

Gauge Invariant Propagators and States in Quantum Electrodynamics

Jordan Wilson-Gerow¹ and P.C.E. Stamp^{1,2,3}

¹*Department of Physics and Astronomy, University of British Columbia,
6224 Agricultural Rd., Vancouver, B.C., Canada V6T 1Z1*

²*Pacific Institute of Theoretical Physics, University of British Columbia,
6224 Agricultural Rd., Vancouver, B.C., Canada V6T 1Z1*

³*Theoretical Astrophysics, Cahill, California Institute of Technology,
1200 E. California Boulevard, MC 350-17, Pasadena CA 91125, USA*

We study gauge invariant states in QED, where states are understood in terms of data living on the boundary of gauge invariant path-integrals. This is done for both scalar and spinor QED, and for boundaries that are either time slices, or the boundaries of a ‘causal diamond’. We discuss both the case where the gauge field falls off to zero at the boundaries, and the case of ‘large gauge transformations’, where it remains finite at the boundaries. The dynamics are discussed using the gauge-invariant propagator, describing motion of both the particles and the field between the boundaries. We demonstrate how the path-integral naturally generates a ‘Coulomb-field’ dressing factor for states living on time-slices, and how this is done without fixing any gauge. We show that the form of the dressing depends only on the nature of the boundaries. We also derive the analogous dressing for states defined on null infinity, showing both its Coulombic parts as well as soft-photon parts.

I. INTRODUCTION

Traditionally, both non-relativistic quantum mechanics (QM) and relativistic Quantum field theory (QFT) have been formulated in terms of states in Hilbert space, upon which operators representing measurements of physical quantities are supposed to act. However the last few decades have seen a growing feeling that one needs to go beyond such a framework - the present paper is an effort in this direction.

A. Background and Rationale

We can begin by recalling some of the reasons for a generalization of the traditional notion of states in QM and QFT. One key motivation has come from quantum gravity, where the difficulties in defining diffeomorphism-invariant physical quantities [1–4] have led to approaches in which states are defined in terms of information residing on boundaries [5–8]. Thus, eg., path integrals can be used to define ground-state wave-functions for different kinds of boundary [9] or even for spacetimes with no boundaries [10]. Much of modern quantum cosmology is also formulated using path integrals [11]

Another motivation for going beyond conventional QM has been to deal with topological field theories, and in the theory of fractional statistics for many-particle systems [12–14]. That path integrals provide a more general formulation of QM has been known for a long time [15].

We also note the extensive efforts to look for generalizations of QM in which, eg., the superposition principle breaks down [16, 17], which in recent years have focused on the possible role of gravity in engineering this breakdown [18–22].

One can argue that even in ordinary one-particle QM, a simple wave-function description is inadequate. Non-

local features of ordinary QM have been known for a long time - in, eg., the Aharonov-Bohm effect [23], and in interaction-free measurements [24]. In such examples, the time evolution of a quantum system \mathcal{S} depends both on what can happen along the paths \mathcal{S} does follow, and also those paths it does *not* follow. In a path integral formulation this seems quite natural - but not if we deal entirely with the wave-function $\langle \mathbf{r} | \Psi_{\mathcal{S}}(t) \rangle$, which is zero in regions where no paths are followed.

In gauge theories, such non-locality is of course essential. A crucial role is then played by constraints, and by the requirement of gauge invariance. This was recognized early on by Dirac, as part of his efforts to quantize constrained theories [25, 26]; he used operator representations of the constraints to annihilate physical states. Dirac was thereby led [27] to introduce gauge-invariant “physical states” in Quantum Electrodynamics (QED); and the constraints were then the generators of the QED gauge transformations.

One can also define gauge-invariant states prepared by a path integral, which are of course non-local objects. These states will satisfy the operator constraints provided that the action, the measure, and the set of summed paths are themselves invariant under the transformations generated by the constraints. A good example is provided in quantum gravity by the Hartle-Hawking “no-boundary” wave-function of the universe [10], where one has a Euclidean path-integral over four-dimensional metrics. This state then satisfies the Hamiltonian constraint of Einstein gravity, in the form of the Wheeler-DeWitt equation [28].

For more general classical spacetimes, and similarly for quantum gravity, one can argue that path integrals are actually unavoidable, because for any non-trivial spacetimes involving wormholes, in which there exist achronal regions [5, 6, 29, 30], it is already known that one must employ path integrals to handle the dynam-

ics of even simple particles. In these situations, conventional Hamiltonian framework is then no longer applicable, and simple Hamiltonian evolution is undefined, whereas path integrals can still compute transition amplitudes/probabilities.

Much of the literature on this whole topic is quite abstract and formal; our approach in this paper will be to start with simple cases and then generalize. Moreover, it seems to us that some of the key issues which still need to be clarified are quite down to earth in nature. These issues and questions include:

(i) Most of the abstract discussion does not explicitly address the role of the - often complicated - dynamics of real charges (in QED) or masses (in quantum gravity). It is often not clear how to separate out the ‘physical’ degrees of freedom from unphysical ones - this is particularly true when one deals with rapidly accelerating objects, where a discussion on terms of ‘near field’ and ‘far field’ zones does not help in making such a separation. One can discuss things entirely in terms of asymptotic states [31, 32], but this is not of much help in dealing with phenomena in bounded regions of spacetime.

(ii) If one is dealing with state superpositions involving a large spatial separation of charge or mass, real confusion arises in discussion of what are the correct physical variables, or how to test, eg., whether or not the gravitational metric field $g^{\mu\nu}(x)$ is quantized [34–36]. The related question of how to properly define notions like decoherence is also unclear, with different results being derived for decoherence rates by different authors [33, 37–40].

(iii) While integrating separately over gauge field and matter variables in a path integral, one needs to deal properly with both the constraints and the gauge redundancy. To deal with the latter one typically uses the Faddeev-Popov technique [41]. This still leaves the problem of implementing the constraints in a manifestly gauge invariant way.

(iv) If one then moves to non-trivial spacetimes, including those containing achronal regions [5, 6, 29, 30], one is immediately faced with the unresolved question of how to define quantum states for physical systems in these spacetimes (including matter coupled to any gauge field).

These and other issues, as well as the inherent non-locality of QM, suggest that it would be very useful to formulate the notion of states in both QM and gauge field theory in a “path-integral-first” manner. We start from the view that such a non-local formulation of quantum states is inevitable in gauge field theories, and is more natural than the traditional Hamiltonian formulation in terms of local states defined in Hilbert space.

Such an idea is of course hardly new. In general it runs into problems, when dealing with non-Abelian theories, connected with the Gribov ambiguity [42], and in quantum gravity, with functional integration over different

metrics. For this reason we stick to QED in the present paper; although gauge invariance in QED is an old topic, we will be deriving here a number of new results.

The main focus here will be to investigate and give explicit expressions for QED states, defined in the path-integral-first approach, for different kinds of boundary condition. We will work out the details for both scalar and spinor QED, and for 2 different flat spacetimes - in one case, boundary information is specified on two time slices, whereas in the other, it is given on a causal diamond. We will look at different kinds of boundary information, depending on whether the EM field A^μ vanishes at infinity or not, and whether or not one needs to allow for ‘large’ gauge transformations.

B. Organization of Paper

The paper is organized as follows. In section 2 we consider a quantum particle coupled to the electromagnetic field and derive the form of the gauge-invariant propagator between time slices. We use this simple example to highlight the gauge independence of the results, demonstrate how the boundary phases emerge without fixing a gauge beforehand, and then sketch an eikonal argument for the dressings coming from the remaining path-integral. We close by introducing the boundary Faddeev-Popov trick, and show how it gives the same results.

After this warm-up exercise, in section 3 we generalize the results to charged matter described by the Dirac field, ie. discuss how the results and methods of the previous section apply to full QED.

In section 4 we start on the much more complicated derivations required for general boundary hypersurfaces. In this section we derive the propagator between states on the future and past regions of a large causal diamond, again for flat space.

Up to this point, all the discussion has been for gauge transformations which vanish at infinity. In section 5 we lift this restriction, and extend all of the previous results to the case of “large gauge transformations”. This leads to an interesting connection with the soft-photon, large gauge transformation, and dressed state literature.

Throughout the paper we will use units in which $\hbar = c = \epsilon_0 = 1$, and a $- + + +$ metric signature; and we will put $\hbar = 1$ and $c = 1$.

II. SCALAR QUANTUM ELECTRODYNAMICS

In this section we discuss scalar QED; the next section will show how the results carry over to spinor QED. We wish to define gauge invariant physical states for scalar QED, starting from the path integral.

After briefly reviewing some of the salient issues, we define gauge invariant propagators for scalar QED, and show how they can be written so that a boundary term separates from the rest. In this section we will assume

the boundaries are defined by time slices. The boundary term is reminiscent of Dirac's well known phase factor, but is only equal to it when a certain boundary condition is imposed.

The calculation is gauge invariant throughout, and is carried out in two different ways. The first relies on a 'natural' transformation of the action in the system, whereas the second involves a generalization of the standard Faddeev-Popov technique [41] to include boundaries. We assume flat spacetime throughout.

A. States and Paths in Quantum Electrodynamics

The question of how to define gauge invariant states in QED has a long history; here we recall some of the key arguments, and set up the calculations to follow.

1. States and Conservation Laws

In classical electrodynamics Gauss' law, that $\nabla \cdot \mathbf{E} = \rho$, obviously does not completely specify the electric field $\mathbf{E}(x)$. One can add to the Coulomb solution any divergence-free field. In quantum theory the Gauss law operator constraint for physical states,

$$(\vec{\nabla} \cdot \hat{\mathbf{E}} - \hat{\rho})|\Psi\rangle = 0, \quad (1)$$

also has no unique solution. In the Schrödinger picture, in the $\hat{A}_j(x)$ field value basis, the electric field operator is conjugate to $\hat{A}_j(x)$, ie.,

$$\hat{E}^j(x) = -i \frac{\delta}{\delta A_j(x)}. \quad (2)$$

Many years ago Dirac argued that one should write the wave-function for a static physical electron in the composite form

$$\begin{aligned} \Psi(\mathbf{r}|A^\mu) &= e^{-i\varphi(\mathbf{r}|A^\mu)}\psi(\mathbf{r}) \equiv \hat{U}_C \psi(\mathbf{r}) \\ \varphi(\mathbf{r}|A^\mu) &= e \int d^3r' f^\mu(\mathbf{r} - \mathbf{r}') A_\mu(\mathbf{r}') \end{aligned} \quad (3)$$

where $\psi(\mathbf{r})$ is the wave-function for the 'bare' charge.

As Dirac recognized, the phase factor $\varphi(\mathbf{r}|A^\mu)$ represents a "dressing" function, describing the change in the field induced by the charge. Dirac chose an intuitively obvious solution for $\varphi(\mathbf{r}|A^\mu)$, writing

$$\varphi(\mathbf{r}|A^\mu) = e \nabla^{-2} \partial_j A^j(\mathbf{r}) = -\frac{e}{4\pi} \int d^3r' \frac{\nabla_{r'} \cdot \mathbf{A}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \quad (4)$$

so that the dressing was that of a Coulomb field, and automatically satisfied (1).

One can always add some divergence-free function to (4), to get a quite different form; a good example is provided by the 'Mandelstam string' solution [43], viz.,

$$\varphi(\mathbf{r}|A^\mu) = e \int_\Gamma^{\mathbf{r}} dz^\mu A_\mu(z) \quad (5)$$

where Γ is some spacelike path terminating at the point \mathbf{r} ; solutions of this form have attracted interest in various contexts [44, 45].

So far this seems straightforward. Suppose however we consider a simple superposition of position eigenstates for the charge. We must then have a superposition of Coulomb fields centred on \mathbf{x}_1 and \mathbf{x}_2 , ie.,

$$\Psi(\mathbf{r}|A^\mu) \rightarrow \frac{1}{\sqrt{2}} (\hat{U}_{C_1}|\mathbf{x}_1\rangle + \hat{U}_{C_2}|\mathbf{x}_2\rangle), \quad (6)$$

where $|\mathbf{x}_1\rangle, |\mathbf{x}_2\rangle$ are bare electronic states with charges at \mathbf{x}_1 and \mathbf{x}_2 . This state is not an eigenstate of the longitudinal electric field operator, and no longer has a well defined electric field, but instead a superposition of Coulomb fields.

These ambiguities are compounded in the general case of a moving charge. Classically, the Liénard-Wiechert solution, which makes explicit reference to the trajectory of the moving charge, can still be used. However, quantum mechanically one would describe a moving charge using a wavepacket peaked on a particular momentum and the Dirac prescription would then generate a continuous superposition of the various Coulomb fields. We clearly cannot isolate out a specific physical Coulomb field – we would have no idea which one to use! Moreover, there is no unique gauge-invariant separation between longitudinal and transverse field degrees of freedom, valid in both near field and far field zones, for a charge following some arbitrary classical trajectory.

Once we sum over multiple paths for a charge, in QED, we must clearly sum over multiple field configurations, some of which will involve radiation, others not - the problem seems hopeless. Nevertheless we will see that, in a path integral formulation, one can give a unique separation between constrained variables and unconstrained radiative variables, which is manifestly gauge-invariant. This then allows unambiguous answers to some of the questions posed in sec. I.A. of the introduction.

2. Paths and Propagators

We know from, eg., the Hartle-Hawking work [10] that path-integrals can be used to prepare gauge-invariant vacuum states. The basic idea is to then generalize this idea to arbitrary states, maintaining gauge invariance throughout, and see what emerges. Some of the questions we would like to see answered include:

- (a) What sort of electromagnetic dressing is "chosen" by states defined in this way?
- (b) How do the physical states so defined depend on the spacetime boundaries and the information specified on them?
- (c) What are the physical degrees of freedom involved in gauge invariant spatial superpositions and in entangled states?
- (d) In defining decoherence and information loss, what states should we average over, how do we distinguish

between real decoherence and “false” decoherence, and what is the correct way to calculate decoherence rates?

This is just a sampling of the kind of concrete problem that one would like to see answered. In all sections of this paper we will use manifestly gauge-invariant path-integrals to address them. We will introduce a novel ‘boundary Faddeev-Popov’ (bFP) trick to derive results - it is a straightforward generalization of the textbook Faddeev-Popov trick to path-integrals with prescribed boundary data. Path integrals will be on the extended configuration space of the gauge field. We prescribe data for all components A_μ , and do not a priori concern ourselves with the non-canonical nature of A_0 , or the implementation of Gauss’ law as a constraint.

However we find that, because all amplitudes are manifestly gauge invariant, the expected constraints are naturally implemented. One finds that boundary phases engender states which are eigenstates of certain parts of the electric field operator. We find that a natural separation of variables occurs, whereby a preferred solution to the constraint equation emerges kinematically, and the additional gauge invariant data is not inserted by hand, but rather it emerges dynamically from the path-integral.

In this section we will look at scalar electrodynamics, wherein a single charged particle and the quantum electromagnetic field propagate between two time slices (surfaces of constant t) in Minkowski space. We will assume here (but not later in the paper) that the gauge field A_μ , and thus the possible gauge transformations of it, will vanish sufficiently quickly at spatial infinity that surface terms generated by spatial integrations by parts can be ignored. In the final section we will consider “large” gauge transformations.

We therefore consider a non-relativistic particle with position q , charge e , in an external potential $V(q, t)$, and coupled to the electromagnetic field A_μ . The extension to multiple particles is trivial. The action for the system evolving from an initial time t_i to a final time t_f is $S = S_M + S_{EM}$, where

$$S_M[X] = \int_{t_i}^{t_f} dt \left[\frac{1}{2} m \dot{q}^2 - V(q, t) \right] \quad (7)$$

describes the particle alone, and

$$S_{EM}[X, A_\mu] = \int_{t_i}^{t_f} d^4x \left[-\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + A_\mu J^\mu \right] \quad (8)$$

describes the electromagnetic field along with the coupling to the matter; here $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the electromagnetic field tensor, and $J^0(x) = e\delta^3(x - q(t))$ and $J^j = e\dot{x}^j\delta^3(x - q(t))$ are components of the charge current. Note here that the current for a charged particle is conserved even when the equations of motion are not satisfied, ie. for a general path in the path integral.

Under gauge transformation this action transforms by a boundary term. We will assume these transformations vanish at spatial infinity, but the contribution from the space-like parts of the boundary will not vanish. We

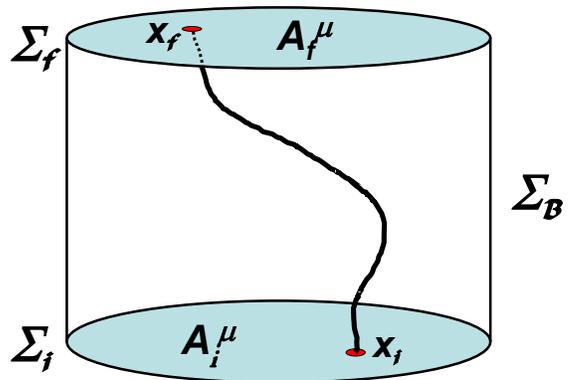


FIG. 1: Depiction of the propagator K_{fi} in eqtn. (10) and the spacetime through which it propagates. The initial configuration is on the timeslice surface $t = t_i$, the final configuration on the surface $t = t_f$. Particle paths propagate between \mathbf{x}_i at t_i and \mathbf{x}_f at t_f ; the gauge field propagates between A_i^μ and A_f^μ on these same two timeslices.

assume the spacetime shown in Fig. 1; we denote the surface of constant time $t = t_f$ by Σ_f , and likewise for t_i, Σ_i . An asymptotic timelike cylinder $S^2 \times \mathbb{R}$ at arbitrarily large radius will be denoted Σ_∞ .

Our path integral is then over field configurations and particle trajectories in a region \mathcal{V} bounded by $\partial\mathcal{V} = \Sigma_f \cup \Sigma_i \cup \Sigma_\infty$. In this notation, under gauge transformation, $A_\mu \rightarrow A^\Lambda = A_\mu + \partial_\mu \Lambda$, the EM action then acquires a boundary term

$$\begin{aligned} \delta_\Lambda S_{EM} &= \int_{\partial\mathcal{V}} d^3x \Lambda n_\mu J^\mu = \int_{\Sigma_f} d^3x \Lambda_f J^0 - \int_{\Sigma_i} d^3x \Lambda_i J^0 \\ &= e (\Lambda_f(q_f) - \Lambda_i(q_i)) \end{aligned} \quad (9)$$

We will choose to quantize the system on the extended configuration space, ie. considering all configurations of $A_\mu(x)$ before quantization rather than imposing constraints and gauge conditions at the classical level and quantizing the remaining degrees of freedom.

The path integral describing the amplitude for transition between configurations $q_i, A_{\mu i}$ and $q_f, A_{\mu f}$ is then

$$\begin{aligned} K_{fi} &\equiv K(q_f, A_{\mu f}; q_i, A_{\mu i}) \\ &= \int_{q_i}^{q_f} \mathcal{D}q e^{iS_M} \int_{A_{\mu i}}^{A_{\mu f}} \mathcal{D}A_\mu e^{iS_{EM}}. \end{aligned} \quad (10)$$

Here and throughout this paper we will absorb field independent constants into the path integral measure. The integral in (10) is usually handled using a Faddeev-Popov (FP) gauge fixing procedure, to divide out the divergent volume from gauge equivalent field configurations. Here we will actually delay performing the FP procedure and perform some formal manipulations of the path-integral before proceeding to fix a gauge. The final result will ultimately be the same; but this order of operations turns out to be illuminating.

B. Gauge Invariant Propagator

The propagator in (10) is manifestly gauge invariant under independent transformations of the initial and final data: provided we simultaneously transform the wavefunction of the particle and transform the gauge field.

To see this explicitly, consider the transformed propagator

$$\begin{aligned} K_{fi}^\Lambda &\equiv K^\Lambda(X_f, A_{\mu f}; X_i, A_{\mu i}) \\ &= e^{-ie\Lambda_f(q_f)} K(X_f, A_{\mu f}^\Lambda; X_i, A_{\mu i}^\Lambda) e^{ie\Lambda_i(q_i)}, \end{aligned} \quad (11)$$

where $\Lambda_{i(f)}$ is the gauge parameter on the initial (final) time slice, and the gauge field transforms as $A_\mu^\Lambda = A_\mu + \partial_\mu \Lambda$. The propagator with transformed boundary data can be expressed simply in terms of the original propagator. We can take the expression

$$K(q_f, A_{\mu f}^\Lambda; q_i, A_{\mu i}^\Lambda) = \int_{q_i}^{q_f} \mathcal{D}q e^{iS_M} \int_{A_{\mu i}^\Lambda}^{A_{\mu f}^\Lambda} \mathcal{D}A_\mu e^{iS_{EM}} \quad (12)$$

and perform a change of variables, $A_\mu = A'_\mu + \partial_\mu \Lambda$, for some time dependent function Λ which takes the value $\Lambda_{i,f}$ on $\Sigma_{i,f}$.

The boundary data for the new variable A'_μ is now just the original configuration, $A_{\mu i,f}$. The action is not invariant under this change of variables, but instead, since it is effectively just a gauge transformation, we know that the action changes by a simple boundary term. The action is expressed in terms of A'_μ as

$$S_{EM}[q, A] = S_{EM}[q, A'] + e\Lambda_f(q_f) - e\Lambda_i(q_i), \quad (13)$$

so that the propagator with transformed boundary data then reads

$$\begin{aligned} K(q_f, A_{\mu f}^\Lambda; q_i, A_{\mu i}^\Lambda) &= \int_{q_i}^{q_f} \mathcal{D}q e^{iS_M[q]} \int_{A_{\mu i}}^{A_{\mu f}} \mathcal{D}A'_\mu e^{iS_{EM}[q, A'] + e\Lambda_f(q_f) - e\Lambda_i(q_i)} \\ &= e^{ie\Lambda_f(q_f)} K_{fi} e^{-ie\Lambda_i(q_i)}. \end{aligned} \quad (14)$$

The boundary phases in (14) generated by the gauge-field action will then precisely cancel the phases in (10) arising from the U(1) transformation of the matter states, and therefore the propagator for the total system is gauge invariant.

We know from Hamiltonian dynamics that the Gauss law constraint is the generator of gauge transformations. The propagator (10) should therefore satisfy Gauss' law as an operator constraint on both Σ_f and Σ_i . To see this,

consider a gauge transformation which vanishes on Σ_i but not on Σ_f , and rewrite the transformed propagator using a linear shift operator, as

$$K_{fi}^\Lambda = e^{-ie\Lambda_f(q_f) + \int_{\Sigma_f} d^3x \partial_\mu \Lambda_f \frac{\delta}{\delta A_{\mu f}}} K_{fi} \quad (15)$$

Since the propagator is gauge invariant, this implies the following simple functional differential equation

$$\begin{aligned} 0 &= \left[-ie\Lambda_f(q_f) + \int_{\Sigma_f} d^3x \partial_\mu \Lambda_f \frac{\delta}{\delta A_{\mu f}} \right] K_{fi} \\ &= \left[\int_{\Sigma_f} d^3x \partial_0 \Lambda_f \frac{\delta}{\delta A_{0f}} - i \int_{\Sigma_f} d^3x \Lambda_f \left(e\delta^3(q_f - x) - i\partial_j \frac{\delta}{\delta A_{jf}} \right) \right] K_{fi} \end{aligned} \quad (16)$$

The remaining functional derivative of the propagator with respect to A_{jf} is just the electric field operator. This is seen in a ‘‘path-integral first’’ treatment by evaluating the functional derivative and using the standard

expression for variations of the action endpoint in mechanics, viz.,

$$\frac{\delta S[x]}{\delta x_f} = \frac{\partial \mathcal{L}}{\partial \dot{x}(t)} \Big|_{t_f}. \quad (17)$$

Written in terms of the electric field operator we then get the constraint equation

$$0 = \left[\int d^3x \partial_0 \Lambda_f \frac{\delta}{\delta A_{0f}} - i \int d^3x \Lambda_f \left(e \delta^3(q_f - x) - \partial_j \hat{E}^j \right) \right] K_{fi} \quad (18)$$

On the surface Σ_f , the functions Λ_f and $\partial_0 \Lambda_f$ are independent, but arbitrary, functions vanishing at spatial infinity. As a result, the propagator then satisfies two separate local constraint equations

$$\frac{\delta}{\delta A_{0f}} K_{fi} = 0 \quad (19)$$

$$\left(\partial_j \hat{E}^j - \hat{J}^0 \right) K_{fi} = 0. \quad (20)$$

so that, as expected, the gauge invariant propagator defined on the extended configuration space satisfies the Gauss law operator constraint. It is also independent of the prescribed data for A_0 .

Since the propagator is independent of the data prescribed for A_0 , up to normalization constants we can then freely integrate over the boundary data for A_0 , to get

$$K_{fi} = \int_{q_i}^{q_f} \mathcal{D}q e^{iS_M} \int \mathcal{D}A_0 \int_{A_{j_i}}^{A_{j_f}} \mathcal{D}A_j e^{iS_{EM}}. \quad (21)$$

showing that A_0 is not a true dynamical variable. This is a well known point, but the way we demonstrated it will be useful for more general amplitudes.

No gauge-fixing is required to make the A_0 integral convergent, and the boundary data is unfixed, so we can directly go ahead and evaluate the integral. As we will see shortly, this rather uniquely determines how we should gauge fix the remaining A_j integral, and consequentially determines the form of the dressing for the states.

Notice that the boundary data for A_0 naturally fell out of the expression as a consequence of gauge invariance - there was no need for a detour through canonical Hamiltonian quantization, or a discussion of the missing conjugate momentum Π^0 to see this point. Geometrically this happens here because we chose to evolve between constant time slices, and the pullback of the 1-form $A_\mu dx^\mu$ to these boundaries is independent of A_0 , making A_0 redundant, i.e., not a true dynamical variable. We will see that in other boundary geometries, the redundant variable will be that part of A_μ normal to the boundary.

C. Extracting the Dressing

We will now take the form (21) as a starting point, and look to evaluate the A_0 integral. Looking at the electromagnetic part of the action we can separate out

the A_0 dependent terms

$$S_{EM} = \int_{t_i}^{t_f} d^4x \left[-\frac{1}{4} F_{jk} F^{jk} + A_j J^j - \frac{1}{2} F^{j0} (\partial_j A_0 - \partial_0 A_j) + A_0 J^0 \right] \quad (22)$$

Integrating the spatial derivatives by parts we obtain

$$S_{EM} = \int_{t_i}^{t_f} d^4x \left[-\frac{1}{4} F_{jk} F^{jk} + A_j J^j + \frac{1}{2} F^{j0} \partial_0 A_j + \frac{1}{2} A_0 (\partial_j F^{j0} + J^0) + \frac{1}{2} A_0 J^0 \right] \quad (23)$$

The variable A_0 appears quadratically in the action, and since its endpoints are being integrated over in (21), it can be integrated out as a simple Gaussian integral. The result of course is to just substitute the saddle point solution \tilde{A}_0 back into (23).

The saddle point equation for A_0 is just the Gauss law Maxwell equation, viz.,

$$(\partial_j F^{j0} + J^0) = -\partial_j \partial^j A_0 + \partial_0 \partial^j A_j + J^0 = 0. \quad (24)$$

for which the solution is

$$\tilde{A}_0 = \nabla^{-2} J^0 + g + h, \quad (25)$$

where g is given by

$$g = \partial_0 \nabla^{-2} (\partial^j A_j), \quad (26)$$

and where h is an undetermined homogeneous solution to the Laplace equation. The only such solution which is both regular at the origin and vanishes at spatial infinity is the trivial solution, $h(x) = 0$, so we set $h = 0$.

The solution (25) is then unique, without needing to impose further gauge fixing to eliminate the homogeneous solutions. If however we allow for large gauge transformations, then non-trivial expressions for $h(x)$ arise, and further gauge fixing is required. We will discuss this sec. 6.

Notice that \tilde{A}_0 is given in terms of a gauge invariant term $\nabla^{-2} J^0$ and a gauge variant term g . Under gauge transformation g transforms as $\delta_\Lambda g = \partial_0 \Lambda$, as it must, so that \tilde{A}_0 transforms appropriately.

Taking inspiration from this, we formally isolate the gauge invariant part of the components A_j by defining

$$A_j = \mathcal{A}_j + \partial_j \Phi, \quad (27)$$

where $\delta_\Lambda \mathcal{A}_j = 0$, and Φ is a functional of A_j with the assumed transformation property $\delta_\Lambda \Phi = \Lambda$. The functions (\mathcal{A}_j, Φ) are just a new choice of field variables for the path-integration. To avoid introducing a field-dependent Jacobian into the integration measure, we will assume the g -potential Φ to be a linear functional of the A_j . Note that Φ is certainly not given uniquely by the required transformation property: for now we leave it unspecified.

At this point one might assume that \mathcal{A}_j and $\partial_j\Phi$ are just the transverse and longitudinal parts of A_j . This is of course a valid decomposition, but not a unique one. Rather than make this assumption, we will instead rewrite the path-integral in terms of the new variables \mathcal{A}_j and $\partial_j\Phi$, and look for a natural decomposition of the path integral. We will see that the action, transformed to the new variables, ends up separating into a non-dynamical boundary term, plus terms which are uniquely associated with the dynamical matter field and the new field variables $\mathcal{A}_j(x)$.

1. New Variables for the Action and Propagator

We begin by writing the propagator K_{fi} in terms of \mathcal{A}_j in (27) and the solution (25) for \tilde{A}_0 , to get

$$K_{fi} = \int_{q_i}^{q_f} \mathcal{D}q e^{iS_M} \int_{\Phi_i}^{\Phi_f} \mathcal{D}\Phi \int_{\mathcal{A}_{j_i}}^{\mathcal{A}_{j_f}} \mathcal{D}\mathcal{A}_j e^{i\tilde{S}_{EM}}, \quad (28)$$

with a new electromagnetic field action \tilde{S}_{EM} given by

$$\tilde{S}_{EM} = \int_{t_i}^{t_f} d^4x \left[-\frac{1}{4} F_{jk} F^{jk} + \mathcal{A}_j J^j + \partial_j \Phi J^j + \frac{1}{2} \tilde{F}^{j0} \partial_0 \mathcal{A}_j + \frac{1}{2} \tilde{F}^{j0} \partial_0 \partial_j \Phi + \frac{1}{2} J^0 \nabla^{-2} J^0 + g J^0 \right], \quad (29)$$

where we've introduced the notation $\tilde{F}_{j0} = \partial_j \tilde{A}_0 - \partial_0 A_j$. Note that F_{jk} is independent of Φ by antisymmetry.

We can now integrate by parts to strip the spatial derivatives off Φ , to get

$$\tilde{S}_{EM} = \int_{t_i}^{t_f} d^4x \left[\frac{1}{2} \tilde{F}^{j0} \partial_0 \mathcal{A}_j - \frac{1}{4} F_{jk} F^{jk} + \mathcal{A}_j J^j + \frac{1}{2} J^0 \nabla^{-2} J^0 - \Phi \partial_j J^j - \frac{1}{2} \partial_j \tilde{F}^{j0} \partial_0 \Phi + g J^0 \right]. \quad (30)$$

We then use the definition $\partial_j \tilde{F}^{j0} = -J^0$ along with the fact that $\partial_\mu J^\mu$ for an arbitrary trajectory of the particle, to rewrite the action as

$$\tilde{S}_{EM} = \int_{t_i}^{t_f} d^4x \left[\frac{1}{2} \tilde{F}^{j0} \partial_0 \mathcal{A}_j - \frac{1}{4} F_{jk} F^{jk} + \mathcal{A}_j J^j + \frac{1}{2} J^0 \nabla^{-2} J^0 + \Phi \partial_0 J^0 + \frac{1}{2} J^0 \partial_0 \Phi + g J^0 \right], \quad (31)$$

This result reveals something remarkable—if we now make the choice $\partial_0 \Phi = g$ for Φ , then the last three terms sum to a total time derivative. There is of course nothing forcing us to choose this form for Φ ; since we are just making a change of path-integration variable, the final result for the propagator cannot depend on which decomposition we choose. We will make this choice, and since g itself is given as the time derivative of $\nabla^{-2}(\partial_j A^j)$, we can simply choose

$$\Phi = \nabla^{-2}(\partial_j A^j). \quad (32)$$

so that our new field variable now becomes

$$\mathcal{A}_j = A_j - \partial_j \nabla^{-2}(\partial^k A_k) \quad (33)$$

We will find this result ultimately corresponds to the transverse-longitudinal decomposition of A_j ; however, instead of assuming this from the start, we will see that this decomposition is simply dictated by the solution to the A_0 saddle point equation. This pattern of logic will be used again in later sections when we consider geometries for which it is much less clear *a priori* how to define a ‘transverse part’ of A_j .

Let us now complete the process of transforming to the new form for the field action. Note first that our choice of decomposition also simplifies the expression for

the electric field, to

$$\begin{aligned} \tilde{F}_{j0} &= \partial_j \tilde{A}_0 - \partial_0 A_j \\ &= \partial_j \nabla^{-2} J^0 - \partial_0 A_j, \end{aligned} \quad (34)$$

ie., a manifestly gauge invariant form; and it renders \mathcal{A}_j divergenceless.

We then find that a simple integration by parts gives

$$\int_{t_i}^{t_f} d^4x \frac{1}{2} \tilde{F}^{j0} \partial_0 \mathcal{A}_j = \int_{t_i}^{t_f} d^4x \frac{1}{2} \partial_0 \mathcal{A}^j \partial_0 \mathcal{A}_j. \quad (35)$$

so that the field action takes the form

$$\begin{aligned} \tilde{S}_{EM} &= \int_{\partial\mathcal{V}} \sigma d^3x J^0 \nabla^{-2}(\partial_j A^j) \\ &+ \frac{1}{2} \int_{t_i}^{t_f} d^4x \left[-\partial_\mu \mathcal{A}^j \partial^\mu \mathcal{A}_j + 2\mathcal{A}_j J^j + J^0 \nabla^{-2} J^0 \right] \end{aligned} \quad (36)$$

with $\sigma = \pm 1$ for the future and past parts of the boundary respectively.

Let us now combine this field action with the original matter action S_M in (7), to get a complete form for the transformed action, as

$$\tilde{S} = \tilde{S}_M + \tilde{S}_C + \tilde{S}_A \quad (37)$$

with the three new terms defined as follows:

(i) We incorporate the ‘Coulomb self-energy’ term from \tilde{S}_{EM} into the matter action, to give

$$\tilde{S}_M = S_M + \frac{1}{2} \int_{t_i}^{t_f} d^4x J^0 \nabla^{-2} J^0 \quad (38)$$

with S_M given by (7) as before. We see that \tilde{S}_M is gauge invariant.

(ii) The boundary term \tilde{S}_C is a pure phase; we have

$$\begin{aligned} \tilde{S}_C &= \int_{\partial V} \sigma d^3r J^0 \nabla^{-2} (\partial_j A^j) \\ &= -\frac{e}{4\pi} \int_{\partial V} \sigma d^3r d^3r' \frac{\nabla_{\mathbf{r}'} \cdot \mathbf{A}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \end{aligned} \quad (39)$$

and we shall see presently in what way this is related to Dirac’s phase.

(iii) The dynamic part of the EM action - including both the free field term and the interaction with the matter current - is now

$$\tilde{S}_A = \frac{1}{2} \int_{t_i}^{t_f} d^4x [-\partial_\mu A^j \partial^\mu A_j + 2A_j J^j] \quad (40)$$

and we see that, like \tilde{S}_M , this is also gauge invariant.

At the risk of future confusion, we will henceforth omit the σ and leave it implicit that a minus sign should be in front of the integral when integrating over the past part of the boundary. All of the variables in the bulk action are gauge invariant, while the boundary term transforms precisely as we determined it ought to in (14). Also, the g-potential Φ is not at all present in the bulk action: it appears only in the boundary term.

We can now write the propagator K_{fi} in the form that we want. Since Φ does not appear in the bulk action, we can freely integrate over it to yield a harmless overall (divergent) normalization. Doing this, and continuing to absorb field independent constants into the measure, we arrive at our final expression for the propagator:

$$K_{fi} = e^{i\tilde{S}_C} \int_{q_i}^{q_f} \mathcal{D}q e^{i\tilde{S}_M} \int_{\mathcal{A}_i^j}^{\mathcal{A}_f^j} \mathcal{D}\mathcal{A}^j e^{i\tilde{S}_A}. \quad (41)$$

where \mathcal{A}_i^j and \mathcal{A}_f^j are the initial and final configurations of the transverse gauge field $\mathcal{A}^j(x)$.

This equation for K_{fi} is one of our key results - we have shown that the original form (10) for K_{fi} can be rewritten as (41), revealing a Coulomb form for the field dressing the matter, and showing that the transverse field will be determined dynamically by the remaining path integral. The arguments above were clearly independent of a gauge choice since we never explicitly chose a gauge. Thus the Coulomb form \tilde{S}_C in (39) arises naturally from boundary terms in the path-integral, and is not a consequence of choosing a Coulomb gauge.

2. Example: Eikonal Approximation

We do not, in this paper, attempt to discuss detailed examples. However, the lowest-order eikonal approximation does serve to illustrate what one can expect. We give a very heuristic treatment here - more detail is found in, eg., refs. [33, 46, 47]). In this approximation, fluctuations in the charge trajectory about the classical saddle point are neglected in the current. Starting from the effective field term \tilde{S}_A in (40), we write it as $\tilde{S}_A = \tilde{S}_A^0 + \tilde{S}_A^{int}$, where the interaction term is

$$\tilde{S}_A^{int} = \int_{t_i}^{t_f} d^4x \mathcal{A}_j(x) J^j(x) = e \int_{t_i}^{t_f} dt \dot{q}^j(t) \mathcal{A}_j(q(t)) \quad (42)$$

We then expand $q(t)$ as $q(t) = q_{cl}(t) + \delta q(t)$, where the classical trajectory q_{cl} is independent of A_μ , so that

$$\begin{aligned} \tilde{S}_A^{int} &= e \int_{t_i}^{t_f} dt (\dot{q}_{cl}^j(t) + \delta \dot{q}^j(t)) \\ &\quad \times \sum_{n=0}^{\infty} \frac{1}{n!} \partial_{k_1 \dots k_n} \mathcal{A}_j(q_{cl}) \delta q^{k_1} \dots \delta q^{k_n} \end{aligned} \quad (43)$$

If we were to isolate only the long wavelength parts of \mathcal{A}_j , we could truncate the above derivative expansion at $n = 0$; moreover, the high frequency trajectory fluctuations would not effectively couple to these long wavelength parts of \mathcal{A}_j , and the term linear in $\delta \dot{q}^j$ would also be negligible. Discarding these terms is of course only valid for the long-wavelength parts of the gauge field, but if we were to simply carry this through for the whole field then we have effectively performed a lowest order eikonal approximation.

The resulting interaction term would then just involve coupling to the classical path

$$S_{int}^{eik} = e \int_{t_i}^{t_f} dt \dot{q}_{cl}^j(t) \mathcal{A}_j(q_{cl}) = \int_{t_i}^{t_f} d^4x \mathcal{A}_j(x) J_{cl}^j(x). \quad (44)$$

The functional integral for the gauge field coupled to an external classical source can be done exactly, and the resulting functional dependence on \mathcal{A}_j is Gaussian. Assuming the initial state of the gauge field is also some Gaussian state, eg. the vacuum, then the Gaussian form remains even after using the propagator as a kernel to evolve the initial state. The Gaussian functional dependence implies that the out state will generally be a squeezed coherent state of the electric field, and if the initial state is the vacuum state then the resulting squeezed coherent state will be peaked on the classical electric field sourced by J_{cl}^j .

This particular eikonal approximation illustrates very nicely that, aside from the universal Coulomb part of the field, coherent dressed states can be understood in terms of quantum state preparation, as noted in a number of papers [48, 49]. It also gives a concrete method for computing the resulting long-wavelength parts of the dressing for a given state preparation mechanism. The resulting state

is an eigenstate of the longitudinal electric field operator, with eigenvalue corresponding to the Coulomb field, and a squeezed coherent state of the transverse electric field, determined by the dynamics of the charged particle.

We also note that the expectation value for the electric field operator in this out state is precisely that expected classically for an electric field sourced by J_{cl}^μ , ie., the Liénard-Wiechert field [50]; however the correlation functions reveal that the longitudinal and transverse electric fields are treated quite differently in quantum theory. For more detail on all these points, see ref. [51].

D. Boundary Faddeev-Popov Trick

We can actually derive (10) in a more interesting manner, without using off-shell current conservation. We will instead use a technique we refer to as the boundary Faddeev-Popov (bFP) trick. Similar manipulations have previously appeared in [52], but we'll generalize the results to include quantum matter, and also make the gauge independence clear.

We start again from the manifestly gauge-invariant (10), and now explicitly perform the Faddeev-Popov trick to fix a gauge. That is, we multiply the path integral by

$$1 = \int \mathcal{D}\Lambda \Delta[A^\Lambda] \delta(\mathcal{G}(A^\Lambda)), \quad (45)$$

where $\Delta[A^\Lambda] = |\det \delta_\Lambda \mathcal{G}(A^\Lambda)|$ is the FP determinant and $\mathcal{G}(A)$ is the gauge fixing function. This expression involves integration over gauge transformations in the region \mathcal{V} over which the path-integral is occurring, and also over transformations on the boundary time slices $\Sigma_i \cup \Sigma_f$. Transformations residing on the boundary time slices are omitted in textbook applications of the FP trick, where one typically considers vacuum generating functionals with no explicit boundaries.

The resulting integral is just

$$K_{fi} = \int \mathcal{D}\Lambda \int_{q_i}^{q_f} \mathcal{D}q e^{iS_M} \times \int_{A_{\mu i}}^{A_{\mu f}} \mathcal{D}A_\mu \Delta[A^\Lambda] \delta(\mathcal{G}(A^\Lambda)) e^{iS_{EM}[A]} \quad (46)$$

As before, we now consider a change of variables to $A'_\mu = A_\mu^\Lambda = A_\mu + \partial_\mu \Lambda$. The FP determinant is gauge invariant, and the action transforms by a boundary term

$$S_{EM}[A] = S_{EM}[A'] - \int_{\partial\mathcal{V}} d^3x \Lambda J^0. \quad (47)$$

The propagator can now be written as

$$K_{fi} = \int \mathcal{D}\Lambda e^{-i \int_{\partial\mathcal{V}} d^3x \Lambda J^0} \int_{q_i}^{q_f} \mathcal{D}q e^{iS_M} \times \int_{A_{\mu i}^{\Lambda_i}}^{A_{\mu f}^{\Lambda_f}} \mathcal{D}A_\mu \Delta[A] \delta^\mathcal{V}(\mathcal{G}(A)) \delta^{\partial\mathcal{V}}(\mathcal{G}(A^\Lambda)) e^{iS_{EM}[A]} \quad (48)$$

where we now omit the primes in the notation, and use superscripts (\mathcal{V}) and ($\partial\mathcal{V}$) to denote quantities evaluated in the bulk and the boundary respectively. Note that both the boundary data for the gauge field, and the delta function fixing the gauge on the boundaries, are still dependent on the gauge parameter Λ – this of course was not changed by a change of integration variables.

In the standard application of the FP trick one would note that there was no remaining dependence in the path-integral on Λ , and the integral over the gauge group would simply be divided out as overall normalization; but clearly we can't quite do that here.

To proceed, recall that the A_0 integral can be performed unambiguously without need for gauge-fixing. We therefore assume a gauge fixing function which does not involve A_0 , and rewrite the transformed boundary data using a linear shift, using functional derivatives as we did in (15). We define the operator

$$\hat{\mathcal{L}}_\Lambda = \int_{\partial\mathcal{V}} d^3x [\Lambda J^0 + i \partial_\mu \Lambda \frac{\delta}{\delta A_\mu}] \quad (49)$$

which now integrates over both past and future boundaries, and get

$$K_{fi} = \int \mathcal{D}\Lambda \delta^{\partial\mathcal{V}}(\mathcal{G}(A^\Lambda)) e^{-i \hat{\mathcal{L}}_\Lambda} \int_{q_i}^{q_f} \mathcal{D}q e^{iS_M} \times \int_{A_{\mu i}}^{A_{\mu f}} \mathcal{D}A_\mu \Delta[A] \delta^\mathcal{V}(\mathcal{G}(A)) e^{iS_{EM}[A]} \quad (50)$$

The boundary delta function depends only on $\Lambda_{i,f}$ and not time derivatives thereof. The gauge transformations of the boundary data for A_0 are then completely decoupled from the transformations of the remaining A_j . In a time-sliced discretization of the path integral, the transformation involving Λ on the slices immediately after Σ_i and Σ_f will only affect the transformation of A_0 on the boundary. Additionally, there is no dependence in the integrand on Λ for any intermediate times. This ‘‘bulk’’ integration over the gauge group can be factored out as usual, leaving a residual integration over boundary gauge transformations.

The net result is that in (50) we can rewrite $\hat{\mathcal{L}}_\Lambda$ as

$$\hat{\mathcal{L}}_\Lambda \rightarrow \int_{\partial\mathcal{V}} d^3x [\Lambda J^0 + i \partial_j \Lambda \frac{\delta}{\delta A_j}] \quad (51)$$

and omit the boundary data for A_0 . The omission of A_0 boundary data dictates that its values on the boundary are integrated over.

We can use the delta function to evaluate the integral over the boundary gauge transformations, and this will fix the boundary phase. Assuming \mathcal{G} is a good gauge fixing function, it will correspond to a unique gauge parameter $\Lambda = \Lambda_{\mathcal{G}}[A]$. Evaluating the integral over the

boundary gauge transformation we then obtain

$$K_{fi} = e^{-i\hat{\mathcal{L}}_{\Lambda_g}} \int_{q_i}^{q_f} \mathcal{D}q e^{iS_M} \times \int_{A_{j_i}}^{A_{j_f}} \mathcal{D}A_\mu \Delta[A] \delta^\nu(\mathcal{G}(A)) e^{iS_{EM}[A]} \quad (52)$$

where now

$$\hat{\mathcal{L}}_{\Lambda_g} = \int_{\partial V} d^3x \Lambda_g[A] \left[J^0 + i\partial_j \Lambda \frac{\delta}{\delta A_j} \right] \quad (53)$$

The difference between the bFP trick and the usual FP technique is clear from (52). While the path integral integrand itself is standard, the additional boundary phase effects a particular gauge transformation of the boundary data, which depends on the choice of bulk gauge fixing function \mathcal{G} . This boundary phase ensures that the resulting propagator remains independent of the choice of gauge fixing; it remains a gauge invariant object.

Since the propagator is independent of gauge choice, we can choose the most convenient gauge. The argumentation is then similar to what we did earlier. We first recall that after the A_0 integration, and the change of variables to the invariant fields \mathcal{A}_j and Φ , we're left with an effective action (30). Great simplification came if we then chose $\partial_0 \Phi = g$, where g given in (26) was the unique gauge-dependent part of the saddle point solution \tilde{A}_0 . Additionally, a few more terms in the effective action which involved g and the current summed to a total derivative after using off-shell current conservation.

We could actually skip the off-shell current conservation argument at this point, by simply choosing the Coulomb gauge $\mathcal{G}(A) = \partial^j A_j$. The particular usefulness of this gauge choice is that it sets $g = \Phi = 0$, considerably simplifying the action. It also makes the FP determinant irrelevant, and implies

$$\Lambda_g[A] = -\nabla^{-2} \partial^j A_j, \quad (54)$$

for our boundary phases.

The resulting expression for the propagator is

$$K_{fi} = e^{i \int_{\partial V} d^3x \nabla^{-2} (\partial^k A_k)} \left[J^0 - i\partial_j \frac{\delta}{\delta A_j} \right] \times \int_{q_i}^{q_f} \mathcal{D}q e^{i\tilde{S}_M} \int_{\tilde{A}_{j_i}}^{\tilde{A}_{j_f}} \mathcal{D}\tilde{A}_\mu e^{i\tilde{S}_A[\tilde{A}]} \quad (55)$$

in which we now write things in terms of the effective actions \tilde{S}_M and \tilde{S}_A , as in (41).

We can show the equivalence of this result to (41) by noting that the remaining path-integral is independent of the longitudinal part of the gauge field. In the shift operator, the functional derivative $\partial_j \frac{\delta}{\delta A_j}$ then vanishes and we're left with an expression for the propagator K_{fi} in the same form as (41) above, but with \tilde{S}_C now written as

$$\tilde{S}_C = \int_{\partial V} d^3x A_k \partial_k (\nabla^{-2} J^0), \quad (56)$$

ie., as an integration by parts over the expression for \tilde{S}_C in (39). Again, we find that the charge is dressed by a Coulomb field.

This concludes our analysis of the propagator K_{fi} . We stress again that the propagator is gauge-independent, and its form is a simple consequence of the gauge invariance of the effective action, rather than being imposed *a priori*.

III. SPINOR QUANTUM ELECTRODYNAMICS

We have devoted considerable space to scalar electrodynamics; it is now fairly straightforward to generalize to real QED, with Dirac spinors coupled to the EM field. Again we will consider the manifestly gauge invariant path-integral for K_{fi} , and we will find the same Coulomb form for the dressing. The manipulations similar to those for scalar electrodynamics, the only difference being that the matter field also changes under gauge transformation, and the $U(1)$ charge current in the boundary phase will become an operator.

The gauge invariant path integral representation of the transition amplitude is now

$$K_{fi} = \int_{\psi_i}^{\psi_f} \mathcal{D}\psi \mathcal{D}\bar{\psi} \int_{A_{\mu_i}}^{A_{\mu_f}} \mathcal{D}A_\mu e^{iS[A, \psi, \bar{\psi}]}, \quad (57)$$

where $\psi, \bar{\psi}$ are Grassmann fields, and the omission of boundary data for $\bar{\psi}$ indicates that this variable is to be integrated over on the boundary—necessary because the Dirac Lagrangian has a first-order form. The action is the QED action with a single Dirac fermion field of charge e , viz.,

$$S[A, \psi, \bar{\psi}] = \int_{t_i}^{t_f} d^4x \left[-\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \bar{\psi} (\gamma^\mu \partial_\mu - ie\gamma^\mu A_\mu + m) \psi \right] \quad (58)$$

This action is completely invariant, without need to discard a boundary term, under the $U(1)$ gauge transformation

$$A_\mu \rightarrow A_\mu + \partial_\mu \Lambda \\ \psi \rightarrow e^{ie\Lambda} \psi, \quad \bar{\psi} \rightarrow e^{-ie\Lambda} \bar{\psi}. \quad (59)$$

One can easily verify that the propagator (57) is gauge invariant in the same way done in the previous section, ie. by transforming its data, undoing the transformation by a change of variables in the path integral, and using the invariance of the action.

The gauge invariance of the propagator implies that it satisfies the condition

$$\int_{\partial V} d^3x \left[ie\Lambda \psi \frac{\delta}{\delta \psi} + \partial_\mu \Lambda \frac{\delta}{\delta A_\mu} \right] K_{fi} = 0. \quad (60)$$

By explicitly differentiating the path integral we can confirm that the functional derivatives are proportional to the conjugate momenta for the fields:

$$\frac{\delta}{\delta\psi_{i,f}} = \pm i\hat{\Pi}_{i,f} = \mp\hat{\psi}_{i,f}^\dagger, \quad (61)$$

$$\frac{\delta}{\delta A_{j i,f}} = \mp i\hat{\Pi}_{i,f}^j = \pm i\hat{E}_{i,f}^j. \quad (62)$$

Together with the expression for the U(1) charge density $J^0 = i\bar{\psi}\gamma^0\psi = -\psi\psi^\dagger$, and the invariance condition (60), this implies the propagator satisfies the operator constraints

$$(\partial_j\hat{E}^j - \hat{J}^0)K(A, \psi) = 0 \quad (63)$$

$$\frac{\delta}{\delta A_0}K(A, \psi) = 0, \quad (64)$$

on both the future and past boundary time slices.

We now proceed to use the bFP trick to see exactly how this constraint is implemented, i.e., how the electric field dressing of the states emerges. Starting from K_{fi} we again insert a gauge fixing function by multiplying by (45), but now we must change variables for both the gauge field and the Dirac field if the action is to be invariant:

$$K_{fi} = \int \mathcal{D}\Lambda \int_{\psi_i^{\Lambda_i}}^{\psi_f^{\Lambda_f}} \mathcal{D}\psi\mathcal{D}\bar{\psi} \int_{A_{\mu i}^{\Lambda_i}}^{A_{\mu f}^{\Lambda_f}} \mathcal{D}A_\mu \times \Delta[A]\delta^\nu(\mathcal{G}(A))\delta^{\partial\nu}(\mathcal{G}(A^\Lambda))e^{iS[A,\psi,\bar{\psi}]} \quad (65)$$

The A_0 integral can again be done without gauge fixing, and we can extract the transformations of the boundary data using exponentiations of the functional derivatives,

$$K_{fi} = \int \mathcal{D}\Lambda \delta^{\partial\nu}(\mathcal{G}(A^\Lambda))e^{\hat{\mathcal{L}}_\Lambda} \times \int_{\psi_i}^{\psi_f} \mathcal{D}\psi\mathcal{D}\bar{\psi} \int_{A_{\mu i}}^{A_{\mu f}} \mathcal{D}A_\mu \Delta[A]\delta^\nu(\mathcal{G}(A))e^{iS[A,\psi,\bar{\psi}]}. \quad (66)$$

in which the operator $\hat{\mathcal{L}}_\Lambda$ now takes the form

$$\begin{aligned} \hat{\mathcal{L}}_\Lambda &= \int_{\partial\nu} d^3x \left[ie\psi \frac{\delta}{\delta\psi} + \partial_\mu \Lambda \frac{\delta}{\delta A_\mu} \right] \\ &= \int_{\partial\nu} d^3x \left[\Lambda (\partial_j \hat{E}^j - \hat{J}^0) - i\partial_0 \Lambda \frac{\delta}{\delta A_0} \right] \end{aligned} \quad (67)$$

where the 2nd expression uses the relations (61), (62).

From this stage onwards, the manipulations are identical to those in the previous section except that the charge density in the boundary phase is an operator rather than a c-number. The resulting expression for the propagator

is

$$K_{fi} = e^{i\hat{\mathcal{L}}_\Lambda \mathcal{G}} \int_{\psi_i}^{\psi_f} \mathcal{D}\psi\mathcal{D}\bar{\psi} \times \int_{A_{j i}}^{A_{j f}} \mathcal{D}A_\mu \Delta^\nu[A]\delta^\nu(\mathcal{G}(A))e^{iS[A,\psi,\bar{\psi}]} \quad (68)$$

where now

$$\hat{\mathcal{L}}_\Lambda = \int_{\partial\nu} d^3x \Lambda \mathcal{G}[A] [\partial_j \hat{E}^j - \hat{J}^0] \quad (69)$$

The final expression for the propagator will of course be independent of choice of $\mathcal{G}(A)$. For formal manipulations the most convenient choice is Coulomb gauge, because this sets the g-potential $\nabla^{-2}\partial^j A_j$ to zero, leaving only the invariant field components, and $\Lambda_G[A] = -\nabla^{-2}(\partial^j A_j)$. Since this choice eliminates the dependence of the integral on the longitudinal part of A_j , the shift operator $\exp(i\int_{\partial\nu} \Lambda \partial_j \hat{E}^j)$ will just give zero, and the propagator is then

$$K_{fi} = e^{i\tilde{S}_C} \int_{\psi_i}^{\psi_f} \mathcal{D}\psi\mathcal{D}\bar{\psi} e^{i\tilde{S}_M} \int_{A_{j i}}^{A_{j f}} \mathcal{D}A_j e^{i\tilde{S}_A}. \quad (70)$$

i.e., of the same form as (41) except that now the matter action is

$$\tilde{S}_M = \int_{t_i}^{t_f} d^4x \left[-\bar{\psi}(\gamma^\mu \partial_\mu + m)\psi + \frac{1}{2}J^0 \nabla^{-2}J^0 \right] \quad (71)$$

and the dynamic gauge field action is as before (cf. eqtn. (33)), except that now the matter current is $J^j = ie\bar{\psi}\gamma^j\psi$.

We emphasize that if we had chosen a different gauge fixing function $\mathcal{G}(A)$, the resulting gauge fixed action would look different, but this difference would only be temporary; the shift operator in (69) would no longer give zero in any other gauge, instead enacting a gauge transformation which would return the action to the form (37), with the three terms given by eqtns. (38)-(40). In this form the theory is not manifestly Lorentz invariant, but this is simply because we evaluated K_{fi} between two constant t surfaces. In principle, one could chose a covariant gauge to compute the path-integral as long as one also evaluates the necessary shift of the longitudinal mode in the final expression.

Note, in this connection, that we could take the expression in eqtn. (70) one step further if we explicitly act with the U(1) transformation sitting outside the path-integral. This locally rotates the boundary data for the Dirac field by an angle which depends on the longitudinal part of the gauge field, giving our final expression for the gauge invariant QED propagator on the extended configuration space,

$$K_{fi} = \int_{e^{-ie\nabla^{-2}\partial^j A_{j i} \psi_i}}^{e^{-ie\nabla^{-2}\partial^j A_{j f} \psi_f}} \mathcal{D}\psi\mathcal{D}\bar{\psi} \int_{A_{j i}}^{A_{j f}} \mathcal{D}A_j e^{i\tilde{S}_M + i\tilde{S}_A}. \quad (72)$$

Expanding the shorthand notation this reads

$$e^{-ie\nabla^{-2}\partial^j A_j}\psi = \exp\left(i\int d^3y A_j(y)\frac{e}{4\pi}\frac{y^j - x^j}{|y - x|^3}\right)\psi(x). \quad (73)$$

The gauge invariant propagator dresses every point excitation of the Dirac field by a Coulomb electric field sourced by the corresponding point charge. This is the central result of applying the bFP trick to QED. As with the particle considered in the first section, the transverse dressing will be determined dynamically by the remaining integral over gauge invariant variables.

IV. FLAT SPACETIME EVOLUTION IN A CAUSAL DIAMOND

Up to now we have dealt with the rather simple problem of QED on a flat background, defined between time slices. However it is clearly crucial to be able to discuss this for much more general kinds of boundary and boundary information. In principle this should extend to spacetimes including achronal regions; as noted in the introduction, discussions of this sort of problem began in the 1980's [5, 6, 10, 29, 30].

To give such a generalization has also been the goal of the “general boundary quantum field theory framework” [7, 8], where one considers general spacetime regions \mathcal{V} bounded by some boundary hypersurface $\partial\mathcal{V}$. A field configuration on $\partial\mathcal{V}$ is then mapped to an amplitude via a path integral over field configurations in \mathcal{V} .

In this formulation, as in the work of Hartle and Hawking [10], states can then be defined as non-local wave functionals, over configurations specified on all of $\partial\mathcal{V}$. If the general boundary hypersurface involves a union of future and past surfaces, then one can still interpret such states on $\partial\mathcal{V}$ as a transition amplitude [7]. However, for more general spacetimes, such an interpretation is not valid, although one can obtain a probabilistic interpretation of the modulus squared of the state on $\partial\mathcal{V}$ in terms of a conditional probability to find a given field configuration on a subregion $\Sigma \subset \partial\mathcal{V}$, and another specified field configuration on the complementary region $\bar{\Sigma}$.

As we have seen, the approach in this paper to defining states is rather different. In our view such an approach is essential for general spacetimes - one of the ultimate motivations in the present work is to set up a technique which can be used for achronal spacetimes, in which information fixed on just the past time slice is not always sufficient to predict quantum evolution [29]. In our opinion such a technique will also be necessary to properly address issues concerning black hole information content.

In pursuit of this end, in the present section we apply the bFP trick to amplitudes for a more general boundary hypersurface. The region we will consider is a finite-sized causal diamond in Minkowski spacetime, where the state is fixed on the null boundary hypersurface. Here there is still a natural splitting into past and future sections, and

so we can define a propagator which represents a transition amplitude between states on the past and future null cones (which tend to null infinity for an infinitely large diamond).

A technical note—there is a subtlety here in the specification of boundary data for the path-integral. Because the conjugate momentum on a null surface involves a derivative along that surface, specifying the field configuration also specifies the conjugate momentum. Specifying this data on both the past and future boundaries would be an over-specification of boundary data for the classical evolution, and the corresponding interpretation as a quantum amplitude is then unclear.

One fixes this by specifying “half” of the field data in some chosen way [53]. We will assume throughout that it is only the positive frequency parts of the field which are specified: a choice which gives this amplitude an interpretation in terms of states in the Bargmann representation, ie. coherent states. In the following discussion we will avoid making this explicit, so as not to clutter the notation.

A. Formulation of the Problem

In the time-slice geometry, the variable A_0 was ultimately unphysical, and the remaining variables A_j split into purely physical transverse and pure gauge longitudinal parts - the transverse part being divergenceless, ie., $\partial^j A_j = 0$. For more general boundary hypersurfaces, a natural idea would be to continue to decompose the field into parts with and without divergence. This is not possible, for 2 reasons. First, as before, there is still the issue of uniqueness - given a transverse-longitudinal decomposition of the vector field, one can freely add some transverse parts onto the longitudinal part and the result still transforms correctly under gauge transformation. Second, on null hypersurfaces there is no unique notion of divergence - the induced metric is degenerate, and so there is no unique inverse metric with which to define the divergence $h^{jk}\nabla_j A_k$.

For these reasons we again use a procedure whereby the path integral is used to generate a unique decomposition into pure gauge and gauge invariant parts of the field.

We recall that for flat timeslice boundaries, the boundary data of the component A_0 was integrated over. The saddle point solution for this Gaussian integral determined the g-potential Φ , ie., the functional of A_j transforming as $\delta_\Lambda \Phi = \Lambda$; the pure gauge part of \tilde{A}_0 was the time derivative of Φ , and the longitudinal part A_j was the gradient of Φ . For more general boundaries we will then need to single out the component of A_μ normal to the boundary hypersurface. This component will play the same role as A_0 , and the pure gauge part of its solution will yield a corresponding g-potential.

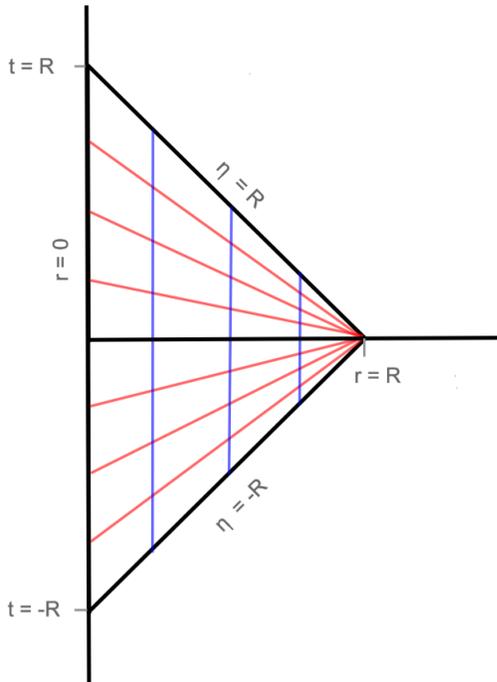


FIG. 2: The r, η -coordinates. Each point represents a two-sphere of radius r . This is a standard Minkowski spacetime diagram, not a conformal diagram. The blue lines are lines of constant r , while the red lines are lines of constant η .

1. Coordination specification

To implement these ideas we need to choose coordinates appropriately. We pick hypersurface adapted coordinates $x^\mu = \{S, y^k\}$ such that $S = \text{const.}$ surfaces foliate the spacetime region \mathcal{V} , and the boundary hypersurface $\partial\mathcal{V}$ is described by particular values, $S = S_i, S_f$. Then, using a coordinate basis it is A_S which is the component generalizing A_0 , because the pullback of $A_\mu dx^\mu$ to $\partial\mathcal{V}$ will be independent of A_S .

For a finite size causal diamond in Minkowski spacetime we then need to construct coordinates adapted to the boundary null cones. The coordinates we will use are rather intuitive. Consider a sphere of radius R at time $t = 0$, and from each solid angle send an inwards going radial null ray to the future and to the past. These null geodesics will converge at $r = 0$ at times $t = R$ and $t = -R$ respectively, and the surface generated by the null rays is the boundary of our causal diamond.

To construct coordinates in the interior we again start from the sphere $r = R$ at $t = 0$, and now send inwards going spacelike rays to the future and past. These spacelike rays converge at $r = 0$ but at times t dictated by their “velocities”. The angles and radii of spheres are still useful coordinates, but now we will replace the time coordinate t with a coordinate parameterizing the “velocity” of each ray.

Each of the rays joining $r = 0$ to $r = R$ is described

by a solid angle and t, r satisfying the simple relation

$$t = \eta f(r), \quad (74)$$

for

$$f(r) = 1 - \frac{r}{R}, \quad (75)$$

and for some $\eta \in [-R, R]$. From this relation we can quickly verify that the surfaces $\eta = \pm R$ are the future and past null boundaries of the causal diamond.

Inside the boundary, η parameterizes spacelike surfaces and thus serves as a useful time coordinate. Thus, as desired, we’ve found hypersurface adapted coordinates where certain values of “time” denote the boundary. We can straightforwardly compute the metric in these coordinates:

$$ds^2 = -f(r)^2 d\eta^2 + 2\frac{\eta}{R} f(r) d\eta dr + \left(1 - \frac{\eta^2}{R^2}\right) dr^2 + r^2 d\Omega^2 \quad (76)$$

where $d\Omega^2$ is the standard line element on the unit 2-sphere.

It is clear from this expression that $\eta = 0$ is just a standard time slice of Minkowski spacetime and that $\eta = \pm R$ are null hypersurfaces. Since $f(r)$ vanishes at $r = R$, there is a coordinate singularity. This is obvious from (2), and indeed several components of the inverse metric will diverge as $r \rightarrow R$.

To deal with this we need to recall why we are interested in this geometry. Ultimately we wish to take R to be larger than all other length scales. The sphere $r = R$ then resembles spatial infinity, and the surfaces $\eta = \pm R$ resemble null infinity. As long as we don’t take the strict limit $R \rightarrow \infty$, we can still specify data for massive fields on the boundary. The boundary considered here then plays a role similar to null infinity, but is not obtained via conformal compactification. Timelike worldlines will be able to connect all points interior to some point on the boundary.

Since the electromagnetic field is massless we expect field excitations to reach null infinity but we do not expect the same for spatial infinity. For this reason we make the assumption that all important quantities will vanish sufficiently fast for $r \rightarrow R$, while allowing for finite limits as $\eta \rightarrow \pm R$. We assume the fall-offs are sufficiently rapid that we can restrict $r \ll R$ throughout, and allow the metric to take the simple form

$$ds^2 = -d\eta^2 + 2\frac{\eta}{R} d\eta dr + \left(1 - \frac{\eta^2}{R^2}\right) dr^2 + r^2 d\Omega^2 \quad (77)$$

For reference, the non-zero inverse metric components are

$$\begin{aligned} g^{\eta\eta} &= -(1 - \frac{\eta^2}{R^2}), & g^{\eta r} &= \frac{\eta}{R}, \\ g^{rr} &= 1, & g^{AB} &= r^{-2} q^{AB} \end{aligned} \quad (78)$$

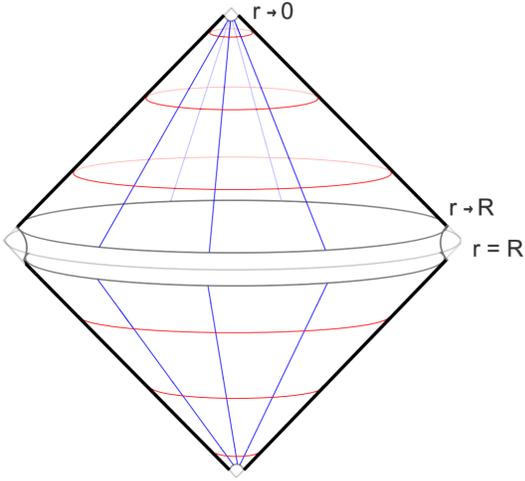


FIG. 3: The blow-up procedure which treats the corners of the causal diamond geometry. The black lines show regions of the boundary, the lightest grey lines show the true causal diamond, and the darker grey lines show the boundaries of the boundary, which in the infinite limit coincide with the true causal diamond. Blue lines follow null generators of the boundary, while red lines denote constant t cuts.

where x^A are sphere coordinates, and q^{AB} is the inverse metric on the unit 2-sphere.

We will formally “blow up” this surface, i.e., excise the sphere $r = R$ from the boundary and consider the boundary as an open set where limits $r \rightarrow R$ can now be η dependent. Note that for all values of η , the boundary region $r = R$ has relative measure zero. Thus when spatially integrating by parts, both $r = 0$ and $r = R$ will be zero volume surfaces, and we can then discard any spatial surface terms.

This deals with the singular behaviour of the spatial “corner” of the boundary hypersurface, but there are still the corners at the top and bottom of the causal diamond, $r = 0, \eta = \pm R$. We will also formally blow up these points to allow fields to take angle dependent limits as $r \rightarrow 0$ on the boundary; see FIG. 3. In doing this, we as-

sume nothing enters or leaves \mathcal{V} through the strict points $r = 0, \eta = \pm R$.

If one now considers a QED propagator with information specified on the boundary of this causal diamond, the transformation of the component A_η involves $\partial_\eta \Lambda$, i.e., a derivative normal to surfaces of constant η and thus independent of the actual pullback of Λ to the surface. Thus any boundary data specified for A_η in the path integral will be superfluous. In addition the QED Lagrangian will be quadratic in A_η , allowing it to be integrated out via Gaussian saddle point substitution.

2. Transformed Effective Action

For brevity we just consider the gauge field coupled to a conserved external source J^μ ; this is easily generalize to scalar charged particles or to a Dirac field by promoting J^μ in the resulting boundary phase to an operator. As before, we first obtain results without explicitly fixing a gauge, then discuss how the bFP trick shortcuts the computation. Expanding the action so as to explicitly write A_η we have

$$S = -\frac{1}{2} \int_{\mathcal{V}} d^4x \sqrt{g} \left[F^{jk} \partial_j A_k - 2A_j J^j + F^{\eta j} \partial_\eta A_j - A_\eta \left(\frac{1}{\sqrt{g}} \partial_j (\sqrt{g} F^{j\eta}) + J^\eta \right) - A_\eta J^\eta \right] \quad (79)$$

where $\sqrt{g} = r^2 \sin \theta$, and $j = \{r, \theta, \phi\}$. In writing this we’ve already freely integrated by parts in spatial directions. To integrate out A_η , we need to solve its saddle point equation, i.e.

$$\frac{1}{\sqrt{g}} \partial_j (\sqrt{g} F^{j\eta}) + J^\eta = 0. \quad (80)$$

Since the metric is non-diagonal, the resulting equation is qualitatively different from the previous equation for A_0 . In terms of A_η the equation of motion reads

$$\partial_r (\sqrt{g} \partial_r A_\eta) - g^{\eta\eta} \partial_A (\sqrt{g} g^{AB} \partial_B A_\eta) = g^{\eta r} \partial_A (\sqrt{g} g^{AB} F_{Br}) + \sqrt{g} J^\eta + \partial_r (\sqrt{g} \partial_\eta A_r) - g^{\eta\eta} \partial_A (\sqrt{g} g^{AB} \partial_\eta A_B) \quad (81)$$

On the RHS the first two terms are obviously gauge invariant, and the last two terms together transform so that the solution to this equation, \hat{A}_η , will transform as $\delta_\Lambda \hat{A}_\eta = \partial_\eta \Lambda$.

Note that $\partial_\eta g^{\eta\eta} = 2\eta/R^2$, a dimensionful quantity of order R^{-1} . By our original assumptions, R is parametrically much larger than any other dimensionful quantity and thus this entire term is sub-leading. With R suffi-

ciently large we can simply assume $\partial_\eta g^{\eta\eta} = 0$, allowing (81) to be written compactly as

$$D^j \partial_j A_\eta = \frac{1}{\sqrt{g}} g^{\eta r} \partial_A (\sqrt{g} g^{AB} F_{Br}) + J^\eta + \partial_\eta D^j A_j, \quad (82)$$

where we’ve defined the divergence-like differential oper-

ator D^j , acting as

$$D^j w_j = \frac{1}{\sqrt{g}} \partial_r (\sqrt{g} w_r) - g^{\eta\eta} \frac{1}{\sqrt{g}} \partial_A (\sqrt{g} g^{AB} w_B). \quad (83)$$

Now (81) can formally be solved by assuming a Green's function G satisfying

$$D^j \partial_j G(x, x') = \frac{\delta^3(x - x')}{\sqrt{g}}, \quad (84)$$

that is,

$$\tilde{A}_\eta = \tilde{A}_\eta^I + g + h, \quad (85)$$

where

$$\tilde{A}_\eta^I = \int_{\Sigma_\eta} d^3 x' \sqrt{g} G \left[\frac{1}{\sqrt{g}} g^{\eta r} \partial_A (\sqrt{g} g^{AB} F_{Br}) + J^\eta \right] \quad (86)$$

$$g = \partial_\eta \int_{\Sigma_\eta} d^3 x' \sqrt{g} G D^j A_j, \quad (87)$$

and h is a homogeneous solution $D^j \partial_j h = 0$. The integration in these expressions is over Σ_η , the constant η hypersurface corresponding to the time η at which \tilde{A}_η is being evaluated.

We don't have a general expression for this Green's function; however the results that we're interested in will ultimately only depend on its value on the null boundary, and one can find G on this boundary as well as at $\eta = 0$. At $\eta = 0$, $g^{\eta\eta} = -1$, and the differential operator simplifies to

$$D^j \partial_j f(x)|_{\eta=0} = \frac{1}{\sqrt{g}} \left[\partial_r (\sqrt{g} \partial_r f(x)) + \partial_A (\sqrt{g} g^{AB} \partial_B f(x)) \right] \quad (88)$$

which is of course just the standard Laplacian in spherical coordinates. This is because the hypersurface $\eta = 0$ is just the hypersurface $t = 0$. Thus at $\eta = 0$ the Green's function is given by

$$G(x, x')|_{\eta=0} = -\frac{1}{4\pi} \frac{1}{|x - x'|}. \quad (89)$$

At the boundary, the operator $D^j \partial_j$ simplifies considerably because $g^{\eta\eta}$ vanishes; we then have

$$D^j \partial_j f|_{\eta=\pm R} = \frac{1}{\sqrt{g}} \partial_r (\sqrt{g} \partial_r f), \quad (90)$$

which can be immediately integrated to find the Green's function

$$G(x, x')|_{\eta=\pm R} = \frac{\delta^2(x^A - x^{A'})}{\sin \theta} \theta(r' - r) \left[\frac{1}{r} - \frac{1}{r'} \right]. \quad (91)$$

which propagates along the null generators of the boundary.

Note that the boundary condition for G is chosen so that influence propagates towards smaller radii, i.e. causally on the future portion of $\partial\mathcal{V}$. When considering the past portion of $\partial\mathcal{V}$ one must flip the argument of the step function appropriately.

More progress can be made when looking at the homogeneous solutions. A general homogeneous solution, $D^j \partial_j h = 0$, will have the form

$$h(x) = \sum_{m,l} Y_l^m(\theta, \phi) \left[c_{ml}^1(\eta) r^{-\frac{1}{2} + \sqrt{\frac{1}{4} - g^{\eta\eta} l(l+1)}} + c_{ml}^2(\eta) r^{-\frac{1}{2} - \sqrt{\frac{1}{4} - g^{\eta\eta} l(l+1)}} \right] \quad (92)$$

with Y_l^m a spherical harmonic and $c_{ml}^{1,2}$ a set of time dependent coefficients.

We can immediately set $c_{ml}^2 = 0$, since it is the coefficient of a term which will never be regular at the origin. The other term will either grow monotonically with r or be constant in r . With our assumptions that the fields vanish at large r , both situations are unacceptable and we can set $c_{ml}^1 = 0$. The solution (85) with $h = 0$ is then the unique solution satisfying the boundary conditions.

As an aside, note that if we relax the asymptotic spatial boundary conditions and simply demand for the fields to be finite as $r \rightarrow \infty$, we can accept solutions that are independent of r . Such solutions satisfy

$$-g^{\eta\eta} l(l+1) = 0. \quad (93)$$

For all spacelike slices, $g^{\eta\eta} < 0$, and the only solution is $l = 0$, i.e. a constant function of θ, ϕ, r . These are the time dependent global $U(1)$ rotations. However on the null boundaries $g^{\eta\eta} = 0$, and the homogeneous solution space is enlarged to include any function on the sphere. This is interesting in the context of large gauge transformations, soft photons, etc, and we will return to this point in section 5.

Returning to the solution (85), note that the gauge-variant part g transforms as $\delta_\Lambda g = \partial_\eta \Lambda$. From (87) we see we can identify it as a g -potential of form $g = \partial_\eta \Phi$ with Φ given by

$$\Phi = \int_{\Sigma_\eta} d^3 x' \sqrt{g} G D^j A_j. \quad (94)$$

For the causal diamond we can now decompose the gauge field into a gauge-invariant part $\mathcal{A}_j = A_j - \partial_j \Phi$, and a pure gauge part $\partial_j \Phi$; the subsequent development then parallels to that for the time slice. We substitute \tilde{A}_η into the action (79) and rewrite the action in the new variables \mathcal{A}_j, Φ . Using current conservation, we then get an effective action

$$\begin{aligned} \tilde{S} = & \int_{\partial\mathcal{V}} d^3x \sqrt{g} \Phi J^\eta - \frac{1}{2} \int_{\mathcal{V}} d^4x \sqrt{g} \left[\tilde{F}^{\mu j} \partial_\mu \mathcal{A}_j - 2\mathcal{A}_j J^j \right. \\ & \left. - J^\eta \int_{\Sigma_\eta} d^3x' \sqrt{g} G \left(J^\eta + \frac{1}{\sqrt{g}} g^{\eta r} \partial_A (\sqrt{g} g^{AB} F_{Br}) \right) \right], \end{aligned} \quad (95)$$

with

$$\tilde{F}^{\mu j} = \partial^\mu \tilde{A}^j - \partial^j \tilde{A}^\mu. \quad (96)$$

Note that all of the terms involving Φ again summed to a total time derivative, and thus formed a boundary term in the action. The remaining bulk action is written in terms of explicitly gauge invariant variables.

We can actually take this expression further because the variable $\mathcal{A}_j = A_j - \partial_j \Phi$ is actually transverse in the sense that $D^j \mathcal{A}_j = 0$. Using this, and a few spatial integrations by parts, we expand the effective action in terms of the gauge invariant variables to get

$$\begin{aligned} \tilde{S} = & \int_{\partial\mathcal{V}} d^3x \sqrt{g} \Phi J^\eta + \frac{1}{2} \int_{\mathcal{V}} d^4x \sqrt{g} \left[\partial_\eta \mathcal{A}_r \partial_\eta \mathcal{A}_r - g^{\eta\eta} g^{AB} \partial_\eta \mathcal{A}_A \partial_\eta \mathcal{A}_B - 2g^{\eta r} g^{AB} (\partial_\eta \mathcal{A}_A) F_{rB} \right. \\ & \left. - F^{AB} \partial_A \mathcal{A}_B + g^{AB} F_{rA} F_{rB} + 2\mathcal{A}_j J^j \right. \\ & \left. + \left(J^\eta + \frac{1}{\sqrt{g}} g^{\eta r} \partial_A (\sqrt{g} g^{AB} F_{Br}) \right) \int_{\Sigma_\eta} d^3x' \sqrt{g} G \left(J^\eta + \frac{1}{\sqrt{g}} g^{\eta r} \partial_C (\sqrt{g} g^{CD} F_{Dr}) \right) \right] \end{aligned} \quad (97)$$

This is the expression we will work with - although we will not actually perform computations with this action. The purpose of the derivation was rather to demonstrate that when propagators are considered for different boundary geometries, we can still unambiguously extract a boundary term describing the dressing required to make charged states gauge invariant.

As expected the action (97) is non-local in space. The ‘‘Coulomb’’ interaction term now contains not just the charge density J^η but also terms describing the magnetic field. These apparent interactions arise because our coordinates are no longer adapted to the isometries of Minkowski spacetime. One can set $\eta = 0$, and thus $g^{\eta r} = 0$, to verify that on a standard constant time slice, this gives the usual Lagrangian.

B. Form of the Propagator

Since Φ doesn’t appear in the integrand, the path integral over Φ can again be removed by a Faddeev-Popov procedure. The propagator

$$K(A_\mu \partial\mathcal{V}) = \int_{A_\mu \partial\mathcal{V}} \mathcal{D}A_\mu e^{i \int_{\mathcal{V}} d^4x \sqrt{g} \left[-\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + A_\mu J^\mu \right]} \quad (98)$$

for evolution of the gauge field coupled to a source J^μ , through a large causal diamond, is then equal to

$$K(A_\mu \partial\mathcal{V}) = e^{i \int_{\partial\mathcal{V}} J^\eta \Phi} \int_{\mathcal{A}_j \partial\mathcal{V}} \mathcal{D}\mathcal{A}_j e^{i \tilde{S}[\tilde{A}_j | J]}, \quad (99)$$

where the effective action \tilde{S} is given by the bulk part of (97), and the prefactor involves the generalized Coulomb dressing, in which Φ is given by eqtns. (94) and (91) evaluated on the boundary. The contribution from the future part reads

$$\Phi|_{\partial\mathcal{V}}(r', x'^A) = \int_{r'}^{\infty} dr \left(\frac{1}{r'} - \frac{1}{r} \right) \partial_r (r^2 A_r(r, x'^A)), \quad (100)$$

whereas on the past part the integration is over all r interior to r' .

This dressing describes the radial electric field at each point on $\partial\mathcal{V}$, with a strength determined by the total charge flux through $\partial\mathcal{V}$ at earlier times. This is our central result for the causal diamond geometry.

We emphasize again that this result is not the result of a specific gauge choice, and that the definition of gauge-invariant variables \mathcal{A}_j again emerged naturally from the path integral. Remarkably, our procedure succeeded even though there is no unambiguous notion of the ‘transverse’ vector field, since one cannot define an intrinsic divergence on a null boundary.

If we now give the matter current J^μ its own dynamics, we can easily generalize the above derivation. This is possible because $U(1)$ charge current is conserved off shell for particles. Alternatively, as before, we can go back and skip the step which invokes current conservation by using the bFP trick. The derivations are as before; for Dirac fermions we then get the gauge invariant QED amplitude on the large causal diamond to be

$$K(A_{\mu\partial\nu}, \psi_{\partial\nu}) = \int_{e^{-i\epsilon\Phi}\psi_{\partial\nu}} \mathcal{D}\psi\mathcal{D}\bar{\psi} \int_{\mathcal{A}_j\partial\nu} \mathcal{D}\mathcal{A}_j e^{iS[A|J]-i\int_{\mathcal{V}} d^4x\sqrt{g}\bar{\psi}(\gamma^\mu\partial_\mu+m)\psi} \quad (101)$$

where Φ is given by (100). Analogous to the time-slice amplitude we see a dressing of each Dirac excitation in the boundary state by a Coulombic electric field.

Since we have skipped the explicit derivation of (101) and foregone the discussion of general boundaries in curved spacetime, we should at least mention that to do the bFP trick for more general boundaries one must necessarily use generalizations of canonical conjugate momenta and commutation relations. To highlight this, for a general path integral with data specified on boundary $\partial\mathcal{V}$, we can consider a variation of this boundary data, viz.,

$$\delta \int_{\phi_{\partial\nu}} \mathcal{D}\phi e^{iS[\phi]} = i \int_{\phi_{\partial\nu}} \mathcal{D}\phi e^{iS[\phi]} \delta S. \quad (102)$$

A general variation of the action is of the form

$$\delta S = \int_{\mathcal{V}} d^4x E(\phi) \delta\phi + \int_{\partial\nu} d^3x (\partial_\mu S) \theta^\mu(\phi, \delta\phi), \quad (103)$$

where $E(\phi)$ is the scalar density equation of motion, the boundary is defined by a constant S hypersurface, and the symplectic potential current density θ^μ is given for a general Lagrangian in ref. [54]. For a Lagrangian density which is a function only of the fields and their first derivatives we have

$$\theta^\mu(\phi, \delta\phi) = \frac{\partial\mathcal{L}}{\partial\nabla_\mu\phi} \delta\phi. \quad (104)$$

For non-null boundaries $\sqrt{g}\partial_\mu S$ can be related to the normal covector and intrinsic volume element for the hypersurface, but the form in (103) is more general and also applies to null boundaries.

For variations with support only on the boundary data we then have the functional derivative

$$\frac{\delta}{\delta\phi_{\partial\nu}(x)} \int_{\phi_{\partial\nu}} \mathcal{D}\phi e^{iS[\phi]} = i \int_{\phi_{\partial\nu}} \mathcal{D}\phi e^{iS[\phi]} \left[\int_{\partial\nu} d^3x' \frac{\delta\theta^S(\phi, \delta\phi)}{\delta\phi(x)} \right] \quad (105)$$

Defining $\frac{\delta}{\delta\phi(x)}\phi(x') = \delta^3(x-x')/\sqrt{g}$, the commutation relation between ϕ and $-i\delta/\delta\phi$ is obviously canonical. The functional derivatives $\frac{-i\delta}{\delta\phi_{\partial\nu}}$ used in the bFP trick will then be operator representations of the generalized conjugate momentum

$$\Pi_{\partial\nu}(x) = \int_{\partial\nu} d^3x' \theta^S(\phi, g^{-1/2}\delta^3(x-x')). \quad (106)$$

This expression was used in deriving (101), and will be explicitly used in the following section.

V. LARGE GAUGE TRANSFORMATIONS AND ADDITIONAL CONSTRAINTS

Up to now we have assumed that both A_μ and the gauge transformations on A_μ vanish sufficiently fast at spatial infinity that one can freely integrate by parts any expression with spatial derivatives. Energy-flux finiteness arguments lead one to expect the field strength $F_{\mu\nu}$ to obey such asymptotic fall-off conditions, at least in many physical situations. However, it is not clear why

either A_μ , or gauge transformations of A_μ , should vanish at infinity.

Gauge transformations which don't fall off as quickly as required for the above manipulations are referred to as large gauge transformations. These have a long history, especially in gravity [55], and have also been widely discussed in recent years [32, 56, 57]. Many different choices of asymptotic fall-off conditions for A_μ have been made in the literature.

Invariance under the set of large gauge transformations implies a further set of constraints, in addition to Gauss' law and $E^0 = 0$. In this section we enlarge the set of allowed gauge transformations to those which are finite and non-vanishing at the spatial boundary, and generalize the techniques used above to handle these. The invariant propagators then shed light on the constraints implied by large gauge invariance; and the path integral gives explicit solutions to the operator constraint equations.

We will treat the spatial boundary as a large sphere or cylinder of radius $R \rightarrow \infty$, and we allow for gauge

transformations which have finite asymptotic limits, viz.,

$$\lambda(t, x^A) \equiv \lim_{r \rightarrow R} \Lambda(t, r, x^A). \quad (107)$$

With finite asymptotic limits for Λ , we must also allow for finite asymptotic limits for the gauge field, viz.,

$$a_\mu(t, x^A) \equiv \lim_{r \rightarrow R} A_\mu(t, r, x^A). \quad (108)$$

We warm up by first discussing large gauge transformations for propagation between time slices; we then proceed to the causal diamond.

A. Large Gauge Transformations: Time Slicing

We would like to compute the propagator

$$K(A_{\mu \partial \mathcal{V}}) = \int_{A_{\mu \partial \mathcal{V}}} \mathcal{D}A_\mu e^{iS}, \quad (109)$$

where the region \mathcal{V} over which we integrate is again part of Minkowski space, bounded by the constant t slices Σ_i , Σ_f and the large cylinder of radius $R \rightarrow \infty$, Σ_∞ , and the action is just (8). Again, for brevity we assume that the source is an external conserved current, but as was the case in the first section, the following manipulations easily generalize to dynamic matter fields. As just discussed, while we fix boundary data $A_{\mu \partial \mathcal{V}}$ on all of $\partial \mathcal{V}$, we now lift the restriction that A_μ vanishes at spatial infinity.

At the technical level, the new challenge is that we can no longer uniquely invert the Laplacian operator when solving the Gauss law equation as in (25): there is now nothing restricting the homogeneous solutions.

To proceed with the integral we need to again use the boundary Faddeev-Popov trick, in the form (45). Suppose now that one tries to fix a Coulomb gauge in the FP path-integral, ie., write $\mathcal{G}(A) = \partial^j A_j$. However in the enlarged gauge group this choice will leave the gauge under-determined, because there are homogeneous solutions, $\nabla^2 \Lambda = 0$ which are non-vanishing at spatial infinity.

If however we restrict ourselves to gauge functions which are finite at spatial infinity, then the only remaining homogeneous solution is $\Lambda(x) = c(t)$. The only residual gauge transformations in the FP integral (45) are then time dependent global $U(1)$ rotations. These leave the spatial components A_j invariant, and only shift the spatially constant part of A_0 . To properly implement the bFP trick we then must supplement the Coulomb gauge fixing delta function with another delta function which eliminates these residual transformations.

A sufficient choice is to gauge fix the $l = 0$ spherical harmonic mode of the asymptotic gauge function $\lambda(t, x^A)$. We refer to the $l = 0$ part of a function on the sphere using a superscript “(0)”. Up to field-independent normalization we may then write

$$1 = \int \mathcal{D}\Lambda \delta(\partial^j A_j^\Lambda) \delta(a_0^{(0)\Lambda}). \quad (110)$$

in place of (45).

In what follows it is more clear if we explicitly separate out the asymptotic $l = 0$ part of all functions. The notation may seem heavier than necessary but it will allow for a much quicker generalization to the later treatment of the causal diamond amplitude. We will therefore write,

$$\Lambda(t, r, x^A) = \bar{\Lambda}(t, r, x^A) + \lambda^{(0)}(t), \quad (111)$$

where $\bar{\Lambda}$ has a finite asymptotic limit $\bar{\lambda}(t, x^A) = \lim_{r \rightarrow \infty} \bar{\Lambda}(t, r, x^A)$, but the function $\lambda^{(0)}(t)$ has a vanishing $l = 0$ mode. We'll use this same notation for the gauge field, in terms of which the action is simply

$$S[A] = S[A_j, \bar{A}_0] + \int_{t_i}^{t_f} dt a_0^{(0)} Q, \quad (112)$$

where $Q = \int d^3x J^0$ is the total charge.

With this, we can now multiply the propagator (109) by a carefully chosen factor of 1, from (110), to obtain

$$K(A_{\partial \mathcal{V}}) = \int \mathcal{D}\bar{\Lambda} d\lambda^{(0)} \int_{A_{\mu \partial \mathcal{V}}} \mathcal{D}\bar{A}_0 \mathcal{D}a_0^{(0)} \mathcal{D}A_j \delta(\partial^j A_j^\Lambda) \delta(a_0^{(0)\Lambda}) e^{iS[A_j, \bar{A}_0] + i \int_{t_i}^{t_f} dt a_0^{(0)} Q}. \quad (113)$$

Now, we implement the bFP trick by changing variables, as done before (cf. eqtns. (46)-(48), and (50)), to get

$$\begin{aligned} K(A_{\partial \mathcal{V}}) &= \int \mathcal{D}\bar{\Lambda} d\lambda^{(0)} \delta^{\partial \mathcal{V}}(\partial^j A_j + \nabla^2 \bar{\Lambda}) \delta^{\partial \mathcal{V}}(a_0^{(0)} + \partial_0 \lambda^{(0)}) \\ &\quad \times e^{-i \int_{\partial \mathcal{V}} [\bar{\Lambda} J^0 + \lambda^{(0)} J^0 + i \partial_0 \bar{\Lambda} \frac{\delta}{\delta \bar{A}_0} + i \partial_0 \lambda^{(0)} \frac{\delta}{\delta a_0^{(0)}} + i \partial_j \bar{\Lambda} \frac{\delta}{\delta A_j}]} \\ &\quad \times \int_{A_{\mu \partial \mathcal{V}}} \mathcal{D}\bar{A}_0 \mathcal{D}a_0^{(0)} \mathcal{D}A_j \delta(\partial^j A_j) \delta(a_0^{(0)}) e^{iS[A_j, \bar{A}_0] + i \int_{t_i}^{t_f} dt a_0^{(0)} Q} \end{aligned} \quad (114)$$

In the bulk part of the path integral we have effectively inserted gauge fixing delta functions as desired. The additional gauge fixing delta function simply sets $a_0^{(0)} = 0$, reducing the action to its usual form. As always in the bFP trick, we've also obtained a number of delta functions and linear shift operators outside the path integral. The crucial observation here is that the delta functions constraining the boundary gauge transformations constrain only $\bar{\Lambda}$ and $\partial_0\lambda^{(0)}$, they do not constrain the other independent functions $\partial_0\bar{\Lambda}$ and λ^0 .

In factoring out the bulk gauge group integral we are then left with residual integrals over $\partial_0\bar{\Lambda}$ and λ^0 . The remaining boundary integrals over $\bar{\Lambda}$ and $\partial\lambda^{(0)}$ are trivially performed using the delta functions. The result is then

$$K(A_{\partial\mathcal{V}}) = \left(\int d\lambda^{(0)} e^{-i \int_{\partial\mathcal{V}} \lambda^{(0)} J^0} \right) e^{i \int_{\partial\mathcal{V}} \nabla^{-2} (\partial^j A_j) J^0} \int_{\bar{A}_j \partial\mathcal{V}} \mathcal{D}\bar{A}_0 \mathcal{D}A_j \delta(\partial^j A_j) e^{iS[A_j, \bar{A}_0]}, \quad (115)$$

where A_j is the transverse component of A_j . We can now perform the \bar{A}_0 integral and there is no ambiguity in its saddle point solution; it is again given by $\bar{A}_0 = \nabla^{-2} J^0$ and the homogeneous solution is necessarily zero because by definition \bar{A}_0 has vanishing asymptotic $l = 0$ mode.

Note the remarkable feature, that the vestige of working on the configuration space for A^μ with non-vanishing asymptote is just the integral over $\lambda^{(0)}$ on the boundary. This does nothing other than add a delta function enforcing charge neutrality on the boundary state. In hindsight it is completely obvious that if we demand the amplitude to be invariant under global $U(1)$ transformations the state must be charge neutral - by enlarging the gauge group, we've simply imposed this new constraint.

If this constraint is physically unacceptable, then we can simply restrict the gauge group. Note however, that we do not need to eliminate all gauge functions which are finite asymptotically, only those which are constant on the sphere at spatial infinity. Gauge functions which approach $l \neq 0$ functions on the asymptotic sphere may still be allowed; however they do not affect time slice amplitudes. In the next subsection we see that allowing such gauge transformations actually has a nontrivial effect on the causal diamond amplitude.

B. Large Gauge Transformations: Causal Diamond Evolution

We would now like to consider the amplitude

$$K(A_{\mu \partial\mathcal{V}}) = \int_{A_{\mu \partial\mathcal{V}}} \mathcal{D}A_{\mu} e^{iS}, \quad (116)$$

where, as before, S is given by the sourced Maxwell action and the integration region \mathcal{V} is the causal diamond of radius $R \rightarrow \infty$, but now we allow the gauge fields to be finite as $r \rightarrow R$. The story is very similar to the treatment of large fields in the time slice propagator, but with an interesting additional feature.

Looking back to (82) and its solution (85), we can see that without the assumption that the gauge field vanishes as $r \rightarrow \infty$, there are infinitely many possible homogeneous solutions. In the bulk, $\eta \in (-R, R)$, the only

acceptable homogeneous solution (92) is a time-varying $h(\eta)$ which is constant in space.

The situation for these time-dependent global $U(1)$ transformations is the same as considered above for A_0 in the time slice amplitude, and the same remedy applies. We must separate off the asymptotic $l = 0$ part, $a_\eta^{(0)}$, and enforce an additional gauge fixing which sets $a_\eta^{(0)} = 0$ in the bulk. This will allow for a unique saddle point solution for the remaining field $\bar{A}_\eta = A_\eta - a_\eta^{(0)}$. The upshot is the same as the previous case; properly treating this asymptotic $l = 0$ part will just introduce delta functions on the boundary which enforce overall charge neutrality $\int_{\partial\mathcal{V}} d^3x \sqrt{g} J^n = 0$.

However there is another, more interesting, result in the causal diamond geometry. When we implement the bFP trick, we aim to introduce the Faddeev-Popov gauge fixing as in (110) above; but the integrand here still does not uniquely fix the gauge. This is because when $\eta = \pm R$, ie. on the boundary of the causal diamond, there are homogeneous solutions $D^j \partial_j \Lambda = 0$ which are arbitrary functions on the sphere. The above delta functions will uniquely determine the gauge function Λ in the bulk, but on the boundary there is still a residual gauge freedom given by all functions Λ approaching a non-constant ($l \neq 0$) function on the sphere, $\lambda^{(l \neq 0)}(x^A)$, as $\eta \rightarrow \pm R$.

Such functions will be discontinuous at $r = 0, \eta = \pm R$, but this is allowed since these singular points have been formally ‘‘blown up’’, allowing for such angle dependent limits as $r \rightarrow 0$ on the boundary.

To uniquely fix the gauge we then append a further gauge fixing term on the boundary; we choose $\nabla^B a_B^\Lambda = 0$, where $\nabla^B a_B$ is the vector divergence on the unit two-sphere. These choices together uniquely fix the gauge, so that we can write

$$1 = \int \mathcal{D}\Lambda \delta(D^j A_j^\Lambda) \delta(a_\eta^{(0)\Lambda}) \delta^{\partial\mathcal{V}}(\nabla^B a_B^\Lambda), \quad (117)$$

up to a field independent constant. We can now multiply this into our path integral representation for the causal diamond propagator. Doing this and implementing again the bFP trick, we obtain

$$\begin{aligned}
K(A_{\partial\nu}) &= \int \mathcal{D}\bar{\Lambda} d\lambda^{(0)} \delta^{\partial\nu}(D^j A_j + D^j \partial_j \bar{\Lambda}) \delta^{\partial\nu}(a_\eta^{(0)} + \partial_\eta \lambda^{(0)}) \delta^{\partial\nu}(\nabla^B a_B + \nabla^B \nabla_B \lambda) \\
&\quad \times e^{-i \int_{\partial\nu} [\bar{\Lambda} J^n + \lambda^{(0)} J^n + i \partial_\eta \bar{\Lambda} \frac{\delta}{\delta \bar{A}_\eta} + i \partial_\eta \lambda^{(0)} \frac{\delta}{\delta a_\eta^{(0)}} + i \partial_j \bar{\Lambda} \frac{\delta}{\delta \bar{A}_j}]} \\
&\quad \times \int_{A_\mu \partial\nu} \mathcal{D}\bar{A}_\eta \mathcal{D}a_\eta^{(0)} \mathcal{D}A_j \delta(D^j A_j) \delta(a_\eta^{(0)}) e^{iS[A_j, \bar{A}_\eta] + i \int_{-R}^R d\eta a_\eta^{(0)} Q} \quad (118)
\end{aligned}$$

The propagator (118) is structurally very similar to (114), except for the new factor we've introduced to gauge fix the residual transformations which are allowed on the causal diamond boundary. Again, we can freely evaluate the integral over $\partial_\eta \bar{\Lambda}$ and $\lambda^{(0)}$ in the boundary gauge transformations since they are not fixed by the boundary gauge fixing delta functions. We can evaluate the integrals over $\bar{\Lambda}$ and $\partial_\eta \lambda^{(0)}$ using the delta functions. The boundary delta functions now constrain $\bar{\Lambda}$ to be

$$\bar{\Lambda} = - \int_{\partial\nu} d^3 x \sqrt{g} G D^j A_j - \int d^2 \Omega \bar{G} \nabla^B a_B \quad (119)$$

where, up to a minus sign, the first term is just Φ given in (100); $d^2 \Omega$ is the area element on the unit two-sphere, and \bar{G} is the Green's function for the Laplacian (less the $l=0$ mode) on the unit two-sphere. Evaluating these integrals we obtain the final expression for the large gauge transformation invariant causal diamond amplitude

$$\begin{aligned}
K(A_{\partial\nu}) &= \left(\int d\lambda^{(0)} e^{-i \int_{\partial\nu} \lambda^{(0)} J^n} \right) e^{i \int_{\partial\nu} d^3 x \sqrt{g} [\int_{\partial\nu} d^3 x' \sqrt{g} G D^j A_j + \int_{\partial\nu} d^2 \Omega' \bar{G} \nabla^B a_B]} J^n \\
&\quad \times \int_{A_j \partial\nu} \mathcal{D}\bar{A}_\eta \mathcal{D}A_j \delta(D^j A_j) e^{iS[A_j, \bar{A}_\eta]} \quad (120)
\end{aligned}$$

If it so happens that the constraint that the net charge flux through each of the past and future boundaries must vanish, then we can simply drop this factor and disallow gauge transformations which approach a constant function on the asymptotic sphere. Alternatively, we could just make this constant the same on the past and future parts of the boundary: this would cause the delta function to enforce charge conservation on the propagator, but not on the states.

We also see that in addition to the total charge flux constraint, we've found that invariance under large gauge transformations with higher spherical harmonics enforces

a new constraint on the system. Since \bar{G} and a_B are just angular functions, the new boundary phase in (120) implies that a certain part of the electric field at each angle is determined solely by the net charge flux of charge through the boundary at each angle. Specifically, if we define

$$\frac{\delta}{\delta a_A(x^A)} a_B(x^{A'}) = \delta_B^A q^{-1/2} \delta^2(x^A - x^{A'}), \quad (121)$$

where q is the determinant of the metric on the unit two-sphere, then the gauge invariant amplitude satisfies

$$-i \frac{\delta}{\delta a_B(x^A)} K(A_{\partial\nu}) = \left(\int_{\partial\nu} d^3 x' \sqrt{g} J^n(x') \nabla^B \bar{G}(x^{A'}, x^A) \right) K(A_{\partial\nu}), \quad (122)$$

on each of the future and past parts of the causal diamond boundary.

It remains to understand what, physically, this functional differential operator represents. We can do so by using the relationship between functional derivatives and the symplectic current density in (105). For the gauge field A_B we have

$$\theta^\eta(A_\mu, \delta A_B) = \frac{\partial \mathcal{L}}{\partial \nabla_\eta A_B} \delta A_B = -\sqrt{g} F^{\eta B} \delta A_B, \quad (123)$$

and if we separate the field as $A_B = \bar{A}_B + a_B$, where a_B is independent of r and \bar{A}_B is vanishing at spatial infinity, then by linearity we have

$$\theta^\eta(A_\mu, \delta a_B) = \frac{\partial \mathcal{L}}{\partial \nabla_\eta A_B} \delta a_B = -\sqrt{g} F^{\eta B} \delta a_B. \quad (124)$$

Invoking (105) we then find

$$\begin{aligned} -i \frac{\delta}{\delta a_B(x^A)} K(A_{\partial\mathcal{V}}) &= \int_{A_{\partial\mathcal{V}}} \mathcal{D}A_\mu e^{iS[A]} \left(\int_{\partial\mathcal{V}} d^3x' \sqrt{g} F^{B\eta} q^{-1/2} \delta^2(x^A - x^{A'}) \right) \\ &= \int_{A_{\partial\mathcal{V}}} \mathcal{D}A_\mu e^{iS[A]} \left(\int_0^\infty dr r^2 F^{B\eta}(r, x^A) \Big|_{\partial\mathcal{V}} \right). \end{aligned} \quad (125)$$

The bFP trick has then illustrated that physical (gauge invariant) states on the boundary of the large causal diamond satisfy the eigenvalue equation

$$\left(\int_0^\infty dr r^2 \hat{F}^{B\eta}(r, x^A) \Big|_{\partial\mathcal{V}} \right) K(A_{\partial\mathcal{V}}) = \left(\int_{\partial\mathcal{V}} d^3x' \sqrt{g} J^\eta(x') \nabla^B \bar{G}(x^{A'}, x^A) \right) K(A_{\partial\mathcal{V}}), \quad (126)$$

at every angle x^A on the sphere, independently of the past and future parts of the boundary. This is an exact relation, irrespective of the data specified for the fields or the dynamics of the charged matter, ie. it is kinematically required. It is a direct consequence of gauge invariance for the causal diamond path-integral on the extended configuration space when the gauge fields are allowed to take finite values at spatial infinity.

This result bears a clear resemblance to results at null infinity which have been widely discussed in the literature [32, 33, 57]. Indeed, since (126) holds at each angle, we can multiply it by $\partial_B \varepsilon(x^A)$, for any function on the sphere $\varepsilon(x^A)$, and integrate over the sphere. We then obtain

$$\int_{\partial\mathcal{V}} d^3x \sqrt{q} \left((q^{AB} \nabla^A \varepsilon(x^A)) \hat{F}_{Br} + \varepsilon(x^A) (r^2 J_r) \right) K(A_{\partial\mathcal{V}}) = 0. \quad (127)$$

If we go to complex stereographic coordinates (z, \bar{z}) on the unit sphere such that the metric is

$$d\Omega^2 = 2\gamma_{z\bar{z}} dz d\bar{z}, \quad (128)$$

with

$$\gamma_{z\bar{z}} = \frac{2}{(1+z\bar{z})^2}, \quad (129)$$

we obtain the operator constraint equation

$$\int_{\partial\mathcal{V}} dr d^2z \left(-\partial_z \varepsilon(z, \bar{z}) \hat{F}_{\bar{z}r} - \partial_{\bar{z}} \varepsilon(z, \bar{z}) \hat{F}_{zr} + \varepsilon(z, \bar{z}) \gamma_{z\bar{z}} (r^2 J_r) \right) K(A_{\partial\mathcal{V}}) = 0. \quad (130)$$

If we recall that in our coordinates r is the affine parameter on the null boundary, then we immediately recognize the operator above as the large gauge charge operator \hat{Q}_ε in (compare ref. [32], section 2.5.11). The electric part of this operator creates soft photon states. When the matter is quantum mechanical the computation can be carried through with no additional complications, and the result is to simply replace J_r by a functional differential operator representation of the $U(1)$ current operator \hat{J}_r .

From our analysis we can see clearly that if we blow up spatial infinity such that the value of the gauge field at spatial infinity, $\lambda(x^A)$, can be different whether approached from the future or past part of the boundary, then the amplitude satisfies (130) on the past and future null boundaries separately. As a consequence, the states the past and future parts are necessarily dressed in the way described originally by Kibble, Chung, and Faddeev and Kulish [58–60]

If however, we do not blow up spatial infinity, then

(130) holds only when integrated over the whole null boundary of the causal diamond, and the dressing involving $\bar{G} \nabla^B a_B$ need only occur on the past boundary. This then implies the infinitely many conservation laws discussed by in the recent literature [32, 33, 57], and is equivalent to Weinberg's soft photon theorem. We won't take a position here on whether such a condition must be imposed on $\lambda(x^A)$, rather we'd just like to highlight how, from the bFP trick on a configuration space with asymptotically finite A_μ , one finds either soft dressing or large gauge charge conservation.

VI. CONCLUSIONS

In this paper we have given a manifestly gauge invariant analysis of QED amplitudes. A primary goal was to define and understand gauge invariant states in QED, where we use the modern understanding of states as data

living the boundaries of path-integrals. Much of the analysis was done for the gauge field coupled to a conserved external current, but we also saw that all our results trivially generalize to the full dynamic theory, in which a gauge field couples to quantum charged particles or to the Dirac field.

We adopted a ‘general boundary QFT framework’, in which the path-integral allows us to go beyond the canonical quantization framework. This framework uses path-integrals, with data fixed on general closed hypersurfaces, to compute amplitudes - the interpretation in terms of states and transition amplitudes is then secondary, and only applies to particular geometries.

To treat the gauge redundancy in the QED path integral we introduced a boundary Faddeev-Popov trick, a natural generalization of the usual Faddeev-Popov procedure to path-integrals with fixed boundary data. Although the bFP trick should be applicable for general boundaries, in this paper we considered two simple examples, viz., (a) when $\partial\mathcal{V}$ consists of two finitely separated constant time slices and a time-like cylinder at spatial infinity, and (b) when $\partial\mathcal{V}$ is the null boundary of a large causal diamond. The former case is then a conventional transition amplitude which has a representation in terms of a Hamiltonian operator, whereas the later is most easily described via the path-integral. In this limit the causal diamond boundary resembles null infinity and the amplitude resembles a scattering amplitude.

We worked in the extended configuration space of $U(1)$ gauge theory, in that we considered the amplitude to ostensibly be a functional of all field configurations A_μ prescribed on the boundary. Using a “path integral first” approach, we did not concern ourselves a priori with identifying canonical variables for quantization; instead we simply prescribed boundary data for the full four-vector potential. As a consequence of the gauge invariance of the QED action, the resulting amplitudes were gauge-invariant and independent of non-canonical variables. The resulting path integrals were written explicitly in terms of gauge invariant variables, and as a consequence of the bFP trick we obtained unique expressions for the dependence of the amplitudes on the gauge-variant parts of A_μ . The dependence arose only as a boundary phase.

The novel result here is that rather than solving the constraint equation, an equation which under-determines the state, we analyzed the path integral itself, and found unique expressions for the boundary phases to see how the constraint equation ought to be satisfied. The amplitude’s dependence on the gauge-variant parts of the field were determined kinematically, whereas the dependence on the gauge-invariant parts of the field remained to be determined dynamically, by a path integral over gauge invariant variables.

For each of the two geometries considered we considered both the case where gauge transformations vanish

at spatial infinity, and the case where they have finite limits. In both cases, when gauge transformations were required to vanish at spatial infinity we obtained Coulombic dressing of the charges in the boundary state. When the gauge group was extended, only the causal diamond amplitude had noteworthy changes. The boundary states were annihilated by the “large-gauge charge” discussed previously in the literature on null infinity [32, 56]. Furthermore, just as the Coulomb field emerged naturally from the path integral in the previous scenario, from the large causal diamond path integral an explicit expression for the soft-photon dressing of states on the null boundary emerges very naturally. Our coordinate system, and the specific limit taken to null infinity, were sufficient to include both null and timelike matter. The resulting expressions were not novel, but the bFP technique used here was, and it provided a manifestly gauge invariant derivation of the result.

One of our main motivations in studying these questions was to develop methods which will allow us to study some rather concrete problems in quantum gravity - in particular, the ongoing debates about how one may test experimentally whether the gravitational field is quantized [34–36], and how to properly define and calculate decoherence rates [33, 37–40]. The answer to both of these problems turns essentially on how one defines physical states for the metric field. The generalization of our methods to linearized gravity - which is all that is necessary to deal with these two problems - is straightforward if somewhat messy, and we will give our results in a future paper.

On a more formal level, it is of considerable interest to generalize the bFP trick to gauge theories beyond QED, as well as to amplitudes in curved spacetime. Additionally if we extended the class of allowed gauge transformations considered here, then gauge invariance would imply a new angle dependent constraints on our asymptotic states, perhaps being related to the sub-leading soft photon theorems. Again, these ideas will be developed elsewhere.

VII. ACKNOWLEDGEMENTS

This work was funded in Canada by the National Science and Engineering Research Council of Canada (NSERC). In the USA, PCES received support at Caltech from the Simons Foundation (Award 568762) and the National Science Foundation (Award PHY-1733907).

We have benefited particularly from extended discussions with C Delisle at UBC. We also thank W.G. Unruh at UBC and Y. Chen at Caltech for illuminating remarks. We would also like to acknowledge the support and hospitality of Y. Chen, T.F. Rosenbaum, and K.S. Thorne at Caltech.

- [1] See, eg., R. Utiyama, Phys. Rev. **101**, 1597 (1956); T.W.B. Kibble, J. Math. Phys. **2**, 212 (1960); P.G. Bergmann, Phys. Rev. **124**, 274 (1961); and J.D. Norton, Rep. Prog. Phys. **56**, 791 (1993) for a survey of some of the earlier discussion.
- [2] K. Kuchar, in “*Proceedings of the 4th Canadian Conference on General Relativity and Relativistic Astrophysics*”, ed. G. Kunstatter et al. (World Scientific, 1992).
- [3] N. C. Tsamis, R. P. Woodard, Ann. Phys. (N.Y.) **215**, 96 (1992).
- [4] S. B. Giddings, D. Marolf, J.B. Hartle, Phys. Rev. **D 74**, 064018 (2006)
- [5] J.B. Hartle, Phys. Rev. **D37**, 2818 (1988); Phys. Rev. **D38**, 2985 (1988)
- [6] J.B. Hartle, Phys. Rev. **D49**, 6543.
- [7] R. Oeckl, Phys. Lett. **B575**, 318 (2003); and Adv. Theor. Math. Phys. **2**, 451 (2015).
- [8] F. Conrady et al., Phys. Rev. **D 69**, 064019 (2004); R. Oeckl, Phys. Rev. **D 73**, 065017 (2006).
- [9] G.W. Gibbons, S.W. Hawking, Phys. Rev. **D15**, 2752 (1977); and S.W. Hawking, pp. 746-789 in “*General Relativity: An Einstein Centenary Survey*”, ed. S.W. Hawking, W. Israel (Cambridge Univ. Press, 1979)
- [10] J.B. Hartle, S.W. Hawking, Phys. Rev. **D28**, 2960 (1983)
- [11] For an introduction to quantum cosmology, see, eg., C. Kiefer, “*Quantum Gravity*” (Oxford Univ. press, 2012)
- [12] M.G.G. Laidlaw, C. Morette-DeWitt, Phys. Rev. **D3**, 1375 (1971); J.M. Leinaas, J. Myrheim, Nuovo Cim. **37B**, 1 (1977)
- [13] Y-S Wu, Phys. Rev. Lett. **52**, 2103 (1984)
- [14] M. Stone, “*Quantum Hall Effect*” (World Scientific, 1992)
- [15] C. Morette-DeWitt, Comm. Math. Phys. **28**, 47 (1972).
- [16] T.W.B. Kibble, Comm. Math. Phys. **64**, 73 (1978); T.W.B. Kibble, pp. 63-80 in “*Quantum Gravity 2*”, ed. C.J. Isham, R. Penrose, D.W. Sciama (O.U.P., 1981)
- [17] S. Weinberg, Phys. Rev. Lett. **62**, 485 (1989); S. Weinberg, Ann. Phys. (NY) **194**, 336 (1989)
- [18] T.W.B. Kibble, S. Randjbar-Daemi, J. Phys A **13**, 141 (1980)
- [19] R. Penrose, Gen. Rel. Grav. **28**, 581 (1996); R. Penrose, Phil. Trans. Roy. Soc. Lond. **A356**, 1927 (1998); W Marshall, C Simon, R Penrose, D Bouwmeester, Phys. Rev. Lett. **91**, 130401 (2003). See also D. Kleckner et al., New J. Phys. **10**, 095020 (2008)
- [20] A. Bassi et al., Rev. Mod. Phys. **85**, 471 (2013)
- [21] P.C.E. Stamp, New J. Phys. **17** 065017 (2015); A. Barvinsky, D. Carney, and P. Stamp, Phys. Rev, **D 98**, 084052 (2018).
- [22] A.O. Barvinsky, J. Wilson-Gerow, P.C.E. Stamp, /arXiv: 2011.03601 [gr-qc]; J. Wilson-Gerow, P.C.E. Stamp, /arXiv XXXXXXX
- [23] Y. Aharonov, D. Bohm, Phys. Rev. **115**, 485 (1959); Y. Aharonov, D. Bohm, Phys. Rev. **123**, 1511 (1961)
- [24] PS Epstein, Am. J. Phys. **13**, 127 (1945); M Renninger, Z. Physik **158**, 417 (1960); RH Dicke, Am. J. Phys. **49**, 925 (1981).
- [25] P.A.M. Dirac, Can. J. Math. **2**, 129 (1950)
- [26] P.A.M. Dirac, “*Lectures on Quantum Mechanics*” (Belfer graduate school of Science, Yeshiva Univ, New York, 1964)
- [27] P.A.M. Dirac, Can. J. Phys. **33**, 650 (1955)
- [28] B.S. DeWitt, Phys. Rev. **160**, 1113 (1967)
- [29] G. Klinkhammer, K.S. Thorne, unpublished; and D.S. Goldwirth, M.J. Perry, T. Piran, K.S. Thorne, Phys. Rev. **D49**, 3951 (1994).
- [30] M. Visser, “*Lorentzian Wormholes*” (Springer Verlag, 1996).
- [31] A. Ashtekar, “*Asymptotic Quantization*” (Bibliopolis, 1987)
- [32] A. Strominger, “*Lectures on the Infrared Structure of Gravity and Gauge Theory*” (Princeton Univ. Press, 2018)
- [33] J. Wilson-Gerow, C. DeLisle, P.C.E. Stamp, Class. Q. Grav. **35**, 164001 (2018)
- [34] C. Marletto, V. Vedral, Phys. Rev. Lett. **119**, 240402 (2017)
- [35] S. Bose et al., Phys. Rev. Lett. **119**, 240401 (2017)
- [36] C. Anastopoulos, B.-L. Hu, /arXiv 1804.11315.
- [37] C. Anastopoulos, B.-L. Hu, Class. Quantum Grav. **32**, 165022 (2015)
- [38] T. Oniga, C.H.-T. Wang, Phys. Rev. **D 93**, 044027 (2016)
- [39] M.P. Blencowe, Phys. Rev. Lett. **111**, 021302 (2013)
- [40] J. Wilson-Gerow, MSc thesis (UBC, Sept. 2017)
- [41] L.D. Faddeev, V.N. Popov, Phys. Lett. **B25**, 29 (1967)
- [42] V.N. Gribov, Nucl. Phys. **B139**, 1 (1978); a more recent review is in N. Vandersickel, D. Zwanziger, Phys. Rep. **520**, 175 (2012)
- [43] S. Mandelstam, Ann. Phys. **19**, 1 (1962), and *ibid.*, 25 (1962)
- [44] S.V. Shabanov, J. Phys. **A23**, 3245 (1990); L.V. Prokhorov, D.V. Fursaev, S.V. Shabanov, Theor. Math. Phys. (USSR), **97**, 1323 (1993)
- [45] W. Donnelly, S.B. Giddings, Phys. Rev. **D 93**, 024030 (2016)
- [46] E.S. Fradkin, Nucl. Phys. **76**, 588 (1966)
- [47] H.M. Fried, “*Green’s Functions and Ordered Exponentials*” (Cambridge Univ. Press, 2002)
- [48] J.-L. Gervais, D. Zwanziger, Phys. Lett. **94B**, 389 (1980)
- [49] D. Carney, L. Chaurette, D. Neuenfeld, G.W. Semenov, Phys. Rev. **D 97**, 025007 (2018)
- [50] J.D. Jackson, “*Classical Electrodynamics*” (Wiley, 1962)
- [51] J. Wilson-Gerow, to be published.
- [52] G.C. Rossi, M. Testa, Nucl. Phys. **B 163**, 109 (1980); and Nucl. Phys. **B 176**, 477 (1980); and F. Mattei, C. Rovelli, S. Speziale, M. Testa, Nucl. Phys. **B 739**, 234 (2006);
- [53] See C. Delisle, to be published. We thank C. Delisle for extensive discussion on this point.
- [54] J. Lee, R.M. Wald, J. Math. Phys. **31**, 725 (1990)
- [55] H. Bondi, A.W.K. Metzner and M.I.G. Van der Berg, Proc. Roy. Soc. **A 269**, 21 (1962); R.K. Sachs, Proc. Roy. Soc. **A 270**, 103 (1962), and Phys. Rev. **128**, 2851 (1962)
- [56] The connection between large gauge transformations and soft theorems was first discussed for QED by R. Ferrari, L.E. Picasso, Nucl. Phys. **B31**, 316 (1971). For more general recent discussions, see, eg., Y. Hamada, M.-S. Sheo, G. Shiu, Phys. Rev. **D 96**, 105013 (2017); Y. Hamada, G. Shiu, Phys. Rev. Lett. **120**, 201601 (2018); and G. Compere, Phys. Rev. Lett. **123**, 021101 (2019).

- [57] For a thorough review of soft photons and gravitons and the connection with conservation laws, see [32]. For the connection with memory effects, see, eg., P. Mao, W.-D. Tan, *Phys. Rev. D* **101**, 124015 (2020), and refs. therein.
- [58] T.W.B. Kibble, *J. Math. Phys.* **9**, 315 (1968); *Phys. Rev.* **173**, 1527 (1968), *Phys. Rev.* **174**, 1882 (1968); and *Phys. Rev.* **175**, 1624 (1968)
- [59] V. Chung, *Phys. Rev.* **140**, B1110 (1965)
- [60] P.P. Kulish, L.D. Faddeev, *Theor. Math. Phys.* **4**, 745 (1970).