d, n'_g \big\rangle = \delta_{n_d,n'_d}\delta_{n_g,n'_g}$. We can express the number operators accordingly:

$$\hat{N}_d = \hat{a}_d^\dagger \hat{a}_d, \tag{85}$$

$$\hat{N}_g = \hat{a}_g^\dagger \hat{a}_g. \tag{86}$$

Since

$$\hat{a}_d^\dagger \hat{a}_d = \frac{1}{2}\big(\hat{a}_{cos}^\dagger \hat{a}_{cos} + \hat{a}_{sin}^\dagger \hat{a}_{sin} - i\hat{a}_{cos}^\dagger \hat{a}_{sin} + i\hat{a}_{cos}\hat{a}_{sin}^\dagger\big), \tag{87}$$

$$\hat{a}_g^\dagger \hat{a}_g = \frac{1}{2}\big(\hat{a}_{cos}^\dagger \hat{a}_{cos} + \hat{a}_{sin}^\dagger \hat{a}_{sin} + i\hat{a}_{cos}^\dagger \hat{a}_{sin} - i\hat{a}_{cos}\hat{a}_{sin}^\dagger\big), \tag{88}$$

we can now rewrite $\tilde{H}$ as

$$\tilde{H} = k\big(\hat{a}_d^\dagger \hat{a}_d + \hat{a}_g^\dagger \hat{a}_g + 1\big) = k\big(\hat{N}_d + \hat{N}_g + 1\big). \tag{89}$$

We can also express $\hat{L}$ as

$$\hat{L} = \hat{a}_d^\dagger \hat{a}_d - \hat{a}_g^\dagger \hat{a}_g = \hat{N}_d - \hat{N}_g. \tag{90}$$

We see that $\hat{a}_d^\dagger$ and $\hat{a}_g^\dagger$ simply has the effect of raising the energy by k and respectively raising and lowering the angular momentum by 1. We know that $\exp(-i\theta\hat{L})$ has the effect of rotating the state by an angle of θ, which has the effect of translating the field by $\theta\lambda/(2\pi) = \theta/k$. This implies

$$\begin{aligned} e^{-i\frac{\theta}{k}\hat{p}_{photons}} = e^{-i\theta\hat{L}} \Rightarrow \\ \hat{L} = k\,\hat{p}_{photons}, \end{aligned} \tag{91}$$

where $\hat{p}_{photons}$ is the photon momentum in the $\boldsymbol{k}$-direction. Changing the angular momentum of the oscillator by 1, therefore means changing the momentum by k. This is exactly the result we could have hoped for. We can immediately interpret n_d as the occupation number of photons with momentum $\boldsymbol{k}$ and n_d as the occupation number of photons with momentum $-\boldsymbol{k}$. The wave functions, ψ_{n_d,n_g}, can be obtained from eq. (84). Some of the wave functions[4], described in polar coordinates (ρ,ϕ), where $\rho = \sqrt{A_{cos}^2 + A_{sin}^2}$, are listed here:

$$\begin{aligned} \psi_{0,0}(\rho,\phi) &= \frac{\sqrt{k}}{\sqrt{\pi}} e^{-\frac{k}{2}\rho^2} \\ \psi_{1,0}(\rho,\phi) &= \frac{\sqrt{k}}{\sqrt{\pi}} \sqrt{k}\rho e^{-\frac{k}{2}\rho^2} e^{i\phi} \\ \psi_{0,1}(\rho,\phi) &= \frac{\sqrt{k}}{\sqrt{\pi}} \sqrt{k}\rho e^{-\frac{k}{2}\rho^2} e^{-i\phi} \\ \psi_{2,0}(\rho,\phi) &= \frac{\sqrt{k}}{\sqrt{\pi}} (\sqrt{k}\rho)^2 e^{-\frac{k}{2}\rho^2} e^{2i\phi} \\ \psi_{1,1}(\rho,\phi) &= \frac{\sqrt{k}}{\sqrt{\pi}} \big((\sqrt{k}\rho)^2 - 1\big) e^{-\frac{k}{2}\rho^2} \\ \psi_{2,0}(\rho,\phi) &= \frac{\sqrt{k}}{\sqrt{\pi}} (\sqrt{k}\rho)^2 e^{-\frac{k}{2}\rho^2} e^{-2i\phi}. \end{aligned} \tag{92}$$

Whether ψ is given in polar or euclidean coordinates in the following should be clear from the context.

To get the full momentum operator for all photons in $A_{\perp_1}$, we first label $\hat{L}$ by $\boldsymbol{k}$, since there are one for each $\boldsymbol{k}$. We will also label it with $A_{\perp_1}$. The full momentum operator, which generates translations in a direction $\boldsymbol{r}$, $\boldsymbol{r}^2 = 1$, is then simply given by

$$\hat{P}_{photons,A_{\perp_1},\boldsymbol{r}} = \sum_{\boldsymbol{k}} (\boldsymbol{r}\cdot\boldsymbol{k}) \hat{L}_{A_{\perp_1},\boldsymbol{k}}. \tag{93}$$

Be sure not to confuse $\hat{L}$ with the actual angular momentum of the photons as the two things are not connected. A rotation in the $(\tilde{A}_{\perp_1\boldsymbol{k},+1}, \tilde{A}_{\perp_1\boldsymbol{k},-1})$-plane describes a phase change, not a rotation of polarization. If we want to include a different polarization, we need to include the space of $\tilde{A}_{\perp_2}$, for which the same results as above apply.

We now have a theory of free photons with the right dispersion relation and the right number of spinor indices, and, most importantly, we know exactly how the photons relate to their fields. We can now get an answer to how the photons interact three-dimensionally, even though they only have two spinor indices. We see that, while two photons obviously do not interfere in the momentum basis, their fields do interfere. This can be seen by considering a localized displacement of the $\boldsymbol{A}$ field, call it $\delta\boldsymbol{A}$, and noting that for two photons, we have $\delta\boldsymbol{A} = \boldsymbol{\delta}_1 + \boldsymbol{\delta}_2$, where the two (independent) distributions of $\boldsymbol{\delta}_1$ and $\boldsymbol{\delta}_2$ are given by the two photon wave function solutions respectively. Since $\boldsymbol{\delta}_1$ and $\boldsymbol{\delta}_2$ are proportional to the individual polarizations, which can only be perpendicular to $\boldsymbol{k}$, the fields of the photons interfere three-dimensionally.

How photons interfere three-dimensionally is explained in other quantum theories as well, and it is not a new feature of this theory. We will now look at result from the theory, which apparently has not been achieved before. Since we have an explicit definition of the phase of a photon state for any given $\boldsymbol{k} \in \mathbb{R}^+ \times \mathbb{R}^2$, it is now straightforward to deduce exactly which operator relates to the phase of a beam of photons. According to Gerry and Knight [5], the search for a quantum phase operator has so far been

[4] See Cohen-Tannoudji et al. [4].

on a trail-and-error basis, as I understand it, and the search has not yielded a fully satisfactory answer so far. One of the examples given in the book is the Susskind-Glogower (SG) operator,

$$\hat{E} = (\hat{N}+1)^{-\frac{1}{2}}\hat{a} = (\hat{a}\hat{a}^\dagger)^{-\frac{1}{2}}\hat{a}, \tag{94}$$

which approximates

$$\hat{E}\,|\phi\rangle = e^{i\phi}\,|\phi\rangle\,, \tag{95}$$

where $|\phi\rangle$ is the eigenvector to the phase operator with eigenvalue ϕ. We can call $\hat{E}$ an approximation to the *phase factor operator* when we want to be specific. As pointed out in the book, $\hat{E}$ is not unitary despite what one might expect from eq. (95). This fact can be seen by calculating

$$\hat{E}^\dagger\hat{E} = \hat{a}^\dagger(\hat{N}+1)^{-\frac{1}{2}}(\hat{N}+1)^{-\frac{1}{2}}\hat{a} = \hat{I} + |0\rangle\langle 0|. \tag{96}$$

It is important to point out that the creation and annihilation operators in definition (95) only creates or annihilates a photon with forward momentum $\boldsymbol{k}$. Therefore, $\hat{a}$ is analogous to $\hat{a}_d$ for $\boldsymbol{k} \in \mathbb{R}^+ \times \mathbb{R}^2$ and to $\hat{a}_g$ for $\boldsymbol{k} \in \mathbb{R}^- \times \mathbb{R}^2$ in our formalism. There are other proposals mentioned for the phase operator, but the SG operator is the preferred one used in Gerry and Knight [5].

The phase factor operator can easily be obtained from the formalism developed above. We will choose it so that it represents $\exp(-i\phi)$ instead of $\exp(i\phi)$. We therefore define

$$\hat{E}\,\psi = (A_{cos}, A_{sin}) \mapsto \left(\frac{A_{cos}}{\sqrt{A_{cos}^2 + A_{sin}^2}} - i\frac{A_{sin}}{\sqrt{A_{cos}^2 + A_{sin}^2}}\right)\psi(A_{cos}, A_{sin}) \tag{97}$$

for all $(A_{cos}, A_{sin}) \neq (0,0)$. Once again, we have suppressed the label of $\boldsymbol{k}$ and of either $\perp_1$ or $\perp_2$. Note that for consistency, both sides of eq. (97) are functions since we take our Hilbert space to be a vector space of functions. For $(A_{cos}, A_{sin}) = (0,0)$, we could choose to define $(\hat{E}\psi)(0,0) = \psi(0,0)$, so that the unitarity of $\hat{E}$ becomes very apparent. Actually, it does not matter what we define $(\hat{E}\psi)(0,0)$ to be as long as it is a finite number. This can be seen if we rewrite $\hat{E}$ as

$$\hat{E} = \sum_{n_d=0}^{\infty}\sum_{n_g=0}^{\infty}\sum_{n'_d=0}^{\infty}\sum_{n'_g=0}^{\infty} |n'_d, n'_g\rangle\,\langle n'_d, n'_g|\hat{E}|n_d, n_g\rangle\,\langle n_d, n_g|\,. \tag{98}$$

It is clear that each of the matrix elements, $\langle n'_d, n'_g|\hat{E}|n_d, n_g\rangle$, which can be calculated from the solutions obtained from eq. (84), is invariant of the choice of $(\hat{E}\psi)(0,0)$. Since we are going to use it later as well, what we will do to make eq. (97) defined for all (A_{cos}, A_{sin}) is to simply remove $(0,0)$ from the plane on which ψ is defined. This does not change the Hilbert space spanned by the $|n_d, n_g\rangle$ states; every matrix element of the operators on the space is unchanged, and every square-integrable wave function still resolves to the same vector in the $|n_d, n_g\rangle$-basis. To remove an isolated point in the domain does not change the differentiability of the wave functions either.

We can simplify eq. (98) by including a Kronecker delta, writing

$$\hat{E} = \sum_{n_d=0}^{\infty}\sum_{n_g=0}^{\infty}\sum_{n'_d=0}^{\infty}\sum_{n'_g=0}^{\infty} \delta_{n'_d-n'_g,n_d-n_g-1}\,|n'_d, n'_g\rangle\,\langle n'_d, n'_g|\hat{E}|n_d, n_g\rangle\,\langle n_d, n_g|\,. \tag{99}$$

This is justified because multiplying $e^{-i\phi}$ to an eigenfunction of $\hat{L}$ with eigenvalue $n_d - n_g$ has the effect of lowering its eigenvalue by 1, and it will therefore be orthogonal to all eigenvectors with eigenvalues different from $n_d - n_g - 1$.

Let us now try to express $\hat{E}$ in terms of $\hat{a}_d$ and $\hat{a}_g$. First, we see that eq. (81) and (82) can be inverted to yield

$$\hat{a}_{cos} = \frac{1}{\sqrt{2}}\big(\hat{a}_d + \hat{a}_g\big), \tag{100}$$

$$\hat{a}_{sin} = \frac{i}{\sqrt{2}}\big(\hat{a}_d - \hat{a}_g\big). \tag{101}$$

It follows that

$$\hat{A}_{cos} = \frac{1}{\sqrt{2k}}\big(\hat{a}_{cos} + \hat{a}^\dagger_{cos}\big) = \frac{1}{2\sqrt{k}}\big(\hat{a}_d + \hat{a}^\dagger_d + \hat{a}_g + \hat{a}^\dagger_g\big), \tag{102}$$

$$\hat{A}_{sin} = \frac{1}{\sqrt{2k}}\big(\hat{a}_{sin} + \hat{a}^\dagger_{sin}\big) = \frac{i}{2\sqrt{k}}\big(\hat{a}_d - \hat{a}^\dagger_d - \hat{a}_g + \hat{a}^\dagger_g\big), \tag{103}$$

and therefore we have

$$\hat{A}_{cos} - i\hat{A}_{sin} = \frac{1}{2\sqrt{k}}\big(\hat{a}_d + \hat{a}^\dagger_d + \hat{a}_g + \hat{a}^\dagger_g + \hat{a}_d - \hat{a}^\dagger_d - \hat{a}_g + \hat{a}^\dagger_g\big) = \frac{1}{\sqrt{k}}\big(\hat{a}_d + \hat{a}^\dagger_g\big) \tag{104}$$

and

$$\hat{A}^2_{cos} + \hat{A}^2_{sin} = \big(\hat{A}_{cos} - i\hat{A}_{sin}\big)\big(\hat{A}_{cos} - i\hat{A}_{sin}\big)^\dagger = \frac{1}{k}\big(\hat{a}_d + \hat{a}^\dagger_g\big)\big(\hat{a}^\dagger_d + \hat{a}_g\big). \tag{105}$$

Since our wave functions are no longer defined in $(0,0)$, it is clear to see that the inverse of $(\hat{A}^2_{cos}+\hat{A}^2_{sin})^{1/2}$, namely $(\hat{A}^2_{cos} + \hat{A}^2_{sin})^{-1/2}$, is defined. We can now finally write

$$\hat{E} = \big(\hat{A}^2_{cos} + \hat{A}^2_{sin}\big)^{-\frac{1}{2}}\big(\hat{A}_{cos} - i\hat{A}_{sin}\big) = \big((\hat{a}_d + \hat{a}^\dagger_g)(\hat{a}^\dagger_d + \hat{a}_g)\big)^{-\frac{1}{2}}\big(\hat{a}_d + \hat{a}^\dagger_g\big). \tag{106}$$

It is remarkable how similar the RHS of this equation looks to the SG operator of eq. (94), just with $\hat{a}$ replaced by $(\hat{a}_d + \hat{a}^\dagger_g)$. It now becomes apparent why one cannot define $\hat{E}$ on the subspace of states only with forward moving photons, that is to say states with $n_g = 0$, and then still expect $\hat{E}$ to be unitary. From eq. (106), we get the idea that $\hat{E}$ will have the effect of exiting backward moving photons. We can show this by turning to the eigenfunctions of eq. (92). As an example, we see that the matrix element $\langle 1,1|e^{-i\hat{\phi}}|1,0\rangle$ is given by

$$\langle 1,1|e^{-i\hat{\phi}}|1,0\rangle = \frac{k\sqrt{k}}{\pi}\int_0^\infty\int_0^{2\pi}(k\rho^2-1)\rho e^{-k\rho^2}e^{-i\phi}e^{i\phi}\rho\,d\phi\,d\rho = \frac{1}{4}\sqrt{\pi} \neq 0, \tag{107}$$

which shows that $\hat{E}$ can excite a backwards moving photon from the $|1,0\rangle$ state. Let us now we define a new operator, call it $\hat{E}_+$, on the subspace spanned by all $|n_d,0\rangle$ states, $n_d \in \{0,1,2,\ldots\}$. We define it by

$$\hat{E}_+ = \sum_{n_d=0}^{\infty} |n_d,0\rangle\,\langle n_d,0|\hat{E}|n_d+1,0\rangle\,\langle n_d+1,0|\,, \tag{108}$$

such that

$$\big\langle n'_d,0\big|\hat{E}_+\big|n_d,0\big\rangle = \big\langle n'_d,0\big|\hat{E}\big|n_d,0\big\rangle \tag{109}$$

for all $n_d, n'_d \in \{1,2,\ldots,\infty\}$. If we then consider for instance

$$\hat{E}_+\,|1,0\rangle = \hat{E}\,|1,0\rangle - \langle 1,1|\hat{E}|1,0\rangle\,|1,1\rangle - \langle 2,2|\hat{E}|1,0\rangle\,|2,2\rangle - \ldots, \tag{110}$$

we see that

$$\langle 1,0|\hat{E}^\dagger_+\hat{E}_+|1,0\rangle = \langle 1,0|\hat{E}^\dagger\hat{E}|1,0\rangle - |\,\langle 1,1|\hat{E}|1,0\rangle\,|^2 - \ldots = 1 - \frac{\pi}{16} - \ldots < 1, \tag{111}$$

and therefore we can conclude that $\hat{E}_+$ is indeed not unitary.

Even though $\hat{E}_+$ is not unitary, it is still useful since it represents the phase factor for all states containing only photons moving in the forward direction. If we want to measure the phase of a photon beam rather than a standing wave, we might as well use $\hat{E}_+$. We can calculate the matrix elements of $\hat{E}_+$ by using the fact (see Cohen-Tannoudji et al. [4]) that the wave function for $|n_d,0\rangle$ is given in polar coordinates by

$$\psi_{n_d,0}(\rho,\phi) = \frac{\sqrt{k}}{\sqrt{\pi}\sqrt{n_d!}}(\sqrt{k}\rho)^{n_d}e^{-\frac{k}{2}\rho^2}e^{in_d\phi}. \tag{112}$$

We therefore have

$$\begin{aligned}\langle n_d,0|e^{i\hat{\phi}}|n_d+1,0\rangle &= \frac{k}{\pi\sqrt{n_d!(n_d+1)!}}\int_0^\infty\int_0^{2\pi}(\sqrt{k}\rho)^{2n_d+1}e^{-k\rho^2}\rho\,d\phi\,d\rho \\ &= \frac{2k}{n_d!\sqrt{n_d+1}}\int_0^\infty(\sqrt{k}\rho)^{2n_d+1}e^{-k\rho^2}\rho\,d\rho \\ &= \frac{1}{n_d!\sqrt{n_d+1}}\int_0^\infty x^{n_d+\frac{1}{2}}e^{-x}\,dx \\ &= \frac{1}{n_d!\sqrt{n_d+1}}\Gamma\Big(n_d+1+\frac{1}{2}\Big),\end{aligned} \tag{113}$$

where we have changed integration variables to $x = k\rho^2$ to get the third equality and then identified the resulting integral as the gamma function. If we write $\Gamma(n_d+1+1/2)$ more compactly as $(n_d+1/2)!$, we see that eq. (108) now becomes

$$\hat{E}_+ = \sum_{n_d=0}^{\infty}\frac{(n_d+\frac{1}{2})!}{n_d!\sqrt{n_d+1}}|n_d,0\rangle\langle n_d+1,0|. \tag{114}$$

The sequence

$$(a_n)_{n\in\{0,1,2,\ldots\}} = \Big(\frac{(n+\frac{1}{2})!}{n!\sqrt{n+1}}\Big)_{n\in\{0,1,2,\ldots\}} \approx (0.8862, 0.9400, 0.9594, 0.9693, 0.9754, \ldots) \tag{115}$$

is convergent, and it converges to 1. A way to see this is to use the duplication formula:

$$\Gamma(n+1)\Gamma(n+1+\frac{1}{2}) = 2^{-2n-1}\sqrt{\pi}\,\Gamma(2n+2) \Rightarrow \frac{(n+\frac{1}{2})!}{n!} = 2^{-2n-1}\sqrt{\pi}(2n+1)\frac{(2n)!}{(n!)^2}. \tag{116}$$

Stirling's approximation tells us that

$$\frac{(2n)!}{(n!)^2} \sim \frac{1}{\sqrt{2\pi}}\frac{(2n)^{2n+\frac{1}{2}}}{(n^{n+\frac{1}{2}})^2} = \frac{2^{2n}}{\sqrt{\pi}}n^{-\frac{1}{2}} = \frac{4^n}{\sqrt{\pi n}} \tag{117}$$

for large n, and therefore we have

$$\lim_{n\to\infty}a_n = \lim_{n\to\infty}\frac{1}{2}\frac{1}{4^n}\sqrt{\pi}\frac{2n+1}{\sqrt{n+1}}\frac{4^n}{\sqrt{\pi n}} = 1. \tag{118}$$

To show that $a_n < 1$ for all n, one can easily calculate $a_{n+1} - a_n$ to show that $a_{n+1} - a_n > 0$ for all n. The strictly increasing sequence will therefore never reach 1.

The similarity between our $\hat{E}_+$ and the SG operator, which is also defined on the space of states with only forward moving photons, is now striking. If we rename all $|n_d,0\rangle$ as $|n\rangle$, $n = n_d \in \{0,1,2,\ldots\}$, for compactness, we can write $\hat{E}_+$ as

$$\hat{E}_+ = \sum_{n=0}^{\infty}a_n|n\rangle\langle n+1|. \tag{119}$$

Meanwhile, the SG operator, as can be seen from eq. (94), can be written as

$$\hat{E}_{SG} = \sum_{n=0}^{\infty}|n\rangle\langle n+1|. \tag{120}$$

We therefore see that the SG operator is a very good approximation to $\hat{E}_+$, especially for large occupation numbers. Just remember that $\hat{E}_+$ represents $\exp(-i\hat{\phi})$, and not $\exp(i\hat{\phi})$ like in eq. (95).

A natural question to ask now is what the uncertainty is on a measurement of the phase factor of a beam of photons. Since $\hat{E}_+$ is not hermitian, we cannot use the commutator to get a lower limit on e.g.

$\Delta N \Delta E_+$. We can, however, still calculate the variance of $\hat{E}$ for any state in the full Hilbert space. It is given by

$$\begin{aligned} |\Delta E|^2 &= \left\langle \left(e^{i\hat{\phi}} - \langle \hat{E} \rangle^* \right)\left(e^{-i\hat{\phi}} - \langle \hat{E} \rangle \right) \right\rangle \\ &= \left\langle \left(\hat{E}^\dagger - \langle \hat{E} \rangle^* \right)\left(\hat{E} - \langle \hat{E} \rangle \right) \right\rangle \\ &= \langle \hat{E}^\dagger \hat{E} \rangle - |\langle \hat{E} \rangle|^2 \\ &= 1 - |\langle \hat{E} \rangle|^2. \end{aligned} \tag{121}$$

When we are interested in a state with only forward moving photons, $|\psi\rangle = \sum_{n=0}^{\infty} c_n |n\rangle$, we can now use eq. (109) and (119) to obtain

$$|\Delta E_+|^2 = 1 - |\langle \psi | \hat{E}_+ | \psi \rangle|^2 = 1 - \Big| \sum_{n=0}^{\infty} a_n c_n^* c_{n+1} \Big|^2. \tag{122}$$

To check that the RHS of this formula is greater than 0, define two new states: $|\psi_1\rangle = \sum_{n=0}^{\infty} a_n c_n |n\rangle$ and $|\psi_2\rangle = \sum_{n=0}^{\infty} c_{n+1} |n\rangle$. Note that $\langle \psi_1 | \psi_1 \rangle < 1 \geq \langle \psi_2 | \psi_2 \rangle$, and therefore $RHS = 1 - |\langle \psi_1 | \psi_2 \rangle|^2 > 0$. Let us now try to search for a state, $|\psi\rangle$, that minimizes $|\Delta E_+|^2$. If we define $|\psi\rangle$ by

$$|\psi\rangle = \sum_{n=l}^{l+m-1} \frac{1}{\sqrt{m}} |n\rangle , \tag{123}$$

we have, for large enough l,

$$|\Delta E_+|^2 = 1 - \Big| \sum_{n=0}^{\infty} a_n c_n^* c_{n+1} \Big|^2 \approx 1 - \Big| \sum_{n=0}^{\infty} c_n^* c_{n+1} \Big|^2 = 1 - \frac{(m-1)^2}{m^2} = \frac{2m-1}{m^2}. \tag{124}$$

This result goes to 0 as m goes to infinity, and we can therefore get an arbitrarily small $|\Delta E_+|^2$ by choosing m large enough. This shows that we can get an arbitrarily precise phase for a photon beam without any backward moving photons.

9 Conclusion

Looking into the subject of trying to find the underlying Hamiltonian for electromagnetism turned out to be very rewarding, even tough we have only discussed free fields so far. It resulted in a theory of photons, in which the photons are explicitly related to the field in terms of the excitations they represent. The theory has provided us with an answer why the phase factor operator is not unitary. It has also provided us with a fully satisfactory phase factor operator, which when defined for a state of only forward moving photons, turned out to be approximately equal to the known SG operator.

References

[1] T. Lancaster and S. J. Blundell, *Quantum Field Theory for the Gifted Amateur* (Oxford University Press, New York, 2014).

[2] M. Srednicky, *Quantum Field Theory* (Cambridge University Press, New York, 2007).

[3] B. Durhuus and J. P. Solovej, *Mathematical Physics*, Lecture notes, Department of Mathematical Sciences, University of Copenhagen, 2014.

[4] C. Cohen-Tannoudji, B. Diu, and F. Laloë, *Quantum Mechanics* (John Wiley & Sons, New York, 1977), Vol. 1, Chap. VI, Compliment D_{VI}.

[5] C. C. Gerry and P. L. Knight, *Introductory Quantum Optics* (Cambridge University Press, Cambridge, 2005).