An alternative derivation of the Hamiltonian of quantum electrodynamics

Mads J. Damgaard*

November 25, 2022

Abstract

We derive a proto-Hamiltonian of quantum electrodynamics (QED) from the coupled Dirac equation by quantizing the electromagnetic field. We then introduce a process of eliminating the gauge symmetry via separation of variables, and argue that this does not break the Lorentz covariance of the theory. From this approach, we obtain a Hamiltonian that is similar to the conventional one of QED. We conclude the paper short of making the Dirac sea reinterpretation, where one would otherwise reinterpret the negative-energy solutions to the Dirac equation as antiparticles.

1 Introduction

In conventional quantum field theory (QFT), derivations of the theories can often be quite involved. As a reader trying to learn about the field for the first time, you can sometimes get the feeling that you need to understand QFT in order to understand QFT.

This paper offers a slightly different approach to deriving such a theory, in particular the theory of quantum electrodynamics (QED). It may not help in understanding the conventional derivations of QED, but it might help to get a clearer understanding of the theory, altogether.

We will thus analyze how to obtain the Hamiltonian of QED in this paper by starting out with the Dirac equation and then quantizing the electromagnetic field. We will then see how the gauge symmetry can be eliminated, not before, but after we have made this quantization. In this process, we will take great care not to break the initial Lorentz covariance of the theory, since this symmetry is of course an important requirement for the desired theory.

2 The goal for an initial theory

In conventional QFT, all particles are said to arise as excitations of quantum fields, both bosons and fermions.[1] But the approach of this paper is instead to introduce the fermions simply as the normal wave functions known from fundamental quantum mechanics, such that only the bosons of the theory arise as excited modes of a quantum field.

---

*B.Sc. at the Niels Bohr Institute, University of Copenhagen. E-mail: fxn318@alumni.ku.dk. GitHub folder: https://www.github.com/mjdamgaard/notes.

[1] See e.g. Srednicki [7].

In order to be able to give the theory a clear mathematical treatment, we want to start with a discretized bosonic field, and then analyze later on what happens in the continuum limit. And for the fermions, we want to start with just a finite number of particles, call it n. Later on we will then extend the theory to a Fock space of fermions.

We thus want to consider an initial theory on a Hilbert space of the form

$$\mathbf{H}_{init} = \mathbf{H}_B \otimes \mathbf{H}_F, \tag{1}$$

where $\mathbf{H}_B = L^2(\mathbb{R}^N)$ is a space of square-integrable wave functions over an N-dimensional space of field configurations for the bosonic field, and where $\mathbf{H}_F = L^2(\mathbb{R}^{3n}; \mathbb{C}^{4^n})$ is a space of 4^n-component spinor functions over a set of $3n$ spatial particle coordinates for the n fermions. Note that $\otimes$ here denotes a tensor product.[2] At some point we will also restrict $\mathbf{H}_F$ to only contain antisymmetric wave functions, but this will only be relevant later on.

We will tend to denote the vectors in $\mathbf{H}_B$, $\mathbf{H}_F$ and $\mathbf{H}_{init}$ respectively by ϕ, ψ and χ in this paper, and we will also tend to let $\mathbf{q} \in \mathbb{R}^N$ and $\bar{\mathbf{x}} = (\mathbf{x}_1, \ldots, \mathbf{x}_n) \in \mathbb{R}^{3n}$ denote the input coordinate vectors of the functions in $\mathbf{H}_B$ and $\mathbf{H}_F$, respectively.

Let us also assume that $\mathbf{H}_{init}$ is space of functions, such that

$$\mathbf{H}_{init} = L^2(\mathbb{R}^N \times \mathbb{R}^{3n}; \mathbb{C}^{4^n}). \tag{2}$$

This is in agreement with Eq. (1) if we use a specific version of the tensor product for this paper where $f \otimes g = (x, y) \mapsto f(x)g(y)$ when the vectors f and g are functions.

Our overall goal is now to find a quantum theory on $\mathbf{H}_{init}$ that will give the predictions known from QED in the continuum limit. Since Lorentz covariance is a very important symmetry, both in classical electrodynamics and in QED, an important part of the goal will be to find a theory that is likely to be Lorentz-covariant in said limit.

If we first consider the spinors in $\mathbf{H}_F$, we know that there is a famous Lorentz-covariant equation for such spinors, namely the Dirac equation. We also know that this equation can be coupled with the electromagnetic field in a Lorentz-covariant way, and that this variant of the theory agrees with physical observations to a great extend (at least once the negative-energy solutions are reinterpreted as antiparticles). When solved for the Dirac Hamiltonian, call it $\hat{H}_D$, this coupled version of the Dirac equation can be written as[3]

$$\hat{H}_D(\varphi, \mathbf{A})\psi(\bar{\mathbf{x}}) = \sum_{j=1}^{n} \Big(\boldsymbol{\alpha}_j \cdot \big(\hat{\mathbf{p}}_j - q_F \mathbf{A}(\mathbf{x}_j)\big) + \beta_j m_F + q_F \varphi(\mathbf{x}_j) \Big) \psi(\bar{\mathbf{x}}) \tag{3}$$

for all $\bar{\mathbf{x}} = (\mathbf{x}_1, \ldots, \mathbf{x}_n) \in \mathbb{R}^{3n}$ and all ψ of some (dense) subspace of $\mathbf{H}_F$. Here, φ and $\mathbf{A}$ are the electric potential and the magnetic vector potential, respectively, and the quantities q_F and m_F are the electric charge and mass of the fermions. The operator $\hat{\mathbf{p}}_j$ is the standard momentum operator for jth fermion, defined by $\hat{\mathbf{p}}_j = -i(\partial/\partial x_{j1}, \partial/\partial x_{j2}, \partial/\partial x_{j3})$ for all $j \in \{1, \ldots, n\}$. If we use the same basis for α_{j1}, α_{j2}, α_{j3} and β_j as in Shankar [8], these matrices are then 4^n-by-4^n extensions of the four-by-four matrices given by

$$\alpha_k = \begin{pmatrix} 0 & \sigma_k \\ \sigma_k & 0 \end{pmatrix}, \quad \beta = \begin{pmatrix} \sigma_0 & 0 \\ 0 & -\sigma_0 \end{pmatrix}, \quad \sigma_\mu = \begin{pmatrix} \delta_{\mu 0} + \delta_{\mu 3} & \delta_{\mu 1} - i\delta_{\mu 2} \\ \delta_{\mu 1} + i\delta_{\mu 2} & \delta_{\mu 0} - \delta_{\mu 3} \end{pmatrix} \tag{4}$$

[2] See e.g. Hall, Appendix A.1, for an introduction to the abstract tensor product, although for this paper, we will mostly be using a specific version of the tensor product where products of (spaces of) functions turn into other (spaces of) functions.

[3] See e.g. Shankar [8].

for all $k \in \{1,2,3\}$, $\mu \in \{0,1,2,3\}$. Here we have used Kronecker's delta, denoted by δ_{ij}, to write up the Pauli matrices in the last equation. For the 4^n-by-4^n extensions, we need all α_{jk} and β_j to only (potentially) change the spin of the jth fermion. This is achieved by having

$$\alpha_{jk} = \overbrace{I \otimes \cdots \otimes I}^{j-1} \otimes \alpha_k \otimes \overbrace{I \otimes \cdots \otimes I}^{n-j}, \quad \beta_j = \overbrace{I \otimes \cdots \otimes I}^{j-1} \otimes \beta \otimes \overbrace{I \otimes \cdots \otimes I}^{n-j} \tag{5}$$

for all $j \in \{1,\ldots,n\}$, $k \in \{1,2,3\}$, where I is the 4-by-4 identity matrix and where $\otimes$ in this context denotes the matrix direct product. And as a last point of clarification for Eq. (3), we also allow ourselves to implicitly reinterpret matrices such as $\boldsymbol{\alpha}_j \cdot \hat{\mathbf{p}}_j$, $\boldsymbol{\alpha}_j \cdot \mathbf{A}$ and β_j as operators in such contexts, namely by letting $\beta_j \psi = \bar{\mathbf{x}} \mapsto \beta_j(\psi(\bar{\mathbf{x}}))$ and $(\alpha_{jk}\hat{p}_{jk})\psi = \bar{\mathbf{x}} \mapsto \alpha_{jk}(\hat{p}_{jk}\psi(\bar{\mathbf{x}}))$, for example.

For the extension of $\hat{H}_D$ onto the combined Hilbert space $\mathbf{H}_{init}$, we then naturally want to have the electromagnetic four-potential $A^\mu = (\varphi, \mathbf{A})$ be parameterized by $\mathbf{q}$. Let us therefore redefine $\hat{H}_D$ as the function $\hat{H}_D : \mathbb{R}^N \to (\mathbf{H}_F \to \mathbf{H}_F)$ given by

$$\hat{H}_D(\mathbf{q})\psi(\bar{\mathbf{x}}) = \sum_{j=1}^{n} \Big(\boldsymbol{\alpha}_j \cdot \big(\hat{\mathbf{p}}_j - q_F \mathbf{A}(\mathbf{q}, \mathbf{x}_j)\big) + \beta_j m_F + q_F \varphi(\mathbf{q}, \mathbf{x}_j) \Big) \psi(\bar{\mathbf{x}}), \tag{6}$$

such that φ and $\mathbf{A}$ are now functions of $\mathbf{q}$. In the next section, 3, we will choose more specifically how $A^\mu = (\varphi, \mathbf{A})$ is parameterized by $\mathbf{q}$. It will turn out that having each entry of $\mathbf{q}$ represent a mode amplitude of the field will be a good choice. This means that φ and $\mathbf{A}$ will be discretized in Fourier space, but will be continuous functions over all $\mathbf{x}_j \in \mathbb{R}^3$ in position space, thus giving us nice C^∞-potentials for $\hat{H}_D(\mathbf{q})$.

Extending $\hat{H}_D(\mathbf{q})$ onto $\mathbf{H}_{init}$ is now straightforward. Let us call this extension $\hat{H}_{IF}$ since it will thus contain the Dirac Interaction as well as the free energy of the Fermions. We will then define $\hat{H}_{IF}$ formally by

$$\hat{H}_{IF}(\phi \otimes \psi)(\mathbf{q}, \bar{\mathbf{x}}) = \phi(\mathbf{q}) \hat{H}_D(\mathbf{q}) \psi(\bar{\mathbf{x}}) \tag{7}$$

for all $\phi \in \mathbf{H}_B$, $\psi \in \mathbf{H}_F$, $\mathbf{q} \in \mathbb{R}^N$ and $\bar{\mathbf{x}} \in \mathbb{R}^{3n}$, where we are again using the specific version of the tensor product described under Eq. (2). Since $\hat{H}_{IF}$ is linear, this defines it for more general (entangled) states as well.

Note that $\hat{H}_{IF}$ is of course not defined on all of $\mathbf{H}_{init}$, since for some $\chi \in \mathbf{H}_{init}$, the function $\hat{H}_{IF}\chi$ will not be normalizable. So by a "formal definition" of an operators such as Eq. (7), we thus mean that we simply forget the fact that not all vectors can be part of the operator's domain. In the following text, we will mainly use formal definitions of operators such as this, and only discuss the domains of the most relevant operators in Appendix B.

We can also define $\hat{H}_{IF}$ formally using Dirac's bra–ket notation instead, which might be more familiar to some readers. We can thus let $|\psi\rangle = \psi$ and $|\mathbf{q}\rangle = \mathbf{q}' \mapsto \delta^3(\mathbf{q}' - \mathbf{q})$, where δ is the Dirac delta function. The latter kind of kets are of course not actual vectors of $\mathbf{H}_B$, but rather so-called "generalized vectors" or "distributions" (see e.g. Hall [4]). With this notation, we can also formally define $\hat{H}_{IF}$ by

$$\hat{H}_{IF} |\mathbf{q}, \psi\rangle = |\mathbf{q}\rangle \otimes \hat{H}_D(\mathbf{q}) |\psi\rangle \tag{8}$$

for all $\psi \in \mathbf{H}_F$ and $\mathbf{q} \in \mathbb{R}^N$, where $|\mathbf{q}, \psi\rangle = |\mathbf{q}\rangle \otimes |\psi\rangle$. We will be using the bra–ket notation more in the following text as well, and often with similar bases for the kets as the ones chosen here.

Now, since we expect $\hat{H}_{IF}$ to contain both the fermionic free energy and the interaction between the bosonic[4] field and the fermions, it would be natural to look for an operator, call it $\hat{H}_B$, which will contain the free energy for the bosonic field. We might thus guess at a full initial Hamiltonian, call it $\hat{H}_{init}$, of the form

$$\hat{H}_{init} = \hat{H}_B + \hat{H}_{IF}, \quad \hat{H}_B = \hat{H}_{EM} \otimes \hat{I}_F, \tag{9}$$

where $\hat{I}_F$ denotes the identity operator on $\mathbf{H}_F$ (and similarly for $\hat{I}_B$ on $\mathbf{H}_B$), first of all, and where $\hat{H}_{EM}$ is a Hamiltonian on $\mathbf{H}_B$ that corresponds to the classical Hamiltonian of the electromagnetic field, A^μ.

The point of having a quantum mechanical Hamiltonian, $\hat{H}$, that corresponds to a classical Hamiltonian, H, which is a Legendre transform of some Lagrangian, L, is that, at least for some class of Lagrangians, we can then derive a path integral for $\hat{H}$, formally given by

$$\langle \mathbf{q}_t | e^{-i\hat{H}t} | \mathbf{q}_0 \rangle = \int_{\mathbf{q}(0)=\mathbf{q}_0}^{\mathbf{q}(t)=\mathbf{q}_t} \mathcal{D}\mathbf{q} \, \exp\Big(i \int_0^t L\big(\mathbf{q}(t), \dot{\mathbf{q}}(t)\big) dt\Big). \tag{10}$$

Here $\mathbf{q}$ is thus turned into a function of time for this formal expression, and $\dot{\mathbf{q}}$ then denotes the time derivative of this function. (The reader is expected to be familiar with this formal notation for path integrals, but if not, see e.g. any of References [6]–[9].) We will let the limits of such path integrals be implicit from now on and simply write $\int \mathcal{D}\mathbf{q}$ instead of $\int_{\mathbf{q}(0)=\mathbf{q}_0}^{\mathbf{q}(t)=\mathbf{q}_t} \mathcal{D}\mathbf{q}$.

From this path integral, we can see that if L is Lorentz-invariant, more precisely meaning that its density $\mathcal{L}$ is a Lorentz-invariant scalar, where $L = \int \mathcal{L}(\mathbf{x}) d\mathbf{x}$, we should then expect $\hat{H}$ to generate Lorentz-covariant dynamics for the quantum system.

We can then ask ourselves what happens for the path integral of $\hat{H}_{init}$ if $\hat{H}_{EM}$ fulfills the conditions of Eq. (10). As we will see in Sect. 4, the path integral for $\hat{H}_{init}$ will then be given by

$$\langle \mathbf{q}_t, \psi' | e^{-i\hat{H}_{init}t} | \mathbf{q}_0, \psi \rangle = \int \mathcal{D}\mathbf{q} \, \exp\Big(i \int_0^t L_B\big(\mathbf{q}(t), \dot{\mathbf{q}}(t)\big) dt\Big) K_{IF}(t, \mathbf{q}, \psi', \psi), \tag{11}$$

where $L_B \equiv L_{EM}$ is the Lagrangian corresponding to $\hat{H}_{EM}$ and $K_F(t, \mathbf{q}, \psi', \psi)$ is the fermion propagator over the (time-dependent) classical field, given by $\mathbf{q}(t)$ at all times t. Using the notation of the time-ordered integral, this propagator can be defined as

$$K_{IF}(t, \mathbf{q}, \psi', \psi) = \langle \psi' | \, T\Big\{ \exp\Big(-i \int_0^t \hat{H}_D\big(\mathbf{q}(t)\big) dt\Big)\Big\} \, |\psi\rangle \,, \tag{12}$$

which is just a formal way of writing $K_{IF}(t, \mathbf{q}, \psi', \psi) = \lim_{M\to\infty} \langle \psi' | \hat{U}_M \hat{U}_{M-1} \cdots \hat{U}_1 | \psi \rangle$, where $\hat{U}_m = \exp[-i\hat{H}_D(\mathbf{q}(m\delta t))\delta t]$, $\delta t = t/M$, for all m.

Equation (11) shows why we are interested in a $\hat{H}_{init}$ of the form given by Eq. (9): We know that the propagator of the Dirac Hamiltonian is Lorentz-covariant, so if we can just find a $\hat{H}_B$ that corresponds to a Lagrangian L_B of the A^μ field that is Lorentz-invariant in the continuum limit, then it is reasonable to expect that $\hat{H}_{init}$ will give a Lorentz-covariant quantum theory in the same limit.

[4] Note that we here allow ourselves to call it the *bosonic* field since we expect bosons to come from it in the continuum limit. But until we get there, the term (in our case of QED) simply refers to the A^μ field, parameterized by a finite-dimensional $\mathbf{q} \in \mathbb{R}^N$.

The heuristic reasoning for this is similar to that of conventional QFT. If we rewrite the path integral formally as

$$\int \mathcal{D}A^\mu\, T\Big\{ \exp\Big(i \int \mathcal{L}_B(A^\mu, x^\mu) - \hat{\mathcal{H}}_D(A^\mu, x^\mu) dx^\mu\Big)\Big\}, \tag{13}$$

where $\mathcal{L}_B$ is the bosonic Lagrangian density and $\hat{\mathcal{H}}_D$ is the (operator-valued) density of the Dirac Hamiltonian, we see that $\mathcal{L}_B(x^\mu) - \hat{\mathcal{H}}_D(A^\mu, x^\mu)$ is locally Lorentz-covariant. We thus see that we can expect each path to give a Lorentz-covariant contribution to the fermion propagator (times a Lorentz-invariant phase factor from the action integral) over the same spacetime volume when viewed in different inertial frames.

Let us thus be motivated by the promising-looking Eq. (11) and look for a $\hat{H}_{init}$ that meets these conditions. This means that we need to derive $\hat{H}_{EM}$ from the Lagrangian of classical electromagnetism, and also decide how $\mathbf{q}$ parameterizes the A^μ field more precisely. We will do this in the following section.

3 Deriving a Hamiltonian of electromagnetism to complete $\hat{H}_{init}$

We want to look for a $\hat{H}_{EM}$ that corresponds to the Lagrangian for classical electromagnetism, as this will lead us to the desired $\hat{H}_{init}$, described in Eqs. (6–9). We know that the Lagrangian density for electromagnetism can be written as

$$\mathcal{L}_{EM} = \frac{1}{2}\big(\nabla\varphi + \frac{\partial}{\partial t}\mathbf{A}\big)^2 - \frac{1}{2}\big(\nabla \times \mathbf{A}\big)^2 + \frac{1}{2\xi}\big(\nabla\cdot\mathbf{A} + \frac{\partial}{\partial t}\varphi\big)^2, \tag{14}$$

where we have used the gauge freedom to add the last term with a positive constant ξ^{-1} in front. This term can thus be seen to only fix the classical equations of motion to the Lorenz gauge, namely where $\partial\varphi/\partial t = -\nabla\cdot\mathbf{A}$. It will also not break the Lorentz symmetry of $\mathcal{L}$, which is important to our goal. The reason we want to add this last term in Eq. (14) is that it will allow us to evaluate the Gaussian integrals in the phase space path integral, which we will derive in Sect. 4, getting us to the configuration space path integral of Eq. (11). Adding $(2\xi)^{-1}(\nabla\cdot\mathbf{A} + \partial\varphi/\partial t)^2$ this way is a well-known trick in QFT (see e.g. Srednicki [7]).

As mentioned, we first of all want to discretize the A^μ field. If we were to make this discretization in position space, we could model it as a lattice with $N/4$ lattice atoms (assuming $N/4 \in \mathbb{N}$), each one allowed to move in four dimension such that each displacement represents the four vector $A^\mu(\mathbf{x})$ at that point in space. We could then have the entries of $\mathbf{q} \in \mathbb{R}^N$ represent all these displacements. This is not what we will do, however, but let us keep this lattice in mind anyway, and let us denote its volume by $\mathcal{V}$.

Instead, we will use a Fourier transform of this lattice as our starting point, since this will allow us to evaluate the spatial derivatives of Eq. (14) nicely. Since the electromagnetic field is real-valued for all its four components, let us make the Fourier transform in terms of (normalized) sine and cosine functions, given for all $\sigma \in \{1, -1\}$ by

$$f_{\mathbf{k}\sigma}(\mathbf{x}) = \frac{i^{(\sigma-1)/2}}{\sqrt{2\mathcal{V}}}(e^{i\mathbf{k}\cdot\mathbf{x}} + \sigma e^{-i\mathbf{k}\cdot\mathbf{x}}) = \begin{cases} \sqrt{\frac{2}{\mathcal{V}}}\cos(\mathbf{k}\cdot\mathbf{x}), & \sigma = 1 \\ \sqrt{\frac{2}{\mathcal{V}}}\sin(\mathbf{k}\cdot\mathbf{x}), & \sigma = -1 \end{cases}. \tag{15}$$

For our discrete case, we take $\mathbf{x}$ and $\mathbf{k}$ to range over finite sets of values.

The exact ranges of $\mathbf{x}$ and $\mathbf{k}$ will not be too important for our purposes, but for readers interested in an example, we could first of all choose $\mathbf{x} \in \{l\delta x\}^3_{l\in\{1,\ldots,N_x\}}$ for the range of $\mathbf{x}$, where we define $\delta x = \sqrt[3]{\mathcal{V}}/N_x$ and $N_x = \sqrt[3]{N/4}$, given that $\sqrt[3]{N/4}/2 \in \mathbb{N}$. And for the range of $\mathbf{k}$, we could then have $k_1, k_2 \in \{(l+1/2)\delta k\}_{l\in\{-N_x/2,\ldots,N_x/2-1\}}$ and have $k_3 \in \{(l+1/2)\delta k\}_{l\in\{0,\ldots,N_x/2-1\}}$, defining $\delta k = 2\pi/(N_x \delta x)$. This gives us a similar range for $\mathbf{k}$ as what we might have chosen for a Fourier transform with complex exponential functions, only cut in half by the $k_3 = 0$ plane, which is necessary in order for the sine and cosine functions to be an orthonormal set. With this range of $\mathbf{k}$, call it $\mathbb{K}$, we have also made sure to always have $k \equiv |\mathbf{k}| \neq 0$, which will be beneficial going forward.

Let us thus assume $k \equiv |\mathbf{k}| \neq 0$ for all $\mathbf{k} \in \mathbb{K}$ from now on. And since it will also be used later on, note that $\mathbb{K}$ should have $N/8$ distinct members and that $\delta k^3 = (2\pi)^3/\mathcal{V}$. Let us also define $N_k = N_x$, such that we can think of $N_k \delta k$ as the "length" of $\mathbb{K}$.

We will then define $\{\widetilde{\varphi}_{\mathbf{k}\sigma}\}$ and $\{\widetilde{\mathbf{A}}_{\mathbf{k}\sigma}\}$ as the sets of Fourier coefficients such that

$$\varphi(\mathbf{x}) = \sum_{\mathbf{k},\sigma} \widetilde{\varphi}_{\mathbf{k}\sigma} f_{\mathbf{k}\sigma}(\mathbf{x}), \quad \mathbf{A}(\mathbf{x}) = \sum_{\mathbf{k},\sigma} \widetilde{\mathbf{A}}_{\mathbf{k}\sigma} f_{\mathbf{k}\sigma}(\mathbf{x}), \tag{16}$$

where we implicitly take the ranges of such summations to be $\mathbf{k} \in \mathbb{K}$ and $\sigma \in \{1, -1\}$. The benefit of going to Fourier space is that we can now almost forget that the range of $\mathbf{x}$ is discretized. We can thus interpolate and extrapolate φ and $\mathbf{A}$ in a straightforward manner, such that $\varphi(\mathbf{x}_j)$ and $\mathbf{A}(\mathbf{x}_j)$ are defined for all particle coordinates $\mathbf{x}_j \in \mathbb{R}^3$. The orthonormality of $\{f_{\mathbf{k}\sigma}\}$ is also preserved (on the volume of size $\mathcal{V}$) when we interpolate and let $\mathbf{x}$ range over a continuous volume. And as mentioned, the Fourier transform also lets us evaluate the spatial derivatives of Eq. (14) in a simple way. Since $\partial f_{\mathbf{k}\sigma}/\partial x_i = -\sigma k_i f_{\mathbf{k}-\sigma}$ for all $i \in \{1, 2, 3\}$, $\mathbf{k} \in \mathbb{K}$, $\sigma \in \{1, -1\}$, we for instance get

$$\nabla\varphi(\mathbf{x}) = \sum_{\mathbf{k},\sigma} -\sigma\mathbf{k}\widetilde{\varphi}_{\mathbf{k}\sigma} f_{\mathbf{k}-\sigma}(\mathbf{x}) = \sum_{\mathbf{k},\sigma} \sigma\mathbf{k}\widetilde{\varphi}_{\mathbf{k}-\sigma} f_{\mathbf{k}\sigma}(\mathbf{x}), \tag{17}$$

where we have flipped the sign of σ in the summation to get the last equality.

For the total Lagrangian $L_{EM} = \int \mathcal{L}_{EM} d\mathbf{x}$, we can thus write

$$L_{EM} = \frac{1}{2}\sum_{\mathbf{k},\sigma}\Big(\big(\sigma\widetilde{\varphi}_{\mathbf{k}-\sigma}\mathbf{k} + \frac{\partial}{\partial t}\widetilde{\mathbf{A}}_{\mathbf{k}\sigma}\big)^2 - \big(\mathbf{k}\times\widetilde{\mathbf{A}}_{\mathbf{k}-\sigma}\big)^2 + \frac{1}{\xi}\big(\sigma\mathbf{k}\cdot\widetilde{\mathbf{A}}_{\mathbf{k}-\sigma} + \frac{\partial}{\partial t}\widetilde{\varphi}_{\mathbf{k}\sigma}\big)^2\Big). \tag{18}$$

To get this formula, we have inserted Eq. (16) in Eq. (14), integrated over $\mathbf{x}$ for both sides, and used the orthonormality of $\{f_{\mathbf{k}\sigma}\}$ to reduce terms of the form $\int[\sum_{\mathbf{k},\sigma} a_{\mathbf{k}\sigma} f_{\mathbf{k}\sigma}(\mathbf{x}) + \sum_{\mathbf{k},\sigma} b_{\mathbf{k}\sigma} f_{\mathbf{k}\sigma}(\mathbf{x})]^2\, d\mathbf{x}$ to $\sum_{\mathbf{k},\sigma}(a_{\mathbf{k}\sigma} + b_{\mathbf{k}\sigma})^2$.

We can simplify this formula further by introducing new coordinates $\{\widetilde{A}_{\mu\mathbf{k}\sigma}\}$, defined by

$$\widetilde{A}_{0\mathbf{k}\sigma} = \widetilde{\varphi}_{\mathbf{k}\sigma}, \quad \widetilde{A}_{i\mathbf{k}\sigma} = \widetilde{\mathbf{A}}_{\mathbf{k}\sigma}\cdot\mathbf{e}_{i\mathbf{k}} \tag{19}$$

for all $i \in \{1, 2, 3\}$, $\mathbf{k} \in \mathbb{K}$, $\sigma \in \{1, -1\}$, where $\mathbf{e}_{3\mathbf{k}} = \mathbf{k}/k$ is always parallel to $\mathbf{k}$, and where $\mathbf{e}_{1\mathbf{k}}$ and $\mathbf{e}_{2\mathbf{k}}$, also normalized, are chosen to always be orthogonal both to $\mathbf{k}$ and to each other. By substituting these relations in Eq. (18), and noting that $\widetilde{\mathbf{A}}_{\mathbf{k}\sigma} = \sum_{i=1}^3 \widetilde{A}_{i\mathbf{k}\sigma}\mathbf{e}_{i\mathbf{k}}$, we get

$$\begin{aligned} L_{EM} = \frac{1}{2}\sum_{\mathbf{k},\sigma}\Big[&\sum_{\lambda=1}^{2}\Big(\big(\frac{\partial}{\partial t}\widetilde{A}_{\lambda\mathbf{k}\sigma}\big)^2 - k^2\widetilde{A}^2_{\lambda\mathbf{k}\sigma}\Big) \\ &+ \big(\sigma k\widetilde{A}_{0\mathbf{k}-\sigma} + \frac{\partial}{\partial t}\widetilde{A}_{3\mathbf{k}\sigma}\big)^2 + \frac{1}{\xi}\big(\sigma k\widetilde{A}_{3\mathbf{k}-\sigma} + \frac{\partial}{\partial t}\widetilde{A}_{0\mathbf{k}\sigma}\big)^2\Big], \end{aligned} \tag{20}$$

where we have also used the fact that $(a\mathbf{k} - \sum_{i=1}^{3} b_i \mathbf{e}_{i\mathbf{k}})^2 = b_1^2 + b_2^2 + (ak - b_3)^2$ for all $a, b_1, b_2, b_3 \in \mathbb{R}$ to rewrite the first term in Eq. (18).

Let us now finally define the basis for $\mathbf{q}$ (disregarding the order of its entries), namely such that each of its entries is a $\widetilde{A}_{\mu\mathbf{k}\sigma}$ coordinate:

$$\mathbf{q} = (\widetilde{A}_{\mu\mathbf{k}\sigma})_{\mu\in\{0,1,2,3\},\mathbf{k}\in\mathbb{K},\sigma\in\{1,-1\}}. \tag{21}$$

Let us also keep the same indexing for $\mathbf{q}$, such that $q_{\mu\mathbf{k}\sigma} = \widetilde{A}_{\mu\mathbf{k}\sigma}$ for all indices. With this definition, we see that the Lagrangian of Eq. (20) has the form

$$L(\mathbf{q}, \dot{\mathbf{q}}) = \frac{1}{2}\big(\dot{\mathbf{q}} - \mathbf{W}(\mathbf{q})\big)^T D\big(\dot{\mathbf{q}} - \mathbf{W}(\mathbf{q})\big) - V(\mathbf{q}), \tag{22}$$

namely if we define $V(\mathbf{q}) = \sum_{\mathbf{k},\sigma}\sum_{\lambda=1}^{2} k^2 q_{\lambda\mathbf{k}\sigma}^2/2$ and $W_{\mu\mathbf{k}\sigma}(\mathbf{q}) = -\sigma k(\delta_{\mu 0} q_{3\mathbf{k}-\sigma} + \delta_{\mu 3} q_{0\mathbf{k}-\sigma})$, and define D as a diagonal matrix with elements on the diagonal given by $D_{\mu\mathbf{k}\sigma\mu\mathbf{k}\sigma} = (\delta_{\mu 0}/\xi + \delta_{\mu 1} + \delta_{\mu 2} + \delta_{\mu 3})$.

The classical Hamiltonian corresponding to the Lagrangian of Eq. (22), i.e. the Legendre transform of said Lagrangian with respect to $\dot{\mathbf{q}}$, is given by

$$H(\mathbf{p}, \mathbf{q}) = \frac{1}{2}\mathbf{p}^T D^{-1}\mathbf{p} + \mathbf{W}(\mathbf{q})\cdot\mathbf{p} + V(\mathbf{q}). \tag{23}$$

A derivation of this is shown in Appendix A. The quantized version of this Hamiltonian would thus be a good guess for $\hat{H}_{EM}$. And as we will confirm in Sect. 4, this choice will indeed lead to the desired path integral of Eq. (11).

To obtain $\hat{H}_{EM}$, let us thus first substitute the definitions used for Eq. (22) back into Eq. (23). This gives us

$$H_{EM} = \sum_{\mathbf{k},\sigma}\Big(\sum_{\lambda=1}^{2}(\frac{1}{2}p_{\lambda\mathbf{k}\sigma}^2 + \frac{1}{2}k^2 q_{\lambda\mathbf{k}\sigma}^2) + \frac{\xi}{2}p_{0k\sigma}^2 + \frac{1}{2}p_{3\mathbf{k}\sigma}^2 - \sigma k q_{3\mathbf{k}-\sigma}p_{0\mathbf{k}\sigma} - \sigma k q_{0\mathbf{k}-\sigma}p_{3\mathbf{k}\sigma}\Big). \tag{24}$$

The process of "quantizing" H_{EM} then specifically means to replace all instances of $q_{\mu\mathbf{k}\sigma}$ and $p_{\mu\mathbf{k}\sigma}$ with $\hat{q}_{\mu\mathbf{k}\sigma}$ and $\hat{p}_{\mu\mathbf{k}\sigma}$, where $\hat{q}_{\mu\mathbf{k}\sigma}\phi(\mathbf{q}) = q_{\mu\mathbf{k}\sigma}\phi(\mathbf{q})$ for all $\mathbf{q} \in \mathbb{R}^N$ and where $\hat{p}_{\mu\mathbf{k}\sigma} = -i\partial/\partial q_{\mu\mathbf{k}\sigma}$. And recalling that $q_{\mu\mathbf{k}\sigma} = \widetilde{A}_{\mu\mathbf{k}\sigma}$, let us also rename these operators as $\hat{q}_{\mu\mathbf{k}\sigma} \to \hat{A}_{\mu\mathbf{k}\sigma}$ and $\hat{p}_{\mu\mathbf{k}\sigma} \to \hat{\Pi}_{\mu\mathbf{k}\sigma}$. So from this point on, let $\hat{A}_{\mu\mathbf{k}\sigma}$ denote the operator that measures the amplitude $\widetilde{A}_{\mu\mathbf{k}\sigma}$ of the relevant mode, and let $\hat{\Pi}_{\mu\mathbf{k}\sigma} = -i\partial/\partial\widetilde{A}_{\mu\mathbf{k}\sigma}$ denote the operator that measures the so-called conjugate momentum of this mode's amplitude. Our guess for $\hat{H}_{EM}$ can thus be written as

$$\begin{aligned}\hat{H}_{EM} = \sum_{\mathbf{k},\sigma}\Bigg[&\sum_{\lambda=1}^{2}(\frac{1}{2}\hat{\Pi}_{\lambda\mathbf{k}\sigma}^2 + \frac{1}{2}k^2\hat{A}_{\lambda\mathbf{k}\sigma}^2)\\ &+ \frac{\xi}{2}\hat{\Pi}_{0k\sigma}^2 + \frac{1}{2}\hat{\Pi}_{3\mathbf{k}\sigma}^2 - \sigma k\hat{A}_{3\mathbf{k}-\sigma}\hat{\Pi}_{0\mathbf{k}\sigma} - \sigma k\hat{A}_{0\mathbf{k}-\sigma}\hat{\Pi}_{3\mathbf{k}\sigma}\Bigg].\end{aligned} \tag{25}$$

Now that we have a preferred basis for $\mathbf{q}$, we can also finally write up $\hat{H}_D(\mathbf{q})$, namely by expanding $\varphi(\mathbf{q}, \mathbf{x}_j)$ and $\mathbf{A}(\mathbf{q}, \mathbf{x}_j)$ on the right-hand side of Eq. (6) in terms of $\{\tilde{A}_{\mu\mathbf{k}\sigma}\}$, interpolating and extrapolating the lattice for all $\mathbf{x}_j \in \mathbb{R}^3$. We can thus define $\hat{H}_D$ formally by

$$\begin{aligned}\hat{H}_D(\mathbf{q})\psi(\bar{\mathbf{x}}) = \sum_{j=1}^{n}\Bigg[&\boldsymbol{\alpha}_j \cdot \hat{\mathbf{p}}_j - q_F \sum_{\mathbf{k},\sigma}\sum_{i=1}^{3}(\boldsymbol{\alpha}_j \cdot \mathbf{e}_{i\mathbf{k}})\widetilde{A}_{i\mathbf{k}\sigma} f_{\mathbf{k}\sigma}(\mathbf{x}_j) + \beta_j m_F \\ &+ q_F \sum_{\mathbf{k},\sigma} \widetilde{A}_{0\mathbf{k}\sigma} f_{\mathbf{k}\sigma}(\mathbf{x}_j)\Bigg]\psi(\bar{\mathbf{x}})\end{aligned} \tag{26}$$

for all $\psi \in \mathbf{H}_F$, $\mathbf{q} \in \mathbb{R}^N$ and $\bar{\mathbf{x}} = (\mathbf{x}_1, \ldots, \mathbf{x}_n) \in \mathbb{R}^{3n}$.

With both $\hat{H}_{EM}$ and $\hat{H}_D(\mathbf{q})$ now defined, we thus get our full $\hat{H}_{init}$ since, as we recall, this is formally given by

$$\hat{H}_{init} = \hat{H}_B + \hat{H}_{IF}, \quad \hat{H}_B = \hat{H}_{EM} \otimes \hat{I}_F, \quad \hat{H}_{IF}\left|\mathbf{q}, \psi\right\rangle = \left|\mathbf{q}\right\rangle \otimes \hat{H}_D(\mathbf{q})\left|\psi\right\rangle \tag{27}$$

for all $\psi \in \mathbf{H}_F$ and $\mathbf{q} \in \mathbb{R}^N$. (We will not define the the actual domain of $\hat{H}_{init}$ in the main part of this paper, but the interested reader can see Appendix B for more details on this domain, including arguments for why $\hat{H}_{init}$ is self-adjoint on it.)

As some concluding remarks for this section, let us look at the derived formula for $\hat{H}_{EM}$ in Eq. (25). We see that the transverse modes with $\mu \in \{1, 2\}$ are simple harmonic oscillators with angular frequency $\omega = k$, meaning that the energy difference between the eigenstates of each oscillator is equal to k. We therefore strongly expect the photons of the theory, which indeed need to have exactly two distinct spin states for each $\mathbf{k}$, to come from these modes in the continuum limit. The ground state energy will grow to infinity in this limit, but since we can always add or subtract a constant energy to L_{EM} without changing the dynamics, we will be able to remove this at every step when approaching said limit.

The modes with $\mu \in \{0, 3\}$, on the other hand, seem somewhat offending at first: We know that the photons should only have two spin states, and we do not expect any other bosons to come from the A^μ field. We therefore naturally want to get rid of these excessive degrees of freedom somehow. It might be tempting to just remove them, but doing so will make us unable to get back to the Lagrangian again, namely if we do the derivation of the path integral in the next section in reverse. And once we have lost the connection to the Lagrangian, it is hard to see how one would prove Lorentz covariance for the theory.

Luckily, however, there is another way to eliminate these excessive degrees of freedom, as we will show in Sect. 5. But before we get to that, let us first derive the path integral for $\hat{H}_{init}$ and confirm Eq. (11), as well as the fact that $\hat{H}_{EM}$ indeed corresponds to the Lagrangian L_{EM} in the path integral formulation. We will do this in the following section.

4 The path integral for the initial Hamiltonian

For this section, we want to show that the path integral for $\hat{H}_{init}$ indeed has the form of Eq. (11), and with L_{EM} as the Lagrangian. To do this, we want to use the Trotter product formula, which allows us to write

$$e^{-i(\hat{A}+\hat{B})t} = \lim_{M\to\infty}(e^{-i\hat{A}t/M}e^{-i\hat{B}t/M})^M \tag{28}$$

for certain kinds of operators, and derive the path integral from there.

We thus want to divide $\hat{H}_{init}$ into two parts, which we can call $\hat{H}'_{init}$ and $\hat{H}''_{init}$. It will be beneficial to make sure that all operator terms coming from $\hat{H}_B$ in both $\hat{H}'_{init}$ and $\hat{H}''_{init}$ commute internally. Let us therefore define $\hat{H}'_B$ and $\hat{H}''_B$ by

$$\begin{aligned}\hat{H}'_B &= \sum_{\mathbf{k},\sigma}\Big(\frac{1}{2}\hat{\Pi}^2_{1\mathbf{k}\sigma} + \frac{1}{2}k^2\hat{A}^2_{2\mathbf{k}\sigma} + \frac{\xi}{2}\hat{\Pi}^2_{0k\sigma} - \sigma k\hat{A}_{3\mathbf{k}-\sigma}\hat{\Pi}_{0\mathbf{k}\sigma}\Big)\otimes \hat{I}_F,\\ \hat{H}''_B &= \sum_{\mathbf{k},\sigma}\Big(\frac{1}{2}\hat{\Pi}^2_{2\mathbf{k}\sigma} + \frac{1}{2}k^2\hat{A}^2_{1\mathbf{k}\sigma} + \frac{1}{2}\hat{\Pi}^2_{3k\sigma} - \sigma k\hat{A}_{0\mathbf{k}-\sigma}\hat{\Pi}_{3\mathbf{k}\sigma}\Big)\otimes \hat{I}_F.\end{aligned} \tag{29}$$

Furthermore, it will also be a good idea to divide $\hat{H}_{IF}$ up into two parts such that one part, call it $\hat{H}'_{IF}$, commutes with $\hat{H}'_B$ and the other, $\hat{H}''_{IF}$, commutes with $\hat{H}''_B$. This is achieved if we define $\hat{H}'_D(\mathbf{q})$ and $\hat{H}''_D(\mathbf{q})$ formally for all $\mathbf{q}\in\mathbb{R}^N$ and $\psi\in\mathbf{H}_F$ by

$$\begin{aligned}\hat{H}'_D(\mathbf{q})\psi(\bar{\mathbf{x}}) &= \sum_{j=1}^{n}\Big[\boldsymbol{\alpha}_j\cdot\hat{\mathbf{p}}_j - q_F\sum_{\mathbf{k},\sigma}\sum_{i=2}^{3}(\boldsymbol{\alpha}_j\cdot\mathbf{e}_{i\mathbf{k}})\widetilde{A}_{i\mathbf{k}\sigma}f_{\mathbf{k}\sigma}(\mathbf{x}_j) + \beta_j m_F\Big]\psi(\bar{\mathbf{x}}),\\ \hat{H}''_D(\mathbf{q})\psi(\bar{\mathbf{x}}) &= \sum_{j=1}^{n}\Big[-q_F\sum_{\mathbf{k},\sigma}(\boldsymbol{\alpha}_j\cdot\mathbf{e}_{1\mathbf{k}})\widetilde{A}_{1\mathbf{k}\sigma}f_{\mathbf{k}\sigma}(\mathbf{x}_j) + q_F\sum_{\mathbf{k},\sigma}\widetilde{A}_{0\mathbf{k}\sigma}f_{\mathbf{k}\sigma}(\mathbf{x}_j)\Big]\psi(\bar{\mathbf{x}}),\end{aligned} \tag{30}$$

and define $\hat{H}'_{IF}$ and $\hat{H}''_{IF}$ accordingly, such that we now have

$$\hat{H}_{init} = \hat{H}'_{init} + \hat{H}''_{init} = \hat{H}'_B + \hat{H}'_{IF} + \hat{H}''_B + \hat{H}''_{IF}. \tag{31}$$

We want to now use the Trotter product formula. According to Hall [4], the formula of Eq. (28) applies whenever $\hat{A}$ and $\hat{B}$ are self-adjoint operators on their domains, $\mathrm{Dom}(\hat{A})$ and $\mathrm{Dom}(\hat{B})$, and when $\hat{A}+\hat{B}$ is self-adjoint on $\mathrm{Dom}(\hat{A})\cap\mathrm{Dom}(\hat{B})$. Now, this might actually not be true for $\hat{H}'_{init}$ and $\hat{H}''_{init}$, as it turns out. We can, however, introduce cutoffs on all $\hat{A}_{\mu\mathbf{k}\sigma}$-operators in $\hat{H}_{init}$ and on the $\hat{\mathbf{p}}_j$-operators, by which the condition becomes true. Introducing these cutoffs will not change the derivation of the path integral below. We will therefore let them be completely implicit until the end of this section, where we will then discuss how to remove them once again.

With these cutoffs introduced, we can make sure that the only parts of $\hat{H}_{init}$ that are not bounded are the $\hat{\Pi}_{\mu\mathbf{k}\sigma}$-operators. From the Kato–Rellich theorem (see e.g. Hall [4], or see Appendix B for details sufficient for our purposes), it then trivially follows that $\hat{H}'_{init}$ is self-adjoint on $\mathrm{Dom}(\sum_{\mathbf{k},\sigma}\sum_{\mu=0}^{1}\hat{\Pi}^2_{\mu\mathbf{k}\sigma})$, that $\hat{H}''_{init}$ is self-adjoint on $\mathrm{Dom}(\sum_{\mathbf{k},\sigma}\sum_{\mu=2}^{3}\hat{\Pi}^2_{\mu\mathbf{k}\sigma})$, and that their sum, $\hat{H}_{init}$, is self-adjoint on $\mathrm{Dom}(\sum_{\mathbf{k},\sigma}\sum_{\mu=0}^{3}\hat{\Pi}^2_{\mu\mathbf{k}\sigma})$.

As is often the case in quantum mechanics (particularly when the coordinate space of the wave functions is not bounded), each of these domains is the subspace of L^2-functions that turn into other L^2-functions when worked on by the relevant operator. (See e.g. Hall [4].) And when Fourier-transformed, all these operators become (quadratic) multiplication operators, which makes it easy to analyze their domains. We can use this to argue that $\mathrm{Dom}(\sum_{\mathbf{k},\sigma}\sum_{\mu=0}^{3}\hat{\Pi}^2_{\mu\mathbf{k}\sigma}) = \mathrm{Dom}(\sum_{\mathbf{k},\sigma}\sum_{\mu=0}^{1}\hat{\Pi}^2_{\mu\mathbf{k}\sigma})\cap\mathrm{Dom}(\sum_{\mathbf{k},\sigma}\sum_{\mu=2}^{3}\hat{\Pi}^2_{\mu\mathbf{k}\sigma})$, allowing us to use the Trotter product formula.

If we simplify this problem, we can see that it amounts to showing that $\mathrm{Dom}(\mathbf{x}^2+\mathbf{y}^2) = \mathrm{Dom}(\mathbf{x}^2)\cap\mathrm{Dom}(\mathbf{y}^2)$ for some Hilbert space of functions over $(\mathbf{x},\mathbf{y})$-coordinates. To do this, we see that

$$\|\mathbf{x}^2\psi + \mathbf{y}^2\psi\|^2 = \|\mathbf{x}^2\psi\|^2 + \|\mathbf{y}^2\psi\|^2 + 2\int \mathbf{x}^2\mathbf{y}^2|\psi|^2\,d\mathbf{x}\,d\mathbf{y} \tag{32}$$

for all ψ in this Hilbert space. And since $0 \le 2\mathbf{x}^2\mathbf{y}^2 \le \mathbf{x}^4 + \mathbf{y}^4$ everywhere, we thus get that

$$\|\mathbf{x}^2\psi\|^2 + \|\mathbf{y}^2\psi\|^2 \le \|\mathbf{x}^2\psi + \mathbf{y}^2\psi\|^2 \le 2\|\mathbf{x}^2\psi\|^2 + 2\|\mathbf{y}^2\psi\|^2. \tag{33}$$

This tells us that $\|(\mathbf{x}^2 + \mathbf{y}^2)\psi\|$ converges exactly when both $\|\mathbf{x}^2\psi\|$ and $\|\mathbf{y}^2\psi\|$ does so, meaning that $\mathrm{Dom}(\mathbf{x}^2 + \mathbf{y}^2) = \mathrm{Dom}(\mathbf{x}^2) \cap \mathrm{Dom}(\mathbf{y}^2)$. (We could in fact have shown this for any potential of the form $f(\mathbf{x}) + g(\mathbf{y})$, where f and g are positive functions.)

It can be thus shown that

$$\mathrm{Dom}(\hat{H}_{init}) = \mathrm{Dom}(\hat{H}'_{init}) \cap \mathrm{Dom}(\hat{H}''_{init}) \tag{34}$$

with our new cutoffs introduced, and we can therefore go ahead and use the Trotter product formula for $\exp(-i\hat{H}_{init}t)$. This gives us

$$\begin{aligned} e^{-i\hat{H}_{init}t} &= \lim_{M\to\infty} (e^{-i\hat{H}'_{init}\delta t} e^{-i\hat{H}''_{init}\delta t})^M, \\ &= \lim_{M\to\infty} (e^{-i\hat{H}'_{IF}\delta t} e^{-i\hat{H}'_{B}\delta t} e^{-i\hat{H}''_{B}\delta t} e^{-i\hat{H}''_{IF}\delta t})^M, \end{aligned} \tag{35}$$

where we have defined $\delta t = t/M$. To get the second equality here, we have also used the fact that $\exp(-i\hat{H}'_B\delta t)$ commutes with $\exp(-i\hat{H}'_{IF}\delta t)$ to separate the exponential operators, and likewise for $\hat{H}''_B$ and $\hat{H}''_{IF}$. The fact that we can do this also follows directly from the Trotter product formula, now that $\hat{H}'_{IF}$ and $\hat{H}''_{IF}$ are bounded, namely since it tells us that for all $t \in \mathbb{R}$, we have

$$e^{-i(\hat{H}'_{IF}+\hat{H}'_B)t} = \lim_{M\to\infty} (e^{-i\hat{H}'_{IF}\delta t} e^{-i\hat{H}'_B\delta t})^M = \lim_{M\to\infty} e^{-i\hat{H}'_{IF}t} e^{-i\hat{H}'_B t} = e^{-i\hat{H}'_{IF}t} e^{-i\hat{H}'_B t}, \tag{36}$$

and similarly for $\exp(-i(\hat{H}''_B + \hat{H}''_{IF})t)$.

We are now ready to proceed with the derivation of the path integral for $\hat{H}_{init}$. Roughly following the derivations found in e.g. References [6]–[9], we will allow ourselves to resolve the identity operator with respect to the well-known generalized bases $\{|\mathbf{q}\rangle\}$ and $\{|\mathbf{p}\rangle\}$, normalized according to $\hat{I} = \int d\mathbf{q}\,|\mathbf{q}\rangle\langle\mathbf{q}| = (2\pi)^{-N}\int d\mathbf{p}\,|\mathbf{p}\rangle\langle\mathbf{p}|$ and with generalized matrix elements given by $\langle\mathbf{q}'|\mathbf{q}\rangle = \delta^N(\mathbf{q}' - \mathbf{q})$, $\langle\mathbf{p}'|\mathbf{p}\rangle = (2\pi)^N\delta^N(\mathbf{p}' - \mathbf{p})$ and $\langle\mathbf{q}|\mathbf{p}\rangle = \exp(i\mathbf{q}\cdot\mathbf{p})$. When these identity operators are extended to our full Hilbert space $\mathbf{H}_{init}$, they can be written as

$$\hat{I} = \sum_j \int d\mathbf{q}\,|\mathbf{q}, j\rangle\langle\mathbf{q}, j| = \sum_j \int \frac{d\mathbf{p}}{(2\pi)^N}\,|\mathbf{p}, j\rangle\langle\mathbf{p}, j|\,, \tag{37}$$

where we have defined $\{|j\rangle\} \equiv \{|\psi_j\rangle\} = \{\psi_j\}$ to be an orthonormal basis of $\mathbf{H}_F$. The integrals are here implicitly taken to be the improper integrals over $\mathbb{R}^N$, meaning that each $\int dp_{\mu\mathbf{k}\sigma}$ is to be interpreted implicitly as $\lim_{A\to\infty}\int_{-A}^{A} dp_{\mu\mathbf{k}\sigma}$, and similarly for each $\int dq_{\mu\mathbf{k}\sigma}$.

Now, if we want to evaluate the propagator from a generalized state $|\mathbf{q}_0, j_0\rangle$ to a new one, $|\mathbf{q}_t, j_t\rangle$, after a time t, it is given by

$$\begin{aligned} K(t, \mathbf{q}_t, \mathbf{q}_0, \psi_{j_t}, \psi_{j_0}) &\equiv \langle\mathbf{q}_t, j_t|e^{-i\hat{H}_{init}t}|\mathbf{q}_0, j_0\rangle \\ &= \lim_{M\to\infty} \langle\mathbf{q}_t, j_t|(e^{-i\hat{H}'_{IF}\delta t} e^{-i\hat{H}'_B\delta t} e^{-i\hat{H}''_B\delta t} e^{-i\hat{H}''_{IF}\delta t})^M|\mathbf{q}_0, j_0\rangle\,, \end{aligned} \tag{38}$$

where we have used Eq. (35) to get the second equality. Let us then define the fixed-M approximation to $K(t, \mathbf{q}_t, \mathbf{q}_0, \psi_{j_t}, \psi_{j_0})$ as $\acute{K}(M, t, \mathbf{q}_t, \mathbf{q}_0, \psi_{j_t}, \psi_{j_0})$, and then expand its formula by inserting the identity operators of Eq. (37). This gives us

$$
\begin{aligned}
\acute{K}(M, t, \mathbf{q}_t, \mathbf{q}_0, \psi_{j_t}, \psi_{j_0}) &\equiv \langle \mathbf{q}_t, j_t | (e^{-i\hat{H}'_{IF}\delta t} e^{-i\hat{H}'_B \delta t} e^{-i\hat{H}''_B \delta t} e^{-i\hat{H}''_{IF}\delta t})^M | \mathbf{q}_0, j_0 \rangle \\
&= \sum_{j_1, \ldots, j_{M-1}} \sum_{j'_1, \ldots, j'_M} \int \prod_{i=1}^{M-1} (d\mathbf{q}_i) \prod_{i=1}^{M} \Big(\frac{d\mathbf{p}_i}{(2\pi)^N} \Big) S_M \cdots S_1,
\end{aligned} \tag{39}
$$

where each expression S_m is given by

$$
\begin{aligned}
S_m &= \langle \mathbf{q}_m, j_m | e^{-i\hat{H}'_{IF}\delta t} e^{-i\hat{H}'_B \delta t} | \mathbf{p}_m, j'_m \rangle \, \langle \mathbf{p}_m, j'_m | e^{-i\hat{H}''_B \delta t} e^{-i\hat{H}''_{IF}\delta t} | \mathbf{q}_{m-1}, j_{m-1} \rangle \\
&= \langle j_m | e^{-i\hat{H}'_D(\mathbf{q}_m)\delta t} | j'_m \rangle \, \langle \mathbf{q}_m | e^{-i\hat{H}'_B \delta t} | \mathbf{p}_m \rangle \, \langle \mathbf{p}_m | e^{-i\hat{H}''_B \delta t} | \mathbf{q}_{m-1} \rangle \, \langle j'_m | e^{-i\hat{H}''_D(\mathbf{q}_{m-1})\delta t} | j_{m-1} \rangle \, .
\end{aligned} \tag{40}
$$

Here we have thus also defined $\mathbf{q}_M = \mathbf{q}_t$ and $j_M = j_t$ for the case of S_M. From this equation, we see that $S_M \cdots S_1$ can also be rewritten as

$$
S_M \cdots S_1 = T_M \cdots T_1 U_M \cdots U_1, \tag{41}
$$

where the expressions T_m and U_m are each given by

$$
T_m = \langle \mathbf{q}_m | e^{-i\hat{H}'_B \delta t} | \mathbf{p}_m \rangle \, \langle \mathbf{p}_m | e^{-i\hat{H}''_B \delta t} | \mathbf{q}_{m-1} \rangle \, , \tag{42}
$$

$$
U_m = \langle j_m | e^{-i\hat{H}'_D(\mathbf{q}_m)\delta t} | j'_m \rangle \, \langle j'_m | e^{-i\hat{H}''_D(\mathbf{q}_{m-1})\delta t} | j_{m-1} \rangle \, . \tag{43}
$$

We can then immediately notice that, since $\{|j\rangle\}$ is an orthonormal basis, each of the inner ket-bras of $U_M \cdots U_1$ can be identified as the identity operator on $\mathbf{H}_F$ when the summations are carried out. We thus have

$$
\sum_{j_1, \ldots, j_{M-1}} \sum_{j'_1, \ldots, j'_M} U_M \cdots U_1 = \langle j_M | e^{-i\hat{H}'_D(\mathbf{q}_M)\delta t} e^{-i\hat{H}''_D(\mathbf{q}_{M-1})\delta t} \cdots e^{-i\hat{H}'_D(\mathbf{q}_1)\delta t} e^{-i\hat{H}''_D(\mathbf{q}_0)\delta t} | j_0 \rangle. \tag{44}
$$

We see that we can interpret the right-hand side of this equation as a finite-M approximation of K_{IF} from Eq. (12), namely where the function $t \mapsto \mathbf{q}(t)$ is replaced by the discrete $\bar{\mathbf{q}}$, defined as $\bar{\mathbf{q}} = (\mathbf{q}_0, \ldots, \mathbf{q}_M)$. Let us therefore define $\acute{K}_{IF}(M, t, \bar{\mathbf{q}}, \psi_{j_M}, \psi_{j_0})$ as the right-hand side of Eq. (44). With this definition we can thus rewrite Eq. (39) as

$$
\acute{K}(M, t, \mathbf{q}_M, \mathbf{q}_0, \psi_{j_M}, \psi_{j_0}) = \int \prod_{i=1}^{M-1} (d\mathbf{q}_i) \prod_{i=1}^{M} \Big(\frac{d\mathbf{p}_i}{(2\pi)^N} \Big) T_M \cdots T_1 \acute{K}_{IF}(M, t, \bar{\mathbf{q}}, \psi_{j_M}, \psi_{j_0}), \tag{45}
$$

where we have also substituted $\mathbf{q}_t = \mathbf{q}_M$ and $j_t = j_M$ for consistency.

To evaluate $T_M \cdots T_1$, we note that

$$
\langle \mathbf{q} | e^{-i\hat{H}'_B \delta t} | \mathbf{p} \rangle = \langle \mathbf{q} | \mathbf{p} \rangle \, e^{-iH'_B(\mathbf{p}, \mathbf{q})\delta t}, \quad \langle \mathbf{q} | e^{-i\hat{H}''_B \delta t} | \mathbf{p} \rangle = \langle \mathbf{q} | \mathbf{p} \rangle \, e^{-iH''_B(\mathbf{p}, \mathbf{q})\delta t}, \tag{46}
$$

where H'_B and H''_B are defined as the classical versions of $\hat{H}'_B$ and $\hat{H}''_B$. This can for instance be seen by making a partial Fourier transformation of $|\mathbf{p}\rangle$ such that $\hat{H}'_B$ or $\hat{H}''_B$, given what case we are looking at, becomes a multiplication operator. The $\langle \mathbf{q}|$ to the left will then yield some delta functions, and the expression can then be evaluated to give the desired result. So for Eq. (42), we thus get

$$\begin{aligned} T_m &= \langle \mathbf{q}_m|\mathbf{p}_m\rangle \langle \mathbf{p}_m|\mathbf{q}_{m-1}\rangle \, e^{-iH'(\mathbf{p}_m,\mathbf{q}_m)\delta t - iH''(\mathbf{p}_m,\mathbf{q}_{m-1})\delta t} \\ &= e^{i\mathbf{q}_m\cdot\mathbf{p}_m - i\mathbf{q}_{m-1}\cdot\mathbf{p}_m - iH'(\mathbf{p}_m,\mathbf{q}_m)\delta t - iH''(\mathbf{p}_m,\mathbf{q}_{m-1})\delta t}. \end{aligned} \tag{47}$$

And we can therefore write $T_M \cdots T_1$ as

$$T_M \cdots T_1 = \exp\Big(i \sum_{m=1}^{M} \big(\mathbf{p}_m \cdot \dot{\mathbf{q}}_m - H'_B(\mathbf{p}_m, \mathbf{q}_m) - H''_B(\mathbf{p}_m, \mathbf{q}_{m-1})\big)\delta t\Big), \tag{48}$$

where we have defined $\dot{\mathbf{q}}_m = (\mathbf{q}_m - \mathbf{q}_{m-1})/\delta t$ for each m.

Inserting Eq. (48) in Eq. (45) will yield the phase space path integral for $\hat{H}_{init}$. But since $T_M \cdots T_1$ is a multidimensional Gaussian function with respect to the $\mathbf{p}$-variables, we can integrate over all these to get the configuration space path integral. And since the matrix D in Eq. (23) is a diagonal matrix, we can simply evaluate the multidimensional Gaussian integral one dimension at a time. According to Hall [4],[5] for any $a, b \in \mathbb{C}$ with $a \neq 0$ and $\mathrm{Re}(a) \geq 0$, we have

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-ap^2/2+bp}\, dp = \Big(\frac{1}{2\pi a}\Big)^{1/2} e^{b^2/(2a)}, \tag{49}$$

when the integral on the left-hand side is understood as an improper integral, and when the square root on the right-hand side is understood as the one with positive real part. Thus, for all $w, v, \dot{q} \in \mathbb{R}$, we have

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ip\dot{q}\delta t - i(\eta p^2/2 + wp + v)\delta t}\, dp = \Big(\frac{1}{2\pi i\eta\delta t}\Big)^{1/2} e^{i(\dot{q}-w)^2\delta t/(2\eta)} e^{-iv\delta t}, \tag{50}$$

where $\eta \in \{1, \xi\}$. We recognize the combined exponent on the right-hand side of this equation to be a one-dimensional version of our Lagrangian of Eq. (22). So when integrating over all the MN p-variables in Eq. (45), we see that we get

$$\int \prod_{i=1}^{M} \Big(\frac{d\mathbf{p}_i}{(2\pi)^N}\Big) T_M \cdots T_1 = C \exp\Big(i \sum_{m=1}^{M} \big(L'_B(\mathbf{q}_m, \dot{\mathbf{q}}_m) + L''_B(\mathbf{q}_{m-1}, \dot{\mathbf{q}}_m)\big)\delta t\Big), \tag{51}$$

where $C = 1/(2\pi i\xi^{1/4}\delta t)^{MN/2}$, and where L'_B and L''_B are defined for all $\mathbf{q}, \dot{\mathbf{q}} \in \mathbb{R}^N$ by

$$\begin{aligned} L'_B(\mathbf{q}, \dot{\mathbf{q}}) &= \sum_{\mathbf{k},\sigma} \Big(\frac{1}{2}\dot{q}_{1\mathbf{k}\sigma}^2 - \frac{1}{2}k^2 q_{2\mathbf{k}\sigma}^2 + \frac{\xi}{2}(\dot{q}_{0\mathbf{k}\sigma} + \sigma k q_{3\mathbf{k}-\sigma})^2\Big), \\ L''_B(\mathbf{q}, \dot{\mathbf{q}}) &= \sum_{\mathbf{k},\sigma} \Big(\frac{1}{2}\dot{q}_{2\mathbf{k}\sigma}^2 - \frac{1}{2}k^2 q_{1\mathbf{k}\sigma}^2 + \frac{1}{2}(\dot{q}_{3\mathbf{k}\sigma} + \sigma k q_{0\mathbf{k}-\sigma})^2\Big). \end{aligned} \tag{52}$$

To abbreviate Eq. (51), let us note that the exponent on the right-hand side is a finite-M approximation to the classical action. Keeping to our notational convention so far, let us

[5] See Eq. (4.8) and Exercise 2 in Chap. 4 of Hall [4] in particular.

therefore define $\acute{S}_B(M,t,\bar{\mathbf{q}})$ to be this approximate action, namely by letting

$$\acute{S}_B(M,t,\bar{\mathbf{q}}) = \sum_{m=1}^{M} \big(L'_B(\mathbf{q}_m, \dot{\mathbf{q}}_m) + L''_B(\mathbf{q}_{m-1}, \dot{\mathbf{q}}_m)\big)\delta t. \tag{53}$$

We can then reduce Eq. (51) and plug it into Eq. (45) to finally obtain the finite-M approximation to the configuration space path integral:

$$\acute{K}(M,t,\mathbf{q}_M,\mathbf{q}_0,\psi_{j_M},\psi_{j_0}) = C\int \prod_{i=1}^{M-1} (d\mathbf{q}_i)\, e^{i\acute{S}_B(M,t,\bar{\mathbf{q}})} \acute{K}_{IF}(M,t,\bar{\mathbf{q}},\psi_{j_M},\psi_{j_0}). \tag{54}$$

The propagator, $K(t,\mathbf{q}_t,\mathbf{q}_0,\psi_{j_t},\psi_{j_0})$, is then given as the $M\to\infty$ limit of this expression, where the first and last $\mathbf{q}$'s in $\bar{\mathbf{q}}$ are fixed at $\mathbf{q}_0$ and $\mathbf{q}_t$, respectively.

We see that this approximation of the configuration space path integral of Eq. (54) does indeed match the desired form of Eq. (11), namely if we formally let

$$\lim_{M\to\infty} \Big(\frac{1}{2\pi i \xi^{1/4}\delta t}\Big)^{MN/2} \int \prod_{i=1}^{M-1} (d\mathbf{q}_i) \to \int_{\mathbf{q}(0)=\mathbf{q}_0}^{\mathbf{q}(t)=\mathbf{q}_t} \mathcal{D}\mathbf{q} \to \int \mathcal{D}\mathbf{q}, \tag{55}$$

$$\acute{S}_B(M,t,\bar{\mathbf{q}}) \to \int_0^t L_B\big(\dot{\mathbf{q}}(t),\mathbf{q}(t)\big)dt \equiv S_B(t,\mathbf{q}), \tag{56}$$

$$\acute{K}_{IF}(M,t,\bar{\mathbf{q}},\psi_{j_t},\psi_{j_0}) \to K_{IF}(t,\mathbf{q},\psi_{j_t},\psi_{j_0}), \tag{57}$$

giving us

$$K(t,\mathbf{q}_t,\mathbf{q}_0,\psi_{j_t},\psi_{j_0}) = \int \mathcal{D}\mathbf{q}\, e^{iS_B(t,\mathbf{q})} K_{IF}(t,\mathbf{q},\psi',\psi). \tag{58}$$

This would be the conclusion of this section if not for the fact that we needed to put cutoffs on $\hat{H}_{init}$ to get here. We have let these be implicit so far, but they take the form of some bounds on the $V(\mathbf{q})$ and $\mathbf{W}(\mathbf{q})$-potentials in H'_B, H''_B, L'_B and L''_B above, and also some bounds that alter $\hat{H}'_D(\mathbf{q})$ and $\hat{H}''_D(\mathbf{q})$ slightly for all $\mathbf{q}$.

To lift these bounds again, let us first define a sequence $(\hat{H}_{init}(\kappa))_{\kappa\in\mathbb{N}}$ of cut-off versions of $\hat{H}_{init}$ where the cutoffs are lifted gradually as $\kappa\to\infty$. We then want to find a subspace $W\subset \mathrm{Dom}(\hat{H}_{init})\cap\mathrm{Dom}(\sum_{\mathbf{k},\sigma}\sum_{\mu=0}^{3}\hat{\Pi}^2_{\mu\mathbf{k}\sigma})$ that is dense in $\mathbf{H}_{init}$ such that

$$\lim_{\kappa\to\infty} \hat{H}_{init}(\kappa)\chi = \hat{H}_{init}\chi \tag{59}$$

for all $\chi\in W$. One can show that $W = C_c^\infty(\mathbb{R}^N\times\mathbb{R}^{3n};\mathbb{C}^{4^n})$ fulfills this, which is the space of all infinitely differentiable functions with compact support on $\mathbb{R}^N\times\mathbb{R}^{3n}$ and with values in $\mathbb{C}^{4^n}$. We argue this in Appendix B, for the interested reader, in addition to arguing that $\hat{H}_{init}$ is self-adjoint on a certain domain.

Given a W that fulfills this, we then get that all "almost eigenvectors"[6] of $\hat{H}_{init}$ in W will also be almost eigenvectors of $\hat{H}_{init}(\kappa)$ when κ is sufficiently large. Since W is dense

[6] Here we use the terminology found in Hall [4] (see Definition 10.24 in particular), where an "almost eigenvector," or an "ε-almost eigenvector," of an operator $\hat{A}$ refers to a vector for which $\|\hat{A}\chi-\lambda\chi\| < \varepsilon\|\chi\|$ for some $\lambda\in\mathbb{C}$, $\varepsilon>0$.

in $\mathbf{H}_{init}$, we can approximate each $\chi \in \mathbf{H}_{init}$ arbitrarily well by some finite sum of such (simultaneous) almost eigenvectors. Suppose that χ_j is a finite sequence of such almost eigenvectors and that λ_j is the corresponding sequence of (approximate) eigenvalues. If we then further suppose that $\sum_j a_j \chi_j \approx \chi$, we therefore see that

$$e^{-i\hat{H}_{init}t}\chi \approx \sum_j e^{-i\lambda_j t} a_j \chi_j \approx e^{-i\hat{H}_{init}(\kappa)t}\chi \tag{60}$$

for sufficiently large κ. And since $\exp(-i\hat{H}_{init}(\kappa)t)$ thus approximates $\exp(-i\hat{H}_{init}t)$ arbitrarily well, we get that the path integral derived above must approximate the propagator of $\hat{H}_{init}$ arbitrarily well, i.e. well when all the cutoffs are lifted sufficiently.

5 Separation of variables to eliminate the excessive degrees of freedom in the A^μ field

As mentioned in the last part of Sect. 3, the dimensions represented by all the $\widetilde{A}_{0\mathbf{k}\sigma}$ and $\widetilde{A}_{3\mathbf{k}\sigma}$-coordinates are excessive to us, meaning that we hope to be able to get rid of them somehow. And we want to be able to do this in a way where we do not obtain any additional bosonic states in the continuum limit.

One approach for trying to achieve this might be to go back and try to fix the gauge for the paths in the path integral. But this has the chance of breaking the Lorentz covariance if we are not careful, and what is more, it stops us from being able to get back to the Hamiltonian the way we came.

Luckily, there is another solution, as we will see in this section. Let us start by transforming $\hat{H}_{init}$ into a new operator, call it $\hat{H}_{tran}$, with the following change of basis:

$$\hat{H}_{tran} = U^{-1}(\mathbf{q}, \bar{\mathbf{x}})\hat{H}_{init}U(\mathbf{q}, \bar{\mathbf{x}}), \tag{61}$$

where

$$U(\mathbf{q}, \bar{\mathbf{x}}) = \exp\Big(-iq_F \sum_{\mathbf{k},\sigma}\sum_{j=1}^{n} \frac{\sigma}{k}\widetilde{A}_{3\mathbf{k}-\sigma} f_{\mathbf{k}\sigma}(\mathbf{x}_j)\Big). \tag{62}$$

(See Appendix C for a note on the original motivation for trying this factor.) We then see that for all $\chi \in \mathrm{Dom}(\hat{H}_{init})$, we formally get

$$\hat{H}_{tran}\chi = U^{-1}\Big[\sum_{\mathbf{k},\sigma}\big(\frac{1}{2}\hat{\Pi}^2_{3\mathbf{k}\sigma}U - \sigma k \hat{A}_{0\mathbf{k}-\sigma}\hat{\Pi}_{3\mathbf{k}\sigma}U\big) + \sum_{j=1}^{n}\boldsymbol{\alpha}_j \cdot \hat{\mathbf{p}}_j U\Big]\chi + \hat{H}_{init}\chi. \tag{63}$$

Here we have thus implicitly extended $\hat{A}_{0\mathbf{k}\sigma}$, $\hat{\Pi}_{3\mathbf{k}\sigma}$ and $\hat{\mathbf{p}}_j$ to $\mathbf{H}_{init}$, such that we formally have $\hat{\mathbf{p}}_j U = -i\nabla_{\mathbf{x}_j}U$, $\hat{\Pi}^2_{3\mathbf{k}\sigma}U = -\partial^2/\partial\widetilde{A}^2_{3\mathbf{k}\sigma}U$ and $\hat{A}_{0\mathbf{k}-\sigma}\hat{\Pi}_{3\mathbf{k}\sigma}U = -i\widetilde{A}_{0\mathbf{k}-\sigma}\partial/\partial\widetilde{A}_{3\mathbf{k}\sigma}U$. Also, the product between two functions such as U and χ is understood as $U\chi = (\mathbf{q}, \bar{\mathbf{x}}) \mapsto U(\mathbf{q}, \bar{\mathbf{x}})\chi(\mathbf{q}, \bar{\mathbf{x}})$.

If we now consider the term $\sum_{j=1}^{n} \boldsymbol{\alpha}_j \cdot \hat{\mathbf{p}}_j U$ in Eq. (63), we see that

$$\begin{aligned} U^{-1}(\mathbf{q},\bar{\mathbf{x}}) \sum_{j=1}^{n} \hat{\mathbf{p}}_j U(\mathbf{q},\bar{\mathbf{x}}) &= q_F \sum_{\mathbf{k},\sigma} \sum_{j=1}^{n} \sigma \mathbf{k} \frac{\sigma}{k} \widetilde{A}_{3\mathbf{k}-\sigma} f_{\mathbf{k}-\sigma}(\mathbf{x}_j) \\ &= q_F \sum_{\mathbf{k},\sigma} \sum_{j=1}^{n} \mathbf{e}_{3\mathbf{k}} \widetilde{A}_{3\mathbf{k}\sigma} f_{\mathbf{k}\sigma}(\mathbf{x}_j), \end{aligned} \tag{64}$$

where we have again used the fact that $\partial f_{\mathbf{k}\sigma}/\partial x_i = -\sigma k_i f_{\mathbf{k}-\sigma}$. The contribution from this term will therefore cancel out with the potential term $-q_F \sum_{j=1}^{n} \sum_{\mathbf{k},\sigma} \boldsymbol{\alpha}_j \cdot \mathbf{e}_{3\mathbf{k}} \widetilde{A}_{3\mathbf{k}\sigma} f_{\mathbf{k}\sigma}(\mathbf{x}_j)$ in $\hat{H}_{init}$, also coming from $\hat{H}_D(\mathbf{q})$. This is not entirely coincidental, as it was in fact the original motivation for trying this factor (see Appendix C for more details).

More remarkable, however, is the fact that the contribution from $-\sum_{\mathbf{k},\sigma} \sigma k \hat{A}_{0\mathbf{k}-\sigma} \hat{\Pi}_{3\mathbf{k}\sigma} U$ in Eq. (63) will cancel out with another potential from the Dirac Hamiltonian as well. To evaluate this term, we first of all note that

$$\hat{\Pi}_{3\mathbf{k}\sigma} U(\mathbf{q},\bar{\mathbf{x}}) = q_F \frac{\sigma}{k} \sum_{j=1}^{n} f_{\mathbf{k}-\sigma}(\mathbf{x}_j) U(\mathbf{q},\bar{\mathbf{x}}), \tag{65}$$

which immediately gives us

$$\begin{aligned} -U^{-1}(\mathbf{q},\bar{\mathbf{x}}) \sum_{\mathbf{k},\sigma} \sigma k \widetilde{A}_{0\mathbf{k}-\sigma} \hat{\Pi}_{3\mathbf{k}\sigma} U(\mathbf{q},\bar{\mathbf{x}}) &= -q_F \sum_{\mathbf{k},\sigma} \sum_{j=1}^{n} \sigma^2 \widetilde{A}_{0\mathbf{k}-\sigma} f_{\mathbf{k}-\sigma}(\mathbf{x}_j) \\ &= -q_F \sum_{\mathbf{k},\sigma} \sum_{j=1}^{n} \widetilde{A}_{0\mathbf{k}\sigma} f_{\mathbf{k}\sigma}(\mathbf{x}_j). \end{aligned} \tag{66}$$

This term thus cancels out with the electric potential term in the Dirac Hamiltonian! We are therefore left with only the transverse potentials in this Hamiltonian. This indeed seems to be a good step if our goal is to decouple the $\widetilde{A}_{0\mathbf{k}\sigma}$ and $\widetilde{A}_{3\mathbf{k}\sigma}$-variables from the fermions!

We still have the term $\sum_{\mathbf{k},\sigma} \hat{\Pi}_{3\mathbf{k}\sigma}^2 U/2$ in Eq. (63) left to consider, though, as this term might add another dependency on $\{\widetilde{A}_{0\mathbf{k}\sigma}, \widetilde{A}_{3\mathbf{k}\sigma}\}$ if we are not lucky. By operating with $\hat{\Pi}_{3\mathbf{k}\sigma}$ once more on both sides of Eq. (65), and multiplying with $U^{-1}/2$, we get

$$U^{-1}(\mathbf{q},\bar{\mathbf{x}}) \frac{1}{2} \sum_{\mathbf{k},\sigma} \hat{\Pi}_{3\mathbf{k}\sigma}^2 U(\mathbf{q},\bar{\mathbf{x}}) = \frac{1}{2} \sum_{\mathbf{k},\sigma} \Big(\frac{q_F}{k} \sum_{j=1}^{n} f_{\mathbf{k}\sigma}(\mathbf{x}_j) \Big)^2 \equiv V(\bar{\mathbf{x}}), \tag{67}$$

where we have once again flipped σ in the summation to get the first equality, and defined $V(\bar{\mathbf{x}})$ as the resulting expression. This term thus contributes with a potential energy $V(\bar{\mathbf{x}})$ that only depends on $\bar{\mathbf{x}}$, and thus not on any of the $\widetilde{A}_{0\mathbf{k}\sigma}$ and $\widetilde{A}_{3\mathbf{k}\sigma}$-variables.

By plugging Eqs. (64), (66) and (67) into Eq. (63), we can now rewrite $\hat{H}_{tran}$ as

$$\hat{H}_{tran} = \hat{H}_{B03} + \hat{H}_{B12} + \hat{H}_{IF12} + \hat{V}, \tag{68}$$

where $\hat{V}$ is the multiplication operator on $\mathbf{H}_{init}$ associated with $V(\bar{\mathbf{x}})$, where

$$\hat{H}_{B03} = \sum_{\mathbf{k},\sigma} \big(\frac{\xi}{2} \hat{\Pi}_{0k\sigma}^2 + \frac{1}{2} \hat{\Pi}_{3\mathbf{k}\sigma}^2 - \sigma k \hat{A}_{3\mathbf{k}-\sigma} \hat{\Pi}_{0\mathbf{k}\sigma} - \sigma k \hat{A}_{0\mathbf{k}-\sigma} \hat{\Pi}_{3\mathbf{k}\sigma} \big) \otimes \hat{I}_F, \tag{69}$$

$$\hat{H}_{B12} = \sum_{\mathbf{k},\sigma} \sum_{\lambda=1}^{2} \big(\frac{1}{2} \hat{\Pi}_{\lambda\mathbf{k}\sigma}^2 + \frac{1}{2} k^2 \hat{A}_{\lambda\mathbf{k}\sigma}^2 \big) \otimes \hat{I}_F, \tag{70}$$

and where

$$\hat{H}_{IF12}\chi(\mathbf{q},\bar{\mathbf{x}}) = \sum_{j=1}^{n}\Big(\boldsymbol{\alpha}_j\cdot\hat{\mathbf{p}}_j - q_F\sum_{\mathbf{k},\sigma}\sum_{\lambda=1}^{2}(\boldsymbol{\alpha}_j\cdot\mathbf{e}_{\lambda\mathbf{k}})\widetilde{A}_{\lambda\mathbf{k}\sigma}f_{\mathbf{k}\sigma}(\mathbf{x}_j) + \beta_j m_F\Big)\chi(\mathbf{q},\bar{\mathbf{x}}) \tag{71}$$

for all $\mathbf{q}\in\mathbb{R}^N$, $\bar{\mathbf{x}}\in\mathbb{R}^{3n}$ and $\chi\in\mathbf{H}_{init}$, formally.

We see that our change of basis thus separates the Hamiltonian into two parts, $\hat{H}_{B03}$ and $\hat{H}_{B12}+\hat{H}_{IF12}+\hat{V}$, such that all dependencies on the $\widetilde{A}_{0\mathbf{k}\sigma}$ and $\widetilde{A}_{3\mathbf{k}\sigma}$-variables are contained in $\hat{H}_{B03}$ and all dependencies on the $\widetilde{A}_{1\mathbf{k}\sigma}$, $\widetilde{A}_{2\mathbf{k}\sigma}$ and $\mathbf{x}_j$-variables are contained in $\hat{H}_{B12}+\hat{H}_{IF12}+\hat{V}$. This means that the solutions to the dynamics generated by $\hat{H}_{tran}$, i.e. solutions to the equation $\chi_t = \exp(-i\hat{H}_{tran}t)\chi_0$, will be separable.[7] We are therefore not far, it seems, from being able to conclude that space of the $\widetilde{A}_{0\mathbf{k}\sigma}$ and $\widetilde{A}_{3\mathbf{k}\sigma}$-coordinates are "unphysical," since they are now effectively decoupled from both the fermions and the photons, and thus do not affect any measurements.

We have to be a bit more careful before we throw $\hat{H}_{B03}$ in this transformed Hamiltonian away, however, since we have to remember that our objective is to derive a Hamiltonian that can be shown, via its corresponding Lagrangian in the path integral formulation, to be Lorentz-covariant in the continuum limit. We can therefore only throw away the $\widetilde{A}_{0\mathbf{k}\sigma}$ and $\widetilde{A}_{3\mathbf{k}\sigma}$-dimensions if we can show that doing so will not break the Lorentz covariance.

Luckily, as we will see in the next section, throwing away $\hat{H}_{B03}$ in $\hat{H}_{tran}$, together with its associated dimensions, does indeed seem to leave the Lorentz covariance intact. If we therefore define

$$\hat{H}_{red} = \hat{H}_{B12} + \hat{H}_{IF12} + \hat{V} \tag{72}$$

on the reduced Hilbert space given by

$$\mathbf{H}_{red} = L^2(\mathbb{R}^{N/2}\times\mathbb{R}^{3n};\mathbb{C}^{4^n}), \tag{73}$$

namely where we have removed the $\widetilde{A}_{0\mathbf{k}\sigma}$ and $\widetilde{A}_{3\mathbf{k}\sigma}$-dimensions, we thus expect $\hat{H}_{red}$ to generate Lorentz-covariant dynamics in the continuum limit. This is provided, however, that $\hat{H}_{red}$ and $\hat{H}_{B03}$ are self-adjoint, and that $\hat{H}_{red}$ turns into a self-adjoint operator in the continuum limit. We argue that $\hat{H}_{red}$ and $\hat{H}_{B03}$ are self-adjoint in Appendix B, and in Sect. 8, we will propose a candidate for the continuum limit of $\hat{H}_{red}$, although we will leave the proof of its existence and self-adjointness for future work.

6 Preservation of the Lorentz covariance for the reduced Hilbert space

Now that the redundant, "unphysical" degrees of freedom are decoupled from the "physical" ones, there is actually a very simple argument for why we do not break the Lorentz covariance by removing said degrees of freedom. This argument does require a somewhat relaxed version of the more natural definition of Lorentz covariance, but we will see that this relaxed

[7] For all $\chi_0\in\mathbf{H}_{init}$ that can be written as $\Phi_0\Psi_0$, where Φ_0 is a function of the $\widetilde{A}_{0\mathbf{k}\sigma}$ and $\widetilde{A}_{3\mathbf{k}\sigma}$-variables and where Ψ_0 is a function of all the other variables, χ_t will thus be equal to $\Phi_t\Psi_t$, where $\Phi_t = \exp(-i\hat{H}_{B03}t)\Phi_0$ and $\Psi_t = \exp[-i(\hat{H}_{B12}+\hat{H}_{IF12}+\hat{V})t]\Psi_0$.

definition is still sufficient to ensure that the theory gives us Lorentz-covariant predictions for any measurement.

A natural definition of Lorentz covariance for a quantum theory is that for any pair of spacelike hyperplanes P and P', there exists a unitary map that Lorentz-transforms any state initially on P from this hyperplane and onto P' such that when the initial and final state, respectively located on P and on P', are time-evolved back or forth along their respective time axes, they give the exact same probabilities for positional measurements in any given point of spacetime.[8]

Suppose now that this is true for the continuum limit $\hat{H}_{init}$,[9] and suppose that there exists a sequence $(\hat{H}_{init}(j))_{j\in\mathbb{N}}$ on Hilbert spaces $(\mathbf{H}_{init}(j))_{j\in\mathbb{N}}$, where $\delta k \to 0$ and $N_k \delta k \to \infty$ when $j \to \infty$, such that the continuum limit theory can be approximated to arbitrary precision by $\hat{H}_{init}(j)$ when $j \to \infty$. (The latter supposition will indeed be how we will define a continuum limit of a theory in Sect. 8, given that it exists.) Then for a sufficiently large j, the measurement probabilities in any given point of spacetime will be approximately the same for any initial state $\chi \in \mathbf{H}_{init}(j)$ and its Lorentz transform $\chi' \in \mathbf{H}_{init}(j)$.

Let us omit the j's again, and let us then consider a state χ that can be written as $\chi = \Phi \otimes \Psi$, where $\Phi \in L^2(\mathbb{R}^{N/2})$ and $\Psi \in \mathbf{H}_{red}$. Its Lorentz transform χ' might then not necessarily have the same separable form, but it can still be written as

$$\chi' = \sum_{i=1}^{\infty} \Phi_i \otimes \Psi'_i \tag{74}$$

for some sequence Ψ'_i belonging to $\mathbf{H}_{red}$, where $\{\Phi_i\}$ is taken to be an orthonormal basis of $L^2(\mathbb{R}^{N/2})$. Given our assumptions, χ' must then yield approximately the same measurement probabilities for the fermion positions as the original Ψ. But since all such physical measurements are independent of the Φ_i's, we see that χ' will yield the same measurements as an ensemble, namely where the density operator is given by

$$\rho' = \sum_{i=1}^{\infty} |\Psi'_i\rangle \langle\Psi'_i| \tag{75}$$

on the reduced Hilbert space. So while this argument does not show that there always exists a pure state Ψ' that is a valid Lorentz transform of Ψ, we do, however, get that all ensembles (including pure states), given by a ρ, will have a valid Lorentz transform ρ' for which the measurement probabilities are (approximately) the same. And it is not hard to see why this relaxed definition of Lorentz covariance of a quantum theory is still enough to ensure that all experiments conducted within the theory will obey the principles of special relativity.

We will thus let this argument suffice for this paper. For any reader that is interested in a proof that there also exists a Lorentz transformation for the continuum limit of $\hat{H}_{red}$ for which pure states are mapped exclusively into other pure states, we will briefly return to the matter of how this can potentially be proven in Sect. 11.

[8] At least, we want the predictions for positional measurements to be the same for the fermions, since their positions are clearly defined for all inertial frames (before we make the Dirac sea reinterpretation, that is). And because all measurements of photons can be reinterpreted as measurements of fermions in practice, we are therefore free to leave the photons out of the definition of Lorentz covariance.

[9] In this paper, we only give the initial heuristic reasoning that we saw in Sect. 2 for why the continuum limit of $\hat{H}_{init}$ should be Lorentz-covariant. In Sect. 11, we will, however, get back to the matter and give a summary of what is needed in order to complete this argument.

7 Rewriting $\hat{H}_{red}$ in terms of ladder operators

For this section, we want to rewrite $\hat{H}_{red}$ in terms of ladder operators in order to prepare it for the continuum limit. For the photons, we thus first of all want to rewrite all the $\hat{A}_{\lambda\mathbf{k}\sigma}$ and $\hat{\Pi}_{\lambda\mathbf{k}\sigma}$-operators in terms of the relevant ladder operators associated with each of the simple harmonic oscillators. We do this so that we can then follow the well-known process from QFT of reinterpreting these ladder operators as creation and annihilation operators.

We also want to introduce creation and annihilation operators for the fermions, partly because this will make the comparison with the conventional theory of QED more clear. And furthermore, this will also help prepare the theory for the Dirac sea reinterpretation, which we will discuss in Sect. 9 to some extent.

To begin rewriting $\hat{H}_{red}$, we first of all recall from Eqs. (70–73) that

$$\hat{H}_{red} = \hat{H}_{B12} + \hat{H}_{IF12} + \hat{V}. \tag{76}$$

Since we no longer have the $\mu \in \{0, 3\}$-dimensions in play, let us first of all get rid of the 12-indices by (re)defining $\hat{H}_B$, $\hat{H}_F$ and $\hat{H}_I$ as

$$\hat{H}_B = \sum_{\mathbf{k},\sigma} \sum_{\lambda=1}^{2} \big(\frac{1}{2}\hat{\Pi}^2_{\lambda\mathbf{k}\sigma} + \frac{1}{2}k^2 \hat{A}^2_{\lambda\mathbf{k}\sigma}\big), \tag{77}$$

$$\hat{H}_F = \sum_{j=1}^{n} (\boldsymbol{\alpha}_j \cdot \hat{\mathbf{p}}_j + \beta_j m_F), \tag{78}$$

$$\hat{H}_I = -q_F \sum_{j=1}^{n} \sum_{\mathbf{k},\sigma} \sum_{\lambda=1}^{2} \hat{A}_{\lambda\mathbf{k}\sigma} \otimes (\boldsymbol{\alpha}_j \cdot \mathbf{e}_{\lambda\mathbf{k}}) \hat{f}_{\mathbf{k}\sigma j}, \tag{79}$$

where $\hat{f}_{\mathbf{k}\sigma j}$ is defined by

$$\hat{f}_{\mathbf{k}\sigma j} \psi(\bar{\mathbf{x}}) = f_{\mathbf{k}\sigma}(\mathbf{x}_j) \psi(\bar{\mathbf{x}}) \tag{80}$$

for all $\bar{\mathbf{x}} \in \mathbb{R}^{3n}$, $\mathbf{k} \in \mathbb{K}$, $\sigma \in \{1, -1\}$, $j \in \{1, \ldots, n\}$ and $\psi \in \mathbf{H}_F$. Since we want these operators to be defined on $\mathbf{H}_{red}$, we should also redefine $\mathbf{H}_B = L^2(\mathbb{R}^{N/2})$ for these equations, and thus let $\hat{A}_{\lambda\mathbf{k}\sigma}$ and $\hat{\Pi}_{\lambda\mathbf{k}\sigma}$ be redefined as operators on this new $\mathbf{H}_B$. All this then allows us to write $\hat{H}_{red}$ simply as

$$\hat{H}_{red} = \hat{H}_B \otimes \hat{I}_F + \hat{I}_B \otimes \hat{H}_F + \hat{H}_I + \hat{V}. \tag{81}$$

We have here chosen to let $\hat{H}_B$ and $\hat{H}_F$ be defined only on $\mathbf{H}_B$ and $\mathbf{H}_F$, respectively, since this will simplify the following analysis.

Now, the well-known ladder operators for our simple harmonic oscillators are given by

$$\begin{aligned} \hat{a}_{\lambda\mathbf{k}\sigma} &= \frac{1}{\sqrt{2k}}(k\hat{A}_{\lambda\mathbf{k}\sigma} + i\hat{\Pi}_{\lambda\mathbf{k}\sigma}), \\ \hat{a}^\dagger_{\lambda\mathbf{k}\sigma} &= \frac{1}{\sqrt{2k}}(k\hat{A}_{\lambda\mathbf{k}\sigma} - i\hat{\Pi}_{\lambda\mathbf{k}\sigma}). \end{aligned} \tag{82}$$

They have commutation relations given by

$$\begin{aligned} [\hat{a}_{\lambda'\mathbf{k}'\sigma'}, \hat{a}_{\lambda\mathbf{k}\sigma}] &= 0, \\ [\hat{a}^\dagger_{\lambda'\mathbf{k}'\sigma'}, \hat{a}^\dagger_{\lambda\mathbf{k}\sigma}] &= 0, \\ [\hat{a}_{\lambda'\mathbf{k}'\sigma'}, \hat{a}^\dagger_{\lambda\mathbf{k}\sigma}] &= \delta_{\lambda'\lambda}\delta_{\mathbf{k}'\mathbf{k}}\delta_{\sigma'\sigma} \end{aligned} \tag{83}$$

for all indices in their respective ranges, where $[\cdot,\cdot]$ denotes the commutator. We therefore see that we can rewrite $\hat{H}_B$ as

$$\hat{H}_B = \sum_{\mathbf{k},\sigma}\sum_{\lambda=1}^{2}\frac{1}{2}k(\hat{a}^\dagger_{\lambda\mathbf{k}\sigma}\hat{a}_{\lambda\mathbf{k}\sigma}+\hat{a}_{\lambda\mathbf{k}\sigma}\hat{a}^\dagger_{\lambda\mathbf{k}\sigma}) = \sum_{\mathbf{k},\sigma}\sum_{\lambda=1}^{2}k\big(\hat{a}^\dagger_{\lambda\mathbf{k}\sigma}\hat{a}_{\lambda\mathbf{k}\sigma}+\frac{1}{2}\hat{I}_B\big). \tag{84}$$

In order for $\hat{H}_B$ to get a well-defined continuum limit, we want to subtract the constant energy coming from $\sum_{\mathbf{k},\sigma}\sum_{\lambda=1}^{2}\hat{I}_B/2$. We are free to do this, since this will not change the actual dynamics of the system; only the overall phase as a function of time. We can therefore change $\hat{H}_B$ going forward by letting

$$\hat{H}_B = \sum_{\mathbf{k},\sigma}\sum_{\lambda=1}^{2}k\hat{a}^\dagger_{\lambda\mathbf{k}\sigma}\hat{a}_{\lambda\mathbf{k}\sigma}. \tag{85}$$

We can also rewrite $\hat{H}_I$ in terms of these ladder operators, which gives us

$$\hat{H}_I = -q_F\sum_{j=1}^{n}\sum_{\mathbf{k},\sigma}\sum_{\lambda=1}^{2}\frac{1}{\sqrt{2k}}(\hat{a}_{\lambda\mathbf{k}\sigma}+\hat{a}^\dagger_{\lambda\mathbf{k}\sigma})\otimes(\boldsymbol{\alpha}_j\cdot\mathbf{e}_{\lambda\mathbf{k}})\hat{f}_{\mathbf{k}\sigma j}. \tag{86}$$

When we encounter creation and annihilation operators for photons in literature, they are commonly those that create and annihilate momentum eigenstates. But if we look at $\hat{a}^\dagger_{\lambda\mathbf{k}\sigma}$ and $\hat{a}_{\lambda\mathbf{k}\sigma}$, we see that these are the ladder operators of sine and cosine modes, which for each $\mathbf{k}\in\mathbb{K}\subset\mathbb{R}^2\times\mathbb{R}_+$ include an equal mix of a $\mathbf{k}\in\mathbb{K}$ and a $-\mathbf{k}\in\mathbb{K}$ momentum state. If we want ladder operators that turn into the creation and annihilation operators of momentum eigenstates, we thus need to change the basis for $\hat{a}^\dagger_{\lambda\mathbf{k}\sigma}$ and $\hat{a}_{\lambda\mathbf{k}\sigma}$. We will see that the change of basis that achieves this is when we define

$$\begin{aligned}
\hat{a}_{\lambda\mathbf{k}} &= \frac{1}{\sqrt{2}}(\hat{a}_{\lambda\mathbf{k}1}-i\hat{a}_{\lambda\mathbf{k}-1}), &\qquad \hat{a}^\dagger_{\lambda\mathbf{k}} &= \frac{1}{\sqrt{2}}(\hat{a}^\dagger_{\lambda\mathbf{k}1}+i\hat{a}^\dagger_{\lambda\mathbf{k}-1}),\\
\hat{a}_{\lambda-\mathbf{k}} &= \frac{1}{\sqrt{2}}(\hat{a}_{\lambda\mathbf{k}1}+i\hat{a}_{\lambda\mathbf{k}-1}), &\qquad \hat{a}^\dagger_{\lambda-\mathbf{k}} &= \frac{1}{\sqrt{2}}(\hat{a}^\dagger_{\lambda\mathbf{k}1}-i\hat{a}^\dagger_{\lambda\mathbf{k}-1}).
\end{aligned} \tag{87}$$

(Note that for the $\hat{a}^\dagger_{\lambda\pm\mathbf{k}}$-operators to the right, the formulas are similar to how complex exponential functions are expressed in terms of cosine and sine functions.) It is easy to check that these ladder operators have the same commutation relations as before.

For $\hat{H}_B$, we then see that since

$$\begin{aligned}
&\hat{a}^\dagger_{\lambda\mathbf{k}}\hat{a}_{\lambda\mathbf{k}}+\hat{a}^\dagger_{\lambda-\mathbf{k}}\hat{a}_{\lambda-\mathbf{k}}\\
&\qquad = \frac{1}{2}(\hat{a}^\dagger_{\lambda\mathbf{k}1}+ia^\dagger_{\lambda\mathbf{k}-1})(\hat{a}_{\lambda\mathbf{k}1}-ia_{\lambda\mathbf{k}-1})+\frac{1}{2}(\hat{a}^\dagger_{\lambda\mathbf{k}1}-ia^\dagger_{\lambda\mathbf{k}-1})(\hat{a}_{\lambda\mathbf{k}1}+ia_{\lambda\mathbf{k}-1})\\
&\qquad = \hat{a}^\dagger_{\lambda\mathbf{k}1}\hat{a}_{\lambda\mathbf{k}1}+a^\dagger_{\lambda\mathbf{k}-1}a_{\lambda\mathbf{k}-1}\\
&\qquad = \sum_{\sigma}\hat{a}^\dagger_{\lambda\mathbf{k}\sigma}\hat{a}_{\lambda\mathbf{k}\sigma},
\end{aligned} \tag{88}$$

we get

$$\hat{H}_B = \sum_{\pm\mathbf{k}\in\mathbb{K}}\sum_{\lambda=1}^{2}k\hat{a}^\dagger_{\lambda\mathbf{k}}\hat{a}_{\lambda\mathbf{k}}. \tag{89}$$

This is the desired result since we expect $\hat{a}^\dagger_{\lambda\mathbf{k}}\hat{a}_{\lambda\mathbf{k}}$ to turn into the occupation number operator for photons with wave vector $\mathbf{k}$ and polarization λ in the continuum limit.

We can also rewrite $\hat{H}_I$ in terms of these new ladder operators by noting that

$$\begin{aligned}\sum_\sigma(\hat{a}_{\lambda\mathbf{k}\sigma}+\hat{a}^\dagger_{\lambda\mathbf{k}\sigma})\sqrt{\mathcal{V}}f_{\mathbf{k}\sigma}(\mathbf{x}) &= \sqrt{2}(\hat{a}_{\lambda\mathbf{k}1}+\hat{a}^\dagger_{\lambda\mathbf{k}1})\cos(\mathbf{k}\cdot\mathbf{x})+\sqrt{2}(\hat{a}_{\lambda\mathbf{k}-1}+\hat{a}^\dagger_{\lambda\mathbf{k}-1})\sin(\mathbf{k}\cdot\mathbf{x})\\ &= \frac{1}{\sqrt{2}}(\hat{a}_{\lambda\mathbf{k}1}+\hat{a}^\dagger_{\lambda\mathbf{k}1}-i\hat{a}_{\lambda\mathbf{k}-1}-i\hat{a}^\dagger_{\lambda\mathbf{k}-1})e^{i\mathbf{k}\cdot\mathbf{x}}\\ &\qquad+\frac{1}{\sqrt{2}}(\hat{a}_{\lambda\mathbf{k}1}+\hat{a}^\dagger_{\lambda\mathbf{k}1}+i\hat{a}_{\lambda\mathbf{k}-1}+i\hat{a}^\dagger_{\lambda\mathbf{k}-1})e^{-i\mathbf{k}\cdot\mathbf{x}}\\ &= (\hat{a}_{\lambda\mathbf{k}}+\hat{a}^\dagger_{\lambda-\mathbf{k}})e^{i\mathbf{k}\cdot\mathbf{x}}+(\hat{a}_{\lambda-\mathbf{k}}+\hat{a}^\dagger_{\lambda\mathbf{k}})e^{-i\mathbf{k}\cdot\mathbf{x}},\end{aligned}\tag{90}$$

where we have used Eq. (87) for the last equality, and of course also used the fact that

$$\cos(\mathbf{k}\cdot\mathbf{x})=\frac{1}{2}(e^{i\mathbf{k}\cdot\mathbf{x}}+e^{-i\mathbf{k}\cdot\mathbf{x}}),\quad \sin(\mathbf{k}\cdot\mathbf{x})=\frac{-i}{2}(e^{i\mathbf{k}\cdot\mathbf{x}}-e^{-i\mathbf{k}\cdot\mathbf{x}}).\tag{91}$$

We can thus write

$$\hat{H}_I=-q_F\sum_{j=1}^n\sum_{\pm\mathbf{k}\in\mathbb{K}}\sum_{\lambda=1}^2\frac{1}{\sqrt{2\mathcal{V}k}}(\hat{a}_{\lambda\mathbf{k}}+\hat{a}^\dagger_{\lambda-\mathbf{k}})\otimes(\boldsymbol{\alpha}_j\cdot\mathbf{e}_{\lambda\mathbf{k}})\hat{g}_{\mathbf{k}j},\tag{92}$$

where $\hat{g}_{\mathbf{k}j}$ is defined by

$$\hat{g}_{\mathbf{k}j}\psi(\bar{\mathbf{x}})=e^{i\mathbf{k}\cdot\mathbf{x}_j}\psi(\bar{\mathbf{x}})\tag{93}$$

for all $\bar{\mathbf{x}}\in\mathbb{R}^{3n}$, $\pm\mathbf{k}\in\mathbb{K}$, $j\in\{1,\dots,n\}$ and $\psi\in\mathbf{H}_F$. Note that since $\hat{g}_{\mathbf{k}j}$ is therefore a boost operator for the jth fermion, one can immediately see from this formula that $\hat{H}_I$ obeys momentum preservation.

As mentioned, we also want to rewrite the fermionic part of $\hat{H}_{red}$ in terms of creation and annihilation operators. To do this, we first of all have to rewrite the fermionic operators in the generalized momentum eigenbasis. It is well-known how the Dirac equation can be "diagonalized" in terms of (spinor-valued) generalized eigenfunctions, namely by looking for solutions of the form $\psi(\mathbf{x})=w_s(\mathbf{p})\exp(i\mathbf{p}\cdot\mathbf{x})$, where $w_s(\mathbf{p})$ is a spinor for all $\mathbf{p}=(p_1,p_2,p_3)\in\mathbb{R}^3$ and s. For fermion number $n=1$, s should take values in $\{1,2,3,4\}$. If we define $E_p=\sqrt{p^2+m_F^2}$, where $p=|\mathbf{p}|$, and define

$$\begin{aligned}\begin{pmatrix}u_1(\mathbf{p}) & u_2(\mathbf{p})\end{pmatrix} &= \frac{1}{\sqrt{E_p+m_F}}\begin{pmatrix}E_p+m_F & 0\\ 0 & E_p+m_F\\ p_3 & p_1-ip_2\\ p_1+ip_2 & -p_3\end{pmatrix},\\ \begin{pmatrix}v_3(\mathbf{p}) & v_4(\mathbf{p})\end{pmatrix} &= \frac{1}{\sqrt{E_p-m_F}}\begin{pmatrix}-E_p+m_F & 0\\ 0 & -E_p+m_F\\ p_3 & p_1-ip_2\\ p_1+ip_2 & -p_3\end{pmatrix},\end{aligned}\tag{94}$$

it can then be shown that $\psi(\mathbf{x})=(2E_p)^{-1/2}w_s(\mathbf{p})\exp(i\mathbf{p}\cdot\mathbf{x})$ will be an eigenfunction if

$(w_s(\mathbf{p}))_{s\in\{1,2,3,4\}} = (u_1(\mathbf{p}), u_2(\mathbf{p}), v_3(\mathbf{p}), v_4(\mathbf{p}))$. See e.g. Shankar [8] for how to show this.[10] The eigenvalues will then be E_p for the u_s-solutions and $-E_p$ for the v_s-solutions.

Note that different definitions of the u_s and v_s-spinors can be found in literature, since one is always free to rotate the spinors due to the degeneracy of the energies. There are also many different choices for the α and β-matrices in literature, which also imply different values for u_s and v_s.

Since these are eigenfunctions of $\hat{H}_F$, i.e. when we still have $n = 1$, one thus gets that

$$\langle s', \mathbf{p}'|\hat{H}_{F(n=1)}|s, \mathbf{p}\rangle = \begin{cases} E_p \delta_{s's}(2\pi)^3\delta^3(\mathbf{p}' - \mathbf{p}), & s \in \{1,2\} \\ -E_p \delta_{s's}(2\pi)^3\delta^3(\mathbf{p}' - \mathbf{p}), & s \in \{3,4\} \end{cases}, \tag{95}$$

where $|s, \mathbf{p}\rangle = (2E_p)^{-1/2} w_s(\mathbf{p}) \exp(i\mathbf{p}\cdot\mathbf{x})$ for all $\mathbf{p} \in \mathbb{R}^3$ and $s \in \{1,2,3,4\}$. (Note that we here use the same normalization for the generalized functions as in Sect. 4.) We can therefore rewrite $\hat{H}_F$, in this special case where $n = 1$, as

$$\hat{H}_{F(n=1)} = \int \frac{d\mathbf{p}}{(2\pi)^3}\, E_p \Big(\sum_{s=1}^{2} |s, \mathbf{p}\rangle \langle s, \mathbf{p}| - \sum_{s=3}^{4} |s, \mathbf{p}\rangle \langle s, \mathbf{p}| \Big). \tag{96}$$

We also see that

$$\begin{aligned} \langle s', \mathbf{p}'|(\boldsymbol{\alpha}\cdot\mathbf{e}_{\lambda\mathbf{k}})\hat{g}_{\mathbf{k}}|s, \mathbf{p}\rangle &= \frac{1}{2E_p} \langle w_{s'}(\mathbf{p}')|\boldsymbol{\alpha}\cdot\mathbf{e}_{\lambda\mathbf{k}}|w_s(\mathbf{p})\rangle \langle \mathbf{p}'|\mathbf{p}+\mathbf{k}\rangle \\ &= \frac{(2\pi)^3}{2E_p} w^\dagger_{s'}(\mathbf{p}')\boldsymbol{\alpha}\cdot\mathbf{e}_{\lambda\mathbf{k}} w_s(\mathbf{p})\delta^3(\mathbf{p}' - \mathbf{p} - \mathbf{k}). \end{aligned} \tag{97}$$

It is common in quantum mechanics/QFT literature[11] to have γ^μ and ϵ^μ_λ defined for $\lambda \in \{1,2\}$ such that $\gamma^\mu = (\beta, \beta\alpha_1, \beta\alpha_2, \beta\alpha_3)$ and $\epsilon^\mu_\lambda(\mathbf{k}) = (0, \mathbf{e}_{\lambda\mathbf{k}})$, which means that $\boldsymbol{\alpha}\cdot\mathbf{e}_{\lambda\mathbf{k}} = \gamma^0\gamma^\mu\epsilon_{\mu\lambda}$ (using Einstein notation). It is also common to use a bar notation for the spinors where $\bar{u}_s(\mathbf{p}) \equiv u^\dagger(\mathbf{p})\gamma^0$ and $\bar{v}_s(\mathbf{p}) \equiv v^\dagger(\mathbf{p})\gamma^0$. With these definitions, we can thus use Eq. (97) to get

$$\begin{aligned} (\boldsymbol{\alpha}\cdot\mathbf{e}_{\lambda\mathbf{k}})\hat{g}_{\mathbf{k}} &= \sum_{s,s'=1}^{4} \int \frac{d\mathbf{p}\, d\mathbf{p}'}{(2\pi)^6} \frac{(2\pi)^3}{2E_p} |s', \mathbf{p}'\rangle\, \bar{w}_{s'}(\mathbf{p}')\gamma^\mu\epsilon_{\mu\lambda}(\mathbf{k})w_s(\mathbf{p})\delta^3(\mathbf{p}' - \mathbf{p} - \mathbf{k}) \langle s, \mathbf{p}| \\ &= \sum_{s,s'=1}^{4} \int \frac{d\mathbf{p}}{(2\pi)^3} \frac{1}{2E_p} |s', \mathbf{p}+\mathbf{k}\rangle\, \bar{w}_{s'}(\mathbf{p}+\mathbf{k})\gamma^\mu\epsilon_{\mu\lambda}(\mathbf{k})w_s(\mathbf{p}) \langle s, \mathbf{p}|\,. \end{aligned} \tag{98}$$

We can now rewrite Eqs. (96) and (98) in terms of creation and annihilation operators. First of all, we can write $\hat{H}_{F(n=1)}$ as

$$\hat{H}_{F(n=1)} = \int \frac{d\mathbf{p}}{(2\pi)^3}\, E_p \Big(\sum_{s=1}^{2} \hat{b}^\dagger_s(\mathbf{p})\hat{b}_s(\mathbf{p}) - \sum_{s=3}^{4} \hat{d}^\dagger_s(\mathbf{p})\hat{d}_s(\mathbf{p}) \Big), \tag{99}$$

[10] To get the eigenvalues, use the fact that the determinant of the equations, namely Eqs. (20.2.7–8) in the case of Shankar [8], has to be 0 in order for them to have a solution, i.e. when one has 0 on all the right-hand sides.

[11] See e.g. Lancaster and Blundell [6].

where $\hat{b}_s^\dagger(\mathbf{p})$ and $\hat{b}_s(\mathbf{p})$ respectively creates and annihilates the state $|s, \mathbf{p}\rangle$ for $s \in \{1, 2\}$ and where $\hat{d}_s^\dagger(\mathbf{p})$ and $\hat{d}_s(\mathbf{p})$ creates and annihilates the state $|s, \mathbf{p}\rangle$ for $s \in \{3, 4\}$. More precisely, we require that $(2\pi)^{-3} \int d\mathbf{p}\, \hat{b}_s^\dagger(A\mathbf{p}+\mathbf{k})\hat{b}_s(\mathbf{p}) = (2\pi)^{-3} \int d\mathbf{p}\; |s, A\mathbf{p}+\mathbf{k}\rangle \langle \mathbf{p}|$ for all $s \in \{1, 2\}$, all $\mathbf{k} \in \mathbb{R}^3$, and all unitary matrices $A \in \mathbb{R}^3 \times \mathbb{R}^3$, and similarly for the $\hat{d}_s$-operators.

With these ladder operators, we see that we can also rewrite $(\boldsymbol{\alpha} \cdot \mathbf{e}_{\lambda\mathbf{k}})\hat{g}_\mathbf{k}$ as

$$(\boldsymbol{\alpha} \cdot \mathbf{e}_{\lambda\mathbf{k}})\hat{g}_\mathbf{k} = \int \frac{d\mathbf{p}}{(2\pi)^3}\, \hat{\bar{\psi}}(\mathbf{p}+\mathbf{k})\gamma^\mu \epsilon_{\mu\lambda}(\mathbf{k})\hat{\psi}(\mathbf{p}), \tag{100}$$

where we formally define $\hat{\psi}(\mathbf{p})$ and $\hat{\bar{\psi}}(\mathbf{p})$ for all $\mathbf{p} \in \mathbb{R}^3$ by

$$\begin{aligned} \hat{\psi}(\mathbf{p}) &= \frac{1}{\sqrt{2E_p}}\Big(\sum_{s=1}^{2} u_s(\mathbf{p})\hat{b}_s(\mathbf{p}) + \sum_{s=3}^{4} v_s(\mathbf{p})\hat{d}_s(\mathbf{p})\Big), \\ \hat{\bar{\psi}}(\mathbf{p}) &= \frac{1}{\sqrt{2E_p}}\Big(\sum_{s=1}^{2} \hat{b}_s^\dagger(\mathbf{p})\bar{u}_s(\mathbf{p}) + \sum_{s=3}^{4} \hat{d}_s^\dagger(\mathbf{p})\bar{v}_s(\mathbf{p})\Big). \end{aligned} \tag{101}$$

This treatment of our two fermionic operators, $\hat{H}_F$ and $\sum_{j=1}^n (\boldsymbol{\alpha} \cdot \mathbf{e}_{\lambda\mathbf{k}})\hat{g}_{\mathbf{k}j}$, has all been for $n = 1$. However, for an arbitrary fermion number $n \in \mathbb{N}$, we can extend these results easily since all the n fermions are completely decoupled for both these operators. We can therefore do this treatment for each individual fermion and then put it all together, namely by letting e.g. $\hat{H}_F = \sum_{j=1}^n \hat{H}_{Fj}$, where

$$\hat{H}_{Fj} = \overbrace{\hat{I}_{F(n=1)} \otimes \cdots \otimes \hat{I}_{F(n=1)}}^{j-1} \otimes \hat{H}_{F(n=1)} \otimes \overbrace{\hat{I}_{F(n=1)} \otimes \cdots \otimes \hat{I}_{F(n=1)}}^{n-j} \tag{102}$$

for all $j \in \{1, \ldots, n\}$. This gives us the same formula for $\hat{H}_F$ as for $\hat{H}_{F(n=1)}$:

$$\hat{H}_F = \int \frac{d\mathbf{p}}{(2\pi)^3}\, E_p\Big(\sum_{s=1}^{2} \hat{b}_s^\dagger(\mathbf{p})\hat{b}_s(\mathbf{p}) - \sum_{s=3}^{4} \hat{d}_s^\dagger(\mathbf{p})\hat{d}_s(\mathbf{p})\Big), \tag{103}$$

but where the $\hat{b}$ and $\hat{d}$-operators are now simply extended to all of $\mathbf{H}_F$ (with an arbitrary n) in the same way as Eq. (102) prescribes. This means that $\hat{b}_s^\dagger(\mathbf{p})\hat{b}_s(\mathbf{p})$ and $\hat{d}_s^\dagger(\mathbf{p})\hat{d}_s(\mathbf{p})$ now measures the occupation number of fermions with momentum $\mathbf{p}$ and spin s. (Such extensions are well-known in the literature, and most readers of this paper will probably be well-familiar.) And for the operator $\sum_{j=1}^n (\boldsymbol{\alpha}_j \cdot \mathbf{e}_{\lambda\mathbf{k}})\hat{g}_{\mathbf{k}j}$, we also get a similar formula as before:

$$\sum_{j=1}^{n} (\boldsymbol{\alpha} \cdot \mathbf{e}_{\lambda\mathbf{k}})\hat{g}_{\mathbf{k}j} = \int \frac{d\mathbf{p}}{(2\pi)^3}\, \hat{\bar{\psi}}(\mathbf{p}+\mathbf{k})\gamma^\mu \epsilon_{\mu\lambda}(\mathbf{k})\hat{\psi}(\mathbf{p}), \tag{104}$$

also with the formal definitions of Eqs. (101), but where the $\hat{b}$ and $\hat{d}$-operators in those formulas are now also the extended ones.

Before moving on to the $\hat{V}$ operator, let us recall that we want the fermions to at some point live up to their name and have antisymmetric wave functions (reducing $\mathbf{H}_F$ to that subspace). And in preparation of the Dirac sea reinterpretation, we also want to extend $\mathbf{H}_F$ to a Fock space. Let us therefore redefine $\mathbf{H}_F$ as a function of n such that

$$\mathbf{H}_F(n) = \{\psi \in L^2(\mathbf{R}^{3n}; \mathbb{C}^{4^n}) \,|\, \psi \text{ is antisymmetric}\} \tag{105}$$

for all $n \in \mathbb{N}$. Note that the formulas of the operators $\hat{H}_F$ and $(\boldsymbol{\alpha} \cdot \mathbf{e}_{\lambda\mathbf{k}})\hat{g}_{\mathbf{k}}$ remain unchanged by this antisymmetrization. We can now construct the fermion Fock space, call it $\mathbf{H}_{FF}$, by having

$$\mathbf{H}_{FF} = \mathbf{H}_F(0) \otimes \mathbf{H}_F(1) \otimes \mathbf{H}_F(2) \otimes \cdots = \bigotimes_{n=0}^{\infty} \mathbf{H}_F(n). \tag{106}$$

The fermionic operators are then extended by a similar infinite tensor product. This allows us to now finally extend the $\hat{b}_s$ and $\hat{d}_s$-operators as the well-known fermion creation and annihilation operators on this Fock space. According to e.g. Srednicki [7],[12] this is achieved by requiring the following anticommutation relations:

$$\begin{aligned}
&\{\hat{b}_{s'}(\mathbf{p}'), \hat{b}_s(\mathbf{p})\} = \{\hat{b}_{s'}(\mathbf{p}'), \hat{d}_s(\mathbf{p})\} = \{\hat{d}_{s'}(\mathbf{p}'), \hat{d}_s(\mathbf{p})\} = 0, \\
&\{\hat{b}^\dagger_{s'}(\mathbf{p}'), \hat{b}^\dagger_s(\mathbf{p})\} = \{\hat{b}^\dagger_{s'}(\mathbf{p}'), \hat{d}^\dagger_s(\mathbf{p})\} = \{\hat{d}^\dagger_{s'}(\mathbf{p}'), \hat{d}^\dagger_s(\mathbf{p})\} = 0, \\
&\{\hat{b}_{s'}(\mathbf{p}'), \hat{b}^\dagger_s(\mathbf{p})\} = \{\hat{b}_{s'}(\mathbf{p}'), \hat{d}^\dagger_s(\mathbf{p})\} = \{\hat{d}_{s'}(\mathbf{p}'), \hat{d}^\dagger_s(\mathbf{p})\} = (2\pi)^3 \delta_{s's} \delta^3(\mathbf{p}' - \mathbf{p})
\end{aligned} \tag{107}$$

for all $\mathbf{p}', \mathbf{p} \in \mathbb{R}^3$ and all $s', s \in \{1, 2, 3, 4\}$, where $\{\cdot, \cdot\}$ denotes the anticommutator for these relations. (One can of course disregard all cases where either s or s' does not fit the b or d symbol in these relations.)

So far in this section, we have let all the ladder operators be defined only on either $\mathbf{H}_B$ or $\mathbf{H}_F$ (or $\mathbf{H}_{FF}$). But they can also easily be extended to the full Hilbert space, letting e.g. $\hat{a}_{\lambda\mathbf{k}} \to \hat{a}_{\lambda\mathbf{k}} \otimes \hat{I}_F$ and $\hat{b}_s(\mathbf{p}) \to \hat{I}_B \otimes \hat{b}_s(\mathbf{p})$. These redefinitions will mean that we can turn all tensor products like the one in Eq. (92) into regular operator products, and thus write

$$\hat{H}_B \otimes \hat{I}_F = \sum_{\pm\mathbf{k} \in \mathbb{K}} \sum_{\lambda=1}^{2} k \hat{a}^\dagger_{\lambda\mathbf{k}} \hat{a}_{\lambda\mathbf{k}}, \tag{108}$$

$$\hat{I}_B \otimes \hat{H}_F = \int \frac{d\mathbf{p}}{(2\pi)^3} E_p \Big(\sum_{s=1}^{2} \hat{b}^\dagger_s(\mathbf{p}) \hat{b}_s(\mathbf{p}) - \sum_{s=3}^{4} \hat{d}^\dagger_s(\mathbf{p}) \hat{d}_s(\mathbf{p}) \Big), \tag{109}$$

$$\hat{H}_I = -q_F \sum_{\pm\mathbf{k} \in \mathbb{K}} \sum_{\lambda=1}^{2} \int \frac{d\mathbf{p}}{(2\pi)^3} \frac{1}{\sqrt{2\mathcal{V}k}} (\hat{a}_{\lambda\mathbf{k}} + \hat{a}^\dagger_{\lambda -\mathbf{k}}) \hat{\bar{\psi}}(\mathbf{p} + \mathbf{k}) \gamma^\mu \epsilon_{\mu\lambda}(\mathbf{k}) \hat{\psi}(\mathbf{p}). \tag{110}$$

Since we can recall that

$$\hat{H}_{red} = \hat{H}_B \otimes \hat{I}_F + \hat{I}_B \otimes \hat{H}_F + \hat{H}_I + \hat{V}, \tag{111}$$

we see that we have now rewritten all the operators of $\hat{H}_{red}$ in terms of ladder operators, except $\hat{V}$.

To analyze and rewrite $\hat{V}$, we first of all recall that it is the multiplication operator associated with the potential $V(\bar{\mathbf{x}})$ defined in Eq. (67), which we have not yet reduced in

[12] See Eq. (3.29) in particular.

any way. Reducing the initial expression for $V(\bar{\mathbf{x}})$, we get that

$$
\begin{aligned}
V(\bar{\mathbf{x}}) &= \frac{1}{2}\sum_{\mathbf{k},\sigma}\Big(\frac{q_F}{k}\sum_{j=1}^{n} f_{\mathbf{k}\sigma}(\mathbf{x}_j)\Big)^2 \\
&= \frac{q_F^2}{2}\sum_{\mathbf{k},\sigma}\frac{1}{k^2}\sum_{i=1}^{n}\sum_{j=1}^{n} f_{\mathbf{k}\sigma}(\mathbf{x}_i) f_{\mathbf{k}\sigma}(\mathbf{x}_j) \\
&= \frac{q_F^2}{\mathcal{V}}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{\mathbf{k}}\frac{1}{k^2}\big(\cos(\mathbf{k}\cdot\mathbf{x}_i)\cos(\mathbf{k}\cdot\mathbf{x}_j)+\sin(\mathbf{k}\cdot\mathbf{x}_i)\sin(\mathbf{k}\cdot\mathbf{x}_j)\big) \\
&= \frac{q_F^2}{\mathcal{V}}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{\mathbf{k}}\frac{1}{k^2}\cos\big(\mathbf{k}\cdot(\mathbf{x}_i-\mathbf{x}_j)\big).
\end{aligned}
\tag{112}
$$

Here we have used one of the addition formulas to get the last equality. We then see that all contributions to $V(\bar{\mathbf{x}})$ with $i = j$ are just constant energies (only depending on $\mathbb{K}$). We can therefore subtract these energies at every step going to the continuum limit, just as we intend to do for the ground state energies of the simple harmonic oscillators of $\hat{H}_B$. But when $i \neq j$, the energy turns out to depend on $\mathbf{x}_i - \mathbf{x}_j \equiv \mathbf{r}_{ij}$. Let us therefore first write

$$
V(\bar{\mathbf{x}}) = \sum_{1\le i<j\le n}\sum_{\mathbf{k}}\frac{2q_F^2}{\mathcal{V}k^2}\cos(\mathbf{k}\cdot\mathbf{r}_{ij}) + E_0, \tag{113}
$$

where $E_0 = q_F^2 n\mathcal{V}^{-1}\sum_{\mathbf{k}} 1/k^2$ is the constant energy. And using the same argument as we did for $\hat{H}_B$ above, let us then immediately remove the uninteresting constant energy E_0 from $V(\bar{\mathbf{x}})$, such that we get

$$
V(\bar{\mathbf{x}}) = \sum_{1\le i<j\le n}\sum_{\mathbf{k}}\frac{2q_F^2}{\mathcal{V}k^2}\cos(\mathbf{k}\cdot\mathbf{r}_{ij}) \tag{114}
$$

instead.

Let us now recall that $\mathcal{V} = (2\pi)^3/\delta k^3$, and also that $\mathbb{K} \subset \mathbb{R}^2 \times \mathbb{R}_+$. We then see that

$$
\sum_{\mathbf{k}}\frac{2q_F^2}{\mathcal{V}k^2}\cos(\mathbf{k}\cdot\mathbf{r}) = \sum_{\mathbf{k}}\frac{\delta k^3}{(2\pi)^3}\frac{2q_F^2}{k^2}\cos(\mathbf{k}\cdot\mathbf{r}) = \sum_{\pm\mathbf{k}\in\mathbb{K}}\frac{\delta k^3}{(2\pi)^3}\frac{q_F^2}{k^2}e^{-i\mathbf{k}\cdot\mathbf{r}}, \tag{115}
$$

and thus that this is just a discretized (inverse) Fourier transform of q_F^2/k^2. It is well-known that $1/k^2$ is the Fourier transform of the Coulomb potential, $V_C(\mathbf{r})/q_F^2 = 1/(4\pi r)$, in three dimensions. Let us confirm this fact by noting that

$$\begin{aligned}
\int_{k\leq A}\frac{d\mathbf{k}}{(2\pi)^3}\frac{e^{-i\mathbf{k}\cdot\mathbf{r}}}{k^2} &= \frac{1}{(2\pi)^3}\int_{k=0}^{A}\int_{\theta=0}^{\pi}\int_{\phi=0}^{2\pi}\frac{e^{-i\mathbf{k}\cdot\mathbf{r}}}{k^2}k^2\sin\theta\,dk\,d\theta\,d\phi \\
&= \frac{1}{4\pi^2}\int_{k=0}^{A}\int_{\theta=0}^{\pi}\sin\theta\,e^{-ikr\cos\theta}\,dk\,d\theta \\
&= \frac{1}{4\pi^2}\int_{0}^{A}\frac{e^{-ikr\cos\theta}}{ikr}\bigg|_{\theta=0}^{\pi}dk \\
&= \frac{1}{2\pi^2 r}\int_{0}^{A}\frac{\sin(kr)}{k}\,dk. \\
&= \frac{1}{4i\pi^2 r}\lim_{B\to 0^+}\left(\int_{-A}^{-B}dz+\int_{B}^{A}dz\right)\frac{e^{izr}}{z}.
\end{aligned} \tag{116}$$

When the cutoff A tends to infinity, the last expression can be evaluated by doing a contour integration in the complex plane and using Jordan's lemma.[13] But alternatively, one can simply look up the fact that $\int_0^\infty \sin(ax)/x\,dx = \pi/2$ for all $a>0$, which is also a well-known result. Either way, we get that

$$\lim_{A\to\infty}\int_{k<A}\frac{d\mathbf{k}}{(2\pi)^3}\frac{e^{-i\mathbf{k}\cdot\mathbf{r}}}{k^2} = \frac{1}{4\pi r} \tag{117}$$

for all $\mathbf{r}$ with $r\equiv|\mathbf{r}|>0$. We thus expect the continuum limit of $V(\bar{\mathbf{x}})$ to be given by

$$V(\bar{\mathbf{x}}) \to \sum_{1\leq i<j\leq n}\frac{q_F^2}{4\pi r_{ij}}, \quad r_{ij}\equiv|\mathbf{x}_i-\mathbf{x}_j|. \tag{118}$$

Remarkably, this suggests that the continuum limit of $V(\bar{\mathbf{x}})$ (with E_0 removed) is a Coulomb potential between each pair of fermions. This might at first seem contrary to the conventional theory of QED, where the only interaction vertex in the Feynman diagrams is that of the Dirac interaction. But as can be seen in Chapters 55–56 of Srednicki [7] and in Sect. 8.5 of Weinberg [1], the Coulomb interaction is still present in the conventional theory despite this. It is just absorbed in the Feynman propagator of the photons instead of getting its own vertex, namely because this allows said propagator to be written in a more covariant form. The results of this paper thus still match the conventional theory of QED.

To conclude this section, let us now go back to the discrete Fourier space and rewrite $\hat{V}$, with $E_0\hat{I}$ now subtracted from it, in terms of ladder operators. We see from Eqs. (114) and (115) that we have

$$\hat{V} = \sum_{1\leq i<j\leq n}\sum_{\pm\mathbf{k}\in\mathbb{K}}\frac{\delta k^3}{(2\pi)^3}\frac{q_F^2}{k^2}\hat{g}_{-\mathbf{k}ij} = \sum_{i,j=1}^{n}\sum_{\pm\mathbf{k}\in\mathbb{K}}\frac{\delta k^3}{(2\pi)^3}\frac{q_F^2}{2k^2}\hat{g}_{\mathbf{k}ij}, \tag{119}$$

where

$$\hat{g}_{\mathbf{k}ij}\psi(\bar{\mathbf{x}}) = e^{i\mathbf{k}\cdot\mathbf{x}_i}e^{-i\mathbf{k}\cdot\mathbf{x}_j}\psi(\bar{\mathbf{x}}) \tag{120}$$

[13] See e.g. Riley and Hobson [17].

for all $\bar{\mathbf{x}} \in \mathbb{R}^{3n}$, $\pm\mathbf{k} \in \mathbb{K}$, $\psi \in \mathbf{H}_F$ and $i, j \in \{1, \ldots, n\}$. We then note that for $n = 2$, we have

$$\begin{aligned}\langle s_1', \mathbf{p}_1', s_2', \mathbf{p}_2' | \sum_{i,j=1}^{n} \hat{g}_{\mathbf{k}ij} | s_1, \mathbf{p}_1, s_2, \mathbf{p}_2 \rangle \\ = \frac{(2\pi)^6}{(2E_p)^2} w_{s_1'}^\dagger(\mathbf{p}_1') w_{s_1}(\mathbf{p}_1) w_{s_2'}^\dagger(\mathbf{p}_2') w_{s_2}(\mathbf{p}_2) \delta^3(\mathbf{p}_1' - \mathbf{p}_1 - \mathbf{k}) \delta^3(\mathbf{p}_2' - \mathbf{p}_2 + \mathbf{k}).\end{aligned} \tag{121}$$

And thus, for $n = 2$, we get that

$$\sum_{i,j=1}^{n} \hat{g}_{\mathbf{k}ij} = \int \frac{d\mathbf{p}_1\, d\mathbf{p}_2}{(2\pi)^6} \big(\hat{\bar{\psi}}(\mathbf{p}_1 + \mathbf{k}) \otimes \hat{\bar{\psi}}(\mathbf{p}_2 - \mathbf{k})\big)(\gamma^0 \otimes \gamma^0)\big(\hat{\psi}(\mathbf{p}_1) \otimes \hat{\psi}(\mathbf{p}_2)\big), \tag{122}$$

where $\otimes$ in this context denotes the matrix direct product (exclusively working on the spinors), such that e.g. $(\bar{w}_{s_1'} \otimes \bar{w}_{s_2'})(\gamma^0 \otimes \gamma^0)(w_{s_1} \otimes w_{s_2}) = w_{s_1'}^\dagger w_{s_1} w_{s_2'}^\dagger w_{s_2}$ (when suppressing the momentum inputs).

It is not hard to see that this formula will also apply for all other n, and we can thus finally write

$$\hat{V} = \sum_{\pm\mathbf{k}\in\mathbb{K}} \frac{\delta k^3}{(2\pi)^3} \frac{q_F^2}{2k^2} \int \frac{d\mathbf{p}_1\, d\mathbf{p}_2}{(2\pi)^6} \big(\hat{\bar{\psi}}(\mathbf{p}_1 + \mathbf{k}) \otimes \hat{\bar{\psi}}(\mathbf{p}_2 - \mathbf{k})\big)(\gamma^0 \otimes \gamma^0)\big(\hat{\psi}(\mathbf{p}_1) \otimes \hat{\psi}(\mathbf{p}_2)\big). \tag{123}$$

We now have all the operators of $\hat{H}_{red}$ expressed in terms of ladder operators, as was the goal for this section. And while it is not conventional to see this $\hat{V}$ in the literature on QED, namely since there exist a more covariant formulation of said theory once the Feynman diagrams are introduced, it is still implicitly present in that theory. In the following section, we will then discuss how to turn all the $\sum_{\pm\mathbf{k}\in\mathbb{K}}$-summations into integrals.

8 The continuum limit

We are now ready to investigate the continuum limit of the theory. In general, a 'continuum limit' refers to a case where a sum over a discrete range is turned into an integral by making the discretization finer and finer. For an operator such as $\hat{H}_{red}$, we thus want to let $\delta k \to 0$ and $N_k \delta k \to \infty$ for the summation range $\mathbb{K}$. If a continuum limit exists for this operator, it thus means that there exists an operator, call it $\hat{H}_{CL}$ in our case, on a Hilbert space, call it $\mathbf{H}_{CL}$, which can be modeled with arbitrary precision by $\hat{H}_{red}$ on $\mathbf{H}_{red}$ when $\delta k \to 0$ and $N_k \delta k \to \infty$. To put this more precisely, we want there to be sequences $(\mathbf{H}_{red}(j))_{j\in\mathbb{N}}$ and $(\hat{H}_{red}(j))_{j\in\mathbb{N}}$ of Hilbert spaces and operators, with the same qualities as we have defined above, but where $\delta k \to 0$ and $N_k \delta k \to \infty$ for $j \to \infty$. We then also want there to be a sequence of functions $(\mathcal{P}_j)_{j\in\mathbb{N}}$, where $\mathcal{P}_j : \mathbf{H}_{CL} \to \mathbf{H}_{red}(j)$ for all $j \in \mathbb{N}$, such that

$$\lim_{j\to\infty} \langle \mathcal{P}_j(\Psi') | e^{-i\hat{H}_{red}(j)t} | \mathcal{P}_j(\Psi) \rangle = \langle \Psi' | e^{-i\hat{H}_{CL}t} | \Psi \rangle \tag{124}$$

for all $\Psi, \Psi' \in \mathbf{H}_{CL}$.

Unfortunately, showing that Eq. (124) holds, or even finding a fitting domain for $\hat{H}_{CL}$, is beyond the scope of this paper, and is thus left for future work. Instead we will simply try to find a $\hat{H}_{CL}$ and a $(\mathcal{P}_j)_{j\in\mathbb{N}}$ which seem to fulfill the property that

$$\lim_{j\to\infty} \langle \mathcal{P}_j(\Psi')|\hat{H}_{red}(j)|\mathcal{P}_j(\Psi)\rangle = \langle \Psi'|\hat{H}_{CL}|\Psi\rangle\,, \tag{125}$$

simply because this seems like a reasonable property to assume for $\hat{H}_{CL}$. We will also simply disregard the fact that we do not yet know $\mathrm{Dom}(\hat{H}_{CL})$ for the following discussion.

With this in mind, we will now look for such a continuum limit of $\hat{H}_{red}$. We want to look for a $\mathbf{H}_{CL}$ where the photons live in a continuous Fock space, and where the ladder operators $\hat{a}^\dagger_{\lambda\mathbf{k}}$ and $\hat{a}_{\lambda\mathbf{k}}$ are thus turned into continuous creation and annihilation operators $\hat{a}^\dagger_\lambda(\mathbf{k})$ and $\hat{a}_\lambda(\mathbf{k})$. We thus want $\mathbf{H}_B$ to turn into a Fock space, call it $\mathbf{H}_{BF}$, given by

$$\mathbf{H}_{BF} = \bigotimes_{n=0}^{\infty} L^2(\mathbb{R}^{3n};\mathbb{C}^{2^n}). \tag{126}$$

The creation and annihilation operators on this Fock space is well-known in literature. We will therefore not detail their exact definitions for all states in $\mathbf{H}_{BF}$ here, but simply note the following basic properties. We first of all note that their commutation relations are given by

$$\begin{aligned}
&[\hat{a}_{\lambda'}(\mathbf{k}'), \hat{a}_\lambda(\mathbf{k})] = 0,\\
&[\hat{a}^\dagger_{\lambda'}(\mathbf{k}'), \hat{a}^\dagger_\lambda(\mathbf{k})] = 0,\\
&[\hat{a}_{\lambda'}(\mathbf{k}'), \hat{a}^\dagger_\lambda(\mathbf{k})] = (2\pi)^3\delta_{\lambda'\lambda}\delta^3(\mathbf{p}' - \mathbf{p})
\end{aligned} \tag{127}$$

for all $\mathbf{k}', \mathbf{k} \in \mathbb{R}^3$ and all $\lambda', \lambda \in \{1, 2\}$. And we also note, as an example, that for the vacuum state $|\,\rangle$, we have

$$\int \frac{d\mathbf{k}}{(2\pi)^3}\, \phi(\mathbf{k})\hat{a}^\dagger_\lambda(\mathbf{k})\,|\,\rangle = \int \frac{d\mathbf{k}}{(2\pi)^3}\, \phi(\mathbf{k})\,|\lambda, \mathbf{k}\rangle \tag{128}$$

for any $\phi \in L^2(\mathbb{R}^3)$. The norm of this state is given, not by the standard L^2 norm, $\|\phi\|$, but by

$$\Big(\int \frac{d\mathbf{k}\, d\mathbf{k}'}{(2\pi)^6}\, \phi^*(\mathbf{k}')\,\langle\lambda, \mathbf{k}'|\lambda, \mathbf{k}\rangle\, \phi(\mathbf{k})\Big)^{1/2} = \Big(\int \frac{d\mathbf{k}}{(2\pi)^3}\, \phi^*(\mathbf{k})\phi(\mathbf{k})\Big)^{1/2} = \frac{\|\phi\|}{(2\pi)^{3/2}}. \tag{129}$$

This is important if we want to get the right factors for the continuum limit.

To analyze what we need for the continuum limit, let us then start by considering operators of the form $\sum_{\pm\mathbf{k}\in\mathbb{K}} \sqrt{\delta k^3}\phi(\mathbf{k})\hat{a}^\dagger_{\lambda\mathbf{k}}$ on $\mathbf{H}_{red}$, where $\phi \in L^2(\mathbb{R}^3)$. Such an operator can be seen to yield a state with norm approximately equal to $\|\phi\|$ when working on the "vacuum state," which for the discretized system is when all oscillators are in the ground state. From Eqs. (128) and (129), we therefore see that we want to have

$$\sum_{\pm\mathbf{k}\in\mathbb{K}} \frac{\delta k^{3/2}}{(2\pi)^{3/2}}\phi(\mathbf{k})\hat{a}^\dagger_{\lambda\mathbf{k}} = \sum_{\pm\mathbf{k}\in\mathbb{K}} \frac{1}{\sqrt{\mathcal{V}}}\phi(\mathbf{k})\hat{a}^\dagger_{\lambda\mathbf{k}} \to \int \frac{d\mathbf{k}}{(2\pi)^3}\, \phi(\mathbf{k})\hat{a}^\dagger_\lambda(\mathbf{k}) \tag{130}$$

in the continuum limit for any $\phi \in L^2(\mathbb{R}^3)$, where we have also used the fact that $\delta k^3/(2\pi)^3 = \mathcal{V}^{-1}$ in order to reduce the left-hand side. And for any ϕ that is not in $L^2(\mathbb{R}^3)$, such as a constant ϕ, e.g., we should of course want the same limit as in Eq. (130) for our operators on $\hat{H}_{red}$.

For operators of the form $\sum_{\pm\mathbf{k}\in\mathbb{K}}(\delta k/2\pi)^{3/2}\phi(\mathbf{k})\hat{a}_{\lambda\mathbf{k}}$, let us consider the fact that we have

$$\langle\,|\sum_{\pm\mathbf{k}\in\mathbb{K}}\Big(\frac{\delta k^{3/2}}{(2\pi)^{3/2}}\phi^*(\mathbf{k})\hat{a}_{\lambda\mathbf{k}}\Big)\sum_{\pm\mathbf{k}\in\mathbb{K}}\Big(\frac{\delta k^{3/2}}{(2\pi)^{3/2}}\phi(\mathbf{k})\hat{a}^\dagger_{\lambda\mathbf{k}}\Big)|\,\rangle=\frac{\delta k^3}{(2\pi)^3}\sum_{\pm\mathbf{k}\in\mathbb{K}}|\phi(\mathbf{k})|^2, \tag{131}$$

which tends toward $\|\phi\|^2/(2\pi)^3$ as we approach the continuum limit (at least when $|\phi|^2$ is a nice, Riemann-integrable function). Comparing this with

$$\int\frac{d\mathbf{k}\,d\mathbf{k}'}{(2\pi)^6}\,\phi^*(\mathbf{k}')\phi(\mathbf{k})\,\langle\,|\hat{a}_\lambda(\mathbf{k}')\hat{a}^\dagger_\lambda(\mathbf{k})|\,\rangle=\int\frac{d\mathbf{k}}{(2\pi)^3}\phi^*(\mathbf{k})\phi(\mathbf{k})=\frac{\|\phi\|}{(2\pi)^3}, \tag{132}$$

where we have used that $\hat{a}_\lambda(\mathbf{k}')\hat{a}^\dagger_\lambda(\mathbf{k})=(2\pi)^3\delta^3(\mathbf{k}'-\mathbf{k})-\hat{a}^\dagger_\lambda(\mathbf{k}')\hat{a}_\lambda(\mathbf{k})$, we can thus see that we similarly want to have

$$\sum_{\pm\mathbf{k}\in\mathbb{K}}\frac{\delta k^{3/2}}{(2\pi)^{3/2}}\phi(\mathbf{k})\hat{a}_{\lambda\mathbf{k}}=\sum_{\pm\mathbf{k}\in\mathbb{K}}\frac{1}{\sqrt{\mathcal{V}}}\phi(\mathbf{k})\hat{a}_{\lambda\mathbf{k}}\to\int\frac{d\mathbf{k}}{(2\pi)^3}\,\phi(\mathbf{k})\hat{a}_\lambda(\mathbf{k}) \tag{133}$$

in the continuum limit for any $\phi\in L^2(\mathbb{R}^3)$. And this result again also extents to all $\phi\notin L^2(\mathbb{R}^3)$. So if we consider $\hat{H}_I$ as written in Eq. (110), we therefore see that we want to have

$$\hat{H}_{ICL}=-q_F\sum_{\lambda=1}^{2}\int\frac{d\mathbf{k}\,d\mathbf{p}}{(2\pi)^6}\,\frac{1}{\sqrt{2k}}\big(\hat{a}_\lambda(\mathbf{k})+\hat{a}^\dagger_\lambda(-\mathbf{k})\big)\hat{\bar{\psi}}(\mathbf{p}+\mathbf{k})\gamma^\mu\epsilon_{\mu\lambda}(\mathbf{k})\hat{\psi}(\mathbf{p}) \tag{134}$$

for its counterpart on $\mathbf{H}_{CL}$. This is of course given that we also extend $\hat{a}_\lambda(\mathbf{k})$ and $\hat{a}^\dagger_\lambda(\mathbf{k})$ to the full Hilbert space in order to replace the tensor product we would otherwise get with a regular operator product.

Moving on to operators of the form $\sum_{\pm\mathbf{k}\in\mathbb{K}}h(\mathbf{k})\hat{a}^\dagger_{\lambda\mathbf{k}}\hat{a}_{\lambda\mathbf{k}}$ for some function h, we naturally want to have

$$\sum_{\pm\mathbf{k}\in\mathbb{K}}h(\mathbf{k})\hat{a}^\dagger_{\lambda\mathbf{k}}\hat{a}_{\lambda\mathbf{k}}\to\int\frac{d\mathbf{k}}{(2\pi)^3}\,h(\mathbf{k})\hat{a}^\dagger_\lambda(\mathbf{k})\hat{a}_\lambda(\mathbf{k}) \tag{135}$$

in the continuum limit. We can see that this fits the other results by noting that

$$\int\frac{d\mathbf{k}\,d\mathbf{k}'}{(2\pi)^6}\,h(\mathbf{k}')\hat{a}^\dagger_\lambda(\mathbf{k}')\hat{a}_\lambda(\mathbf{k}')\phi(\mathbf{k})\hat{a}^\dagger_\lambda(\mathbf{k})\,|\,\rangle=\int\frac{d\mathbf{k}}{(2\pi)^3}\,h(\mathbf{k})\phi(\mathbf{k})\hat{a}^\dagger_\lambda(\mathbf{k})\,|\,\rangle\,. \tag{136}$$

A state created from the vacuum with the operator of Eq. (130) will thus turn into itself multiplied by h in Fourier space when operated on by the right-hand side of Eq. (135), which indeed corresponds to the action of the operator on the left-hand side of Eq. (135) in the discrete Fourier space. We thus want the continuum limit of $\hat{H}_B$ to be given by

$$\hat{H}_{BCL}=\sum_{\lambda=1}^{2}\int\frac{d\mathbf{k}}{(2\pi)^3}\,k\hat{a}^\dagger_\lambda(\mathbf{k})\hat{a}_\lambda(\mathbf{k}). \tag{137}$$

For the operator $\hat{V}$ in Fourier space, we only need to note that for the continuum limit, we should have $(2\pi)^{-3}\sum_{\pm\mathbf{k}\in\mathbb{K}}\delta k^3\to(2\pi)^{-3}\int d\mathbf{k}$. We thus want Eq. (123) to turn into

$$\hat{V}_{CL}=\int\frac{d\mathbf{k}}{(2\pi)^3}\frac{d\mathbf{p}_1\,d\mathbf{p}_2}{(2\pi)^6}\,\frac{q_F^2}{2k^2}\big(\hat{\bar{\psi}}(\mathbf{p}_1+\mathbf{k})\otimes\hat{\bar{\psi}}(\mathbf{p}_2-\mathbf{k})\big)\big(\gamma^0\otimes\gamma^0\big)\big(\hat{\psi}(\mathbf{p}_1)\otimes\hat{\psi}(\mathbf{p}_2)\big). \tag{138}$$

This was the last piece of $\hat{H}_{CL}$, and we can now write

$$\hat{H}_{CL} = \hat{H}_{BCL} + \hat{H}_F + \hat{H}_{ICL} + \hat{V}_{CL} \tag{139}$$

for our expected continuum limit of $\hat{H}_{red}$.

The function $\mathcal{P}_j$ for each j can then be defined as the function that resolves and approximates any $\Psi \in \mathbf{H}_{CL}$ by step functions in a discretized Fourier space, essentially by writing Ψ in terms of creation operators and then rewriting this expression according to Eq. (130), only in reverse. Because of how the commutation relations of the discrete ladder operators correspond to those of the continuous ones, it is not hard to see that Eq. (125) will be fulfilled for all pairs of Ψ, Ψ'.

9 The Dirac sea reinterpretation

Before we conclude this paper, we should also briefly discuss the last step needed in order to finally obtain a theory of QED. This step is of course the (well-known) Dirac sea reinterpretation, where we want to reinterpret all the negative-energy solutions as "holes" in a so-called Dirac sea, yielding us the antiparticles of the theory. More specifically, this means that we want to replace all instances of $\hat{d}_3(\mathbf{p})$ and $\hat{d}_4(\mathbf{p})$ in the formulas for $\hat{H}_{CL}$ with $\hat{d}_3^\dagger(-\mathbf{p})$ and $\hat{d}_4^\dagger(-\mathbf{p})$, respectively, and vice versa.

We might hope that this would yield us a functioning Hamiltonian for QED right away, but it turns out that we get some divergences from the process, some of which are particularly troublesome. These particular divergences come from the fact that the resulting Hamiltonian has the potential to perturb the empty vacuum, causing infinitely many vacuum fluctuations at all times when we approach the continuum limit (in which $\mathcal{V}$ also tends to infinity). These fluctuations make the problem quite hard to solve, and we will thus leave it for future work. See Appendix D for further details on how these divergences appear and why they are troublesome.

We will not deal with this problem further in this paper, but simply leave it for future work.

10 Conclusion

We have derived a Hamiltonian for a proto-theory of QED, where the negative-energy solutions are not yet reinterpreted as antiparticles. To do this, we used an approach for eliminating the gauge symmetry by separating the variables of the initially derived Hamiltonian. The result is a Hamiltonian on a reduced Hilbert space, which in the continuum limit of the theory seems to be exactly equal to the conventional Hamiltonian of QED, namely when accounting for the fact that the Coulomb interaction is simply absorbed into the photon propagator in that formulation of the theory.

We argued why this approach should not break the initial Lorentz covariance, and thus that the theory ought to be Lorentz-covariant given that a fitting domain can be found on which the Hamiltonian is self-adjoint.

11 Future work

11.1 Foreword about what has changed from the first version of this paper

In the first version of this paper, the author (I/me/my) did not manage to realize that the Coulomb interaction was actually already present in the conventional theory. I was therefore very excited about seeing the approach of this paper through to the end, and used this section to summarize what was needed in order to finish the work, and also to ask for help with it. But now it seems that this approach is only slightly different from the conventional one, and that it leads to exactly the same result.

But *maybe* there is a chance that this work can do more than just provide a fresh way of deriving the same theory. It is my general understanding (but do not at all take my word for it) that conventional QFT is not the most mathematically well-defined field in physics, and if this is true, *perhaps* this work could open up for a way to develop more mathematically well-defined quantum field theories.

So even though I am no longer sure that these tasks are worth looking into, let me still end this paper by listing the future tasks that need completion in order for the approach of this paper to be carried through to the end.

11.2 Tasks to complete

In order to show that $\hat{H}_{CL}$ is a valid continuum limit of $\hat{H}_{red}$, one first of all has to show that $\hat{H}_{CL}$ **is self-adjoint** on some domain. If we look at the interaction terms, in particular the Dirac interaction, we might be worried that the ultraviolet divergence makes it hard to even find vectors that turn into normalizable vectors when the formula for $\hat{H}_{CL}$ is applied. It might, however, be possible to find a set of such vectors, namely by constructing these such that all infinite "tails" produced by photon emissions are canceled by some product of photon absorption by other parts of the state. But one can then see that $\hat{H}_{CL}$ will probably not be symmetric on its most inclusive domain possible. So to prove that $\hat{H}_{CL}$ is self-adjoint on some domain, one must therefore find a domain on which $\hat{H}_{CL}$ is symmetric and make sure that this domain is just inclusive enough to make the operator self-adjoint.

If one can find domains on which $\hat{H}_{CL}$ is self-adjoint, the next task will then be to show that **Eq. (124) can be fulfilled** for $\hat{H}_{CL}$ when it is defined on such a domain. If this task can be completed, one will thus get that the propagator for $\hat{H}_{CL}$ can be approximated to arbitrary precision by the path integral of Sect. 4. This will then suggest that $\hat{H}_{CL}$ is Lorentz-covariant.

To actually *prove* that $\hat{H}_{CL}$ **is Lorentz-covariant** will require a lot more work, however. Here, one can hope to succeed with an analytical approach to this where one finds a similar result as the Feynman–Kac formula (see e.g. Hall [4], Chap. 20) for the theory. But it might be much easier in our case to take a more brute-force approach where one keeps the path integrals discretized. In the latter approach, one might then first of all Fourier-transform the path integral along the time axis and thus parameterize it in terms of $(\omega, \mathbf{k})$-components. Then one might try to show that the propagator can be approximated arbitrarily well with a finite set of Fourier components, fixed even as $\delta t, \delta x \to 0$. In this limit, the A^μ fields of the paths become smooth, namely *if* we can exclude the high-frequency Fourier components thus. Let $\widetilde{X}_\kappa$ denote a sequence of such finite (but growing) sets of Fourier components for which the path integral becomes more and more precise when $\kappa \to \infty$. If one can then find a $\widetilde{X}'_\kappa$ with the same property for the second inertial frame where each $(\omega, \mathbf{k}) \in \widetilde{X}_\kappa$ Lorentz-transforms to a $(\omega', \mathbf{k}') \in \widetilde{X}'_\kappa$ and vise versa, it is then not hard to show from there that the theory will be Lorentz-covariant.

With these three tasks completed, we will get that $\hat{H}_{CL}$ defines a Lorentz-covariant quantum theory, at least with our somewhat relaxed definition of Lorentz covariance from Sect. 6. There are, however, also a rather simple argument which shows that pure states will **Lorentz-transform into other pure states**. This argument uses the gauge symmetry of the Dirac equation and the remaining gauge symmetry of $\mathcal{L}_{EM}$ as defined in Eq. (14), which is the freedom to make a gauge transformation with a scalar λ that fulfills $\Box^2\lambda = 0$, where $\Box^2 = \nabla^2 - \partial^2/\partial t^2$. This gauge symmetry for the paths of the path integral then allows one to show that $\chi = \Phi \otimes \Psi$-states with Φ approximately equal to a delta function around 0 in Fourier space will Lorentz-transform into similar states, thus giving us a process for Lorentz-transforming any pure state Ψ into another pure state Ψ'.

There are of course one more (big) task left from there, since $\hat{H}_{CL}$ still needs to undergo a **Dirac sea reinterpretation** if we want to find a Hamiltonian that describes QED fully. See Appendix D for more on this topic. As mentioned in that appendix, this task might require a solution to the perturbed vacuum in order for one to be able to "renormalize" the reinterpreted Hamiltonian appropriately. Note that the exact process of this "renormalization" will probably only be clear once the solution to the vacuum is found, and we can thus only hope that this second step of the task is possible. And as a third step, one might also have to show that the solution to the perturbed vacuum Lorentz-transforms into itself (or a physically equivalent state) in order to complete the argument for the Lorentz covariance of that theory.

Appendices

A The Legendre transform of Lagrangians of the class to which L_{EM} belongs

We want to show that any Lagrangian of the form

$$L(\mathbf{q},\dot{\mathbf{q}}) = \frac{1}{2}\big(\dot{\mathbf{q}} - \mathbf{W}(\mathbf{q})\big)^T M \big(\dot{\mathbf{q}} - \mathbf{W}(\mathbf{q})\big) - V(\mathbf{q}), \tag{140}$$

where M is a real, symmetric and positive-definite matrix, corresponds to a classical Hamiltonian given by

$$H(\mathbf{p},\mathbf{q}) = \frac{1}{2}\mathbf{p}^T M^{-1}\mathbf{p} + \mathbf{W}(\mathbf{q})\cdot\mathbf{p} + V(\mathbf{q}). \tag{141}$$

The classical Hamiltonian $H(\mathbf{p},\mathbf{q})$ is obtained from the Lagrangian by making a Legendre transformation of $L(\mathbf{q},\dot{\mathbf{q}})$ with respect to $\dot{\mathbf{q}}$. According to e.g. Durhuus and Solovej [12], a Legendre transformation of any $L(\mathbf{q},\dot{\mathbf{q}})$ that is strictly convex with respect to $\dot{\mathbf{q}}$ can be carried out by solving

$$\mathbf{p} = \nabla_{\dot{\mathbf{q}}} L(\mathbf{q},\dot{\mathbf{q}}) \tag{142}$$

for $\dot{\mathbf{q}}$ and substituting the result in

$$H(\mathbf{p},\mathbf{q}) = \mathbf{p}^T\dot{\mathbf{q}} - L(\mathbf{q},\dot{\mathbf{q}}). \tag{143}$$

First we see that since M is symmetric, we have that

$$\begin{aligned}\frac{\partial}{\partial \dot{q}_i} L(\mathbf{q},\dot{\mathbf{q}}) &= \frac{1}{2}\mathbf{e}_i^T M\big(\dot{\mathbf{q}} - \mathbf{W}(\mathbf{q})\big) + \frac{1}{2}\big(\dot{\mathbf{q}} - \mathbf{W}(\mathbf{q})\big)^T M\mathbf{e}_i \\ &= \mathbf{e}_i^T M\big(\dot{\mathbf{q}} - \mathbf{W}(\mathbf{q})\big),\end{aligned} \tag{144}$$

where $\dot{q}_i$ denotes the ith entry of $\dot{\mathbf{q}}$, and where $\{\mathbf{e}_i\}$ is the standard basis. Equation (142) thus becomes

$$\mathbf{p} = M\big(\dot{\mathbf{q}} - \mathbf{W}(\mathbf{q})\big). \tag{145}$$

And since M is positive-definite, it has an inverse, which we can multiply on both sides of the equation to get

$$\dot{\mathbf{q}} = M^{-1}\mathbf{p} + \mathbf{W}(\mathbf{q}). \tag{146}$$

We now have to substitute this in Eq. (143). This gives us

$$\begin{aligned}H(\mathbf{p},\mathbf{q}) &= \mathbf{p}^T\big(M^{-1}\mathbf{p} + \mathbf{W}(\mathbf{q})\big) - L\big(\mathbf{q}, M^{-1}\mathbf{p} + \mathbf{W}(\mathbf{q})\big) \\ &= \mathbf{p}^T M^{-1}\mathbf{p} + \mathbf{p}^T\mathbf{W}(\mathbf{q}) - \frac{1}{2}\big(M^{-1}\mathbf{p}\big)^T M(M^{-1}\mathbf{p}) + V(\mathbf{q}) \\ &= \mathbf{p}^T M^{-1}\mathbf{p} + \mathbf{p}^T\mathbf{W}(\mathbf{q}) - \frac{1}{2}\mathbf{p}^T(M^{-1})^T\mathbf{p} + V(\mathbf{q}) \\ &= \frac{1}{2}\mathbf{p}^T M^{-1}\mathbf{p} + \mathbf{p}^T\mathbf{W}(\mathbf{q}) + V(\mathbf{q}),\end{aligned} \tag{147}$$

where we have used the fact that M^{-1} is also symmetric to get the last equality. This is what we wanted to show.

B Self-adjointness of the initial Hamiltonian

In this appendix, we want to argue that $\hat{H}_{init}$ from Eq. (27) is self-adjoint on a (dense) subspace of $L^2(\mathbb{R}^N\times\mathbb{R}^{3n};\mathbb{C}^{4^n})$. At the end of this appendix, we will also show that $C_c^\infty(\mathbb{R}^N\times\mathbb{R}^{3n};\mathbb{C}^{4^n})$ is a subspace of $\mathrm{Dom}(\hat{H}_{init})\cap\mathrm{Dom}(\sum_{\mathbf{k},\sigma}\sum_{\mu=0}^{3}\hat{\Pi}^2_{\mu\mathbf{k}\sigma})$, and that

$$\lim_{\kappa\to\infty}\hat{H}_{init}(\kappa)\chi = \hat{H}_{init}\chi \tag{148}$$

for all $\chi\in C_c^\infty(\mathbb{R}^N\times\mathbb{R}^{3n};\mathbb{C}^{4^n})$, since we use this at the end of Sect. 4.

Before we can do this, let us note the following definitions and basic properties, which can all be found in Hall [4].

The **adjoint** of an operator $\hat{A}$ is defined to be the unique operator $\hat{A}^*$[14] for which $\langle\phi|\hat{A}\psi\rangle = \langle\hat{A}^*\phi|\psi\rangle$ for all $\psi\in\mathrm{Dom}(\hat{A})\subset\mathbf{H}$ and $\phi\in\mathrm{Dom}(\hat{A}^*)$, where (importantly) $\mathrm{Dom}(\hat{A}^*)$ is defined to be the space of all $\phi\in\mathbf{H}$ for which $\psi\mapsto\langle\phi|\hat{A}\psi\rangle$ is a bounded functional on $\mathrm{Dom}(\hat{A})$.

An operator $\hat{A}$ is said to be **symmetric** if $\langle\phi|\hat{A}\psi\rangle = \langle\hat{A}\phi|\psi\rangle$ for all $\psi,\phi\in\mathrm{Dom}(\hat{A})$, and it is said to be **self-adjoint** if $\mathrm{Dom}(\hat{A}^*) = \mathrm{Dom}(\hat{A})$ and $\hat{A}^*\psi = \hat{A}\psi$ for all $\psi\in\mathrm{Dom}(\hat{A})$. (A self-adjoint operator is always symmetric, but a symmetric operator $\hat{A}$ is not always self-adjoint since $\mathrm{Dom}(\hat{A})$ might be a proper subset of $\mathrm{Dom}(\hat{A}^*)$.)

The **graph** of an operator $\hat{A}$ is the set given by $\{(\psi,\phi)\in\mathrm{Dom}(\hat{A})\times\mathbf{H}\,|\,\hat{A}\psi=\phi\}$. If the closure of this graph is still a graph of a function (i.e. if it keeps having only one ϕ for each ψ), then $\hat{A}^{cl}$ denotes the operator corresponding to this function and $\hat{A}$ is said to be **closable** with $\hat{A}^{cl}$ as its **closure**. Note that the domain of $\hat{A}^{cl}$ is then the set of all $\psi\in\mathbf{H}$ for which there exist a Cauchy sequence ψ_j belonging to $\mathrm{Dom}(\hat{A})$ that converges to ψ such that the sequence $\hat{A}\psi_j$ also converges to a vector $\phi\in\mathbf{H}$ (which is unique for each such ψ if and only if $\hat{A}$ is closable).

It can be proven that a symmetric operator $\hat{A}$ is always closable, and that the domain of its closure $\mathrm{Dom}(\hat{A}^{cl})$ it a subset of $\mathrm{Dom}(\hat{A}^*)$. And a symmetric operator $\hat{A}$ is then said to be **essentially self-adjoint** if its closure $\hat{A}^{cl}$ is self-adjoint, which means that $\mathrm{Dom}(\hat{A}^{cl}) = \mathrm{Dom}(\hat{A}^*)$ since the adjoint of a symmetric operator always coincides with the adjoint of its closure. We can also note that the graph of $\hat{A}^*$ is always closed, which is true for any operator, not just for symmetric ones.

Before we move on, let us also quickly note the **Kato–Rellich theorem**, which can also be found in Hall [4], since we refer to this in Sect. 4. It states that if $\hat{A}$ and $\hat{B}$ are self-adjoint operators with $\mathrm{Dom}(\hat{A})\subset\mathrm{Dom}(\hat{B})$, and if there exist positive constants a and b with $a<1$ such that $\|\hat{B}\psi\|\leq a\|\hat{A}\psi\|+b\|\psi\|$ for all $\psi\in\mathrm{Dom}(\hat{A})$, then $A+B$ is a self-adjoint operator on $\mathrm{Dom}(\hat{A})$, and it is essentially self-adjoint on any subspace of $\mathrm{Dom}(\hat{A})$ on which $\hat{A}$ is essentially self-adjoint. The theorem also gives a lower bound on the spectrum of $\hat{A}+\hat{B}$ (which for a self-adjoint operator is the set of generalized eigenvalues, so to speak) if $\hat{A}$ is non-negative, but we do not need this part for our purposes. (See Hall [4] for the rest of this theorem, and for proofs.) A special case of this theorem is when $\hat{B}$ is bounded, for which the conditions of the theorem are always fulfilled.

We can now go on to discuss why $\hat{H}_{init}$ is self-adjoint on a subspace of $L^2(\mathbb{R}^N\times\mathbb{R}^{3n};\mathbb{C}^{4^n})$.

[14] It is also common to use daggers, †, instead in physics literature (partly because it is also common to have a^* denote the complex conjugate of a number instead of $\overline{a}$). But for this appendix, we will use the notation that is more common in mathematical physics literature.

In Sect. 5, we show that a certain change of basis turns the operator into

$$\hat{H}_{init} \to \hat{H}_{tran} = \hat{H}_{B03} + \hat{H}_{red}, \tag{149}$$

with definitions of $\hat{H}_{B03}$ and $\hat{H}_{red}$ given in Eqs. (69–72). In this form, the operator is completely separated, in the sense that we can reinterpret it as

$$\hat{H}_{tran} = \hat{H}_{B03} \otimes \hat{I}_{red} + \hat{I}_{B03} + \hat{H}_{red}, \tag{150}$$

where $\hat{H}_{B03}$ and $\hat{H}_{red}$ are now reinterpreted as operators on $L^2(\mathbb{R}^{N/2})$ and on $L^2(\mathbb{R}^{N/2} \times \mathbb{R}^{3n}; \mathbb{C}^{4^n})$, respectively. If $\hat{H}_{B03}$ and $\hat{H}_{red}$ can then be shown to be self-adjoint on some domains, $\mathrm{Dom}(\hat{H}_{B03})$ and $\mathrm{Dom}(\hat{H}_{red})$, it follows that $\hat{H}_{tran}$ will be self-adjoint on their tensor product, $\mathrm{Dom}(\hat{H}_{B03}) \otimes \mathrm{Dom}(\hat{H}_{red})$.[15] And $\hat{H}_{init}$ will then be self-adjoint on the domain obtained from the reverse change of basis for this subspace.

We will start by arguing why $\hat{H}_{B03}$ is self-adjoint on some subspace of $L^2(\mathbb{R}^{N/2})$, and move on to $\hat{H}_{red}$ afterwards. Recall that $\hat{H}_{B03}$ is given by

$$\hat{H}_{B03} = \sum_{\mathbf{k},\sigma} \big(\frac{\xi}{2} \hat{\Pi}^2_{0\mathbf{k}\sigma} + \frac{1}{2} \hat{\Pi}^2_{3\mathbf{k}\sigma} - \sigma k \hat{A}_{3\mathbf{k}-\sigma} \hat{\Pi}_{0\mathbf{k}\sigma} - \sigma k \hat{A}_{0\mathbf{k}-\sigma} \hat{\Pi}_{3\mathbf{k}\sigma} \big), \tag{151}$$

i.e. when we limit it to $L^2(\mathbb{R}^{N/2})$. Let us now flip the σ's in the first and in the third term, and then rename all these operators by letting

$$\hat{p}_{x\mathbf{k}\sigma} = \hat{\Pi}_{0\mathbf{k}-\sigma}, \quad \hat{p}_{y\mathbf{k}\sigma} = \hat{\Pi}_{3\mathbf{k}\sigma}, \quad \hat{x}_{\mathbf{k}\sigma} = \hat{A}_{0\mathbf{k}-\sigma}, \quad \hat{y}_{\mathbf{k}\sigma} = \hat{A}_{3\mathbf{k}\sigma}. \tag{152}$$

This gives us

$$\hat{H}_{B03} = \sum_{\mathbf{k},\sigma} \big(\frac{\xi}{2} \hat{p}^2_{x\mathbf{k}\sigma} + \frac{1}{2} \hat{p}^2_{y\mathbf{k}\sigma} + \sigma k \hat{y}_{\mathbf{k}\sigma} \hat{p}_{x\mathbf{k}\sigma} - \sigma k \hat{x}_{\mathbf{k}\sigma} \hat{p}_{y\mathbf{k}\sigma} \big). \tag{153}$$

We hereby see that $\hat{H}_{B03}$ can be written as a sum of N_k^3 decoupled Hamiltonians of two-dimensional one-particle systems, each having the form:

$$\hat{H} = \frac{\xi}{2} (\hat{p}_x \pm \xi^{-1} k \hat{y})^2 + \frac{1}{2} (\hat{p}_y \mp k \hat{x})^2 - \frac{k^2 \hat{y}^2}{2\xi} - \frac{k^2 \hat{x}^2}{2}. \tag{154}$$

To get a physical understanding of this Hamiltonian, let us compare it with the Pauli equation (see e.g. Shankar [8]):

$$\hat{H}_P \psi = \frac{1}{2m} (\hat{\mathbf{p}} - q\mathbf{A})^2 \psi - \frac{q}{2m} \boldsymbol{\sigma} \cdot \mathbf{B} \psi + qV\psi, \tag{155}$$

where we also include a potential V. Let us also choose $\xi = 1$ for this comparison, and let $m = q = 1$ in the Pauli equation as well. Since $\mathbf{B}(x, y) = \nabla \times (\mp ky, \pm kx, 0) = (0, 0, \pm 2k)$, we see that Eq. (154) in this case is similar to a spin-up particle moving in a two-dimensional plane with a magnetic field perpendicular to that plane and with a potential given by $V(x, y) = -k^2(x^2 + y^2)/2 \pm k$.

[15] This follows from the definition above of the adjoint of an operator, and it also follows from Propositions 9.21–23 in Hall [4].

From Reed and Simon [3], we know that Schrödinger Hamiltonians can be self-adjoint whenever $V(\mathrm{x})$ is bounded from below by $-a\mathbf{x}^2+b$ for some $a, b \in \mathbb{R}$. (See the first corollary of Theorem X.38 in particular, and see perhaps Theorems X.5–9 for some intuition of what goes wrong when potentials goes to $-\infty$ too fast.) From Theorems X.34–35 in Reed and Simon [3], we also know that Hamiltonians of particles moving in a magnetic field, such as the one we have in Eq. (155), can be self-adjoint if V is locally L^2 and bounded from below. But none of these theorems apply to Eq. (154), however.

Luckily, Ikebe and Kato [2] provide us with what we are looking for. Their Theorem 1 thus tells us that $\hat{H}$ of Eq. (154) is essentially self-adjoint on $C_c^\infty(\mathbb{R}^2)$,[16] i.e. the space of all infinitely differentiable functions with compact support on $\mathbb{R}^2$ (and with values in $\mathbb{C}$).

More generally, this theorem gives a condition for which the operator given by

$$\hat{T}\psi = \sum_{i,j}^{m} \left[\big(\hat{p}_i - b_i(\mathbf{x})\big)a_{ij}(\mathbf{x})\big(\hat{p}_i - b_j(\mathbf{x})\big)\right]\psi + q(\mathbf{x})\psi \tag{156}$$

is essentially self-adjoint on $C_c^\infty(\mathbb{R}^m)$, where a_{ij} and b_i are real-valued and smooth functions for each i, j, and where $(a_{ij}(\mathbf{x}))$ is a symmetric and positive-definite matrix for all $\mathbf{x} \in \mathbb{R}^m$. For our case where the a_{ij}'s are constant and where q is locally bounded, we can let $q(\mathbf{x}) = q_1(\mathbf{x}) + q_2(\mathbf{x})$ and have $q_2(\mathbf{x}) = 0$ everywhere, which means that the requirements put on q_2 for the theorem are trivially met. The remaining condition of the theorem then simply reduces to requiring that there exists a positive, non-decreasing function q' such that $q(\mathbf{x}) \geq -q'(|\mathbf{x}|)$ everywhere and

$$\int^\infty \frac{dr}{\sqrt{a^+ + q'(r)}} = \infty, \tag{157}$$

where a^+ (which can of course also be omitted for this equation) is the greatest eigenvalue of the matrix (a_{ij}).

For $\hat{H}$ of Eq. (154), we see that we can choose $q'(r) = a^+k^2r^2/\xi$ for this condition to be fulfilled, where $a^+ = \max(1, \xi)/2$. We thus have that $\hat{H}$ of Eq. (154) with $C_c^\infty(\mathbb{R}^2)$ as its domain is essentially self-adjoint (for either of the choices of signs). And when we take the closure of this operator (which will have the same formula as in Eq. (154) but with a more inclusive domain), we therefore get a self-adjoint version of $\hat{H}$.

By the same argument that we can treat $\hat{H}_{B03}$ and $\hat{H}_{red}$ individually and then combine their domains afterwards via tensor products, we can also combine all these N_k^3 $\hat{H}$-operators back into $\hat{H}_{B03}$. We thus get that there exists a domain on which $\hat{H}_{B03}$ is self-adjoint, and more precisely, we get that $\hat{H}_{B03}$ is essentially self-adjoint on $C_c^\infty(\mathbb{R}^{N/2})$.

We can now move on to the self-adjointness of $\hat{H}_{red}$, for which we will let the following heuristic argument suffice for our purposes.

In Sect. 7, we show how $\hat{H}_{red}$ can be rewritten in terms of ladder operators. This shows how $\hat{H}_{red}$ can be reinterpreted as an operator on a Fock space, where all the terms of the operator thus simply has the effect of "transporting" vectors from one part of the Fock space to other parts of the space. It is therefore easy to see that $\hat{H}_{red}$ is symmetric if we choose $\mathrm{Dom}(\hat{H}_{red})$ to be the subspace of all L^2-functions that turn into other L^2-functions when the formula for $\hat{H}_{red}$ is applied. Note that this predicate is invariant whether we have $\hat{H}_{red}$ expressed in the Fock space basis (via reinterpreted ladder operators) or in any other basis.

[16] Here we choose to use the same notation as in Hall [4], but in both Reed and Simon [3] and in Ikebe and Kato [2], C_c^∞ is denoted by C_0^∞ instead.

Furthermore, it is also not hard to see from the definition above of the adjoint of an operator, that $\mathrm{Dom}(\hat{H}^*_{red})$ cannot include any L^2 function that does not turn into a L^2 function by the formula of $\hat{H}_{red}$. We thus have that $\mathrm{Dom}(\hat{H}^*_{red}) \subset \mathrm{Dom}(\hat{H}_{red})$, and since $\hat{H}_{red}$ is symmetric, we also have the converse, giving us $\mathrm{Dom}(\hat{H}^*_{red}) = \mathrm{Dom}(\hat{H}_{red})$. In other words, we have that $\hat{H}_{red}$ is self-adjoint on said domain.

We have now found domains for both $\hat{H}_{B03}$ and $\hat{H}_{red}$ on which the operators are (essentially) self-adjoint, and it follows that $\hat{H}_{tran}$ is (essentially) self-adjoint on $\mathrm{Dom}(\hat{H}_{B03}) \otimes \mathrm{Dom}(\hat{H}_{red})$. We can then transform this domain back, i.e. with the reverse change of basis from Sect. 5, to get a $\mathrm{Dom}(\hat{H}_{init})$ on which $\hat{H}_{init}$ is (essentially) self-adjoint.

To conclude this appendix, we also need to argue why $C_c^\infty(\mathbb{R}^N \times \mathbb{R}^{3n}; \mathbb{C}^{4^n})$ is a subspace of $\mathrm{Dom}(\hat{H}_{init}) \cap \mathrm{Dom}(\sum_{\mathbf{k},\sigma} \sum_{\mu=0}^{3} \hat{\Pi}^2_{\mu\mathbf{k}\sigma})$, and why

$$\lim_{\kappa\to\infty} \hat{H}_{init}(\kappa)\chi = \hat{H}_{init}\chi \tag{158}$$

for all $\chi \in C_c^\infty(\mathbb{R}^N \times \mathbb{R}^{3n}; \mathbb{C}^{4^n})$ (and for some appropriate cutoff schema). Given that all C_c^∞-functions turn into L^2-functions when we apply $\sum_{\mathbf{k},\sigma} \sum_{\mu=0}^{3} \hat{\Pi}^2_{\mu\mathbf{k}\sigma}$ or $\hat{H}_{red}$, we immediately get that $C_c^\infty(\mathbb{R}^N \times \mathbb{R}^{3n}; \mathbb{C}^{4^n})$ is a subspace of both $\mathrm{Dom}(\sum_{\mathbf{k},\sigma} \sum_{\mu=0}^{3} \hat{\Pi}^2_{\mu\mathbf{k}\sigma})$ and $\mathrm{Dom}(\hat{H}_{tran})$. And since the change of basis from Sect. 5 leaves $C_c^\infty(\mathbb{R}^N \times \mathbb{R}^{3n}; \mathbb{C}^{4^n})$ invariant, we also get that $C_c^\infty(\mathbb{R}^N \times \mathbb{R}^{3n}; \mathbb{C}^{4^n})$ is a subspace of $\mathrm{Dom}(\hat{H}_{init})$ as well.

To show that Eq. (158) holds for this subspace, we first of all see that this is trivially true if we remove the kinetic part of the Dirac Hamiltonian, i.e. the part that includes the cut-off $\hat{\mathbf{p}}_j$-operators. And to see that these $\hat{\mathbf{p}}_j$-operators also cause us no trouble, we can use the fact that, according to Proposition A.16 in Hall [4], any Schwartz function, which includes all C_c^∞-functions, will Fourier-transform into another Schwartz function. (See e.g. Hall [4], Definition A.15, for the definitions of Schwartz functions and Schwartz spaces.) The χ's we need to consider for Eq. (158) will thus all be Schwartz functions in Fourier space, meaning that the original (not cut-off) version of $\hat{\mathbf{p}}_j\chi$ will converge in the distribution sense for all $j \in \{1, \ldots, n\}$. It is then easy to see that we will have

$$\lim_{\kappa\to\infty} \hat{\mathbf{p}}_j(\kappa)\chi = \hat{\mathbf{p}}_j\chi \tag{159}$$

for each j, where $\hat{\mathbf{p}}_j(\kappa)$ thus denotes the appropriate sequence of cut-off versions of $\hat{\mathbf{p}}_j$, i.e. where the cutoffs are imposed in Fourier space. And since we can therefore write up similar equations as Eq. (159) for all the terms in $\hat{H}_{init}(\kappa)$ and $\hat{H}_{init}$, we thus see that Eq. (158) is true for all $\chi \in C_c^\infty(\mathbb{R}^N \times \mathbb{R}^{3n}; \mathbb{C}^{4^n})$.

C The original motivation for the change of basis in Section 5

When adding the $(2\xi)^{-1}(\nabla \cdot \mathbf{A} + \partial\varphi/\partial t)^2$ term to $\mathcal{L}_{EM}$, one might initially hope that this will fix the gauge completely, at least perhaps when ξ is sent to 0. But as one can see from the resulting Hamiltonian, this is not the case, and we still get a theory with some redundant, "unphysical" degrees of freedom, along which the wave functions are free to propagate. As mentioned in the beginning of Sect. 5, one might then consider going back to the path integral and try to impose restrictions on the paths that fix the gauge more completely. But the idea that has lead to the approach of this paper is to instead consider the option of simply letting the wave functions be free to propagate along these redundant degrees of freedom and then look for generalized solutions that are completely and evenly "spread out" over these.

As an example to illustrate this idea, let us imagine that we have some well-defined physical system such as a set of simple harmonic oscillators, and that we now add a free particle to the system that is thus completely decoupled from these oscillators. We can compare the freedom of this extra particle to having a gauge symmetry for the theory: it adds some extra dimensions to the coordinate space but does not change the measurement outcomes of the oscillators. And while having this extra free particle now means that the Hamiltonian will no longer have any (normalizable) eigenstates, we can still look for generalized solutions to it for which the wave function of the free particle is simply allowed to "spread out" over all the dimensions in which it can move. By looking at the set of all such generalized solutions, where the particle is thus in a specific momentum eigenstate, we therefore effectively get back to the original system, so to speak, and can still derive all the relevant physics of the oscillators from this new Hamiltonian.

We then want to apply this same idea for $\hat{H}_{init}$, which means that we might look for generalized solutions to it of the form $\chi(\mathbf{q}, \bar{\mathbf{a}}) = \Phi(\mathbf{q}, \bar{\mathbf{x}})\Psi(\mathbf{q}, \bar{\mathbf{x}})$, where Φ depends only on the $\widetilde{A}_{0\mathbf{k}\sigma}$ and $\widetilde{A}_{3\mathbf{k}\sigma}$-variables and on $\bar{\mathbf{x}}$, and where Ψ, on the other hand, is constant with respect to all the $\widetilde{A}_{0\mathbf{k}\sigma}$ and $\widetilde{A}_{3\mathbf{k}\sigma}$-variables. We thus hope for Ψ to encapsulate the physical part of the quantum system, and hope for Φ to be the part that is "spread out" over the redundant degrees of freedom.

The formula for $\Phi(\mathbf{q}, \bar{\mathbf{x}})$ that solves this problem turns out to be exactly the formula for $U(\mathbf{q}, \bar{\mathbf{x}})$ in Eq. (62), namely where

$$\Phi(\mathbf{q}, \bar{\mathbf{x}}) = \exp\Big(-iq_F \sum_{\mathbf{k},\sigma} \sum_{j=1}^{n} \frac{\sigma}{k} \widetilde{A}_{3\mathbf{k}-\sigma} f_{\mathbf{k}\sigma}(\mathbf{x}_j)\Big). \tag{160}$$

This can easily be checked, which is in fact essentially what we do in Sect. 5, namely since said check requires exactly the same calculations as we go through there. But since this appendix is about the motivation for said factor, let us here briefly discuss how one can derive it.

The original motivation for the $\Phi(\mathbf{q}, \bar{\mathbf{x}})$ factor of Eq. (160), and hence also the $U(\mathbf{q}, \bar{\mathbf{x}})$ factor of Eq. (62), was to consider the well-known gauge symmetry of the Dirac equation. This gauge symmetry specifically implies that if we add $\partial\lambda/\partial t$ to φ and subtract $\nabla\lambda$ from $\mathbf{A}$ such that $(\varphi, \mathbf{A}) \to (\varphi + \partial\lambda/\partial t, \mathbf{A} - \nabla\lambda)$, where λ is an arbitrary real function of $(t, \mathbf{x})$, we can then transform any previous solution ψ by

$$\psi(t, \mathbf{x}) \to e^{-iq_F\lambda(t,\mathbf{x})}\psi(t, \mathbf{x}), \tag{161}$$

and get a solution for the transformed φ and $\mathbf{A}$ (see e.g. Shankar [8]).

Let us then start by considering a $\psi(t, \mathbf{x})$ (when there is only one fermion in the system) that solves the Dirac equation for some $A^\mu(t, \mathbf{x})$ with $\{\widetilde{A}_{0\mathbf{k}\sigma}(t), \widetilde{A}_{3\mathbf{k}\sigma}(t)\} = \{0\}$. From there we can then try to use the mentioned gauge symmetry to extend this solution to a space where the $\widetilde{A}_{3\mathbf{k}\sigma}$-parameters are now free to take on any value. When we change the $\widetilde{A}_{3\mathbf{k}\sigma}$-parameters from $\{0\}$ to an arbitrary set, $\{\widetilde{A}_{3\mathbf{k}\sigma}\}$, this is equivalent of adding a (time-independent) vector field to the original $\mathbf{A}(t, \mathbf{x})$ equal to $\sum_{\mathbf{k},\sigma} \widetilde{A}_{3\mathbf{k}\sigma}\mathbf{e}_{3\mathbf{k}} f_{\mathbf{k}\sigma}(\mathbf{x})$. We furthermore see that such a change is equivalent of adding a certain (time-independent) $-\nabla\lambda(\mathbf{x})$ to $\mathbf{A}(t, \mathbf{x})$, obtained by first resolving $\lambda(\mathbf{x})$ into $\sum_{\mathbf{k},\sigma} \widetilde{\lambda}_{\mathbf{k}\sigma} f_{\mathbf{k}\sigma}(\mathbf{x})$ and seeing that we can thus write $-\nabla\lambda(\mathbf{x}) = \sum_{\mathbf{k},\sigma} \sigma\mathbf{k}\widetilde{\lambda}_{\mathbf{k}\sigma} f_{\mathbf{k}-\sigma}(\mathbf{x})$, and then equating this to $\sum_{\mathbf{k},\sigma} \widetilde{A}_{3\mathbf{k}\sigma}\mathbf{e}_{3\mathbf{k}} f_{\mathbf{k}\sigma}(\mathbf{x}) = \sum_{\mathbf{k},\sigma} \widetilde{A}_{3\mathbf{k}-\sigma}\mathbf{e}_{3\mathbf{k}} f_{\mathbf{k}-\sigma}(\mathbf{x})$. We see that this is solved by having $\sigma k\widetilde{\lambda}_{\mathbf{k}\sigma} = \widetilde{A}_{3\mathbf{k}-\sigma}$, which gives us

$$\lambda(\mathbf{x}) = \sum_{\mathbf{k},\sigma} \frac{\sigma}{k} \widetilde{A}_{3\mathbf{k}-\sigma} f_{\mathbf{k}\sigma}(\mathbf{x}). \tag{162}$$

So if we plug this into Eq. (161), we should get a solution of this new (less restricted) $\mathbf{A}(t,\mathbf{x})$.

When we do this, we first of all see that in this case where $n = 1$, this give us exactly our $\Phi(\mathbf{q}, \bar{\mathbf{x}})$ factor of Eq. (160). And when we check that this factor indeed does manage to extend the original $\psi(t,\mathbf{x})$ to the less restricted space, which is in fact done by the calculation shown in Eq. (64) in Sect. 5, it is then easy to see how the result can be extended to arbitrary n.

From here one can move on to figuring out how to remove the $\{\widetilde{A}_{0\mathbf{k}\sigma}\} = \{0\}$ restriction as well, and be pleased to discover that this is now trivial to do, namely because all the terms involving the $\widetilde{A}_{0\mathbf{k}\sigma}$-variables will now cancel out, which is what we essentially show in Eq. (66).

This explains the original motivation behind the U factor in Sect. 5. In that section, we then simply use this factor to make a change of basis instead. With this, Φ gets a trivial solution, namely $\Phi(\mathbf{q}, \bar{\mathbf{x}}) = 1$, which makes the redundant degrees of freedom trivial to spot in the formula for the transformed $\hat{H}_{init}$, i.e. $\hat{H}_{tran}$. This is a much easier approach since it does not require any reasoning about generalized solutions. And what is more, it also makes it quite easy to show that the Lorentz covariance is preserved when we cut out the redundant degrees of freedom, which is what we argue in Sect. 6.

The reader might be interested to know, however, that there is another argument for the preservation of the Lorentz covariance, where one shows that the generalized wave functions conforming to Eq. (160) will Lorentz-transform into generalized wave functions with the same exact property. We do not show this in this paper, but in Sect. 11 we briefly mention how this argument goes.

As a last, potentially interesting point of this appendix, let us briefly note the similarity between the **Weyl gauge** of classical electrodynamics and the generalized quantum states of $\mathbf{H}_{init}$ where Φ fulfill Eq. (160). The Weyl gauge is defined by having $\varphi = 0$. And while φ takes on all values for these quantum states, we can still interpret φ as effectively being equal to 0 due to the fact that the electric potential energy is canceled everywhere. Additionally, when we interpret the energy as being stored in the fields, the classical electric potential energy is given by $\mathbf{E}^2/2 = (\partial\mathbf{A}/\partial t)^2/2$ in the Weyl gauge. And when we analyze the quantum mechanical analogue to this, namely $\sum_{\mathbf{k},\sigma} \hat{\Pi}^2_{3\mathbf{k}\sigma}/2$, which we do in Eqs. (112–118) in Sect. 7, we see that we indeed get a potential energy equal to a Coulomb potential between all the fermions!

D The divergences of a naive Dirac sea reinterpretation

As mentioned in Sect. 9, simply replacing $\hat{d}_s(\mathbf{p})$ in the formula for $\hat{H}_{CL}$ with $\hat{d}_s^\dagger(-\mathbf{p})$ and vise versa leads to some divergences. In this appendix, we expand a bit on why this is.

First of all, we see that if we carry out said replacements for $\hat{H}_F$, we would get

$$\begin{aligned}\hat{H}_F &\to \int \frac{d\mathbf{p}}{(2\pi)^3}\, E_p\Big(\sum_{s=1}^{2}\hat{b}_s^\dagger(\mathbf{p})\hat{b}_s(\mathbf{p}) - \sum_{s=3}^{4}\hat{d}_s(-\mathbf{p})\hat{d}_s^\dagger(-\mathbf{p})\Big)\\ &= \int \frac{d\mathbf{p}}{(2\pi)^3}\, E_p\Big(\sum_{s=1}^{2}\hat{b}_s^\dagger(\mathbf{p})\hat{b}_s(\mathbf{p}) + \sum_{s=3}^{4}\hat{d}_s^\dagger(\mathbf{p})\hat{d}_s(\mathbf{p}) - (2\pi)^3\delta(0)\Big),\end{aligned} \tag{163}$$

which can be seen to blow up. This divergence can be cured quite easily, however, if we simply take a step backwards and discretize $\hat{H}_{CL}$ once more, this time where we also turn the fermionic ladder operators into discretized versions of themselves. It is then easy to see that the $\delta(0)$ term in Eq. (163) will become a constant energy for this discretized case. If we then remove this energy at every step while going to the continuum limit once again, we get

$$\hat{H}'_F = \int \frac{d\mathbf{p}}{(2\pi)^3}\, E_p\Big(\sum_{s=1}^{2}\hat{b}_s^\dagger(\mathbf{p})\hat{b}_s(\mathbf{p}) + \sum_{s=3}^{4}\hat{d}_s^\dagger(\mathbf{p})\hat{d}_s(\mathbf{p})\Big). \tag{164}$$

(Here we thus use the prime to denote that the operator is in its Dirac sea-reinterpreted form.) We can therefore see that we get a well-defined operator as a result, and one which has the desired property that all the energies of the antiparticles are now positive.

The other parts of $\hat{H}_{CL}$ that contain the $\hat{d}_s(\mathbf{p})$ and $\hat{d}_s^\dagger(-\mathbf{p})$-operators are the formulas for $\hat{H}_{ICL}$ and $\hat{V}_{CL}$. If we simply interchange said ladder operators in these formulas, we can still write

$$\hat{H}'_{ICL} = -q_F\sum_{\lambda=1}^{2}\int \frac{d\mathbf{k}\,d\mathbf{p}}{(2\pi)^6}\,\frac{1}{\sqrt{2k}}\big(\hat{a}_\lambda(\mathbf{k}) + \hat{a}_\lambda^\dagger(-\mathbf{k})\big)\hat{\bar{\psi}}(\mathbf{p}+\mathbf{k})\gamma^\mu\epsilon_{\mu\lambda}(\mathbf{k})\hat{\psi}(\mathbf{p}), \tag{165}$$

$$\hat{V}'_{CL} = \int \frac{d\mathbf{k}}{(2\pi)^3}\frac{d\mathbf{p}_1\,d\mathbf{p}_2}{(2\pi)^6}\,\frac{q_F^2}{2k^2}\big(\hat{\bar{\psi}}(\mathbf{p}_1+\mathbf{k})\otimes\hat{\bar{\psi}}(\mathbf{p}_2-\mathbf{k})\big)(\gamma^0\otimes\gamma^0)\big(\hat{\psi}(\mathbf{p}_1)\otimes\hat{\psi}(\mathbf{p}_2)\big), \tag{166}$$

as before, but where $\hat{\psi}(\mathbf{p})$ and $\hat{\bar{\psi}}(\mathbf{p})$, previously defined by Eq. (101), are now redefined for this context as

$$\begin{aligned}\hat{\psi}(\mathbf{p}) &= \frac{1}{\sqrt{2E_p}}\Big(\sum_{s=1}^{2}u_s(\mathbf{p})\hat{b}_s(\mathbf{p}) + \sum_{s=3}^{4}v_s(\mathbf{p})\hat{d}_s^\dagger(-\mathbf{p})\Big),\\ \hat{\bar{\psi}}(\mathbf{p}) &= \frac{1}{\sqrt{2E_p}}\Big(\sum_{s=1}^{2}\hat{b}_s^\dagger(\mathbf{p})\bar{u}_s(\mathbf{p}) + \sum_{s=3}^{4}\hat{d}_s(-\mathbf{p})\bar{v}_s(\mathbf{p})\Big).\end{aligned} \tag{167}$$

If one analyzes all these terms, most of them seem to not cause any divergences, and thus seem to be easily defined on the desired Fock space without any changes needed. But there are one type of terms that seems to cause trouble, namely that of all the terms that consist purely of either creation or annihilation operators. These are for example the terms that involve $\hat{a}_\lambda^\dagger(-\mathbf{k})\hat{b}_s^\dagger(\mathbf{p}+\mathbf{k})\hat{d}_{s'}^\dagger(-\mathbf{p})$ or $\hat{a}_\lambda(\mathbf{k})\hat{d}_{s'}(-\mathbf{p}-\mathbf{k})\hat{b}_s(\mathbf{p})$ in the case of $\hat{H}'_{ICL}$, either creating a set of particles from the empty vacuum (with zero momentum combined) or annihilating such a set.

If one tries to do what we did for $\hat{H}_F$ and discretize all the creation and annihilation operators, one finds that there is no clear continuum limit of such operators on the desired

Fock space, and that the operators seem to cause more and more vacuum fluctuations in said limit. This makes intuitive sense if we consider the fact that $\hat{H}_{CL}$, or rather $\exp(-i\hat{H}_{CL}t)$, in all local volumes of space can cause transitions from negative-energy fermion states to positive-energy states (and vice versa). If we now interchange all the $\hat{d}_s(\mathbf{p})$ and $\hat{d}_s^\dagger(-\mathbf{p})$-operators, we expect to get the same formula for $\exp(-i\hat{H}'_{CL}t)$ as for $\exp(-i\hat{H}_{CL}t)$ when expressed in terms of ladder operators, only with said operators interchanged also. We thus expect $\exp(-i\hat{H}'_{CL}t)$ to cause pair productions (and annihilations) at every local volume of space, and if we thus start out with the bare vacuum, $\exp(-i\hat{H}'_{CL}t)$ should therefore immediately cause an infinite number of pair productions if we let $\mathcal{V}$ tend to infinity.[17]

Now, it would be great if we could simply remove all these "offending" terms from $\hat{H}'_{CL}$, i.e. the terms that perturb the empty vacuum, and let the result be our proposed Hamiltonian for QED. And in fact, such a Hamiltonian might very well still yield us good predictions for most experiments. But if we do this, we might very well break the Lorentz covariance slightly for the theory (especially when pair productions or annihilations are part of the experiments). So if we truly believe that our universe is relativistic, which certainly seem to be the case from experiments thus far, we should want to continue the search for a Hamiltonian that takes the perturbed vacuum into account.

To solve this problem, one would then likely have to solve the perturbed vacuum and use this solution to "renormalize" $\hat{H}'_{CL}$, i.e. by making the perturbed vacuum the new "offset" for all states, so to speak. One might then also have to show that this new vacuum state Lorentz-transforms into itself, or at least into states that are physically equivalent to it. And if this is achievable, one might then finally be able to prove Lorentz covariance for such a full theory of QED.

[17] This argument becomes even more clear when realizing that we are free to change the basis of the $\hat{d}_s(\mathbf{p})$ and $\hat{d}_s^\dagger(-\mathbf{p})$-operators at any time, with the same outcome whether we do it before or after we interchange them. We are thus free to express both $\exp(-i\hat{H}_{CL}t)$ and $\exp(-i\hat{H}'_{CL}t)$ in terms of ladder operators that cause transitions only between states that localized close to each other in position space.

References

[1] S. Weinberg, *The Quantum Theory of Fields*, Vol. I: Foundations (Cambridge University Press, Cambridge, 1995).

[2] T. Ikebe and T. Kato, Uniqueness of the self-adjoint extension of singular elliptic differential operators, Arch. Rat. Mech. Anal. **9**, 77–92 (1962).

[3] M. Reed and B. Simon, *Methods of Modern Mathematical Physics*, Vol. II: Fourier Analysis, Self-adjointness (Academic Press, New York, 1975).

[4] B. C. Hall, *Quantum Theory for Mathematicians* (Springer, New York, 2013).

[5] M. J. Damgaard, B.Sc. thesis, Niels Bohr Institute, University of Copenhagen, 2017.

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

[7] M. Srednicki, *Quantum Field Theory* (Cambridge University Press, Cambridge, 2007).

[8] R. Shankar, *Principles of Quantum Mechanics*, 2nd ed. (Springer, New York, 1994).

[9] E. S. Abers, *Quantum Mechanics* (Pearson Education, London, 2004).

[10] C. Cohen-Tannoudji, B. Diu, and F. Laloë, *Quantum Mechanics*, Vol. 1 (John Wiley & Sons, New York, 1977).

[11] C. C. Gerry and P. L. Knight, *Introductory Quantum Optics* (Cambridge University Press, Cambridge, 2004).

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

[13] P. Hedegaard, *Statistisk Fysik*, 2015 version, Lecture notes, Niels Bohr Institute, University of Copenhagen, 2015.

[14] D. J. Griffiths, *Introduction to Quantum Mechanics*, 2nd ed. (Pearson Education, Edinburgh, 2014).

[15] D. J. Griffiths, *Introduction to Electrodynamics*, 4th ed. (Pearson Education, Edinburgh, 2014).

[16] A. L. Fetter and J. D. Walecka, *Theoretical Mechanics of Particles and Continua* (Dover, New York, 2003).

[17] K. F. Riley and M. P. Hobson, *Essential Mathematical Methods for the Physical Sciences* (Cambridge University Press, Cambridge, 2011).