I ain't afraid of no ghost*

John Dougherty†

May 25, 2021

Abstract

This paper criticizes the traditional philosophical account of the quantization of gauge theories and offers an alternative. On the received view, gauge theories resist quantization because they feature distinct mathematical representatives of the same physical state of affairs. This resistance is overcome by a sequence of ad hoc modifications, justified in part by reference to semiclassical electrodynamics. Among other things, these modifications introduce "ghosts": particles with unphysical properties which do not appear in asymptotic states and which are said to be purely a notational convenience. I argue that this sequence of modifications is unjustified and inadequate, making it a poor basis for the interpretation of ghosts. I then argue that gauge theories can be quantized by the same method as any other theory. On this account, ghosts are not purely notation: they are coordinates on the classical configuration space of the theory—specifically, on its gauge structure. This interpretation does not fall prey to the standard philosophical arguments against the significance of ghosts, due to Weingard. Weingard's argumentative strategy, properly applied, in fact tells in favor of ghosts' physical significance.

1 Introduction

Our current best theories of high-energy particle physics model most particle interactions with gauge theories, so it's no surprise that the interpretation of gauge theories is a matter of major concern to philosophers of physics. More surprising is the philosophical literature's emphasis on the interpretation of *classical* gauge theories: whether their symmetries are observable (Brading and Brown, 2004; Greaves and Wallace, 2014), whether and how they're deterministic (Belot, 1998; Earman, 2003), how most perspicuously to formulate them (Healey, 2007; Rosenstock and Weatherall, 2016), whether they prompt deep revisions of our background metaphysics (Gilton, 2020; Maudlin, 2007), and so on. This emphasis is surprising because high-energy physics uses quantum field theory, not classical field theory. But quantum theories are often obtained by quantizing classical ones, and the quantum has interpretational challenges of its own. So we can justify interpretive work on classical gauge theories with the thought that it

*This is a preprint of an article appearing in *Studies in the History and Philosophy of Science* (doi:10.1016/j.shpsa.2021.05.005). Please only cite the final version.

†This paper benefited from comments from Joshua Eisenthal, Casey McCoy, Porter Williams, and two anonymous reviewers, as well as audiences at Oxford, Caltech, and the Munich Center for Mathematical Philosophy.

will help us to understand quantum ones, at least if our work in the classical case is appropriately informed by quantization. In this paper I argue that our work in the classical case has not been appropriately informed by quantization, because the philosophical literature's account of quantizing gauge theories is deficient.

The particular deficiency that will be my focus is the treatment of "ghosts" in the perturbative quantization of gauge theories. I have three aims: to criticize the received view on ghosts in the philosophical literature, to offer a better story, and to argue that ghost fields are an indispensable—indeed, constitutive—feature of gauge theories. According to the received view, we cannot apply standard quantization procedures to classical gauge theories because these theories contain distinct mathematical representatives of the same physical state of affairs. The remedy for this problem is to "fix the gauge" by eliminating the redundancy. Further ad hoc manipulations introduce ghost fields, and at this point we can quantize. It's often said that this last preparatory step is mere convenience: for example, Weingard argues that ghost fields are "just an artefact of our notation" that we may dispense with in principle (1988, 57).[1]

In Section 2 I argue that this story is inadequate for philosophical purposes; it fails to informatively justify the use of ghost fields in quantization. From the start, it's too vague about the alleged problems with applying standard quantization procedures to a theory with mathematical redundancy. Sometimes the problem is taken to be conceptual: it would be "naive" to try to quantize the classical gauge theory directly, and so we must modify the quantization procedure to "make physical sense of the theory" (Redhead, 2003, 135). Others take the problem to be mathematical, claiming that the predictions of the unmodified theory diverge due to the "massive 'over-counting'" induced by the redundancy (Healey, 2007, 145). Neither of these diagnoses can be quite right. The received view must allow for theories with mathematical redundancy, because it takes classical gauge theories to be of this sort. And while you can cook up a mathematical problem with certain ingredients, this problem is only connected to the multiplicity of representations under a sequence of controversial background assumptions. Spelling this all out in more detail shows that the received view's treatments of classical and quantum gauge theories are at odds with one another. These problems mean that the received view cannot give a satisfying account of the origin of ghost fields.

I think the problems with the received view can be rectified without too much revision to our classical interpretive projects, and in Section 3 I give an alternative account of ghosts meant to do this. Ghost fields do not arise in the process of quantization; they are a feature of the classical theory, wherein they coordinatize the gauge structure of that theory. The standard quantization procedure applies straightforwardly to theories with gauge structure, though injudicious conventions can lead to a coordinate singularity. The received view misses this because it is committed to the superfluity of gauge structure. That is, on the received view gauge structure is mathematical excess to be avoided or eliminated when interpreting the theory. This means that when gauge structure is physically relevant, the received view must introduce ad hoc replacements to compensate for the structure it has eliminated. Ghost fields were first introduced

[1]This story is a synthesis of philosophical sources including Guay (2008), Healey (2007), Redhead (2003), Rickles (2008), and Weingard (1988). Something like it also appears in many quantum field theory textbooks.

as one of these ad hoc replacements. This alternative story explains the problems encountered by the received view while avoiding them. It's also more compatible with the received view's treatment of classical gauge theories than the received view itself.

The last part of this paper replies to two arguments for the claim that ghost fields are dispensable in principle. Weingard notes that a clever choice of coordinates can simplify computations involving ghost fields and argues from this that ghost fields are "purely a result of our notation" (1988, 58). In Section 4 I argue that Weingard's premises cannot secure his conclusion. Weingard gives an explicit criterion to identify some mathematical feature as purely a result of our notation: roughly, we can suppose it to have any value we like with no consequence. I argue that this criterion does not classify ghost fields as purely notation. Indeed, some violations of Weingard's criterion have empirical interpretations; for example, it follows from one that the electric charge of the electron must be exactly thrice the electric charge of the down quark.

2 The received view

The received view claims that the standard quantization procedure is inapplicable to gauge theories because such theories contain multiple representatives for the same physical state of affairs. In this section I reconstruct and criticize this claim. Two variants appear in the literature. According to the first, the existence of multiple representations is itself a problem that makes the standard quantization procedure inappropriate for some theories; according to the second, the multiplicity of representations produces divergences in the standard quantization procedure. In this section I argue that a plausible version of the first claim relies on the truth of the second and that the second claim is too quick to identify the multiple representations as the cause of divergence. Both claims rely on a particular conception of the classical field theories at issue, and this is what I dispute in the rest of the paper.

2.1 The received view's problem

On the standard philosophical account, ghosts arise during a modification of the standard quantization procedure, and this modification is prompted by mathematical redundancy. On one reading, the modification is required by the mere fact of multiplicity, and this fact follows from an antecedent interpretation of the classical theory. But this reasoning is self-defeating. A theory that multiply represents physical states of affairs can't be incoherent, since the received view takes a classical gauge theory to have just this feature. And the justification for this interpretation of the classical theory relies on the in-principle applicability of the unmodified quantization procedure to any classical theory. So if there is a problem with quantizing gauge theories then it must be a problem in practice, rather than principle.

The standard quantization procedure expresses a quantum theory of some system using integrals built from the data of a classical theory of that system.[2]

[2]I will restrict attention to path integral quantization. Nothing hangs on this choice; my main claims apply to constrained Hamiltonian quantization as well. See Henneaux and Teitelboim (1992, §18.4) for a discussion of how the formalism below relates to constrained

For example, given some classical configuration space X with coordinate ϕ, some action $S(\phi)$ on X, and some classical observable $\mathcal{O}(\phi)$ on X, the time-ordered quantum expectation value of $\mathcal{O}$ is given by

$$\langle \mathcal{O} \rangle = \frac{1}{Z} \int_X \mathcal{D}\phi \, e^{\frac{i}{\hbar} S(\phi)} \, \mathcal{O}(\phi) \qquad\qquad Z = \int_X \mathcal{D}\phi \, e^{\frac{i}{\hbar} S(\phi)}$$

So for the purposes of quantization, a classical theory mathematically consists of a configuration space X and a real-valued function S on X. The space X can be any space that supports the appropriate notion of integration and S any function on it; in practice we're interested in particular infinite-dimensional supermanifolds that provide the configuration space for classical field theories and classical actions for these theories given by integrating a Lagrangian density over a spacetime manifold.

The usual philosophical story claims that this quantization procedure does not apply to gauge theories. This claim has three parts: first, that the natural configuration space of a gauge theory has a certain form; second, that two elements of this configuration space represent the same physical state of affairs if they are related by one of a distinguished set of "gauge transformations"; third, that the existence of nontrivial gauge transformations poses a problem for the standard quantization procedure. For example, Redhead has it that

> a naive approach would involve integrating over paths which are connected by gauge transformations. To make physical sense of the theory, the obvious move is to 'fix the gauge', so that each path intersects each gauge orbit in just one point. (2003, 135)

The naive approach to quantizing a gauge theory would take the domain of integration to be a classical configuration space containing distinct elements related by gauge transformations, but integrating over this space wouldn't "make physical sense". Redhead doesn't elaborate on the senselessness here, but this passage does suggest that it would be eliminated by choosing a different classical configuration space—one in which each physical state of affairs has exactly one mathematical representative. It also suggests that we somehow know we ought to fix the gauge before we've tried to quantize the theory. Guay echoes this sentiment, claiming that "[a]t least formally we know that the right way to quantize" is to apply the standard quantization procedure to the theory whose configuration space is the set of gauge-equivalence classes (2008, 361).

The argument for the first two parts of this claim undermines the third part, at least on the reading I've given of Redhead and Guay. The desire to fix the gauge comes from studying the results of quantization, so it can't conceptually precede the quantization of the field theory. The argument for fixing the gauge takes Yang–Mills theory as a paradigm gauge theory, and it takes the Yang–Mills model of electromagnetism as a paradigm Yang–Mills theory. The full mathematical characterization of Yang–Mills theory is part of my disagreement with the received view, but some features are uncontested. Any Yang–Mills theory has an associated Lie algebra $\mathfrak{g}$.[3] In local coordinates, some portion of

Hamiltonian quantization.

[3]Some conventions: Roman indices are Lie algebra components, raised and lowered freely; Greek indices are spacetime components, raised and lowered with the Minkowski metric. For the one-dimensional Lie algebra we drop the Roman indices. Repeated indices are implicitly

the Yang–Mills configuration space is coordinatized by a family of real scalar functions A^a_μ on spacetime, which assemble into a $\mathfrak{g}$-valued one-form. That is, any $\mathfrak{g}$-valued one-form determines a local configuration of the Yang–Mills field, but distinct $\mathfrak{g}$-valued one-forms might coordinatize the same point of the local configuration space. In these coordinates, the theory's Lagrangian is

$$\mathcal{L}_{\mathrm{YM}}(A) = -\frac{1}{4}(F^a_{\mu\nu})^2 \qquad\qquad F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g f^a{}_{bc} A^b_\mu A^c_\nu$$

where g is a coupling constant and $f^a{}_{bc}$ are the structure constants of $\mathfrak{g}$.

The desire to fix the gauge arises from a particular analysis of classical electromagnetism.[4] The only classical particle phenomenon aptly described by Yang–Mills theory is the electromagnetic interaction of Newtonian matter. The worldline x of a particle moving in an electromagnetic field satisfies the Lorentz force law

$$qF_{\mu\nu}\dot{x}^\nu = m\ddot{x}_\mu$$

with m and q the mass and charge of the particle, respectively, and $F_{\mu\nu}$ a closed two-form. That is, the force some particle experiences is proportional to the parallel component of a Lorentz-covariant electromagnetic configuration $F_{\mu\nu}$ that is the same for all particles. The motion of Newtonian matter is fully determined by the forces to which it's subject, so if we can only probe the electromagnetic configuration using such matter then the accessible electromagnetic facts are fully captured by the tensor $F_{\mu\nu}$. Moreover, the dynamics of the electromagnetic configuration in the presence of a current j^ν are classically described by Maxwell's equation

$$\partial_\mu F^{\mu\nu} = j^\nu$$

So the electromagnetic facts accessible to Newtonian matter are encoded by $F_{\mu\nu}$, and the dynamics of these facts are expressed in terms of $F_{\mu\nu}$ as well. Therefore we plausibly ought to take the configuration space of Maxwell electrodynamics to be the space of closed two-forms.

We can also use Yang–Mills theory to model the electromagnetic interaction, and this suggests a first-pass interpretation of Yang–Mills theories in general. For the Yang–Mills theory associated with the Lie algebra $\mathfrak{u}(1)$, the configuration space is coordinatized by an ordinary one-form A_μ, and because the structure constants vanish the corresponding field strength

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$$

is a closed two-form. Conversely, for any closed two-form $F_{\mu\nu}$ over a contractible manifold there is some one-form A_μ whose field strength is $F_{\mu\nu}$. Moreover, the Euler–Lagrange equation of the Yang–Mills action is Maxwell's equation in vacuum. So we can use the Newtonian interpretation of Maxwell electromagnetism

summed over after raising and lowering to give one upper and one lower index of each pair. We assume that $\mathfrak{g}$ is reductive. Decorated actions are given by integrating the corresponding decorated Lagrangian over $\mathbb{R}^4$—so that, for example, we have $S_{\mathrm{YM}} = \int_{\mathbb{R}^4} d^4x\, \mathcal{L}_{\mathrm{YM}}$—with the exception of the action S_{AB} in the Aharonov–Bohm experiment, which is obtained by integrating the Lagrangian $\mathcal{L}_{\mathrm{AB}}$ over the topologically nontrivial exterior of the apparatus.

[4]The following four paragraphs are modelled most closely on Healey (2007, Ch. 2), but the essentials of this analysis also appear in Belot (1998, §4), Guay (2008, §3.1), Redhead (2003, §6), Rickles (2008, Ch. 3), and Weingard (1988, §III), as well as most other philosophical discussions of the Aharonov–Bohm effect.

to interpret Yang–Mills theories as well: the configuration space of the theory is the space of closed $\mathfrak{g}$-valued two-forms, and a $\mathfrak{g}$-valued one-form A^a_μ coordinatizes the point

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g f^a{}_{bc} A^b_\mu A^c_\nu$$

of the configuration space.

This first-pass interpretation is generally thought untenable, even for electromagnetism. But the arguments against it and for a replacement have the same logic as the initial argument for it. If we probe the electromagnetic configuration with quantum particulate matter then we can distinguish two one-forms with the same field strength tensor. The wavefunction $\psi(t, x)$ of a quantum particle moving in an electromagnetic field satisfies

$$\psi(t_1, x_1) = \int dx_0 \int \mathcal{D}x \, e^{\frac{i}{\hbar} S_{\mathrm{AB}}(A,x)} \, \psi(t_0, x_0)$$

where the domain of the inner integral is the collection of all paths x such that $x(t_0) = x_0$ and $x(t_1) = x_1$ and

$$\mathcal{L}_{\mathrm{AB}}(A, x) = \frac{1}{2} m \dot{x}^2 + q \dot{x}^\mu A_\mu$$

with A_μ a one-form describing the electromagnetic configuration. As in the classical case, we can interpret A_μ as the coordinate on the configuration space of a $\mathfrak{u}(1)$ Yang–Mills theory. However, unlike the classical case, there are experimental setups in which two Yang–Mills potentials A_μ and B_μ give rise to distinct dynamics and thus distinct actions, even though they induce the same two-form (Aharonov and Bohm, 1959). Since the action is a function on the Yang–Mills configuration space, the one-forms A_μ and B_μ must coordinatize different points. So the Yang–Mills configuration space is not the space of closed two-forms.

Applying the logic of the Newtonian first pass leads to the interpretation of Yang–Mills theory behind Redhead's and Guay's claims. In the first pass, we took the configuration space of Yang–Mills theory to be the space of closed $\mathfrak{g}$-valued two-forms because it's impossible in principle for Newtonian particulate matter to distinguish electromagnetic potentials with the same field strength. Quantum particulate matter discriminates more finely, but as a second pass we can take two one-forms to coordinatize the same point of configuration space if they are in principle indistinguishable to this matter, as well. This leads to the standard story's interest in gauge transformations. An infinitesimal gauge transformation is determined by a $\mathfrak{g}$-valued function on spacetime—in coordinates, a family c^a of scalar fields. This transformation acts on the space of $\mathfrak{g}$-valued one-forms as

$$A^a_\mu \mapsto A^a_\mu + D_\mu c^a \qquad\qquad D_\mu c^a = \partial_\mu c^a + g f^a{}_{bc} A^b_\mu c^c$$

Exponentiating this transformation gives a gauge transformation. Since the action is invariant under all gauge transformations, any two potentials related by a gauge transformation will produce the same dynamics for a quantum particle. According to the received view, applying the interpretive principle from the classical case means taking the configuration space of Yang–Mills theory to be the space of equivalence classes of $\mathfrak{g}$-valued one-forms, where two potentials belong to the same equivalence class if they are related by a gauge transformation.

This reasoning supports the first two parts of the received view, but not the third—at least if we understand it as claiming that we have reason to fix the gauge before even trying to quantize. The first part of the story is that the naive configuration space of Yang–Mills theory is the space of $\mathfrak{g}$-valued one-forms. All hands agree that this space is a candidate for the configuration space of Yang–Mills theory, though few think it's the correct one: Guay claims above that we ought to move to a quotient space; methods of "gauge fixing" are meant to choose a particular model of this quotient; I give a third alternative in Section 3. The second part of the story claims that gauge-equivalent one-forms represent the same physical state of affairs, and in this section I've briefly reconstructed the reasoning for this claim. But this reasoning is incompatible with the version of the standard story that claims gauge fixing as the "obvious move", or that we know gauge fixing to be the "right way" to quantize. In the cases of Newtonian or quantum particulate matter, it was a live possibility that distinct one-forms could generate distinct matter dynamics. As Weingard puts it, "while in classical electrodynamics $[A_\mu]$ can be regarded as a calculational device, it is able to take a physical significance when a new theory like quantum mechanics comes along" (1988, 58). In both cases, we gave a positive argument for interpreting two one-forms as corresponding to the same point of the Yang–Mills configuration space: differences between them are invisible in principle to matter. Maintaining this interpretive principle in the field-theoretic case means giving a positive argument that gauge-equivalent one-forms must produce the same matter dynamics in quantum field theory. And this means looking at the results of quantization.[5]

If the standard story is right, there must be some technical obstruction to quantizing the naive Yang–Mills configuration space, not some conceptual obstruction. It certainly seems that there could be a theory whose configurations correspond to $\mathfrak{g}$-valued one-forms, and the standard reasoning reviewed in this section treats this as an open possibility to be dealt with by positively arguing that some differences are undetectable. And indeed, some presentations of the standard story take the problem to be technical. As Healey puts it,

> path-integral quantization of gauge fields [involves] a difficulty associated with gauge invariance. The path integrals are functional integrals over the space of potentials A_μ^a, and since many such potentials are gauge equivalent to one another, this involves massive 'over-counting.' Not surprisingly, the resulting functional integrals diverge! (2007, 145).

Healey briefly reviews textbook techniques for dealing with this divergence, but says nothing more about its character or how it comes about due to over-counting. More is needed for two reasons. First, quantum field theory is traditionally full of divergences, so we need some reason to think that there are divergences specifically associated with gauge theories. Second, some justification is needed

[5] I intend to leave unspecified the precise content and justification of the interpretive principle at play here. There are really two worries: how can the semiclassical Aharonov–Bohm effect teach us something about classical electromagnetism, and why does the in-principle indistinguishability of gauge-equivalent potentials give us reason to interpret them as representing the same physical state of affairs? On the former, see Belot (1998); on the latter, Dasgupta (2018). I won't address either. My argument is only that if you think that the results of quantization can inform our interpretation of the classical theory, as in the case of the Aharonov–Bohm effect, then by parity of reasoning you ought to think that quantizing the electromagnetic field might lead us to distinguish gauge-equivalent one-forms.

for the claim that over-counting leads to divergence. I will take these in order, indicating a divergence that's distinctive of gauge theories and then trying to connect it to over-counting.

2.2 Three sources of divergence in quantization

Quantization takes as input a classical theory in the form of a classical configuration space and an action on that configuration space. It gives as output perturbative expressions for the quantum expectation value of observables on the classical configuration space. The inner workings of this quantization machine notoriously fall short of most standards of rigor, and the perturbative expansions it produces pose interpretive difficulties that compound the usual interpretive problems of quantum theory—in part because they are plagued by divergences. But we can still seek local justification for particular steps in the quantization process and fend off specific divergences. I take the received view's technical problem with gauge theories to be concerned with a particular divergence and the justification for getting around it. This divergence is associated with the quadratic part of the action and should be distinguished from other divergences that have recently interested philosophers.

The quantum expectation value $\langle \mathcal{O} \rangle$ of some classical observable $\mathcal{O}$ is given by an integral over the classical configuration space, weighted by a phase depending on the action of the classical theory. In certain nice contexts the path integral can be interpreted as a measure-theoretic integral (Johnson and Lapidus, 2000). But in general the equation defining $\langle \mathcal{O} \rangle$ above indicates that the left hand side is to be assigned the power series in $\hbar$ that (loosely speaking) would describe the small-$\hbar$ limit of the integral on the right, if only it existed. It's not obvious what it means, conceptually speaking, to say that the expectation value of an observable is a power series, especially since the series describing the small-$\hbar$ asymptotics of these integrals often diverge.[6] But it's clear enough what it means numerically: just replace $\hbar$ with $10^{-34}\,\mathrm{J\,s}$ and sum up the first couple terms in the series to obtain predictions that agree with experiment to extreme precision.

We can usually overlook a series' divergence by simply ignoring large powers of $\hbar$. But if we're careless about quantizing a gauge theory then every term in the series will itself diverge. These terms are produced by formally generalizing the asymptotic description of analogous finite-dimensional integrals. In the finite-dimensional case, this description is an algebraic expression with the same asymptotic behavior as a computationally intractable analytic integral expression.[7] For example, consider an integral of the form

$$I(\hbar) = \int d^n x \, d^m \theta \, e^{-\frac{1}{\hbar} S(x,\theta)} \, f(x,\theta)$$

where the domain of integration is a supermanifold with n even (commuting) coordinates x^i and m odd (anticommuting) coordinates θ^i, and where f and S

[6] Fraser (2020) suggests that we understand these series as approximations with no underlying exact model, while Miller (2021) develops a novel semantics for theories that assign divergent perturbative expansions to observables.

[7] My exposition of the finite-dimensional model follows Mnev (2019). Under appropriate conditions this approach can even reproduce the exact value of the integral (Johnson-Freyd, 2015).

are supersmooth functions. Suppose that S is even, so that it's of the form

$$S(x,\theta) = S_0(x) + \frac{1}{2}\sum_{i,j} S_{ij}(x)\,\theta^i\theta^j + \cdots$$

where S_0 and S_{ij} are smooth functions of x with S_{ij} skew-symmetric. We approximate $I(\hbar)$ with an algebraic combination of Gaussian integrals, giving a series with the same small-$\hbar$ behavior. This reduces the transcendental content of the integral to the volumes of finitely many Gaussian integrals, which are computable. We obtain the terms of the series assigned to $\langle\mathcal{O}\rangle$ by mimicking the algebra of the finite-dimensional case.

The approximation of $I(\hbar)$ by a series in Gaussians requires the action S to be appropriately non-degenerate. One use of this non-degeneracy pulls out an overall volume factor from the asymptotic series in two steps. The first step uses the Morse lemma to massage the action $S(x,\theta)$ into quadratic terms plus higher-order perturbations. Using partitions of unity, we can write $I(\hbar)$ as a linear combination of integrals with the same form as $I(\hbar)$ and in which S_0 has a single critical point at the coordinate origin, where it also vanishes. Supposing that S satisfies these conditions, we can Taylor expand it to write

$$I(\hbar) \simeq \int d^n x\, d^m\theta\, e^{-\frac{1}{2\hbar}Q_e(x,x)} e^{-\frac{1}{2\hbar}Q_o(\theta,\theta)} e^{\frac{1}{\hbar}p(x,\theta)} f(x,\theta)$$

where

$$Q_e(x,x) = \sum_{i,j} \frac{\partial^2 S_0}{\partial x^i\,\partial x^j}\bigg|_{x=0} x^i x^j \qquad\qquad Q_o(\theta,\theta) = \sum_{i,j} S_{ij}(0)\,\theta^i\theta^j$$

The zeroth-order term in the Taylor expansion of S vanishes by our hypothesis on S_0, and the first-order term vanishes because we are expanding around a critical point. The second step uses the fact that a Gaussian is sharply peaked around its center to replace the first two terms in the integrand with their volumes and the rest of the integrand with its value near the coordinate origin. This results in the following perturbative expression for $I(\hbar)$:

$$I(\hbar) \simeq \sqrt{\frac{(2\pi\hbar)^n}{\hbar^m}}\,\frac{\mathrm{pf}(Q_o)}{\sqrt{|\det Q_e|}}\,\left\langle\!\!\left\langle e^{\frac{1}{\hbar}p(x,\theta)} f(x,\theta)\right\rangle\!\!\right\rangle$$

where pf and det are the Pfaffian and determinant, respectively, and $\langle\!\langle - \rangle\!\rangle$ is the expectation value with respect to the Gaussian. This is not an equality; rather, the two sides of this expression have the same asymptotics: in the small-$\hbar$ limit their difference is smaller than any power of $\hbar$. And when these integrals appear in physics—in optics, for example—it's this asymptotic behavior that meets experiment (Batterman, 2002; Miller, 2021).

We'll see in Section 2.3 that the determinant of Q_e vanishes in naive treatments of gauge theories, making the volume term diverge—or, if $\mathrm{pf}(Q_o)$ also vanishes, making the volume term a $0/0$ indeterminate form. This also causes problems for the expectation value $\langle\!\langle - \rangle\!\rangle$. The expectation value of two coordinate functions—the "propagator" for these coordinates—is a matrix element

$$\langle\!\langle x^i x^j \rangle\!\rangle = (Q_e^{-1})^{ij} \qquad\qquad \langle\!\langle \theta^i\theta^j \rangle\!\rangle = (Q_o^{-1})^{ij}$$

The expectation value of four even coordinate functions is

$$\langle\!\langle x^i x^j x^k x^\ell \rangle\!\rangle = \langle\!\langle x^i x^j \rangle\!\rangle \langle\!\langle x^k x^\ell \rangle\!\rangle + \langle\!\langle x^i x^k \rangle\!\rangle \langle\!\langle x^j x^\ell \rangle\!\rangle + \langle\!\langle x^i x^\ell \rangle\!\rangle \langle\!\langle x^j x^k \rangle\!\rangle$$

and so is determined by the propagators. This is also true for odd coordinates after inserting appropriate signs, and this fact generalizes: the expectation value of any polynomial in x and θ is a polynomial in propagators. If $\det Q_e$ vanishes then Q_e isn't invertible, and so any expectation value involving an even coordinate is ill-defined.

The mathematical issues with gauge theories arise already in the finite-dimensional case, but the physically relevant examples are field-theoretic. The infinite-dimensional generalization of this process mimics the algebraic features of the asymptotic approximation. For simple theories the generalization is relatively direct. For example, consider a theory containing a real scalar field ϕ and a spinor field ψ, interacting via a Yukawa coupling:

$$\mathcal{L}_{\mathrm{Y}}(\phi, \psi) = -\frac{1}{2}(\partial_\mu \phi)^2 - \overline{\psi}\partial\!\!\!/\psi - g\overline{\psi}\phi\psi$$

In the finite-dimensional case the propagator $\langle\!\langle x^i x^j \rangle\!\rangle$ is the inverse of the quadratic form in the even sector. The field-theoretic analogue of this statement says that the propagator $\langle\!\langle \phi(x)\, \phi(y) \rangle\!\rangle$ is the inverse of $-i\partial^\mu \partial_\mu$, which is to say that it's the unique solution to

$$-i\frac{\partial}{\partial x_\mu}\frac{\partial}{\partial x^\mu}\langle\!\langle \phi(x)\, \phi(y) \rangle\!\rangle = \delta(x-y)$$

Likewise, the propagators in the odd sector are the inverses of the corresponding blocks of the quadratic form in the odd sector: the unique solutions to

$$\langle\!\langle \psi(x)\, \psi(y) \rangle\!\rangle = \langle\!\langle \overline{\psi}(x)\, \overline{\psi}(y) \rangle\!\rangle = 0 \qquad\qquad i\gamma^\mu \frac{\partial}{\partial x^\mu}\langle\!\langle \psi(x)\, \overline{\psi}(y) \rangle\!\rangle = \delta(x-y)$$

For any observable $\mathcal{O}$, every term in the expectation value $\langle \mathcal{O} \rangle$ is built from these propagators and the coupling g. So, at least in principle, predicting any expectation value in the quantum theory of these fields reduces to solving these two equations for $\langle\!\langle \phi(x)\, \phi(y) \rangle\!\rangle$ and $\langle\!\langle \psi(x)\, \overline{\psi}(y) \rangle\!\rangle$.

The transition to infinite-dimensional linear algebra invites new divergences, but we will assume these are taken care of since they arise in every quantum field theory, not just gauge theories. For example, we need a determinant for operators on infinite-dimensional vector spaces to make sense of the overall volume factor in the perturbative expansion of $I(\hbar)$. Any definition should reproduce the basic algebraic properties of the finite-dimensional determinant; most importantly for our purposes, the determinant $\det Q$ should vanish if and only if Q is not invertible. A second problem concerns products of propagators. A solution for the above equation for $\langle\!\langle \phi(x)\, \phi(y) \rangle\!\rangle$ will be a distribution singular at $x = y$. This prevents us from defining the expectation value $\langle\!\langle \phi(w)\, \phi(x)\, \phi(y)\, \phi(z) \rangle\!\rangle$ as a product of propagators, since the product of distributions with overlapping singular support is indeterminate. Traditionally this problem is solved in two steps: the indeterminate product of distributions is replaced with a determinate but divergent expression, then this divergence is eliminated by some normalization procedure. For example, the divergence in the propagator can be removed by ignoring its Fourier modes above some large cutoff Λ, allowing us to define products of propagators and thus expectation values $\langle \mathcal{O} \rangle_\Lambda$ that depend on the cutoff. The expectation value $\langle \mathcal{O} \rangle_\Lambda$ diverges in the large-Λ limit, but we can compensate for this divergence by introducing Λ-dependence to the couplings in the action.[8]

Philosophers have recently discussed some divergences of perturbative quantum field theories, but the divergences due to a degenerate action haven't been among them. Much of the literature on perturbative quantum field theory has been concerned with divergences that arise in regularization and renormalization and whether they undercut the coherence of the quantization process (Fraser, 2009, 2011; Halvorson, 2006; Hancox-Li, 2015; Wallace, 2011). The divergences at issue in this literature arise even for very well-behaved quadratic terms, as in the Yukawa theory above. And poorly-behaved quadratic terms can lead to problems even in the finite-dimensional case, where issues of renormalization don't arise. So the divergences related to renormalization are distinct from those related to the quadratic terms. More recent work attends to the interpretation of the divergent series produced by the quantization procedure (Fraser, 2020; Miller, 2021). These divergences do arise in the finite-dimensional case; indeed, the special character of such divergent series was first identified by Poincaré (1892) in the context of celestial mechanics. But they are also distinct from the divergences associated with singular quadratic terms: we can avoid the divergence of asymptotic series by truncation, or attempt to cure it with resummation, but if Q_e has vanishing determinant then the series diverges at every order.

In sum, there is a distinctive divergence associated with gauge theories due to the singularity of the quadratic terms Q_e and Q_o. These quadratic terms arise in the perturbative expansion of Gaussian integrals like $I(\hbar)$, contributing an overall factor to each order of the series through their scalar invariants and to expectation values through their inverses. They produce divergences or indeterminacies when their scalar invariants vanish, which are distinct from those usually addressed by the philosophical literature. This divergence at every order is characteristic of an incautious treatment of gauge theories.

2.3 Over-counting and indeterminism

The divergence isolated in Section 2.2 can be connected to over-counting, but only through certain controversial assumptions about counting and determinism. On the naive approach, the quadratic part of the Yang–Mills Lagrangian is not invertible, so its determinant vanishes and integrals weighted by it have divergent asymptotics. Physically speaking, you might say that the problem is one of determinism: since gauge transformations are spacetime-dependent there are many futures for any initial datum, each pair of which are related by a gauge transformation that leaves the initial datum alone, and this spoils the invertibility of the free dynamics. And if you suppose that determinism is a matter of possibility-counting, then you can blame the divergence on over-counting. But the ways to avoid the divergence offered by the received view aren't related to over-counting. Nor do they adequately deal with the divergence problem.

[8]See Butterfield and Bouatta (2015) and Hancox-Li (2015) for more detailed philosophical discussions of this method, and Duch et al. (2021) for discussion of one alternative.

The gauge invariance of the Yang–Mills action makes it degenerate, causing the asymptotics of gauge-theoretic path integrals to diverge and expectation values to be ill-defined. Consider the $\mathfrak{u}(1)$ case. Naively, the Yang–Mills configuration space is the space of $\mathfrak{u}(1)$-valued one-forms, and the action is the Yang–Mills action above. An inverse for the quadratic part of the action would produce the unique solution to the equation

$$-i(\eta^{\mu\nu}\partial^{\lambda}\partial_{\lambda} - \partial^{\mu}\partial^{\nu})A_{\mu} = j^{\nu}$$

for any current j^{ν}. But if A_{μ} is a solution to this equation then so is $A_{\mu} + \partial_{\mu}c$ for any $\mathfrak{u}(1)$-valued function c. Interpreted electromagnetically, this is Maxwell's equation, and its singularity amounts to the fact that any gauge transformation of a solution is also a solution. In the nonabelian case this interpretation is complicated by the non-linearity of the Yang–Mills equation, but the mathematical point is the same: the quadratic part of the action cannot be inverted, so the propagator $\langle\!\langle A^{a}_{\mu}(x)\, A^{b}_{\nu}(y)\rangle\!\rangle$ is ill-defined and the determinant of the quadratic part vanishes, making each term in the asymptotic expansion of any integral diverge.

This mathematical glitch has a natural interpretation as a failure of determinism. The spacetime-dependence of gauge transformations makes them "spoilers" in the sense of Belot (2008). That is, for any solution A^{a}_{μ} to the equations of motion and any initial data surface Σ we can choose some $\mathfrak{g}$-valued function c^{a} with compact support to the future of Σ, giving two solutions A^{a}_{μ} and $A^{a}_{\mu} + D_{\mu}c^{a}$ that coincide on Σ. In yet other words, there can be no Cauchy surface for the Yang–Mills equation for $\mathfrak{g}$-valued one-forms—no surface Σ such that restriction to Σ exhibits an equivalence between instantaneous states on Σ and solutions to the Yang–Mills equation, even with boundary conditions in place. This amounts to a certain kind of indeterminism: any instantaneous state has infinitely many possible futures. In the philosophical literature, this failure of determinism is often taken as a mark against the idea that the configuration space of Yang–Mills theory is the space of $\mathfrak{g}$-valued one-forms (Belot, 2003, 2008; Healey, 2007; Lyre, 2004).

But our divergence's connection to determinism and possibility-counting relies on substantial accounts of the latter, and regaining determinism won't necessarily resolve the divergence. As Belot puts it, the connection between Cauchy surfaces and determinism "obtains only if we assume that distinct solutions of our theory always represent physically distinct situations" (2008, 200). And as Section 2.1 reviewed, the received view of Yang–Mills theories denies this assumption. There has been extensive philosophical work on how one ought to count physical possibilities in this context—and about whether possibility-counting is relevant to determinism at all—and on these accounts Yang–Mills theory counts as deterministic if we take gauge-equivalent one-forms to represent the same physical state of affairs (Belot, 1995, 2008; Brighouse, 1997; Butterfield, 1989; Dasgupta, 2011; Melia, 1999). But none of this makes the quadratic part of the action invertible. So while there may be some connection between over-counting and divergent asymptotics, any such connection must go through some principles connecting determinism and possibility-counting with the non-degeneracy of the action.

Setting aside the interpretation of the divergence, the philosophical literature's mathematical response to the problem is to "fix the gauge", as Redhead suggests above. This is accomplished in one of two ways; in both approaches, ghost fields appear as a mathematical trick. Because the Yang–Mills action is degenerate on the one-form interpretation, that theory has no perturbative quantization by the standard route. A quantum version of Yang–Mills theory must therefore have a different configuration space, a different action, or be produced by a different quantization procedure. Philosophical discussions tend to avoid the details at this point, but allude to replacing the action or modifying the quantization procedure.

The action approach attempts to fix the gauge by adding a Lagrange multiplier to the action. For example, consider the Lagrangian

$$\mathcal{L}_\xi(A) = -\frac{1}{4}(F^a_{\mu\nu})^2 - \frac{1}{2\xi}(\partial^\mu A^a_\mu)^2$$

with Lagrange multiplier $1/\xi$. Classically, the effect of this modification is to add the condition $\partial^\mu A^a_\mu = 0$ to the equations of motion. The critical surface of S_ξ is therefore a strict subset of the critical surface of the Yang–Mills action, so there are fewer classical solutions over which we must integrate when computing the asymptotic expansion of an integral weighted by S_ξ. Moreover, the quadratic part of S_ξ is invertible: in the case $\xi = 1$ its inverse is

$$\langle\!\langle A^a_\mu(x)\, A^b_\nu(y) \rangle\!\rangle = \int \frac{d^4k}{(2\pi)^4}\, \frac{-i\delta^{ab}\eta_{\mu\nu}}{k^2}\, e^{ik\cdot(x-y)}$$

So the quadratic part of S_ξ is invertible, its determinant doesn't vanish, and the divergence is avoided. The resulting predictions are even empirically adequate in some cases.[9]

This approach to gauge fixing requires ever more modifications to the action; one of these is the addition of ghost fields. The action S_ξ allows for perturbative quantization, but in the nonabelian case the resulting quantum theory is pathological. As Feynman (1963) first realized, quantizing S_ξ for nonabelian $\mathfrak{g}$ gives a non-unitary theory. The problem comes from the Yang–Mills potential's self-interaction. For example, the term $gf^a{}_{bc}(\partial^\mu A^{a\nu})A^b_\mu A^c_\nu$ in the action allows one quantum of the A^a_μ field to decay into two, and the outgoing quanta might have disallowed polarizations.[10] Feynman's solution to this problem was to add an "artificial, dopey particle" (1963, 710) to give the Lagrangian[11]

$$\mathcal{L}_{\mathrm{FP}}(A, c, \overline{c}) = -\frac{1}{4}(F^a_{\mu\nu})^2 - \frac{1}{2\xi}(\partial^\mu A^a_\mu)^2 - \overline{c}_a\, \partial^\mu D_\mu c^a$$

Here c^a is an anticommuting $\mathfrak{g}$-valued scalar field, and $\overline{c}_a$ is its conjugate, an anticommuting $\mathfrak{g}^*$-valued scalar field. This new field is a "ghost" in the technical sense that its kinetic term has the wrong sign. Its spin–statistics relation is also dopey: it's anticommuting, but has integer spin.[12] The practical effect of these new terms is to compensate for the timelike- and longitudinally-polarized modes

[9]For example, this propagator is good enough for much of quantum electrodynamics; for textbook treatments see Peskin and Schroeder (1995, Ch. 5), Schwartz (2014, Part II), or Weinberg (1995, §8.7).

[10]See Peskin and Schroeder (1995, §16.3) for a pedagogical discussion of this problem.

[11]More precisely, Feynman devised appropriate rules for the one-loop level, and DeWitt (1967) for every level of perturbation theory, which ensure that the optical theorem holds. These are the rules obtained from $\mathcal{L}_{\mathrm{FP}}$ by the method discussed in Section 4.1.

[12]Since the connection between integer spin and Bose–Einstein statistics is a theorem in

created by the Yang–Mills potential's self-interaction. The last term in $\mathcal{L}_{\mathrm{FP}}$ is dictated by the requirement that it have this effect.

The action-modification approach has both conceptual and computational shortcomings. The addition of ghost fields is facially ad hoc, meant to formally patch a failure of unitarity. This makes it difficult to say anything at all about their conceptual role—though, as I'll return to in Section 4.1, you can try to argue that they're a coordinate artifact if you can find a gauge-fixed action that doesn't require them. There's also conceptual problems with the first step. Adding a "gauge-fixing term" doesn't change of the domain of integration, as Redhead's and Healey's glosses suggest. It does eliminate some of the action's critical points, but reducing the number of classical solutions does nothing to avoid our divergence, because the asymptotic expansion around *each* critical point diverges. The "gauge-fixing term" repairs the divergence by changing the quadratic part of the action; the domain of the integral is still the space of $\mathfrak{g}$-valued one-forms. The action-modification approach has little to do with "over-counting". So it's hard to see how this approach can get off the ground, conceptually speaking. Practically speaking, this method will always work—as ad hoc modifications are wont to do—but it gives no systematic recipe for identifying and fixing failures of unitarity or renormalization.

The second approach to gauge fixing modifies the quantization procedure by inserting two forms of the identity and formally applying Fubini's theorem. Following 't Hooft (1971), preface the S_{YM} path integral with the identity

$$1 = \frac{1}{\mathcal{N}(\xi)} \int \mathcal{D}\omega^a \, \exp\!\left(-\frac{i\xi}{2\hbar}\omega_a\omega^a\right)$$

where ω^a is a $\mathfrak{g}$-valued scalar field and $\mathcal{N}(\xi)$ a (divergent) constant defined to make this equality hold. To fix the $\partial^\mu A^a_\mu = 0$ gauge, follow Faddeev and Popov (1967) by inserting the scaling property of the δ distribution

$$1 = \int \mathcal{D}\alpha^a \, \delta\big(\partial^\mu(A^a_\mu + D_\mu\alpha^a) - \omega^a\big)\, \det(\partial^\mu D_\mu)$$

into the integrand, where the domain of integration is all $\mathfrak{g}$-valued scalar fields. Integrating over ω^a and shifting A^a_μ while holding the determinant constant gives

$$\frac{1}{\mathcal{N}(\xi)} \int \mathcal{D}\alpha^a \, \mathcal{D}A^a_\mu \, e^{\frac{i}{\hbar}S_\xi} \det(\partial^\mu D_\mu)$$

The final trick, also due to Popov and Faddeev (1967), writes the determinant as a Gaussian integral over odd coordinates, leaving us with

$$\frac{1}{\mathcal{N}(\xi)} \int \mathcal{D}\alpha^a \, \mathcal{D}A^a_\mu \, \mathcal{D}c^a \, \mathcal{D}\bar{c}_a \, e^{\frac{i}{\hbar}S_{\mathrm{FP}}}$$

These three tricks reproduce the gauge–ghost action of the ad hoc method, so they avoid the loss of unitary that Feynman identified. Moreover, they give a

structural analyses of relativistic quantum field theory, ghosts must violate some plausible assumption. And they do: they are states of negative norm, so the space of quantum states isn't a Hilbert space. However, the subspace of in and out states is; see Section 3.2 for more on this.

systematic method for avoiding this loss, applicable to more—but not all—choices of gauge-fixing conditions.[13]

The Faddeev–Popov method has clear advantages over the ad hoc addition of ghost fields and interaction terms, but it's ultimately still inadequate for both conceptual and practical purposes. The meaning of the new volume $\int \mathcal{D}\alpha^a/\mathcal{N}(\xi)$ is opaque, and it diverges even for the finite-dimensional model reviewed in Section 2.2, preventing any interpretation by analogy. But if we can set this problem aside then there are at least two remarks we can make. First, as Redhead puts it, the introduction of the ghost field c^a is "a purely mathematical manoeuvre" (2003, 135). The Faddeev–Popov method introduces ghosts for the purposes of computing the determinant $\det(\partial^\mu D_\mu)$, but any other method of computation would work just as well. So we can bypass ghost fields entirely, at least in principle, by computing the determinant using some other method. Second, the Faddeev–Popov method has no obvious relation to "over-counting". Two insertions of the identity and a re-writing don't make a difference to which configurations we integrate over. Indeed, up to the ∞/∞ indeterminacy in front, this integral is (by design) precisely the same as the one produced by the ad hoc insertion of ghosts, and as we've already seen this integral doesn't fit the "gauge fixing" gloss. I don't think we can set aside the new volume term, even for qualitative purposes. But even if we follow the traditional philosophical account in doing so, its interpretive claims do not follow.

In addition to its conceptual issues, the Faddeev–Popov method is inadequate for practical purposes. There are problems of at least two kinds, illustrated by the following two examples:

1. Because the ghost fields are introduced when writing a determinant of a Yang–Mills-dependent operator as a Gaussian, any Lagrangian produced by the Faddeev–Popov method will be at most quadratic in the ghost fields.[14] But some gauge-fixing conditions require ghost self-interactions, even in Yang–Mills theory (Zinn-Justin, 1975, §4). For example, a maximal abelian gauge condition writes the gauge algebra $\mathfrak{g}$ as $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{a}$ for some maximal abelian subalgebra $\mathfrak{a}$ of $\mathfrak{g}$, then fixes the $\mathfrak{h}$-components of A^a_μ with a maximality condition and treats the $\mathfrak{a}$-components like electromagnetic potentials. If the Faddeev–Popov procedure worked, then applying it to this gauge-fixing condition would produce the same integral as it did when applied to the $\partial^\mu A^a_\mu = 0$ condition. But it doesn't: the action S_{FP} is perturbatively renormalizable, while the Faddeev–Popov procedure applied to maximal abelian gauge produces a $\bar{c}cAA$ term that generates a $\bar{c}c\bar{c}c$ counterterm at the first loop level. So the Faddeev–Popov method does not correctly implement maximal abelian gauge.

2. A second source of ghost interactions appears in higher gauge theories, which include gauge transformations between gauge transformations (Henneaux and Teitelboim, 1992, Ch. 10; Weinberg, 1995, §§8.8, 15.8). For

[13]For a more detailed treatment of this procedure, see Peskin and Schroeder (1995, §9.4), Schwartz (2014, §25.4), or Weinberg (1995, §15.5).

[14]This is also true of Guay's (2008) treatment of BRST quantization, which follows DeWitt (2005). As such, I take Guay's analysis to ultimately apply only to Faddeev–Popov quantization, which happens to compute the BRST quantization in special cases.

example, consider a two-form potential $B_{\mu\nu}$ with Lagrangian

$$\mathcal{L}(B) = -\frac{1}{6} H_{\lambda\mu\nu} H^{\lambda\mu\nu} \qquad\qquad H_{\lambda\mu\nu} = \partial_\lambda B_{\mu\nu} + \partial_\nu B_{\lambda\mu} + \partial_\mu B_{\nu\lambda}$$

Any one-form A_μ gives a transformation

$$B_{\mu\nu} \mapsto B_{\mu\nu} + \partial_\mu A_\nu - \partial_\nu A_\mu$$

that preserves the Lagrangian, and any scalar field c gives a transformation $A_\mu \mapsto A_\mu + \partial_\mu c$ that preserves the action of A_μ on $B_{\mu\nu}$. On the ad hoc approach, gauge fixing leads to a loss of unitarity in processes involving the $B_{\mu\nu}$ field, which are patched up by coupling it to an anticommuting vector field c_μ. Fixing the gauge of the new vector field again breaks unitarity, which is fixed by coupling c_μ to a commuting scalar c. More generally, the p-form gauge fields appearing throughout supergravity and string theory lead to $(p-1)$-form ghost fields whose non-unitarity is fixed by $(p-2)$-form ghost fields, and so forth.

These cases indicate that the Faddeev–Popov procedure can't be whole story. It works when it works, but it doesn't generally resolve the characteristic divergence of gauge theories. So we should avoid drawing any conclusions about ghosts from the Faddeev–Popov procedure without some account of how it is a special case of a more general picture. A more general story isn't on offer in the extant philosophical literature, so in the next section I provide one.

The standard philosophical story has it that ghosts are introduced as a technical convenience in the process of gauge fixing—a process by which we restrict the domain of the path integral—which we ought to do for conceptual or mathematical reasons. In this section I've disputed each part of this story. Most philosophical treatments of gauge theory are set against an account of gauge equivalence that focuses on classical electromagnetism: first we interpret Yang–Mills theory as a model of Maxwell electromagnetism, then the Aharonov–Bohm effect prompts us to revise our interpretation, taking the potential A_μ to coordinatize its gauge-equivalence class rather than its field strength $F_{\mu\nu}$. We have reason to doubt that a quantum Yang–Mills theory distinguishes gauge-equivalent one-forms, but we don't have dispositive reason to think that it won't. Investigating this possibility brings us to a mathematical problem: gauge theories exhibit a characteristic divergence in perturbative quantization, identified in Section 2.2. This divergence can be connected to over-counting under a sufficiently simplistic conception of determinism, but the techniques used to circumvent the divergence that were mooted in this section don't change which potentials get counted. They also don't give any reliable information about how we ought to interpret ghost fields. Feynman's ad hoc approach is a good guide to when ghosts and their interactions are necessary for a well-behaved theory with a modified Yang–Mills action, but it gives no theory to guide interpretation of these interactions. The Faddeev–Popov procedure gives an account of ghosts on which they are a calculational device, but this account doesn't cover every gauge theory or even every part of Yang–Mills theory. The philosophical literature's gloss on the technical facts doesn't hold up, so we aren't warranted in drawing any conclusions about ghosts from that gloss.

3 An alternative story

The philosophical literature is wrong about ghosts, but it's right that the one-form interpretation of Yang–Mills theory has no perturbative quantization. As I'll argue in this section, it's even right to say that identifying gauge-equivalent one-forms gives a good quantum theory. But in this new theory ghost fields are more than a mere calculational device: they are used to identify gauge-equivalent one-forms and there are coordinate-invariant facts about them. On this alternative interpretation, the apparent divergence of Section 2.2 is merely a coordinate singularity born of neglecting part of the classical configuration space. A change of coordinates resolves the singularity and explicitly introduces ghost dependence to the action. This approach avoids the two kinds of problems faced by the Faddeev–Popov method. It's also compatible with the logic behind the usual philosophical approach to gauge theories in Section 2 and the work on classical gauge theories referenced in this paper's introduction.

3.1 Stacky Yang–Mills

We cannot perturbatively quantize one-form Yang–Mills theory, so we need an alternative classical theory. One alternative is suggested by the work of Becchi et al. (1976) and Tyutin (1975). Becchi–Rouet–Stora–Tyutin (BRST) quantization can be understood as standard perturbative quantization applied to a classical configuration space that contains gauge structure—that is, nontrivial equivalences between its points. Because the Yang–Mills action doesn't depend on the (odd) gauge dimensions, the integral over this configuration space has a 0/0 indeterminacy. We can see this as a coordinate singularity: the gauge structure of the configuration space means that there are many gauge-equivalent representations of the action, and choosing one that explicitly depends on the gauge coordinates results in a well-defined perturbative quantization. These gauge coordinates are precisely the ghost fields.

In Section 2.3 I argued that reconceiving possibility-counting and determinism doesn't directly solve the divergence problem, since it need not make the quadratic part of the action invertible. But meditation on these concepts can suggest alternative choices for the Yang–Mills configuration space. For example, philosophers often want to take a set-theoretic quotient, making the configuration space the set of gauge-equivalence classes of one-forms.[15] This choice of configuration space is interesting and pursuit-worthy, but I think it's implausible. Taking the set-theoretic quotient restores determinism in the naive sense, and it gets the possibility-counting right, but it destroys important information about the gauge structure of the theory. Among other things, the set-theoretic quotient erases the fact that the possible future of an initial datum is not just unique up to gauge equivalence, but up to unique gauge equivalence (Benini et al., 2018). This implies that the set-theoretic quotient satisfies the naive definition of determinism, but it's a strictly stronger condition that's inexpressible in the quotiented theory.

A less destructive approach with similar motivations takes the homotopy quotient of the space of $\mathfrak{g}$-valued one-forms by the action of gauge transformations. The result is the category in which an object is a $\mathfrak{g}$-valued one-form and an

[15] This popular desire is made explicit—though not necessarily endorsed—by Belot (2001, 2003), Butterfield (2006), Caulton (2015), and Dewar (2019).

arrow from A^a_μ to B^a_μ is a gauge transformation that sends A^a_μ to B^a_μ. Because gauge transformations are invertible this category is a groupoid, meaning that every arrow is an isomorphism. And because everything in sight is smooth, this groupoid is also smooth; in other words, Lie groupoids are good finite-dimensional analogues for our spaces of interest.[16] This construction is a type of quotient because it identifies gauge-equivalent one-forms—treating this groupoid category-theoretically means treating isomorphic objects as the same—but it doesn't eliminate the gauge structure as the set-theoretic quotient would. A real-valued function on the space of $\mathfrak{g}$-valued one-forms descends to a map on the quotient groupoid just in case it's gauge invariant; in particular, the Yang–Mills action gives such a map. For brevity, call this configuration space and action the "stacky interpretation" of Yang–Mills theory.[17]

The stacky interpretation of Yang–Mills theory has a good perturbative quantization, up to a $0/0$ indeterminacy. Perturbative integrals depend only on an infinitesimal neighborhood of the action's critical points. For a Lie groupoid, this means the integral depends on the Lie algebroid obtained by differentiating the Lie groupoid. And Lie algebroids can be modelled by structured supermanifolds, to which we can apply the perturbative integration theory of Section 2.2. Differentiating the configuration space of stacky Yang–Mills theory gives the Lie algebroid coordinatized by a pair (A^a_μ, c^a), where A^a_μ is a family of even coordinates that assemble into a $\mathfrak{g}$-valued one-form and c^a a family of odd coordinates that assemble into a $\mathfrak{g}$-valued function. The first entry coordinatizes the collection of objects of the stacky Yang–Mills configuration space, while the second entry coordinatizes the "infinitesimal arrows" with domain A^a_μ. Since the Yang–Mills action doesn't explicitly dependent on the arrow coordinates, the quadratic part of the odd sector vanishes. And since it isn't invertible in the configuration coordinates, the quadratic part of the even sector has vanishing determinant. Each term in the perturbative expansion of an integral over this supermanifold therefore includes a volume term

$$\frac{\mathrm{pf}(0)}{\sqrt{|\det(Q_e)|}} = \frac{0}{0}$$

Though this indeterminacy is a problem, it's more tractable than a divergence: an indeterminacy can be fixed by a change in coordinates, while a divergence can only be fixed by a change in definition.

To resolve the indeterminacy due to the quadratic part of the Yang–Mills action we can change coordinates; permissible coordinate changes are controlled by the structure of the Lie algebroid. Indeterminate forms can often be resolved by rewriting; for example, the function x/x is of the form $0/0$ in the small-x limit, but simplifying the division shows that its limit is 1. We have a lot of freedom when rewriting functions on a Lie groupoid, since these functions are

[16] See Crainic and Fernandes (2011) for an overview of Lie groupoids, Lie algebroids, and their relation. For the purposes of perturbative quantization we use Vaĭntrob's (1997) formulation of Lie algebroids as differential graded manifolds; see also Mnev (2019, 4.2.28–29).

[17] The name refers to the fact that on this interpretation, configurations of the Yang–Mills field are given by sections of a stack, or homotopy sheaf of groupoids, in contrast with the space of states in many other fields theories, which is given by a section of a sheaf of sets. For a more detailed treatment, see Benini et al. (2018). What follows is an interpretation of BRST quantization that combines the cohomological interpretation of BRST quantization with the higher-categorical interpretation of cohomology; I learned these interpretations from Schreiber (2017).

only defined up to isomorphism. Differentiating the isomorphism structure of a Lie groupoid gives a vector field on its Lie algebroid in the form of a derivation δ on the space of real-valued supersmooth functions; because this vector field comes from differentiating a Lie groupoid, it satisfies $\delta^2 = 0$. A function f on the Lie algebroid is infinitesimally gauge invariant if δf vanishes, and two functions are infinitesimally gauge equivalent if their difference is δh for some h. So to remove the $0/0$ indeterminacy in the perturbative expansion of the path integral, we can try to replace the Yang–Mills Lagrangian $\mathcal{L}_{\mathrm{YM}}$ with a gauge-equivalent Lagrangian $\mathcal{L}_{\mathrm{YM}} + \delta h$ whose quadratic part is invertible.

In the case of stacky Yang–Mills there's a hiccup in this strategy. The distinguished vector field on the Yang–Mills algebroid acts as

$$\delta A^a_\mu = D_\mu c^a \qquad\qquad \delta c^a = -\frac{1}{2} g f^a{}_{bc} c^b c^c$$

The action of δ on the configuration coordinates is given by differentiating the action of gauge transformations on these coordinates, while the action on the gauge transformation coordinates is given by differentiating the adjoint action of gauge transformations on infinitesimal gauge transformations. These expressions show that δ introduces one new gauge coordinate when applied to any monomial. In physicists' terminology, any monomial function on the Lie algebroid has a "ghost number", which the vector field δ increments. This blocks the action-replacement strategy: to replace $\mathcal{L}_{\mathrm{YM}}$ with another Lagrangian $\mathcal{L}_{\mathrm{YM}} + \delta h$ of ghost number zero, the function h must have ghost number -1. Since our only coordinates have ghost number zero (A^a_μ) or one (c^a), there are no candidates for h.

Though the stacky Yang–Mills configuration space has no coordinates with ghost number -1, it's equivalent to one that does. An equivalence of Lie algebroids is a supersmooth map that induces a bijection on gauge-equivalence classes of gauge-invariant functions. For example, consider the Lie algebroid with coordinates $(A^a_\mu, c^a, \bar{c}_a, b_a)$ and whose distinguished vector field acts as

$$\begin{aligned} \delta A^a_\mu &= D_\mu c^a & \qquad \delta \bar{c}_a &= b_a \\ \delta c^a &= -\frac{1}{2} g f^a{}_{bc} c^b c^c & \qquad \delta b_a &= 0 \end{aligned}$$

Here A^a_μ and c^a are the same as the coordinates on the stacky Yang–Mills configuration space. The new coordinate $\bar{c}_a$, which we'll call the "antighost" coordinate, is an odd $\mathfrak{g}^*$-valued scalar field with ghost number -1, while the Nakanishi–Lautrup field b_a is an even $\mathfrak{g}^*$-valued scalar field of ghost number 0. Since δb_a vanishes, the auxiliary coordinate b_a is a gauge-invariant function, but because $b_a = \delta \bar{c}_a$, the function b_a is gauge equivalent to the constant zero function. So the new coordinate b_a adds no gauge-invariant functions. The antighost coordinate $\bar{c}_a$ isn't gauge invariant, since $\delta \bar{c}_a$ is nonzero, and it adds no gauge-invariant functions either. In other words, the Lie algebroid coordinatized by $(A^a_\mu, c^a, \bar{c}_a, b_a)$ carries the same gauge-equivalence classes of gauge-invariant functions as the stacky Yang–Mills configuration space, despite the different coordinate functions. That is, these Lie algebroids are equivalent and hence have the same perturbative quantization.[18]

[18] More precisely, a map of Lie algebroids is an equivalence, or quasi-isomorphism, if it

With these rewritings out of the way, we can quantize stacky Yang–Mills theory by choosing an appropriate gauge-equivalent replacement of the Yang–Mills action. For example, consider the function

$$h_{\mathrm{F}} = \overline{c}_a \left(\frac{1}{2} b_a + \partial^\mu A_\mu^a \right)$$

The Yang–Mills Lagrangian $\mathcal{L}_{\mathrm{YM}}$ is gauge equivalent to the Lagrangian $\mathcal{L}_{\mathrm{F}} = \mathcal{L}_{\mathrm{YM}} + \delta h_{\mathrm{F}}$, which is

$$\mathcal{L}_{\mathrm{F}}(A, c, \overline{c}, b) = -\frac{1}{4}(F^a_{\mu\nu})^2 - \frac{1}{2}(\partial^\mu A^a_\mu)^2 - \overline{c}_a \partial^\mu D_\mu c^a + \frac{1}{2}\left(b_a + \partial^\mu A^a_\mu\right)^2$$

This Lagrangian reproduces the results of the Faddeev–Popov procedure while also providing an interpretation of the ∞/∞ indeterminacy of that theory: it's the volume of the Gaussian in b_a. This indeterminacy will therefore be taken care of by whatever regularization method we use for functional determinants.

Stacky Yang–Mills avoids the two technical shortcomings of the Faddeev–Popov procedure. Because the stacky Yang–Mills configuration space has gauge structure from the start, there's no limitation on the number or kind of ghost coordinates or their self-interactions. This gives it two kinds of flexibility:

1. We can add any term of the form δh to the Yang–Mills Lagrangian to give a different representative, as long as h has ghost number -1. As such, we can have interaction terms of arbitrarily high degree. For example, we can add a term of the form
$$h = f^a{}_{bc} c^a \overline{c}_b \overline{c}_c$$
Then δh includes a term quartic in the (anti)ghost fields, which cannot be generated by the Faddeev–Popov procedure. Faddeev–Popov quantization works in those cases that h is of the form $\overline{c}_a\, G(A)$ for some function G that's linear in A^a_μ and its derivatives, but generally fails for non-linear gauge-fixing conditions.

2. Higher gauge theories may be quantized along exactly the same lines, taking the homotopy quotient to produce a higher Lie groupoid and then differentiating to produce a higher Lie algebroid. For the two-form theory in Section 2.3, this is the Lie algebroid with coordinates $(B_{\mu\nu}, c_\mu, c)$, where c_μ has ghost number 1 and c has ghost number 2, and the distinguished vector field acts as
$$\delta B_{\mu\nu} = \partial_\mu c_\nu - \partial_\nu c_\mu \qquad \delta c_\mu = \partial_\mu c \qquad \delta c = 0$$
Quantization of this theory proceeds as in the one-form case, adding terms of the form δh to the action as needed to give an invertible quadratic part.

Because the ghost coordinates are part of the stacky Yang–Mills configuration space itself, rather than a tool for computing a particular determinant, we have

induces an isomorphism on the cohomology of δ—that is, on the gauge-equivalence classes of gauge-invariant functions. If the Berezinian of the integral is gauge invariant, then the integral depends only on the cohomology class of the integrand (Mnev, 2019, §4.3.2). On the stacky interpretation of Yang–Mills the Berezinian must be gauge invariant, so equivalent Lie algebroids are interchangeable domains of integration.

the kind of direct control required to explain why the Faddeev–Popov procedure works when it does and to quantize those theories for which it does not.

The perspective on BRST quantization sketched in this section gives a general, unitary, renormalizable quantization procedure for classical gauge theories that's empirically adequate when applied to Yang–Mills theory and has an appropriate classical limit. Interest in the stacky Yang–Mills configuration space is not only motivated by the same reasoning as the ad hoc or Faddeev–Popov quantization methods, but executes that reasoning more faithfully: the stacky Yang–Mills configuration space actually identifies gauge-equivalent one-forms, unlike procedures that simply add gauge-fixing terms to the action. Differentiating the resulting Lie groupoid gives a Lie algebroid to which we can apply the standard perturbative quantization procedure. Applying this procedure to the Yang–Mills action produces a $0/0$ indeterminacy, since S_{YM} is independent of the gauge coordinates, and this can be resolved by a change of coordinates. The results of this coordinate change can include terms of arbitrarily high degree in c^a and $\bar{c}^a$, and for higher gauge theories they can also include higher ghosts. The ad hoc and Faddeev–Popov quantization procedures work insofar as they compute the results of BRST quantization.

3.2 The stacky story's virtues

The conception of BRST quantization described in Section 3.1 is technically superior to the quantization methods usually discussed in the philosophical literature. It also has philosophical advantages; I will mention two. First, ad hoc treatments of ghosts can make them look like surplus mathematical structure that plays "a mysterious, even mystical, Platonist–Pythagorean role" (Redhead, 2003, 138). But according to the stacky interpretation, ghost fields are coordinates on a configuration space obtained by eliminating structure, not adding it, and they're not surplus. Second, stacky Yang–Mills theory is compatible with the logic of the philosophical literature on the interpretation of classical Yang–Mills theory in a way the received view is not.

Most philosophical discussions of BRST quantization treat it as a novel type of symmetry that is either mysterious or a technical convenience. This is a natural consequence of the received view. As I outlined in Section 2.3, the received view responds to a gauge theory's characteristic divergence by adding some terms to the action to give it an invertible quadratic part. This theory is generally non-unitary and perturbatively non-renormalizable, but can be repaired with the further ad hoc addition of ghost terms. On the received view, the role of the auxiliary field b_a is to gives the action "BRST symmetry" in the form of the equation $\delta S = 0$, which can be used to show that scattering will not produce disallowed polarizations and perturbative renormalization will generate no new counterterms. Because this new symmetry arises from a sequence of ad hoc modifications, it has no obvious physical interpretation; it's merely "chosen in such a way as to ease quantization" (Rickles, 2008, 178).

Redhead interprets this sequence of modifications as a successive increase in surplus structure, or mathematical structure with no physical correlate. On Redhead's account, gauge theories are rife with surplus structure. In classical Yang–Mills theory, gauge-equivalent one-forms represent the same physical state of affairs, so Redhead takes the gauge-variant differences to be surplus in this sense. But when we quantize this theory by ad hoc methods, we seem to appeal to just those gauge-variant features when adding a gauge-fixing term to the action. And after we've removed the classical surplus by gauge fixing, we add entirely new surplus in the form of (anti)ghost and auxiliary fields. Redhead describes this situation as "mysterious". I take it that the mystery is about how purely mathematical structure like ad hoc ghost fields can have physical consequences, like unitarity.

If we adopt the approach to quantization in Section 3.1 then the mystery dissipates. We might indeed say that the space of $\mathfrak{g}$-valued one forms has surplus mathematical structure with respect to the stacky Yang–Mills configuration space. After all, the latter is obtained by identifying gauge-equivalent one-forms, so it has less structure, and it is also adequate for quantization purposes, so the structure that was eliminated was surplus. But the (anti)ghost and auxiliary fields aren't surplus. The ghost field is a coordinate function on the stacky configuration space; it represents the gauge structure of the theory, and this structure has physical content. Because the ghost field implements the homotopy quotient that eliminates the surplus structure from the space of $\mathfrak{g}$-valued one-forms, it represents a subtraction of structure, not a surplus. The antighost and auxiliary fields don't add any structure, either. Indeed, they are chosen precisely so that the Lie algebroid coordinatized by $(A^a_\mu, c^a, \bar{c}_a, b_a)$ has the same structure as the one coordinatized by (A^a_μ, c^a)—that is, so that these Lie algebroids are equivalent. On the picture of quantization I offered in Section 3.1 there's no mystery of surplus structure; all of the structure is accounted for.[19]

To be clear: ghost fields aren't surplus on the stacky interpretation, but this doesn't mean that we should go hunting for them in particle colliders. Redhead provides a particular formal account of surplus structure, but I claim—and argue, in future work—that stacky Yang–Mills theory violates the assumptions of his account. So we must fall back on the ordinary language meaning of "surplus": "eliminable", "excess", "superfluous", "unnecessary", and so on. The odd dimensions of the classical configuration space aren't surplus in this sense, because they can't be removed. But we shouldn't expect to find ghost quanta in incoming and outgoing states any more than we should expect to find quanta of the equally indispensable path integral measure (whatever that could mean). Ghosts coordinatize gauge transformations, not configurations, so it would make no physical sense for them to appear in asymptotic states. Of course, unitarity prevents us from simply throwing out states we don't like; we bumped up against this constraint in Section 2.3, where one-form Yang–Mills theory with the Lagrangian $\mathcal{L}_\xi$ gave nonzero amplitude to produce timelike- and longitudinally-polarized quanta of the Yang–Mills field. So the stacky interpretation of Yang–Mills theory is only consistent with unitarity if there is no amplitude to create ghost quanta in scattering. If we assume that the matter charges cooperate,

[19] Redhead identifies three kinds of surplus structure in the quantization of gauge theories: ghost fields, ghost-of-ghost fields, and the antifields of the Batalin–Vilkovisky formalism (2003, 137). On the stacky interpretation of gauge theories, ghost fields coordinatize gauge dimensions of the classical configuration space, while ghost-of-ghost fields coordinatize gauge-of-gauge dimensions. In both cases these stacky features implement the homotopy quotient by gauge transformations, which plays the conceptual role of the quotient for actions that aren't free (and agrees with the set-theoretic quotient for actions that are). Antifields, on the other hand, implement the derived intersection of the graph of dS with the zero section, which plays the conceptual role of the intersection for submanifolds that aren't transverse—in physics terminology, when the gauge algebra doesn't close off-shell. See Calaque (2015) for more on this distinction.

then the quartet mechanism identified by Kugo and Ojima (1978) ensures a vanishing amplitude for ghost creation. And it is generally assumed that the matter charges do cooperate. This leads to predictions discussed in Section 4.2.

Stacky Yang–Mills theory is able to avoid the mystery of surplus structure in part because it conforms to the logic of Section 2.1, unlike approaches that modify the quantization procedure. Recall that the analyses of Yang–Mills models of classical and semiclassical electromagnetism had the same form: we take two $\mathfrak{g}$-valued one-forms to coordinatize the same point of configuration space if we can argue that they are empirically indistinguishable in principle. The ad hoc approach to quantization inverts this logic, presupposing that the correct configuration space is the set of gauge-equivalence classes of $\mathfrak{g}$-valued one-forms and asking how quantum Yang–Mills theory may be interpreted so as to make this true. This inversion generates a mystery when faced with the fact that neither the space of one-forms nor the space of gauge-equivalence classes can be perturbatively quantized in general; the new goal is to explain why the modifications needed for quantization aren't needed or aren't modifications. The stacky Yang–Mills interpretation doesn't have these problems. It treats classical theories uniformly and lets the quantum chips fall as they may. One-form Yang–Mills theory with the action S_{YM} is not the same theory as one-form Yang–Mills with the action S_ξ, and neither has a physically adequate quantization. Stacky Yang–Mills with the action represented by S_{YM} gives a well-defined and empirically adequate theory, and we can give positive arguments to show that gauge-equivalent one-forms represent the same element of the stacky configuration space and that gauge-equivalent actions on the stacky configuration space give the same quantum expectation values.

In Section 2 I argued that philosophical discussions of gauge theories rely on a picture of quantization with technical and conceptual problems. According to this picture, a gauge theory must be modified before it can be quantized: terms must be added to the action and determinants must be introduced and expanded in terms of odd coordinates. These manipulations can sometimes produce a good quantum theory. But justifications involving gauge fixing aren't borne out by the details, and also do not suffice in general. In this section I've offered an alternative account that avoids these problems. The same quantization procedure applies just as well to gauge theories and non-gauge theories, as long as we use the appropriate classical configuration spaces. It's also more satisfying philosophically—indeed, it even realizes the conceptual motivations of the usual philosophical story better than the ad hoc or Faddeev–Popov approaches do.

4 Ghosts can't be busted

In the previous two sections I've argued that the received philosophical view has shortcomings that the stacky view avoids. But setting aside the contrastive claim, you might think that the stacky view just takes ghosts too seriously. Following Weingard, it is often said in the philosophical literature that ghost fields are "just an artefact of our notation" and can be "transformed away" (Weingard, 1988, 57). This argument has a specific and a general form. The specific version attends to the computation of quantum expectation values. For appropriate choices of h, we can neglect ghost terms when doing computations involving $\mathcal{L}_{\mathrm{YM}} + \delta h$; Weingard concludes from this that ghosts are an artefact of a convenient choice of h. More generally, Weingard argues that ghost fields are eliminable in a way that the Yang–Mills potential is not because the former can be "gauged away". Neither of these arguments succeeds. The specific argument does show that certain features of ghost fields can be ignored in certain contexts, but it does not show that ghosts are wholly dispensable in the case Weingard considers. And the criterion Weingard offers in the general argument in fact tells in favor of the ghost field's significance, not against it. In particular, the violation of Weingard's criterion has physical consequences for the charges of particles.

4.1 Feynman diagrams

In Section 2.2 I outlined the procedure for computing perturbative integrals over a supermanifold. The computation was reduced to an expectation $\langle\!\langle - \rangle\!\rangle$, which itself reduces to a polynomial in propagators like $\langle\!\langle x^i x^j \rangle\!\rangle$. Feynman first identified the need for ghosts in the last step of this computation, where they are needed to go beyond leading order in the Yang–Mills self-coupling while retaining unitarity. But this need only arises in certain coordinates: for an appropriate choice of h, the ghost terms produced by the Lagrangian $\mathcal{L}_{\mathrm{YM}} + \delta h$ drop out of the calculation. Weingard argues that this makes ghosts dispensable, but this is too fast. Ghosts aren't just needed for unitarity, we also need them to avoid the divergent volume factor at each order and to make sense of the propagator $\langle\!\langle A^a_\mu(x)\, A^b_\nu(y) \rangle\!\rangle$.

Weingard is primarily concerned with the interpretation of Feynman diagrams, which are used in the quantization procedure of Section 2.2 when computing Gaussian expectation values.[20] Recall that a Gaussian integral is asymptotically proportional to an expectation value of the form $\langle\!\langle f(x,\theta)\, e^{p(x,\theta)/\hbar} \rangle\!\rangle$. Taylor expanding the argument and using the linearity of $\langle\!\langle - \rangle\!\rangle$ reduces this expression to a polynomial in expectation values of monomials, and each of these expectation values reduces to a product of propagators, which are matrix elements of the inverse of the quadratic part of the action. The combinatorics of this expansion are conveniently organized by a graphical calculus due to Feynman. The asymptotic expansion of a Gaussian expectation value can be expressed by a sum indexed by decorated graphs, where each edge of the graph is assigned a type of field and contributes a propagator for that field and each internal vertex of the graph is assigned a term of $p(x,\theta)$ and contributes the coefficient of that term.

As an example, consider the Lagrangian $\mathcal{L}_{\mathrm{F}}$ for some nonabelian Yang–Mills theory, and call the quanta of the Yang–Mills field "gluons". Applying Feynman's graphical calculus to the amplitude for two gluons to scatter to two gluons means summing over diagrams like those in Fig. 1. Each internal curly edge in these diagrams contributes a gluon propagator $\langle\!\langle A^a_\mu(x)\, A^b_\nu(y) \rangle\!\rangle$. The two kinds of vertices correspond to the two gluon self-interaction terms,

$$-g f^a{}_{bc} (\partial_\mu A^a_\nu) A^{b\mu} A^{c\nu} \qquad\qquad -\frac{1}{4} g^2 f^a{}_{bc} f^a{}_{de} A^b_\mu A^c_\nu A^{d\mu} A^{e\nu}$$

[20] In a 1982 paper, Weingard considers the role of Feynman diagrams in a Fock space formalism, arguing that we should not interpret ladder operators as creating and annihilating virtual particles. His 1988 paper adopts the perturbative integration framework of Section 2.2, arguing that Feynman diagram edges represent propagators, which do not represent virtual particles.

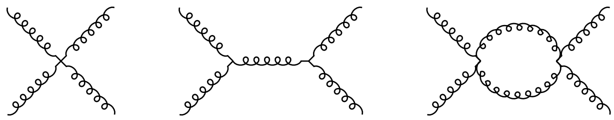

Figure 1: Gluon–gluon scattering diagrams

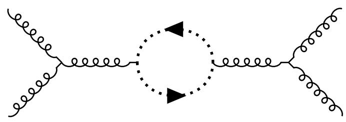

Figure 2: A ghost loop diagram

A trivalent vertex, corresponding to the term on the left, contributes an antisymmetrized product of $gf^a{}_{bc}$, the metric, and the momenta of each of the three edges. A quadrivalent vertex corresponds to the term on the right and contributes an antisymmetrized product of $-ig^2 f^a{}_{bc} f^a{}_{de}$ with two copies of the metric. Summing over all appropriate graphs and all possible decorations gives the full amplitude for gluon–gluon scattering, with each graph's contribution suppressed by the number of vertices.

Weingard's main aim is to warn us off from a too-literal interpretation of Feynman diagrams. It can be tempting to think that Feynman diagrams depict the interactions of "virtual" particles, so that the center diagram in Fig. 1 depicts two gluons annihilating into a virtual gluon which then decays into two non-virtual gluons. This temptation is bolstered by comparison with a Fock space formulation of quantum field theory, where a Feynman diagram corresponds to a sequence of operators that create and annihilate modes of the gluon field. But as Weingard argues, and as the majority of commentators agree, talk of virtual particles doesn't really hold up. Feynman diagrams are a bookkeeping device useful for keeping track of all the factors accrued in the reduction to a polynomial of propagators.[21] In particular, the amplitude assigned to each diagram can vary with our conventions. The physically significant fact is the total asymptotic behavior of the integral.

The convention-dependence of each diagram's amplitude undercuts one argument for the necessity of ghost fields. Ghosts originated as an ad hoc fix for non-unitarity, but we can avoid this problem by other means. The source of non-unitary is neglect of Feynman diagrams with nonzero amplitude. For example, at higher order in perturbation theory, the gluon–gluon scattering amplitude computed by the Lagrangian $\mathcal{L}_\mathrm{F}$ includes diagrams like the one in Fig. 2, which includes a ghost–antighost loop. Each dotted edge contributes a propagator

$$\langle\!\langle c^a(x)\,\bar{c}_b(y)\rangle\!\rangle = \int \frac{d^4k}{(2\pi)^4}\frac{i\delta^a_b}{k^2}e^{ik\cdot(x-y)}$$

while the ghost–gluon vertex

$$-gf^a{}_{bc}(\partial^\mu \bar{c}_a)c^b A^c_\mu$$

[21]For a review of the arguments for this position, see Passon (2019) and Bacelar Valente (2011), the latter dissenting.

contributes a product of $-gf^a{}_{bc}$ with the momentum of the antighost. Neglecting this diagram leads to a loss of unitarity. From the perspective of stacky Yang–Mills theory this is obvious: by leaving out diagrams that include ghosts we're leaving out some of the terms in the asymptotic expansion of our integral. But you can also demonstrate the problem directly. For without the ghost diagrams, there is a nonzero amplitude for gluon scattering to produce states with negative norm. So ghost diagrams are necessary, either from first principles or from reflection on Feynman diagrammatics.

The argument from Feynman diagrammatics doesn't go through if we can find conventions in which the amplitude of every ghostly diagram vanishes. As Weingard points out, we can. Instead of adding δh_{F} to $\mathcal{L}_{\mathrm{YM}}$ we can choose some fixed vector n^μ and modify $\mathcal{L}_{\mathrm{YM}}$ with

$$h_{\mathrm{A}} = \bar{c}_a \left(\frac{1}{2} b_a + n^\mu A^a_\mu \right)$$

giving the Lagrangian

$$\mathcal{L}_{\mathrm{A}}(A, c, \bar{c}, b) = -\frac{1}{4}(F^a_{\mu\nu})^2 - \frac{1}{2}(n^\mu A^a_\mu)^2 - \bar{c}_a n^\mu D_\mu c^a + \frac{1}{2}\big(b_a + n^\mu A^a_\mu\big)^2$$

with gluon propagator

$$\langle\!\langle A^a_\mu(x)\, A^b_\nu(y) \rangle\!\rangle = -\int \frac{d^4k}{(2\pi)^4} \frac{i\delta^{ab}}{k^2} \left[\eta_{\mu\nu} - \frac{n_\mu k_\nu + n_\nu k_\mu}{n \cdot k} + \frac{n^2 + k^2}{(n \cdot k)^2} k_\mu k_\nu \right] e^{ik\cdot(x-y)}$$

and ghost propagator

$$\langle\!\langle c^a(x)\, \bar{c}_b(y) \rangle\!\rangle = \int \frac{d^4k}{(2\pi)^4} \frac{\delta^a_b}{n \cdot k} e^{ik\cdot(x-y)}$$

Because ghosts can't appear in incoming or outgoing states and the only vertex involving a ghost corresponds to the cubic term

$$-gn^\mu f^a{}_{bc} \bar{c}_a A^b_\mu c^c$$

any ghost edge must appear as an arc in a loop consisting of ℓ ghost edges connecting ℓ trivalent vertices. The amplitude of such a loop is proportional to $\int d^4k\,(n\cdot k)^{-\ell}$, which dimensional regularization sets to zero. So with a particular choice of Lagrangian and regulation prescription any diagram that contains a ghost will have zero amplitude. And this removes the need for ghosts as a fix for non-unitarity.

I agree with Weingard that this result suggests the physical insignificance of ghost fields, at least on something like the received view. The physically significant quantities are expectation values of operators (and partition functions, and things like this), which can be expressed as integrals. The expansion in terms of Feynman diagrams is physically contentful insofar as it computes the asymptotics of these integrals. Choices about how to count diagrams, say, or a choice of renormalization prescription, are purely conventional unless they make a difference to the total sum describing the integral's asymptotic behavior. The received view uses an informal appeal to gauge invariance and unitarity to claim that $\mathcal{L}_{\mathrm{F}}$ and $\mathcal{L}_{\mathrm{A}}$ give the same integral, so any physically significant structures must appear in the Feynman expansions of both Lagrangians. Virtual particles don't exist, but edges in a Feynman diagram do track something—namely, propagators appearing in the asymptotic approximation of a perturbed Gaussian integral. The Lagrangian $\mathcal{L}_{\mathrm{F}}$ simplifies the expression for the Yang–Mills propagator by splitting some of it off into a new particle. But $\mathcal{L}_{\mathrm{A}}$ shows that ghosts needn't appear in the Feynman diagrammatics at all, so they "are just an artefact of our notation in a way that the other virtual 'processes' are not" (Weingard, 1988, 57).

However, this argument fails as an objection to the stacky interpretation of Yang–Mills theory. On the stacky interpretation, ghost fields are in the first place coordinates on the classical configuration space, and the appearance of ghost vertices in some nonzero diagrams is only one of their downstream effects. We don't need to survey different Feynman expansions to determine whether ghost fields should be introduced into our integrals, because there are ghost dimensions in the domain of integration from the start. Even when the action doesn't explicitly depend on the ghost coordinates it's a function on the stacky configuration space; this is what explains the $0/0$ indeterminacy for such actions. So a strategy like Weingard's, which takes the value of the integral as primary, can't show that ghost fields are mere notation in the stacky interpretation. And even if we focus on Feynman diagrams, we can detect ghosts in the expansion of $\mathcal{L}_{\mathrm{A}}$. Ghost fields are important because they allow us to choose a nondegenerate representative of the Yang–Mills action. If it weren't for the ghost fields, the gluon propagator would be ill-defined. So ghost fields reveal themselves even in the Feynman diagrams produced by $\mathcal{L}_{\mathrm{A}}$, because these can include gluon edges without diverging.

The fact that ghostly diagrams vanish for certain choices of Lagrangian shows that we can ignore them yet still live a unitary life of calculating Feynman diagrams. On the received view, Weingard is therefore right to say that ghosts are "purely a result of our notation" (1988, 58), for they are only introduced to patch a failure of unitarity. But this isn't why they're introduced on the stacky interpretation, and showing that ghostly diagrams vanish doesn't show that ghosts make no difference to the integrals that compute expectation values. This is because ghosts do make a difference to these integrals, and even to Feynman diagrams, on the stacky interpretation.

4.2 Weingard's general criterion

Though Weingard's argument from Feynman diagrams doesn't pose a problem for stacky Yang–Mills theory, the general principle behind his argument remains plausible. If it were the case that ghost fields could always be "gauged away", as Weingard puts it, then we would have some reason to think that the set of gauge-equivalence classes is an adequate configuration space for Yang–Mills theory. But it's not the case. Given Weingard's framing, the question is essentially mathematical: is there a nontrivial gauge-invariant function with nonzero ghost number? In the case of Yang–Mills theory—and many other gauge theories—the answer is "yes". Indeed, one such function is used to constrain the charges of the matter appearing in the Standard Model. Weingard's criteria in fact speak against his conclusion that ghosts are mere notation.

Generalizing his argument from Feynman diagrams, Weingard articulates a criterion for identifying pure notation using an analogy with the Yang–Mills model of the Aharonov–Bohm effect. As recounted in Section 2.1, the Aharonov–Bohm effect shows that two potentials with the same field strength can induce distinct dynamics even if they have the same field strength tensor, and this shows that we cannot take the set of field strength tensors to be the configuration space of Yang–Mills theory. Generalizing this case, Weingard extracts a test for determining that some mathematical structure is pure notation

> while an arbitrary $[A_\mu]$ can always be transformed to zero at any given point, it cannot, as we have seen, be transformed to zero along an arbitrary closed curve (or finite area). Thus, some of the degrees of freedom of $[A_\mu]$ depend on our notation—on our choice of gauge, but not all do. (1988, 58).

If we could always choose some conventions that made the potential vanish then we could write it off as a computational convenience, but we can't always choose such conventions. Weingard goes on to claim that this criterion shows ghost fields to be purely notation in a way that the potential isn't, because ghostly Feynman diagrams can always be gauged away.

Mathematically speaking, Weingard's criterion takes the same form for both the Yang–Mills potential and the ghost fields; in both cases the criterion formally reduces to classifying antiderivatives. Weingard expresses the violation for the Yang–Mills case as follows:

> we cannot make the connection $[A_\mu]$ zero throughout a non-simply-connected region R, even if the curvature $[F_{\mu\nu}]$ is zero throughout R. (1988, fn. 3)

The exterior derivative d_μ over the region R sends any p-form over R to its antisymmetrized derivative, a $(p+1)$-form. In p-form electromagnetism, two p-forms are gauge equivalent if their difference is the exterior derivative of a $(p-1)$-form. In particular, the exterior derivative sends any one-form A_μ to its field strength $F_{\mu\nu} = d_\mu A_\nu$, and a gauge transformation between one-forms A_μ and A'_μ is a smooth function c such that $A'_\mu - A_\mu = d_\mu c$. The gauge-invariance of the electromagnetic field strength then follows from the fact that $d_\mu d_\nu$ vanishes:

$$F'_{\mu\nu} - F_{\mu\nu} = d_\mu(A'_\nu - A_\nu) = d_\mu d_\nu c = 0$$

According to Weingard's criterion, the potential A_μ is more than mere notation for $F_{\mu\nu}$ because there are gauge-inequivalent potentials with the same field strength. By linearity, this amounts to the existence of a potential A_μ such that $d_\mu A_\nu = 0$ but $A_\mu \neq d_\mu c$ for all smooth functions c, which is just a rephrasing of Weingard's geometric gloss in coordinates. There are some R for which such one-forms exist, as in the Aharonov–Bohm effect, so the potential is more than pure notation.

The vector field δ in stacky Yang–Mills is formally analogous to the exterior derivative, and the analogue of Weingard's criterion in this context classifies the ghost coordinates as more than mere notation. The vector field δ sends any function on the stacky configuration space with ghost number p to a function with ghost number $p+1$. Two functions with ghost number p are gauge equivalent if their difference is δh for some function h with ghost number $p-1$. Applying Weingard's criterion, ghost fields are more than mere notation if there is some function f with nonzero ghost number such that $\delta f = 0$ but $f \neq \delta h$ for all h.

And in general there are such functions. For example, on the configuration space of stacky $\mathfrak{su}(n)$ Yang–Mills theory there is the function

$$\mathcal{A} = -\frac{g^3}{24\pi^2} d_{abc} \int d^4x \, \epsilon^{\mu\nu\alpha\beta} c^a \partial_\mu \Big(A^b_\nu \, \partial_\alpha A^c_\beta + \frac{g}{4} f^c{}_{de} A^b_\nu A^d_\alpha A^e_\beta \Big)$$

where d_{abc} is the symmetrized trace of hermitian basis elements in the defining representation of $\mathfrak{su}(n)$. A computation shows that $\delta\mathcal{A} = 0$, meaning that $\mathcal{A}$ is gauge invariant, and because it's linear in c^a it's a function of ghost number 1. We also have $\mathcal{A} \neq \delta h$ for all functions h on the configuration space.[22] So by Weingard's criterion, the ghost field is more than pure notation. We can gauge away some quantities on the stacky configuration space, like the function assigning an amplitude to the diagram in Fig. 2. But we can't gauge away all of them.

The ghost field and the Yang–Mills potential are mathematically analogous with respect to Weingard's criterion, but the criterion also has a physical aspect. The mathematical fact that connections aren't uniquely determined by their curvature over topologically nontrivial regions is only physically relevant insofar as we use connections to model the electromagnetic configuration and use topologically nontrivial regions to model aspects of some electromagnetic experimental setups. For Weingard's criterion to ratify the ghost field as more than mere notation, we need a nonzero gauge-invariant function with nonzero ghost number that also has some physical relevance.

The nonvanishing of $\mathcal{A}$ constrains the charges of matter coupled to Yang–Mills fields, so it also satisfies the physicality condition of Weingard's criterion. Consider a family ψ of left-handed fermions charged under an $\mathfrak{su}(n)$ Yang–Mills field. The Lagrangian for this theory is

$$\mathcal{L}_{\chi\mathrm{YM}}(A, \psi, \overline{\psi}) = -\frac{1}{4}(F^a_{\mu\nu})^2 - \overline{\psi} \not{D} \psi \qquad\qquad D_\mu \psi = \partial_\mu \psi - iQgA^a_\mu \tau_a \psi$$

where Q is the charge of the fermions and τ_a the hermitian basis elements of the defining representation of $\mathfrak{su}(n)$. Variations on this Lagrangian are found many times in the Standard Model, where they describe families of quarks or leptons charged under the electroweak and strong forces. As with any integral over a supermanifold, the path integral for this theory is evaluated by first integrating over the odd coordinates, giving a theory of the Yang–Mills field with action

$$S_{\mathrm{eff}}(A) = -\frac{1}{4} \int d^4x \, (F^a_{\mu\nu})^2 + W(A) \qquad\qquad e^{\frac{i}{\hbar} W(A)} = \int \mathcal{D}\psi \, \mathcal{D}\overline{\psi} \, e^{-\frac{i}{\hbar} \overline{\psi} \not{D} \psi}$$

where $W(A)$ is defined by the equation on the right. Physically speaking, the function $W(A)$ encodes the integral over the fermions into an effective source for the Yang–Mills field, and the total path integral reduces to the path integral for a gauge theory. For this theory to be well-defined, the function S_{eff} must descend to a function on the stacky Yang–Mills configuration space—that is, it must be gauge invariant. The pure Yang–Mills term is gauge invariant, but we have

$$W(A^a_\mu + D_\mu c^a) = W(A^a_\mu) + Q^3 \mathcal{A}$$

[22]More precisely, the integrand defining $\mathcal{A}$ represents a nontrivial first cohomology class of δ modulo the exterior derivative; see Weinberg (1995, §22.6) or Bertlmann (1996, §8.3) for textbook treatments.

So this action only gives a well-defined quantum gauge theory if Q vanishes and the fermions decouple from the Yang–Mills field.[23]

Physicists often take it to be a prediction of the theory that the fermion charges cooperate to make the effective action gauge invariant. Replacing the left-handed fermions in the previous paragraph with right-handed fermions gives an effective action such that

$$W(A^a_\mu + D_\mu c^a) = W(A^a_\mu) - Q^3\mathcal{A}$$

and if more than one family of fermions couples to the Yang–Mills field then the gauge variation of the effective action is the sum of the contributions of each family. For example, the effective action for the hypercharge sector of the Standard Model satisfies

$$W(A^a_\mu + D_\mu c^a) = W(A^a_\mu) + \big(6Y_Q^3 + 2Y_L^3 - Y_e^3 - Y_\nu^3 - 3Y_u^3 - 3Y_d^3\big)\mathcal{A}$$

where the Ys are the hypercharges of left-handed quarks and leptons and the right-handed electrons, neutrinos, and up- and down-type quarks. The observed values of these charges are consistent with the gauge invariance of the effective action. And if the demand for gauge invariance is justified then the vanishing of this linear combination of cubed charges is in fact a prediction of the Standard Model. Similar reasoning implies

$$2Y_Q - Y_u - Y_d = 0 \qquad\qquad Y_L + 3Y_Q = 0$$

the latter implying the relation between the electron and down quark mentioned in the introduction (Schwartz, 2014, 633). These predictions rely on the fact that $\mathcal{A}$ doesn't vanish, so by Weingard's criterion applied to $\mathcal{A}$, ghosts are more than mere notation.

The justification for this prediction deserves further attention from philosophers, but its status isn't important for my argument in this paper. The effective action isn't a well-defined function on the stacky Yang–Mills configuration space, but it could perhaps be well-defined on another configuration space. The usual claim is that a theory with a gauge-variant effective action isn't "coherent" (Dawid, 2013, 12) or "consistent" (Schwartz, 2014, 627). This is either because gauge variance is a threat to unitarity or because it will "destroy the renormalizability, and thus the consistency, of the gauge theory" (Bertlmann, 1996, 245). This kind of appeal to consistency is common in high-energy theory, but the meaning of "consistent" in this sense isn't obvious. Non-renormalizable theories certainly exist and can be useful, like the chiral perturbation theory describing low-energy hadron physics. And Preskill (1991) argues that we can obtain such a non-renormalizable quantum gauge theory for nonvanishing $Q^3\mathcal{A}$ as well. But even if the demand for a gauge-invariant effective action can't be justified and any charges are possible, the nonvanishing of $\mathcal{A}$ has less dramatic consequences as well. It appears in the anomalous Ward–Takahashi identity, for example. So regardless of how we sort out these issues, the function $\mathcal{A}$ shows that ghost fields are more than mere notation.

Weingard's general argument against the physical significance of ghost fields appeals to a plausible principle: if there is some obstruction to eliminating

[23] As noted in Section 3.2, BRST invariance of the action is also required for the Kugo–Ojima quartet mechanism to ensure unitarity.

a mathematical structure by an appropriate choice of conventions, then that structure is more than purely notation. The formal incarnation of this principle in the case of ghost fields is the same as in Weingard's motivating example of the Aharonov–Bohm effect. And, as in that case, there is an obstruction to transforming away dependence on ghost fields. This obstruction has good physical credentials: it is used to constrain the charges of quarks and leptons through its modification of the Ward–Takahashi identities. So Weingard's criterion, properly applied, does not say that ghosts are purely notation.

5 Conclusion

The goal of this paper was to tell a story about ghost fields that's better than the story currently found in the philosophical literature. On the story I'm offering, quantum gauge theories are built like most other quantum field theories: by asymptotically approximating a perturbed Gaussian integral over the classical configuration space. Different interpretations of the classical theory—that is, different choices of configuration spaces and actions thereon—will lead to different quantum theories. In Section 3 I described one interpretation of Yang–Mills theory and its quantization. One notable feature of this interpretation is that it has gauge structure: odd dimensions related to the even dimensions so as to mathematically implement the informal idea that configurations related by a gauge transformation are "the same". Ghost fields are coordinates on this gauge structure. Another notable feature is the empirical success of its quantization as one ingredient in the Standard Model.

The account of ghosts in Section 3 both avoids and explains problems with the received view. Classical theories aren't just the raw materials for quantum theories; they can themselves model some phenomena. But it's generally agreed that classical and semiclassical models ought to be interpreted so that they give the correct quantum theory upon quantization. This is why the Aharonov–Bohm effect is taken to be evidence against the two-form interpretation of Maxwell electromagnetism. As I argued in Section 2, the received view inverts this reasoning, insisting that quantum Yang–Mills theory be interpreted such that it's obtained by quantizing the Yang–Mills action on the space of $\mathfrak{g}$-valued one-forms. And this leads to trouble, because that classical theory has a degenerate action and thus no perturbative quantization. The received view tries to get around this problem with a sequence of ad hoc modifications to the quantization process and takes ghosts to be one feature of these modifications. But these modifications aren't adequate in general, nor do they comport with the conceptual gloss that's usually offered as justification.

The stacky interpretation of Yang–Mills theory also accounts for the data that the received view is meant to capture—indeed, the stacky interpretation fits this data better. Both views agree that the divergence of the naive perturbative quantization of Yang–Mills theory can be avoided by replacing the Lagrangian with one of many equivalent substitutes, such as $\mathcal{L}_{\mathrm{F}}$ or $\mathcal{L}_{\mathrm{A}}$. On the received view, the equivalence of these replacements is a conjecture supported by an informal notion of gauge equivalence and case-by-case checks, while on the stacky view the equivalence of these substitutes follows from their construction over the stacky configuration space.[24] Weingard has used this equivalence between $\mathcal{L}_{\mathrm{F}}$

[24]It is sometimes even claimed that these case-by-case checks fail. For example, in the recent

and $\mathcal{L}_{\text{A}}$ to argue that ghosts are an artefact of our notation: on the received view they are introduced to patch a failure of unitarity, and $\mathcal{L}_{\text{A}}$ doesn't fail in this way, so ghosts are not a feature of the convention-independent integral they may be used to calculate. But on the stacky interpretation, ghosts aren't a fix for non-unitarity; they coordinatize dimensions of the configuration space that are relevant no matter which gauge-equivalent action we choose. More generally, the existence of nontrivial gauge-invariant functions of the ghost coordinates makes a difference to the predictions of the quantum theory, so the general criterion Weingard proposes does not show ghosts to be purely notational.

Finally, what are the consequences of this better understanding of quantization for the interpretation of classical gauge theories? Thankfully, nothing too radical. Indeed, the account of quantization in Section 3 enables a uniform continuation of reasoning from the semiclassical to the quantum case. If the Aharonov–Bohm effect bears on the interpretation of classical Yang–Mills theory, then so do ghosts, and for the same reason. The one-form interpretation has an inadequate quantization and so should be rejected, just as the two-form interpretation should be rejected in light of the quantum dynamics it produces. We can also conclude as usual that gauge-equivalent one-forms coordinatize the same point of configuration space, but now the logic points in the right direction: we don't appeal to the semiclassical description to make ad hoc modifications to the quantization process, but instead argue from the empirically successful quantization of the stacky configuration space to a conclusion about the classical theory.

However, the results of quantization do show that we must be more careful about what it means to say that gauge-equivalent one-forms are "the same". Quantization shows that we should reject the one-form interpretation in favor of the stacky interpretation, in which gauge-equivalent one-forms represent the same point of configuration space. But a parallel argument shows that we should also reject an interpretation in terms of gauge-equivalence classes, since this also has no perturbative quantization. Moreover, a configuration space without gauge structure—and thus without ghost coordinates—will *ipso facto* have no functions of nonzero ghost number. In particular, a Yang–Mills configuration space without gauge structure will lack the function $\mathcal{A}$ of Section 4.2 and the empirical constraints it provides. So gauge structure isn't surplus, eliminable, excess, redundant, superfluous, unnecessary, or whatever. At the level of mathematics there's no ambiguity: the stacky configuration space and the collection of gauge-equivalence classes have the same set of isomorphism classes of objects, but disagree about the arrows between them. There's more to a category than its isomorphism classes of objects, so two categories can agree that two objects are the same without agreeing on everything. Some recent philosophical work has been attuned to this further structure (Dougherty, 2017; Nguyen et al., 2020; Weatherall, 2016), but for the most part debates over the interpretation of gauge theories have restricted themselves to classical theories whose configuration spaces lack it. In this paper I've argued that this restriction

controversy over diphotonic Higgs decay, Wu and Wu (2017) argue that the R_ξ and unitary gauges give different amplitudes for a Higgs to decay to two photons via a W boson loop. As many have pointed out, this discrepancy is an artefact of subtleties surrounding regularization of the loop (Duch et al., 2021). But this example shows that dissatisfaction with the received view's informal conjecture is more than a fetish for rigor: the received view makes gauge non-invariance an open question with practical consequences.

is harmful.

References

Aharonov, Y. and Bohm, D. (1959). Significance of electromagnetic potentials in the quantum theory. *Physical Review*, 115(3):485.

Bacelar Valente, M. (2011). Are virtual quanta nothing but formal tools? *International Studies in the Philosophy of Science*, 25(1):39–53.

Batterman, R. W. (2002). *The Devil in the Details: Asymptotic Reasoning in Explanation, Reduction, and Emergence*. Oxford University Press.

Becchi, C., Rouet, A., and Stora, R. (1976). Renormalization of gauge theories. *Annals of Physics*, 98(2):287–321.

Belot, G. (1995). New work for counterpart theorists: Determinism. *The British Journal for the Philosophy of Science*, 46(2):185–195.

Belot, G. (1998). Understanding electromagnetism. *The British Journal for the Philosophy of Science*, 49(4):531–555.

Belot, G. (2001). The principle of sufficient reason. *The Journal of Philosophy*, 98(2):55–74.

Belot, G. (2003). Symmetry and gauge freedom. *Studies in History and Philosophy of Modern Physics*, 34(2):189–225.

Belot, G. (2008). An elementary notion of gauge equivalence. *General Relativity and Gravitation*, 40(1):199–215.

Benini, M., Schenkel, A., and Schreiber, U. (2018). The stack of Yang–Mills fields on Lorentzian manifolds. *Communications in Mathematical Physics*, 359(2):765–820.

Bertlmann, R. A. (1996). *Anomalies in Quantum Field Theory*. Clarendon Press.

Brading, K. and Brown, H. R. (2004). Are gauge symmetry transformations observable? *The British Journal for the Philosophy of Science*, 55(4):645–665.

Brading, K. and Castellani, E., editors (2003). *Symmetries in Physics*. Cambridge University Press.

Brighouse, C. (1997). Determinism and modality. *The British Journal for the Philosophy of Science*, 48(4):465–481.

Butterfield, J. (1989). The hole truth. *The British Journal for the Philosophy of Science*, 40(1):1–28.

Butterfield, J. (2006). On symplectic reduction in classical mechanics. In Butterfield, J. and Earman, J., editors, *The Handbook of Philosophy of Physics*, pages 1–131. North Holland.

Butterfield, J. and Bouatta, N. (2015). Renormalization for philosophers. In Bigaj, T. and Wüthrich, C., editors, *Metaphysics in Contemporary Physics*, pages 437–485. Brill.

Calaque, D. (2015). A derived and homotopical view on field theories. In Calaque, D. and Strobl, T., editors, *Mathematical Aspects of Quantum Field Theories*, pages 1–14. Springer.

Caulton, A. (2015). The role of symmetry in the interpretation of physical theories. *Studies in History and Philosophy of Modern Physics*, 52:153–162.

Crainic, M. and Fernandes, R. L. (2011). Lectures on integrability of Lie brackets. *Geometry & Topology Monographs*, 17:1–107.

Dasgupta, S. (2011). The bare necessities. *Philosophical Perspectives*, 25:115–160.

Dasgupta, S. (2018). Symmetry and superfluous structure. To appear in Knox, E. and Wilson, A., editors, *The Routledge Companion to Philosophy of Physics*.

Dawid, R. (2013). *String Theory and the Scientific Method*. Cambridge University Press.

Dewar, N. (2019). Sophistication about symmetries. *The British Journal for the Philosophy of Science*, 70(2):485–521.

DeWitt, B. S. (1967). Quantum theory of gravity. II. The manifestly covariant theory. *Physical Review*, 162:1195–1239.

DeWitt, B. S. (2005). The space of gauge fields: Its structure and geometry. In 't Hooft (2005), pages 15–32.

Dougherty, J. (2017). Sameness and separability in gauge theories. *Philosophy of Science*, 84(5):1189–1201.

Duch, P., Dütsch, M., and Gracia-Bondía, J. M. (2021). Diphoton decay of the higgs from the Epstein–Glaser viewpoint. *The European Physical Journal C*, 81(2):131.

Earman, J. (2003). Tracking down gauge: an ode to the constrained Hamiltonian formalism. In Brading and Castellani (2003), pages 140–162.

Faddeev, L. D. and Popov, V. N. (1967). Feynman diagrams for the Yang–Mills field. *Physics Letters B*, 25:29–30.

Feynman, R. P. (1963). Quantum theory of gravitation. *Acta Physica Polonica*, 24:697–722.

Fraser, D. (2009). Quantum field theory: Underdetermination, inconsistency, and idealization. *Philosophy of Science*, 76:536–567.

Fraser, D. (2011). How to take particle physics seriously: A further defence of axiomatic quantum field theory. *Studies in History and Philosophy of Modern Physics*, 42(2):126–135.

Fraser, J. D. (2020). The real problem with perturbative quantum field theory. *The British Journal for the Philosophy of Science*, 71(2):391–413.

Gilton, M. J. R. (2020). Could charge and mass be universals? *Philosophy and Phenomenological Research*, 00:1–21.

Greaves, H. and Wallace, D. (2014). Empirical consequences of symmetries. *The British Journal for the Philosophy of Science*, 65(1):59–89.

Guay, A. (2008). A partial elucidation of the gauge principle. *Studies in History and Philosophy of Modern Physics*, 39(2):346–363.

Halvorson, H. (2006). Algebraic quantum field theory. In Butterfield, J. and Earman, J., editors, *Handbook of Philosophy of Science: Philosophy of Physics*, pages 731–922. Elsevier.

Hancox-Li, L. (2015). Coarse-graining as a route to microscopic physics: the renormalization group in quantum field theory. *Philosophy of Science*, 82(5):1211–1223.

Healey, R. (2007). *Gauging What's Real*. Oxford University Press.

Henneaux, M. and Teitelboim, C. (1992). *Quantization of Gauge Systems*. Princeton University Press.

Johnson, G. W. and Lapidus, M. L. (2000). *The Feynman Integral and Feynman's Operational Calculus*. Clarendon Press.

Johnson-Freyd, T. (2015). Homological perturbation theory for nonperturbative integrals. *Letters in Mathematical Physics*, 105(11):1605–1632.

Kugo, T. and Ojima, I. (1978). Manifestly covariant canonical formulation of the Yang–Mills field theories. I: —general formalism—. *Progress of Theoretical Physics*, 60(6):1869–1889.

Lyre, H. (2004). Holism and structuralism in $U(1)$ gauge theory. *Studies in History and Philosophy of Modern Physics*, 35:643–670.

Maudlin, T. (2007). Suggestions from physics for deep metaphysics. In *The Metaphysics Within Physics*, pages 78–103. Oxford University Press.

Melia, J. (1999). Holes, haecceitism and two conceptions of determinism. *The British Journal for the Philosophy of Science*, 50(4):639–664.

Miller, M. (2021). Mathematical structure and empirical content. Forthcoming in *The British Journal for the Philosophy of Science*.

Mnev, P. (2019). *Quantum Field Theory: Batalin–Vilkovisky Formalism and Its Applications*. American Mathematical Society.

Nguyen, J., Teh, N. J., and Wells, L. (2020). Why surplus structure is not superfluous. *The British Journal for the Philosophy of Science*, 71(2):665–695.

Passon, O. (2019). On the interpretation of Feynman diagrams, or, did the LHC experiments observe $H \to \gamma\gamma$? *European Journal for Philosophy of Science*, 9(2):20.

Peskin, M. E. and Schroeder, D. (1995). *An Introduction to Quantum Field Theory*. Addison-Wesley.

Poincaré, H. (1892). *Les méthodes nouvelles de la mécanique céleste*, volume 1. Gauthier-Villars.

Popov, V. N. and Faddeev, L. D. (1967). Perturbation theory for gauge-invariant fields. Technical Report ITP-67-36, Institute for Theoretical Physics, Kiev, UkrSSR. Translated into English by Benjamin W. Lee and David Gordon as Technical Report NAL-THY-57, National Accelerator Laboratory, USA (1972). English translation reprinted in 't Hooft (2005), pages 40–60.

Preskill, J. (1991). Gauge anomalies in an effective field theory. *Annals of Physics*, 210(2):323–379.

Redhead, M. (2003). The interpretation of gauge symmetry. In Brading and Castellani (2003), pages 124–139.

Rickles, D. (2008). *Symmetry, Structure, and Spacetime.* Elsevier.

Rosenstock, S. and Weatherall, J. O. (2016). A categorical equivalence between generalized holonomy maps on a connected manifold and principal connections on bundles over that manifold. *Journal of Mathematical Physics*, 57:102902.

Schreiber, U. (2017). Mathematical quantum field theory. `https://ncatlab.org/nlab/show/geometry+of+physics+--+perturbative+quantum+field+theory`.

Schwartz, M. D. (2014). *Quantum Field Theory and the Standard Model.* Cambridge University Press.

't Hooft, G. (1971). Renormalization of massless Yang–Mills fields. *Nuclear Physics B*, 33(1):173–199.

't Hooft, G., editor (2005). *50 Years of Yang–Mills Theory.* World Scientific.

Tyutin, I. V. (1975). Gauge invariance in field theory and statistical physics in operator formalism. Technical Report LEBEDEV-75-39, Lebedev Physical Institute.

Vaĭntrob, A. Yu. (1997). Lie algebroids and homological vector fields. *Russian Mathematical Surveys*, 52(2):428–429.

Wallace, D. (2011). Taking particle physics seriously: A critique of the algebraic approach to quantum field theory. *Studies in History and Philosophy of Modern Physics*, 42(2):116–125.

Weatherall, J. O. (2016). Understanding gauge. *Philosophy of Science*, 83(5):1039–1049.

Weinberg, S. (1995). *The Quantum Theory of Fields.* Cambridge University Press.

Weingard, R. (1982). Do virtual particles exist? In *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, volume 1982, pages 235–242. Philosophy of Science Association.

Weingard, R. (1988). Virtual particles and the interpretation of quantum field theory. In Brown, H. R. and Harré, R., editors, *Philosophical Foundations of Quantum Field Theory*, chapter 3, pages 43–58. Clarendon Press.

Wu, T. T. and Wu, S. L. (2017). Comparing the R_ξ gauge and the unitary gauge for the standard model: An example. *Nuclear Physics B*, 914:421–445.

Zinn-Justin, J. (1975). Renormalization of gauge theories. In Rollnik, H. and Dietz, K., editors, *Trends in Elementary Particle Theory*, pages 1–39. Springer.