

# Abandoning Galileo’s Ship: The quest for non-relational empirical significance

Forthcoming in *The British Journal for the Philosophy of Science*

Sebastián Murgueitio Ramírez

Nicholas J. Teh

## Abstract

The recent debate about whether gauge symmetries can be empirically significant has focused on the possibility of ‘Galileo’s ship’ types of scenarios, where the symmetries effect relational differences between a subsystem and the environment. However, it has gone largely unremarked that apart from such Galileo’s ship scenarios, Greaves and Wallace (2014) proposed that gauge transformations can also be empirically significant in a ‘non-relational’ manner that is analogous to a Faraday-cage scenario, where the subsystem symmetry is related to a change in a charged boundary state. In this paper, we investigate the question of whether such non-relational scenarios are possible for gauge theories. Remarkably, the answer to this question turns out to be closely related to a foundational puzzle that has driven a host of recent developments at the frontiers of theoretical physics. By drawing on these recent developments, we show that a very natural way of elaborating on Greaves and Wallace’s claim of non-relational empirical significance for gauge symmetry is incoherent. However, we also argue that much of what they suggest is correct in spirit: one can indeed construct non-relational models of the kind they sketch, albeit ones where the empirical significance is not witnessed by a gauge symmetry but instead by a superficially similar boundary symmetry. Furthermore, the latter casts doubt on whether one really abandons Galileo’s ship in such scenarios.

# Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Type II empirical significance</b>	<b>7</b>
2.1	Preliminaries . . . . .	7
2.2	An interlude: The Faraday cage . . . . .	12
2.3	The $U(1)$ Type II scenario . . . . .	15
<b>3</b>	<b>Constructing the Core Type II scenario</b>	<b>19</b>
3.1	A reminder: the covariant phase space framework . . . . .	20
3.2	A link to the frontiers of theoretical physics . . . . .	22
3.3	A dynamical completion of DF’s account . . . . .	26
3.4	Core Type II embodied . . . . .	28
<b>4</b>	<b>Whence Type II empirical significance?</b>	<b>30</b>
4.1	Assessing (Empirical Symmetry) . . . . .	30
4.2	Non-relationality and Galileo’s ship again . . . . .	32
<b>5</b>	<b>Conclusion</b>	<b>33</b>

## 1 Introduction

The recent philosophy of symmetry literature has seen huge advances in the discussion and interpretation of the notion of ‘gauge symmetry’, which—roughly speaking—refers to the *spacetime-dependent* transformations which act on the states (gauge fields) of many of our best physical theories, and one of whose various functions is to encode the ‘representational

equivalence’ of a certain set of symmetry-related states.<sup>1</sup>

One strand of the literature has investigated the question of whether gauge symmetry can be said to have ‘empirical significance’ in the sense of Galileo’s ship, viz. in the sense that certain subsystem symmetries can be used to effect a relational difference between a subsystem (e.g. the ship system) and an environment (e.g. the shore). A schema for an affirmative answer to this question was first put forward by Greaves and Wallace [2014] (‘GW’ henceforth), which was followed by two further developments.<sup>2</sup> First, Teh [2015] demonstrated that one way of understanding GW’s schema is in terms of ‘asymptotic’ formulations of such scenarios. In such models, (i) physicists impose asymptotic ‘isolated system’ boundary conditions on the subsystem, thus implying that the environment state can be thought of as a fixed frame at the asymptotic boundary; and (ii) one can then argue that the ‘redundant’ symmetries of the subsystem states obtained by solving for these boundary conditions are the asymptotically trivial gauge symmetries, which cannot be used to effect relational differences with an environment. On the other hand, the residual ‘large’ (asymptotically non-trivial) gauge transformations can effect such differences.<sup>3</sup> Second, Gomes [2019a] has recently explored a non-asymptotic interpretation of GW’s schema, in which the subsystems are defined over finite regions and the environment subsystem plays a genuinely dynamical role.

---

<sup>1</sup>We note that in contrast with our use of the term ‘gauge symmetry’, some authors take this term to simply mean ‘a symmetry that relates physically indistinguishable states’. Our use is the dominant one in the ‘Galileo’s ship in gauge theory’ literature, as well as large parts of the mainstream theoretical physics literature. We further clarify our use of the term in Section 2.1.

<sup>2</sup>Prior to Teh’s [2015], the rich literature investigating whether or not gauge transformations could be empirically significant did not clearly distinguish between the asymptotic and the non-asymptotic cases (i.e., Kosso [2000], Brading and Brown [2004], Healey [2009]). We think that a failure to distinguish between asymptotic and non-asymptotic transformations might explain why these authors did not think that gauge transformations could be empirically significant (after all, as Teh [2015] and Gomes [2019a] show, the asymptotic character of these transformations is crucial to their empirical significance).

<sup>3</sup>We note that in this nomenclature, we take ‘non-contractible symmetries’ (symmetries that cannot be smoothly deformed to the identity map) to be non-trivial.

In this paper, we will confine ourselves to the asymptotic models when we are discussing relational differences between a subsystem and an environment. We will say that these kinds of scenarios have ‘Type I empirical significance’ and we will use ‘Type I case/scenario’ to refer to a case of a symmetry exhibiting Type I empirical significance. Although initially heterodox with respect to the philosophy literature, these Type I scenarios for gauge symmetry have long been accepted within physics as an orthodox explanation for how asymptotically-defined gauge field subsystems can possess a meaningful empirical ‘charge’.

On the other hand, a second strand of the literature has investigated the question of why the full structure of gauge symmetry is necessary in order to express relationships between subsystems on overlapping regions of spacetime (for example, see Nguyen et al. [2018], Mathieu et al. [2019] and many others).<sup>4</sup> Roughly speaking, the moral of this body of work is that the rich structure of gauge symmetry represents the myriad ways in which such subsystems can be composed or coupled; for instance, gauge symmetry ‘witnesses’ the different ways in which the field content of a gauge subsystem can ‘mesh’ with some set of boundary conditions. Since such boundary conditions are central to the understanding of ‘empirical subsystem symmetry’ in Greaves and Wallace [2014] and in Teh [2015], any attempt to grapple with their claims must also take this second strand of the literature into consideration.

Now, in the heat of the debate over whether gauge symmetry can exhibit Type I empirical significance, it seems to have gone largely unremarked that Type I is only one of two kinds of empirical significance that GW attribute to subsystem gauge symmetries. The other kind—which we will call ‘Type II’—is supposed to have the remarkable feature that (in GW’s own words) it is ‘non-fully relational’, meaning that—unlike Galileo’s Ship (Type I)—its empirical significance cannot be cashed out in terms of relational differences be-

---

<sup>4</sup>Gomes [2019b] also discusses this topic, albeit in terms of a covariant gauge-fixing.

tween a subsystem and an environment. Furthermore, GW claim that these Type II cases are tightly linked to the concept of ‘boundary charges’ at the spatial boundary of a gauge subsystem, because the ‘empirically significant’ gauge symmetries are related to physically distinct charged environment states. This aspect of their claim is especially tantalizing because the topic of ‘boundary charges and boundary degrees of freedom in gauge theory’ is one of the most enduring sources of conceptual puzzles and theoretical innovations in field theory; indeed, in just the last three years, the work of Donnelly and Freidel [2016] on this topic has driven a whole host of developments at the frontiers of theoretical physics.

In brief, much is at stake in the quest to understand GW’s Type II empirical significance for gauge symmetry: the topic brings together two rich themes in the philosophy of symmetry literature (i.e., the two strands mentioned before) and potentially provides a living connection between the philosophy of physics and cutting-edge developments in physics.

One possible reason that philosophers have been slow to discuss Type II empirical significance is the highly schematic manner in which GW define a Type II scenario, and the lack of a fully fleshed-out example of what such a scenario would look like.<sup>5</sup> However, we will argue that if one takes GW’s description of a Type II scenario seriously, one already has sufficient materials to appreciate how radical Type II scenarios are—so radical, in fact, that they are likely to induce *aporia*, as well as attempts at novel theorizing in response to this *aporia*.

In this paper, we set out to perform two related tasks. First, we will show that the deep source of this *aporia* is a puzzle in the foundations of gauge theory that has recently been highlighted and addressed in the work of Donnelly and Freidel [2016], Mathieu et al. [2019], Gomes [2019a,b], and others. In the course of doing so, we not only develop a

---

<sup>5</sup>One of GW’s schematic examples of such a scenario is (what they take to be) the ‘ $U(1)$  gauge theory version of the Faraday cage’. But as we will see in Section 2.2, the important details of this example have largely been black-boxed.

detailed analysis of what a ‘filled in’ Type II scenario might look like, but also provide a philosophical interpretation of what Donnelly and Freidel [2016] have accomplished in their seminal work on the foundations of gauge symmetry. Second, having appreciated its source, we show how the *aporia* can be dissolved: the key is to realize that a ‘subsystem gauge symmetry’ is not doing the empirical work in a Type II scenario, although there is a formally similar symmetry that is doing this work; furthermore, although there is a (spatio-temporal) sense in which the latter empirical significance is ‘non-relational’—thus making partial sense of GW’s claim—there is also a different sense in which it should be understood in the spirit of Galileo’s ship, i.e., relationally.

We proceed as follows. Section 2 reviews GW’s notion of Type II empirical significance for subsystem gauge symmetries. Section 2.1 recalls the abstract schema for Type II scenarios and decomposes such scenarios into two parts, viz., a cluster of features that we call ‘Core Type II features’, and a claim that we call ‘(Empirical Symmetry)’. Even at this abstract level, one already encounters the key radical feature of such scenarios: there are *no* ‘field boundary conditions’ placed on the subsystem fields. Since GW say that one of their main examples of a Type II scenario is the  $U(1)$  gauge theory analog of the Faraday cage, Section 2.2 briefly reminds the reader of the details of the Faraday cage. Section 2.3 then interrogates GW’s  $U(1)$  Type II example and shows that it is still too schematic relative to the kind of fine-grained analysis that we wish to engage in. We explain what would need to be accounted for in order to adequately fill in the  $U(1)$  Core Type II features. It will emerge from our exploration that in one crucial respect, GW’s  $U(1)$  Type II example is significantly disanalogous from the Faraday cage.

Section 3 sets about the task of filling in the  $U(1)$  Core Type II features. We do so by taking the following steps to argue for a link between Type II scenarios and recent advances in the foundations of gauge theory: first, Section 3.1 presents a philosophical interpretation

of the work of Donnelly and Freidel [2016]. Second, Section 3.2 argues that when the picture developed in Section 3.1 is supplemented with the recent results of Mathieu et al. [2019], we have precisely the materials that we need to ‘fill in’ the structure of a Type II scenario. Section 3.3 then summarizes the consistently ‘filled in’  $U(1)$  Core Type II features.

Finally, Section 4 uses the material developed in previous sections to assess the coherence of Type II empirical significance. As we show in Section 2.1, Type II scenarios amount to the conjunction of the Core Type II features and (Empirical Symmetry). Thus, our assessment of the  $U(1)$  case proceeds by considering whether the  $U(1)$  Core Type II features are consistent with (Empirical Symmetry). In Section 4.1, we show that the former is in fact inconsistent with the latter. Section 4.2 then explores the scenario in which GW give up (Empirical Symmetry). In this case, we find that there is a sense in which one has ‘non-relational’ empirical significance, but that the specter of Galileo’s ship still looms large.

## 2 Type II empirical significance

### 2.1 Preliminaries

In order to bring the Type II case of empirical significance for gauge symmetries into view, it will help to first issue a reminder about some features of Type I cases, viz., cases in which a gauge symmetry is used in a way that is straightforwardly analogous to Galileo’s Ship.

Recall that the ‘dynamically allowed’ states of a subsystem are the set of kinematical states that satisfy the theory’s equations of motion. Alternatively (and in a formulation that will be useful in Section 3), they are the set of kinematical states that extremize the theory’s action. Roughly speaking, GW work with the following schema for symmetries: let  $\mathcal{S}$  and  $\mathcal{E}$  be the set of dynamically allowed states for a subsystem and its environment respectively; the set of dynamically allowed states for the ‘universe’ (or total system) is

then described as  $\mathcal{U} \subset \mathcal{S} \times \mathcal{E}$ . A subsystem state  $s \in \mathcal{S}$  and an environment state  $e \in \mathcal{E}$  are dynamically compatible just in case there exists a universe state  $s \odot e \in \mathcal{U}$  that restricts to  $s$  and  $e$  respectively, where the ‘ $\odot$ ’ notation is used to indicate the composite interpretation of the universe state.

GW assume that the universe carries a set of dynamical symmetries  $\Sigma$ , and for any universe symmetry  $\sigma_{\mathcal{U}} \in \Sigma$ , a subsystem symmetry  $\sigma$  is defined by restricting  $\sigma_{\mathcal{U}}$  to  $\mathcal{S}$ ; the set of such subsystem gauge symmetries will be denoted  $\Sigma_{\mathcal{S}}$ . We note that GW assume that the subsystem symmetry  $\sigma$  *leaves the subsystem  $\mathcal{S}$  invariant* (whatever goes into the definition of  $\mathcal{S}$ ), a consideration that will have special relevance for us in Section 3.

GW then provide the following informal characterization of the Type I case: a subsystem gauge symmetry  $\sigma \in \Sigma_{\mathcal{S}}$  has Type I empirical significance when some subset of subsystem states  $\{s\}$  has boundary conditions that are preserved by  $\sigma$ , and the composite universe state  $\sigma(s) \odot e$  is physically distinct from  $s \odot e$ , where  $e$  is any compatible environment state. Notice that—just as with the classic Galileo’s ship scenario—the subsystem gauge symmetry is here being used to effect a *relational* difference between the subsystem and the environment [Greaves and Wallace, 2014, 73].

Due to recent work, we have acquired a good understanding of how this schematic Type I scenario can be instantiated in various gauge theories, including Yang-Mills theory and General Relativity. In such scenarios, asymptotic reasoning is invoked to justify and implement boundary conditions that are characteristic of certain isolated systems (because, e.g., the fields ‘die off’ in a suitable sense at the asymptotic boundary dividing the subsystem and the environment).

As Teh [2015] points out, in order to understand how such examples are consistent with one common use of the term ‘gauge symmetry’, one must first disambiguate between (A) a generic—possibly non-dynamical—notation of gauge symmetry (in which any map from



spacetime to a structure group acting on fields in a certain way is a gauge symmetry); and (B) a more refined dynamical notion according to which an (A)-type gauge symmetry is further required to encapsulate redundancy for a particular dynamical system, whose states can only be defined after fixing specific boundary conditions (it follows that (B)-type gauge symmetries leave the boundary conditions of this subsystem invariant). In fact, in Section 3 we will have occasion to remind the reader of a useful geometric characterization of (B)-type gauge symmetries, and to exhibit a model in which these symmetries are asymptotically trivial (or ‘small’) according to this characterization. Furthermore, in this model, certain (A)-type transformations transform subsystem states into physically distinct states. As discussed in Teh [2015], these ‘large’ transformations are precisely the ones that can exhibit Type I empirical significance.

Remarkably, GW claim that the Type I case is only one of two paradigmatic ways in which (subsystem) gauge symmetries may be said to have empirical significance; we shall call this putative second way ‘Type II empirical significance’. In Type II scenarios, the subsystem region  $R_S$  is stipulated to be the *interior* (call this the ‘bulk’) of a spacetime  $M$  with boundary  $\partial M$ , the environment region  $R_E$  is  $\partial M$ , and the universe region is  $M$ .

To review Type II scenarios efficiently, let us follow GW in introducing the notion of a ‘boundary class’ for subsystem states that are dynamically compatible with the boundary state  $e \in \mathcal{E}$ , denoted  $C_e$ . More specifically, the ‘boundary class’ of  $e$ ,  $C_e$ , is the set of states  $s \in \mathcal{S}$  such that  $s \odot e$  is well-defined [2014, 84].<sup>6</sup> GW further stipulate that  $e, e' \in \mathcal{E}$  are two physically distinct charged states of the environment (i.e. they are boundary states) with respect to some very general notion of ‘boundary charge’. Given the above, Type II scenarios take the following schematic form: we suppose there is a subsystem gauge

---

<sup>6</sup>We note that we have opted to use ‘boundary class’ here in place of GW’s term ‘boundary condition’ because a boundary class is not a set of boundary conditions in the ordinary sense, see fn. 7 below.

symmetry  $\sigma$  such that  $\sigma(C_e) = C_{e'}$  (in other words,  $\sigma$  changes the boundary class of  $e$  to that of the physically distinct state  $e'$ ). Clearly, in order for this supposition to hold,  $\sigma$  must be non-trivial in a small ‘collar’ of the subsystem boundary  $\partial M$ , and by continuity, it has a non-trivial extension to a universe symmetry that *includes* a boundary symmetry supported on  $\partial M$  (apart from Section 4.2—where we shall introduce explicit notation for such an extension—in what follows, we will simply speak of the ‘subsystem symmetry’ in contexts where it is clear that we mean to include the boundary extension). Then,  $s \odot e \mapsto \sigma(s) \odot e'$  is a change in the physical state of the universe, and GW say that  $\sigma$  is an ‘empirically significant’ subsystem gauge symmetry. Because of this reliance on a non-trivial change  $e \mapsto e'$  in the environment state (in contrast to Type I scenarios which only rely on a change in the subsystem state), GW consider Type II to be ‘not fully relational’.

GW offer two clarifications about how the transition  $s \odot e \mapsto \sigma(s) \odot e'$  should be understood. First, they remind us that the transition cannot be effected by a universe symmetry, since (by their lights) states connected by a universe symmetry are physically indistinguishable, whereas (by hypothesis) the states  $s \odot e$  and  $\sigma(s) \odot e'$  are physically distinct in a Type II scenario. Second, they note that invoking the transition  $e \mapsto e'$  might lead to the concern (voiced, for example, by Friederich [2015, 544]) that the empirical content they wish to ascribe to the subsystem gauge symmetry is in fact *completely* attributable to the change in the environment state, thus leaving no room for the gauge symmetry to do any work (that is, one might worry that the subsystem symmetry is not empirically significant; only the change in the environment state is). GW react to this concern by requiring a further condition of adequacy on Type II scenarios: there should be a ‘principled connection’ between  $e \mapsto e'$  and the gauge symmetry  $\sigma$  [2014, 68], although they say nothing further about what such a connection might be. In Section 4, we will see that on the most natural and straightforward way of cashing out ‘principled connection’, the concern is indeed well-

founded. However, it will take a bit of work to fully appreciate this point.

The above is the extent of what GW say in their schematic characterization of Type II empirical scenarios. For the purposes of our discussion, it will be useful to more carefully regiment and comment on the salient list of features in their characterization. First, we will lay out what we call the four **Core Type II** features:

**F1:** Recall that in Type II scenarios, the subsystem region  $R_S$  is the *interior* (call this the ‘bulk’) of a spacetime  $M$  with boundary  $\partial M$ , the environment region  $R_E$  is  $\partial M$ , and the universe region is  $M$ . In Type II scenarios there is no condition of ‘dynamical isolation’ imposed between the subsystem and the environment. Thus, *the dynamical compatibility between the subsystem (bulk) states and the environment (boundary) states is a non-trivial consequence of the dynamics of the universe  $\mathcal{U}$ .*

**F2:** The environment (boundary) system is equipped with a notion of state and a standard according to which  $e$  and  $e'$  are physically distinct charged states with respect to some notion of boundary charge.

**F3:** We recall from standard accounts of dynamical systems that one of the key conceptual roles of a ‘charge’ is to generate transformations between charged states. Thus, since (by hypothesis) the environment of Type II scenarios carries a boundary charge and correspondingly charged states, we should take the transition  $\hat{\sigma} : e \mapsto e'$  to be generated by the corresponding boundary charge (we provide more details on this in Section 2.3).

**F4:** The subsystem states are free of any boundary conditions, because the environment (boundary) system is supposed to be truly dynamical and imposing such conditions would quite generally trivialize (or at least unduly constrain) the dynamics of the

environment system.<sup>7</sup>

GW’s full picture can then be arrived at by supplementing these Core Type II features with the following claim:

(Empirical symmetry) The non-trivial-on-the-boundary subsystem gauge symmetry  $\sigma$  (such that  $\sigma(C_e) = C_{e'}$ ) exhibits a ‘principled connection’ with the transition  $\hat{\sigma} : e \mapsto e'$ .

Evidently, the above combination of features (F1—F4 and (Empirical Symmetry)) is highly schematic and understanding what it really amounts to—or if the combination is even coherent—will only be possible upon adding more detail to the schema. One of GW’s main attempts to add such detail is by means of a  $U(1)$  gauge theory Type II scenario, which they say is simply the gauge-theoretic (and dynamical) analog of the Faraday cage. In order to get a better grip on this analogy (and any potential disanalogies) we now proceed to a brief review of the Faraday cage. Before we continue, let us stress that this paper is not about the Faraday cage, nor about the empirical significance of gauge transformations in electrodynamics more generally (for such a discussion, see Healey [2009]). The reason we will now briefly discuss the Faraday cage is to clarify the extent to which it is analogous to Type II scenarios, and the extent to which it is not—such discussion will help us add some details to the characterization of Type II cases.

## 2.2 An interlude: The Faraday cage

We now turn to a brief review of the Faraday cage in a form that will be useful for our discussion. Consider the (kinematic) electromagnetic field content  $A = (\phi, \vec{A})$  in a finite

---

<sup>7</sup>By ‘boundary conditions’ here, we mean conditions on the kinematic states at the boundary of the subsystem region (for example, Dirichlet or Neumann boundary conditions) that serve as input parameters for the problem of solving the equations of motion to obtain the dynamical states of the subsystem.

spatial subsystem region  $R_S$ , on whose boundary surface  $\partial R_S \equiv S$  we place the boundary conditions for a perfect electric conductor with surface charge  $\sigma$  (for this subsection only,  $\sigma$  will denote surface charges and not subsystem transformations).<sup>8</sup> We then obtain subsystem states by requiring that the field content solves the electrostatic equations of motion.

We now review how one can construct an explicit scalar potential solution (in terms of the surface charge) in the interior of  $R_S$ . In the standard way of proceeding, we do so by first gauge-fixing  $A$ , i.e., we perform a gauge-transformation  $A \mapsto A^f = A + d\chi^f(x, t)$  (where the superscript  $f$  stands for ‘gauge-fixed’) such that the resultant field  $A^f$  satisfies the Coulomb gauge:

$$f(A) = \nabla \cdot \vec{A}^f = \nabla \cdot (\vec{A} + \nabla\chi^f(x, t)) = 0. \quad (1)$$

As it is well known, the above data leads to Poisson’s equation for the scalar potential (and to a somewhat complicated equation for the vector potential).<sup>9</sup> For the purposes of the present discussion, we will not need the explicit solution of these equations, but instead wish to emphasize some generic features of the model that follow from two considerations. First, it is clear from equation (1) that the gauge-parameter  $\chi^f(x, t)$  is field-dependent because it depends on the initial vector potential  $\vec{A}$  via  $\nabla \cdot \vec{A} = -\nabla^2\chi^f(x, t)$ . Second, it is a trivial fact that by fixing a gauge, we end up shifting the scalar potential both in the interior and at the boundary:  $\phi(x) \mapsto \phi^f(x) = \phi(x) - \partial_t\chi^f(x, t)$ .

From the above two considerations and the boundary conditions involved in a conductor,

---

<sup>8</sup>There are two relevant boundary conditions: (1) the electric field  $\vec{E}$  at the boundary vanishes in the tangential direction, and (2) its magnitude equals  $E_\perp = 4\pi\rho$  in the perpendicular direction. From this, it follows that the potential at the boundary is constant and its normal derivative in the boundary is proportional to the surface charge.

<sup>9</sup>For a good introductory discussion of the topics discussed in this section, see Zangwill [2013, Ch. 15].

one can easily show that a Faraday cage is a system in which (i) the boundary charge  $\sigma$  is defined in terms of a particular boundary condition, viz. a fixed gauge parameter  $\chi_s^f(x, t)$  on the boundary (which comes from the gauge-fixing required to solve the equations of motion); and (ii) shifts in the potential at the boundary, coming from a boundary symmetry that is non-trivial and rigid at the boundary, result in changes in the boundary charge. Notice that (ii) corresponds to the familiar result that in a Faraday cage, rigid shifts of the potential at the boundary are related to changes in the surface charge.<sup>10</sup>

Now, the Faraday cage case just discussed is analogous to a Type II scenario with respect to (ii): for here we have the Type II-like phenomenon of a boundary symmetry that does not preserve the boundary class of  $e$ , where  $e$  is in this case the (constant) values of the potential at the surface that are associated with a given charge  $Q$ . In other words, the rigid symmetry in the subsystem maps a given scalar field  $\phi(x)$  (associated with a given charge  $Q$ ) to another scalar field  $\phi'(x)$  (associated with a different charge  $Q'$ ). However, as we can already see from the Core Type II feature F4, there is also a significant disanalogy between Type II scenarios and the Faraday cage with respect to (i) above: unlike the Faraday cage scenario, the boundary charge of a Type II scenario cannot be defined in terms of (subsystem) field boundary conditions, because the environment states are themselves taken to be dynamical (by contrast, recall that in the Faraday cage scenario, the boundary conditions force the potential to be constant at the boundary). In what follows, we will see how this disanalogy plays out in the  $U(1)$  Type II context.

---

<sup>10</sup>Both of these points can be easily illustrated in the following manner. In the Coulomb gauge, the scalar potential  $\phi^f(x)$  satisfies Poisson's equation. At the boundary of the conductor, this means that  $\epsilon_0 \nabla^2 \phi^f(x_s) = -\sigma$  (where  $x_s$  corresponds to locations at the boundary). Given that  $\phi_s^f(x) = \phi_s(x) - \partial_t \chi_s^f(x, t)$  (coming from the gauge fixing), we then get an implicit relation between  $\chi_s^f(x, t)$  and the surface charge  $\sigma$ . Furthermore, in a conductor, a specification of  $\phi^f(x_s)$  (the potential on the boundary) uniquely determines the surface charge via  $Q_s = -\epsilon_0 \int_s dS \hat{n} \cdot \nabla \phi^f(x_s)$ . Hence, shifts in  $\phi^f(x_s)$  coming from  $\partial_t \chi_s^f(x, t)$  induce shifts on the charge.

### 2.3 The $U(1)$ Type II scenario

We begin by fixing terminology for the subsystem and environment regions that we will consider in the rest of the paper. Let  $M = \Sigma \times \mathbb{R}$  be a 4-dimensional flat spacetime with timelike boundary  $\partial M = \partial\Sigma \times \mathbb{R}$ , where  $\Sigma$  is a 3-dimensional spatial Cauchy surface with boundary.<sup>11</sup> Our universe region will be  $M$ , our subsystem region  $R_S$  (the bulk) will be the interior of  $M$ , and our environment region  $R_E$  will be  $\partial M$ .<sup>12</sup>

We now describe how GW fill in the Type II schema in the context of a  $U(1)$  gauge theory. Recall that the schema is

$$s \odot e \mapsto \sigma(s) \odot e', \tag{2}$$

where  $\sigma$  is a subsystem gauge symmetry such that  $\sigma(C_e) = C_{e'}$ , the environment carries a boundary charge, and  $e$  and  $e'$  are physically distinct charged environment states.

GW take the subsystem  $\mathcal{S}$  to be the space of on-shell gauge fields  $\{A\}$  on the region  $R_S$ , and the symmetries of  $\mathcal{S}$  to be the gauge transformations  $A \mapsto A + d\chi$ , where  $\chi$  is an arbitrary smooth function. Unfortunately, GW's description of environment (boundary) states in the  $U(1)$  context is essentially black-boxed; apart from including the boundary value  $A_\partial$  of the subsystem gauge field  $A$ , the only thing we know about this state is that its dynamical degrees of freedom should give rise to a boundary charge.<sup>13</sup> For the moment, we will thus describe the environment states schematically as  $e = (A_\partial, \Phi)$ , where  $\Phi$  denotes the relevant 'black box' boundary degrees of freedom that give rise to a charge. According

---

<sup>11</sup>All the spacetime manifolds in this paper will be assumed to topologically trivial or 'contractible'.

<sup>12</sup>GW's use of the 'interior' here to individuate subsystems is somewhat idiosyncratic [Greaves and Wallace, 2014, 69-70], but we adhere to this set-up in order to facilitate a direct comparison with their claims.

<sup>13</sup>The reason we know this is that GW stipulate that the environment/boundary system has its own dynamical 'field' degrees of freedom, and that it carries a boundary charge—thus, the charge must be underwritten by the dynamics of the environment system.

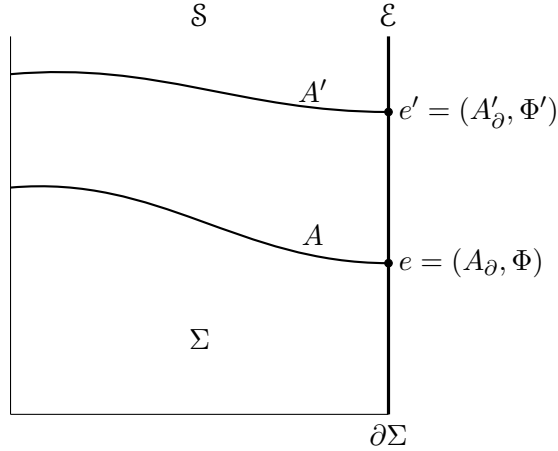


Figure 1: A schematic illustration of a spatial slice of a  $U(1)$  Type II case. The universe spatial region is represented by  $\Sigma$  with spatial boundary  $\partial\Sigma$ . Our subsystem's region is the interior of  $\Sigma$ , and the environment's region is  $\partial\Sigma$ . The subsystem  $\mathcal{S}$  is the space of on-shell gauge fields  $\{A\}$  defined on the interior of  $\Sigma$ . We take  $e = (A_\partial, \Phi)$  and  $e' = (A'_\partial, \Phi')$  to represent physically distinct states of the environment.

to GW, the putatively ‘empirically significant’ subsystem symmetry  $\sigma : A \mapsto A + d\chi$  is non-trivial on the boundary and sends  $C_e$  (the set of subsystem gauge fields that are dynamically compatible with  $e$ ) to  $C_{e'}$ . We refer the reader to Figure 1 for an intuitive visualization of two subsystem states that fall into these distinct boundary classes.

Finally, GW tell us that the above subsystem gauge symmetry  $\sigma$  should have a ‘principled connection’ to the transition  $\hat{\sigma} : e \mapsto e'$  between physically distinct charged environment (boundary) states. We already know (see F3 in Section 2.1) that we should think of  $\hat{\sigma}$  as being generated by the corresponding boundary charge. Thus, the ‘principled connection’ that GW are after is precisely one between a non-trivial-on-the-boundary subsystem symmetry  $\sigma$  and the boundary charge  $Q$ . What might this relationship be? In this context, we propose to interpret this relationship as (what physicists would uncontroversially take to be) the paradigm case of a non-accidental connection between a symmetry and a charge



$Q$ , viz., the relationship expressing how the symmetry is *generated* by the charge in the context of the theory's phase space:

$$\delta Q_\xi = \Omega(X_\xi, \cdot) \tag{3}$$

where  $\delta$  denotes the variation of the charge,  $X_\xi$  is the symmetry vector field on phase space representing the infinitesimal symmetry  $\xi$ , and  $\Omega(\cdot, \cdot)$  is the symplectic form of the theory's phase space.<sup>14</sup>

Though the above is certainly a slender description of the  $U(1)$  Type II scenario, we can gain a better understanding of what GW are proposing by again isolating the philosophically salient features in this more concrete context. In the labeling scheme of Section 2.1, we have the following  $U(1)$  **Core Type II** features:

**F1:** The dynamics of the universe gives rise to the dynamical compatibility between the bulk states  $\{A\}$  and the boundary states  $\{(A_\partial, \Phi)\}$ .

**F2:**  $\Phi$  represents degrees of freedom that give rise to a notion of boundary charge for  $e$ , and there is a standard according to which  $e$  and  $e'$  count as distinct charged states.

**F3:** The transition  $\hat{\sigma} : e \mapsto e'$  is *generated* by the boundary charge, in the sense of equation (3). Thus, we can deduce that this transition should be thought of as a kind of boundary symmetry acting on boundary states.<sup>15</sup>

---

<sup>14</sup>We note that since GW are quite vague about what they mean by a 'principled connection', the general notion is open to different interpretations. We have adopted an interpretation based on their specific illustration of this notion in their electrostatic analogy (where the environment carries a charge), but there may be ways of understanding the general notion that do not require charges in the environment (and GW do not specifically mention charges in their discussion of the general notion). At any rate, we believe that our interpretation is a fruitful one in the context of this paper because it provides a clear and plausible account of what GW might mean, and it uncovers a non-trivial connection with recent themes in the theoretical physics literature.

<sup>15</sup>Recall that the transition cannot be implemented by means of a universe symmetry. That is,  $\sigma(s) \odot \hat{\sigma}(e)$  is not equivalent to  $\sigma_U(s \odot e)$  for some universe symmetry  $\sigma_U$ .

**F4:** The subsystem gauge fields  $\{A\}$  are not subject to boundary conditions on  $\partial M$ .<sup>16</sup>

Finally, the full GW picture is arrived at by supplementing F1–F4 with the following claim:

(Empirical Symmetry) The subsystem gauge symmetry  $\sigma : A \mapsto A + d\chi$  (such that  $\sigma(C_e) = C_{e'}$ ) exhibits a ‘principled connection’ with  $\hat{\sigma} : e \mapsto e'$ , in the sense that  $\sigma$  is generated by the corresponding boundary charge.

Now that we have regimented the above features, we are in a position to make two important observations about the  $U(1)$  Type II scenario. First, we repeat in this more concrete context what we already said at the end of Section 2.2, namely that there is a profound disanalogy between the Faraday cage scenario and the  $U(1)$  Type II scenario: in the Faraday cage scenario, boundary charges are tied to the imposition of appropriate boundary conditions on the static fields, whereas—as we have just seen in F4—the  $U(1)$  Type II scenario places no boundary conditions on subsystem gauge fields. Thus, a significant consequence of this disanalogy for GW is that they cannot rely on the Faraday cage as a guiding heuristic for spelling out the Type II ‘principled connection’ between a subsystem gauge symmetry and a boundary charge—such a connection must come about as a matter of ‘boundary dynamics’, and not (as in the Faraday cage case) as a matter of ‘boundary conditions’.

Second, even putting aside the claim that we have labeled (Empirical Symmetry), it is far from clear how to construct a theory that exhibits the above combination of Core Type II features. Two radical features of such a theory are (1) that it should include a dynamical theory of boundary charge that gives rise to the right ‘meshing’ conditions between subsystem and environment (boundary) fields; and (2) that the subsystem gauge

---

<sup>16</sup>As we said when presenting F4 for the first time, imposing boundary conditions on the subsystem would unduly constrain the dynamics of the environment system.

fields should be subject to no boundary conditions whatsoever.<sup>17</sup> As we have just seen above, even when GW are presenting a *particular* Type II model, viz. the  $U(1)$  case, these features are ‘black boxed’ and so one gains no insight into how—or if—they can be coherently combined in a specific model.

In what follows, we will decompose our assessment of the  $U(1)$  Type II scenario into two parts. First, in Section 3, we will discuss how the Core Type II features can be consistently combined within a concrete model; and second, in Section 4, we will discuss whether this combination of features is compatible with (Empirical Significance). Our strategy has the added dividend of forging a link between the philosophical literature on gauge symmetry and a cutting edge topic in theoretical physics, viz., the subject of ‘gauge symmetry on a spacetime with boundary’ (and the related issue of constructing ‘edge modes’). Together, these two parts of our assessment will help us address the two main tasks of our paper, namely, to explain why exactly Type II scenarios pose a deep puzzle, and second, to explain how that puzzle can be dissolved.

### 3 Constructing the Core Type II scenario

This section addresses the problem of ‘filling in’ the details of GW’s  $U(1)$  Type II scenario in such a way that all of the Core Type II features can be consistently satisfied. By way of preface, subsection 3.1 offers a brief summary of the covariant phase space formalism, which provides a convenient framework for theorizing about this problem. In subsection 3.2, we argue that Donnelly and Freidel’s (2016) (‘DF’ henceforth) novel treatment of ‘gauge theory on spacetimes with corners’ provides a strategy for filling in Core Type II features

---

<sup>17</sup>We note that there is another weaker—because purely ‘topological’—sense in which the subsystem might still be said to have boundary conditions; such topological boundary conditions will enter our story in Sections 3.2 and 3.3.

F2–F4. We then round out this strategy in subsection 3.3 by showing that the recent work of Mathieu et al. [2019] (‘MSTW’ henceforth) offers a ‘dynamical completion’ (in the sense of variational dynamics) of DF’s results, thus clarifying their approach and providing a precise way of filling in the Core Type II feature F1. Finally, subsection 3.4 draws on the insights of DF and MSTW to construct a detailed and consistent model that satisfies the four Core Type II features.

### 3.1 A reminder: the covariant phase space framework

In this subsection, we review the basic ingredients of the covariant phase space framework in the context of a  $U(1)$  gauge theory over a spacetime region  $M$  with boundary  $\partial M$ .<sup>18</sup>

Recall that in a  $U(1)$  gauge theory, the space of kinematic fields is  $\mathcal{F} = \{A\}$ , and the Lagrangian is  $L = -\frac{1}{2}F \wedge *F$  (we assume there are no sources). In what follows, we will use ‘ $\delta$ ’ to denote a field space exterior derivative on  $\mathcal{F}$  and ‘ $\wedge$ ’ to denote the wedge product between differential forms on the space of fields.

We recall (the standard fact) that when there is a sufficiently strong field boundary condition placed on  $\partial M$ , e.g. the Dirichlet boundary condition  $A|_{\partial M} = A^*$  (where  $A^*$  is a fixed value), one obtains an unambiguous formula from varying the Lagrangian in the bulk:<sup>19</sup>

$$\delta L = \delta A \wedge (d * F) + d\theta[A, \delta A], \tag{4}$$

where the first term contains the Euler-Lagrange expression  $d * F$  and thus vanishes on-shell.<sup>20</sup>

---

<sup>18</sup>We direct the reader to Gomes [2019a] for a review in the philosophical literature, and Khavkine [2014] for a more technical treatment.

<sup>19</sup>The formula is also unambiguous in the case where  $M$  is a closed manifold.

<sup>20</sup>By ‘on shell’, we mean that we restrict the fields to the critical locus of the variation; equivalently, the fields are taken to satisfy the equations of motion.

$\theta$  in the second term is called the *pre-symplectic potential current* and has the explicit form  $\theta = -(\delta A \wedge *F)$ . When it is integrated over a spatial Cauchy surface  $\Sigma$ , it yields the *pre-symplectic potential* (a field space 1-form)

$$\Theta := \int_{\Sigma} \theta = - \int_{\Sigma} (\delta A \wedge *F). \quad (5)$$

We can then define the *pre-symplectic form* (a field space 2-form) by taking the field space exterior derivative of  $\Theta$ :

$$\Omega := \int_{\Sigma} \delta\theta = - \int_{\Sigma} (\delta A \wedge \delta(*F)). \quad (6)$$

In what follows, we will use two facts about the physical interpretation of  $\Omega$  and  $\Theta$ . First, since we are dealing with a gauge theory,  $\Omega$  will have a non-trivial kernel; the ‘redundant’ or (B)-type gauge symmetries can then be characterized as ‘the (A)-type gauge symmetry vector fields (on field space) which lie in the kernel of  $\Omega$ ’ (we refer the reader to Section 2.1 for a review of the (A)-type and (B)-type terminology). For instance, if we are assuming the Dirichlet boundary condition  $A|_{\partial M} = A^*$  (where  $A^*$  is constant) on our subsystem fields, then the gauge transformations in the kernel of  $\Omega$  (i.e. the ‘redundant ones’) will be the ones that are trivial on the spatial boundary  $\partial\Sigma$  (we note that the (B)-type gauge symmetries discussed in the context of Type I scenarios in Section 2.1 fall into this category, albeit interpreted asymptotically).

Note that specifying which directions in field space are ‘redundant’ and which directions are ‘physical’ constitutes part of the *specification* of a particular gauge subsystem. Thus, specifying  $\Omega$  is necessary in order to provide the full physical information about the subsystem space of states  $\mathcal{S}$ . Furthermore, and as GW note, ‘subsystem symmetries’ should leave the relevant subsystem invariant, and thus  $\Omega$  and  $\Theta$  should be invariant under the subsystem gauge symmetries; it is easy to check that this is true (on-shell) when equation

(4) is unambiguously defined (for example, when there are sufficiently strong boundary conditions for the fields, or when there is no boundary).

Second, let us observe that  $\Theta$  is closely related to the construction of Noether (or ‘symplectic’) charges for a subsystem: when there is an (infinitesimal) symmetry  $\xi$  that is a stabilizer of the field configuration on the boundary, then the charge is defined as  $Q_\xi := \iota_\xi \Theta$  (where ‘ $\iota$ ’ denotes the interior product).<sup>21</sup> For instance, in the case of the gauge field Dirichlet boundary conditions considered above, there is a set of rigid (constant) automorphisms of the boundary condition  $A^*$  that do not lie in the kernel of  $\Omega$  (these constitute examples of the ‘empirically significant’ (A)-type gauge transformations referred to in Section 2.1). From Cartan’s magic formula  $L_\xi \Theta = \delta \iota_\xi \Theta + \iota_\xi \delta \Theta$  and the fact that  $\Theta$  is invariant (i.e.  $L_\xi \Theta = 0$ ), we obtain the relation  $\delta Q_\xi = -\iota_\xi \Omega$ .<sup>22</sup> And this is nothing other than a geometric form of equation (3), i.e. the familiar statement that the rigid boundary symmetries are generated by the charge  $Q$ .

### 3.2 A link to the frontiers of theoretical physics

The covariant phase space framework introduced above has been widely used by physicists since at least Crnkovic and Witten [1987]. More recently, Donnelly and Freidel [2016] applied this framework to the following question in the foundations of gauge theory: how should one construct the pre-symplectic form of a gauge subsystem for which (i) the subsystem region has a boundary; (ii) the subsystem gauge fields are not subject to boundary conditions; and (iii) there is a notion of boundary charge (and boundary symmetry) asso-

---

<sup>21</sup>This notion of ‘stabilizer symmetry’ is just the  $U(1)$  analog of a Killing vector symmetry in General Relativity.

<sup>22</sup>Recall that  $Q_\xi = \iota_\xi \Theta$  and  $\Omega = \delta \Theta$ . The reason that  $\Theta$  is invariant here is that we do not allow boundary variations of the gauge parameter  $\chi$  (due to the restrictive boundary condition); later in section 3.2, we will lift the restriction of field space boundary conditions and treat the space of fields as a ‘stack’ – in that scenario it is clear that  $\chi$  is field-dependent, i.e.  $\delta \chi \neq 0$ .

ciated with ‘edge mode’ degrees of freedom on the boundary? It should be immediately apparent to the reader that this set of desiderata is very closely related to what we called the Core Type II features in Section 2.3—as we will see in this section, satisfying these essentially corresponds to satisfying features F2–F4. Indeed, when supplemented with an account of the meshing between bulk and boundary states (condition F1), we will see that DF end up with the full set of Core Type II features.

To spell out how DF construct an ‘adequate’ pre-symplectic potential and form (call these  $\Theta'$  and  $\Omega'$  respectively) that accounts for (i), (ii) and (iii), we will need to first provide a philosophical reconstruction of their argument. Officially, DF seem to argue as follows in their [2016] paper: because we are in a situation where there are no boundary conditions on subsystem gauge fields, we should allow arbitrary non-trivial gauge transformations on the subsystem boundary. A quick calculation shows that  $\Theta$  (from equation (5)) is not invariant under all of these transformations. However, any reasonable pre-symplectic potential should be invariant under such transformations, and so we should introduce additional ‘edge mode’ fields on  $\partial\Sigma$  that contribute to  $\Theta$ , making the resulting structure  $\Theta'$  gauge-invariant.

We think that DF have made an important foundational contribution through their construction of  $\Theta'$  and the associated edge modes. However, we also think that the above argumentative strategy requires some clarification, because it is not clear what role  $\Theta$  (from equation (5)) is playing in the argument. Here is our preferred version: start with the goal of constructing some gauge-invariant pre-symplectic potential  $\Theta'$  (and thus a gauge-invariant pre-symplectic form  $\Omega'$ ) for a subsystem whose gauge fields are not subject to any boundary conditions and for a subsystem for which there is a notion of boundary charge. Now, we already know that  $\Theta$  in equation (5) (stemming from e.g. the variational problem for a subsystem with Dirichlet boundary conditions) will not help us meet this goal—it does not stem from the right variational problem for our (boundary condition-free) scenario, and

there is no reason to expect that it will be invariant with respect to gauge symmetries that are non-trivial on the boundary of the subsystem. Nonetheless, this obviously ‘wrong’  $\Theta$  can still be effectively employed in an attempt to construct the suitable  $\Theta'$ . Indeed, this is precisely how we understand DF’s strategy.

The use of the wrong pre-symplectic potential  $\Theta$  to construct the right one ( $\Theta'$ ) rests on a simple observation concerning how  $\Theta$  fails to be gauge invariant under a generic (and thus possibly field-dependent)  $U(1)$  gauge transformation  $\chi$ . Recall that the action of such a transformation on  $A$  is given by  $\chi^*A = A - d\chi$ . It is easy to see that  $\Theta$  transforms as follows:

$$\chi^*\Theta_\Sigma = \Theta_\Sigma + \int_{\partial\Sigma} \delta\chi \wedge (*F). \quad (7)$$

In other words, the failure of gauge invariance is due only to a boundary integral over  $\partial\Sigma$  (the second term on the RHS of equation (7)).

In light of this observation, it is also straightforward to ‘correct’ for this non-invariance by directly introducing a ‘compensating’ boundary symplectic potential, thereby producing a gauge-invariant  $\Theta'$  that is appropriate for our system (this is precisely what DF do in Section 2 of their [2016] paper). However, we will take a slightly different (and more systematic) approach to implementing this correction because we wish to emphasize a connection here with recent philosophical work [François, 2019] on the ‘dressing field method’ in gauge theory.

A dressing field for a  $U(1)$  theory is a scalar field  $\varphi$  that transforms under gauge transformations as  $\chi^*\varphi = \varphi - \chi$ . François notes that, if a dressing field  $\varphi$  exists, one can always introduce the gauge-invariant gauge field  $A^\varphi \equiv d_A\varphi := A - d\varphi$  (where  $d_A$  is an affine covariant differential acting on  $\varphi$ ). Now, by defining



$$\Theta' = \Theta[A^\varphi, \delta A^\varphi] = \Theta + \int_{\partial\Sigma} \delta\varphi \wedge (*F) \quad (8)$$

we immediately see (by computing  $\chi^*\Theta'$ ) that not only is  $\Theta'$  gauge-invariant, but also that in order to construct  $\Theta'$  one only needs to assume that the dressing field  $\varphi$  *exists on the boundary*  $\partial\Sigma$ . As it turns out,  $\Theta'$  so defined is exactly what DF claim is the appropriate pre-symplectic potential for a system satisfying (i)–(iii), and the boundary dressing field  $\varphi$  is what they call the ‘edge mode’ whose contribution ( $\int_{\partial\Sigma} \delta\varphi \wedge (*F)$ ) to the pre-symplectic potential (as defined in equation (5)) cancels out the term responsible for the non-invariance of  $\Theta$ .<sup>23</sup> At this point, DF also introduce the phase space momentum mode  $\pi$  that is conjugate to  $\varphi$ .

DF then observe that one can introduce a new symmetry on  $\partial\Sigma$ , which acts trivially on the gauge fields and transforms  $\varphi$  in the opposite manner to a gauge symmetry—they call this a ‘physical boundary symmetry’. So, to sum up, in DF’s setup, we have a gauge symmetry which acts on fields as

$$A \mapsto A - d\chi, \quad \varphi \mapsto \varphi - \chi \quad (9)$$

as well as a physical boundary symmetry that acts as

$$A \mapsto A, \quad \varphi \mapsto \varphi + \alpha. \quad (10)$$

Furthermore, DF then note that the generator—or *charge*—of the physical boundary sym-

---

<sup>23</sup>We note that, strictly speaking, there is a problem with regarding  $*F$  and  $\delta\varphi$  as symplectic ‘pairs’ in this formulation of  $\Theta'$ , viz. that they are not functionally independent of the other terms in  $\Theta$ ; this problem is rectified in our Section 3.3, where one directly derives the correct pre-symplectic potential from a boundary Lagrangian. We thank a referee for pushing us to clarify this point.

metry is the phase space function

$$E[\alpha] := \int_{\partial\Sigma} \pi\alpha, \tag{11}$$

as one can see explicitly from the formula  $\delta E[\alpha] = \iota_{\xi_\alpha} \Omega'$ , where  $\Omega' = \delta\Theta'$  and  $\xi_\alpha$  is the tangent vector (on field space) representing the infinitesimal symmetry transformation (10).

The above is our retelling of how DF construct what they take to be the right pre-symplectic form  $\Omega'$  for a system satisfying (i), (ii) and (iii). And it is easy to see from equation (8) that all the naive (A)-type symmetries lie in the kernel of  $\Omega'$  (i.e. they are of (B)-type), as is appropriate for a scenario where we do not impose any boundary conditions on the subsystem gauge fields.

### 3.3 A dynamical completion of DF's account

Despite its significant achievements, DF's account is not quite the whole story for four reasons. First, as standardly construed, the pre-symplectic form and potential are the outputs of a dynamical variational principle, viz. formula (4). But what is the corresponding Lagrangian that gives rise to  $\Omega'$  and  $\Theta'$ ? DF do not say. Second, recall that according to the Core Type II feature F1, the universe dynamics should give rise to a meshing condition between the environment (boundary) states and the subsystem states. While DF in fact write down such a condition, i.e.  $(*F = \pi)$ , they do not provide a dynamical account of how it arises.<sup>24</sup> Third, the existence of the edge modes  $\varphi$  is still somewhat mysterious: Why do they exist on the boundary and why is it natural to invoke them in the dynamics of Type II scenarios? Fourth, and relatedly, why do the edge modes 'carry' the action of a boundary symmetry that acts in the opposite manner from that of a gauge symmetry?

---

<sup>24</sup>They motivate their prescription by appealing to 'gauge-invariance' considerations, as explained in the previous section. But F1 demands a dynamical underpinning of the meshing conditions.

An explanation of all four issues has recently been given by Mathieu et al. [2019] (MSTW henceforth). They address the first point by showing that  $\Omega'$  can be directly derived by varying the following action with respect to bulk and boundary variations:

$$S(A, \varphi) = \frac{1}{2} \int_M F \wedge *F + \frac{1}{2} \int_{\partial M} d_A \varphi \wedge *_{\partial} d_A \varphi \quad (12)$$

(where  $*_{\partial}$  is the Hodge star operator on  $\partial M$ ). And they address the second point by showing that the variation yields three equations of motion:

$$d * F = 0 \quad (\text{bulk equation of motion}) \quad , \quad (13a)$$

$$d *_{\partial} d_A \varphi = 0 \quad (\text{inhomogeneous Klein-Gordon equation on the boundary}) \quad , \quad (13b)$$

$$*F = *_{\partial} d_A \varphi \quad (\text{bulk-to-boundary meshing condition}) \quad , \quad (13c)$$

where we note that  $\pi := *_{\partial} d_A \varphi$  can be regarded as the conjugate momentum to  $\varphi$  in the symplectic form (and thus relative to a choice of Cauchy surface). In other words, MSTW's equation (13c) accounts for the Core Type II feature F1, because it explains how the universe dynamics (12) gives rise to the meshing between bulk and boundary dynamical states.

On the third point (the somewhat mysterious and artificial presence of the edge modes), MSTW start by arguing that even though DF do not impose any *gauge field* boundary conditions on the subsystem, there is still a certain sense in which one is imposing a ‘topological boundary condition’. This is because one still requires that the bulk fields in the sense of ‘principal bundles with connection’ restrict to a ‘fixed background’ trivial principal bundle on the boundary. MSTW then use a homotopy-theoretic argument to show that  $\varphi$

is necessary in order to encode or ‘witness’ the different ways in which the bulk degrees of freedom can be related to this fixed background structure. In this respect,  $\varphi$  is not artificial or mysterious but rather arises naturally from the homotopy-theoretic interpretation of ‘topological boundary conditions’.

Fourth, based on their explanation of the third point, MSTW show that there are two kinds of symmetries that naturally act on  $\varphi$ : one kind is the familiar gauge symmetry of the bulk fields which transforms  $\varphi$  as in equation (9), but we can also consider automorphisms of the fixed background structure itself, which turn out to constitute the novel boundary symmetry introduced by DF. Indeed, it is easy to see from MSTW’s construction that the automorphisms of the background structure act on  $\varphi$  in the opposite manner from the gauge symmetry, and only act trivially on the gauge fields. In this way, MSTW recover the transformation rules (equation (10)) introduced by DF.

Finally, we note that if  $\alpha$  is a constant shift, then the boundary symmetry (in equation (10)) is actually a symmetry of the boundary action  $\int_{\partial M} d_A \varphi \wedge *_{\partial} d_A \varphi$  in equation (12). In what follows, we will restrict ourselves to this rigid symmetry. Thus the boundary symmetry and charge should be interpreted as properties of a ‘defect’ matter field lying on the boundary of  $M$ .

### 3.4 Core Type II embodied

We can now combine the insights of DF and MSTW to see how the  $U(1)$  Core Type II features can really be filled in, and not merely black-boxed as in GW’s original formulation. First, we assume that our universe theory is given by the MSTW action presented in equation (12). In other words, in addition to the usual bulk gauge field degrees of freedom,

we have a charged Klein-Gordon ‘defect matter field’ on the timelike boundary  $\partial M$ .<sup>25</sup> The subsystem states are solutions to  $d * F = 0$ , the environment states are solutions to  $d\pi = 0$ , and the subsystem and environment states are required to mesh by means of the equation of motion  $*F = \pi$  (these are equations (13a), (13b) and (13c) respectively).

We can now provide the full details of the Core Type II features F1–F4:

- F1:** The dynamics of the universe (encoded in the Lagrangian presented in equation (12)) gives rise to the dynamical compatibility between the bulk states  $\{A\}$  and the boundary states  $\{(A_\partial, \varphi)\}$ ; this is expressed in the bulk-to-boundary meshing condition (13c).
- F2:** The boundary degrees of freedom  $\varphi$  give rise to a notion of boundary charge for environment states, and there is a standard (e.g. some externally fixed  $\varphi$  frame) such that  $\varphi$  and  $\varphi'$  count as distinct charged states.
- F3:** The dynamical boundary symmetry  $\hat{\sigma} : \varphi \mapsto \varphi + \alpha$  (equation (10), where  $\alpha$  is a rigid parameter) is generated by the boundary charge  $E[\alpha]$  (equation (11)).
- F4:** The subsystem gauge fields  $\{A\}$  are not subject to any boundary conditions on  $\partial M$ , and this feature is reflected in the pre-symplectic potential  $\Theta'$  and the pre-symplectic form  $\Omega'$  of the subsystem.

---

<sup>25</sup>Two comments here: (i) One can also think of the boundary theory as a nonlinear sigma model that arises through the Stueckelberg mechanism on the boundary; (ii) the terminology of ‘defects’ is used by physicists to refer to a theory on a codimension 1 boundary of a theory on a codimension 0 region.

## 4 Whence Type II empirical significance?

### 4.1 Assessing (Empirical Symmetry)

In the previous section, we drew on the insights of DF and MSTW to spell out how GW’s  $U(1)$  model needs to be filled in so that it satisfies all four of the Core Type II features. We are thus finally in a position to assess whether GW’s claim of (Empirical Symmetry) is compatible with this fully fleshed out model, and thus whether  $U(1)$  Type II scenarios are consistent.

First, recall that in order for (Empirical Symmetry) to hold, the boundary extension of the subsystem gauge symmetry (9) needs to be generated by the boundary charge (this was our uncontroversial construal of the ‘principled connection’ between symmetries and charges). Recall too from equations (9) and (10) that the subsystem gauge symmetry  $\sigma$  acts on the boundary according to  $\sigma : A \mapsto A - d\chi$ ,  $\varphi \mapsto \varphi - \chi$ , whereas the physical boundary symmetry  $\hat{\sigma}$  acts as  $\hat{\sigma} : A \mapsto A$ ,  $\varphi \mapsto \varphi + \alpha$ .

We now claim that from the results of the previous section, it is clear why (Empirical Symmetry) cannot hold in a  $U(1)$  Core Type II scenario, i.e. a scenario with features F1–F4. Since the vector field (on the space of fields) representing the subsystem gauge symmetry  $\sigma$  lies in the kernel of  $\Omega'$ , it immediately follows from Cartan’s magic formula ( $L_\xi\Theta = \delta\iota_\xi\Theta + \iota_\xi\delta\Theta$ ) that  $\sigma$  cannot be generated by a boundary charge—thus, (Empirical Symmetry) does not hold. By contrast, we recall that  $\hat{\sigma}$  is indeed generated by the boundary charge according to (11). The key moral here is that once one has completely abandoned the project of imposing any gauge field boundary conditions on the boundary of a finite subsystem (as Type II scenarios suppose), then in a scenario with non-trivial boundary charge (as Type II scenarios also suppose), the subsystem gauge symmetries can no longer satisfy the functional role of ‘being generated by the boundary charge’. Instead, that role

gets handed off to a symmetry of the defect matter field  $\varphi$  on the subsystem boundary. It follows that, given our precisification, *Type II scenarios are incoherent*.

Our account also puts us in a position to explain the following curious feature of GW's story about Type II scenarios. Recall that, on the one hand, GW hold that (a) the transition  $s \odot e \mapsto \sigma(s) \odot e'$  cannot be effected by a universe symmetry, because the related states are physically distinct (which can in turn be traced back to the assumption that  $e$  and  $e'$  are physically distinct). On the other hand, they also hold that (b) the transition  $\hat{\sigma} : e \mapsto e'$  must be matched by the boundary extension  $\bar{\sigma}$  of a subsystem gauge symmetry  $\sigma$  (i.e.  $\sigma = \bar{\sigma}|_S$ ) that sends the boundary class  $C_e$  to  $C_{e'}$  (see Figure 1 for a visualization). But then it seems that  $\bar{\sigma}$  is precisely a universe transformation that sends  $s \odot e$  to  $\sigma(s) \odot e'$ !

How are we to resolve the tension between (a) and (b)? With the analysis of Section 3.2 in hand the answer is clear. First, consider the boundary symmetry  $\hat{\sigma}$  that sends  $e = (A_\partial, \varphi)$  to  $e' = (A_\partial, \varphi' = \varphi + \alpha)$ , where  $\alpha$  is a constant shift. How can the result of this transformation be ‘matched’ or ‘simulated’ by the boundary part of universe gauge symmetry  $\bar{\sigma}$ ? From (9), we see this can be done if we set  $\chi = -\alpha$ , which yields  $\bar{\sigma}|_S : (A_\partial, \varphi) \mapsto (A_\partial, \varphi' = \varphi - (-\alpha))$ . Thus, since the boundary parts match and  $\bar{\sigma}|_S = \sigma$ , applying the universe symmetry  $\bar{\sigma}$  results in the same *field value* as the result of transforming by  $\sigma \odot \hat{\sigma}$ . However, we also know from the previous analysis that according to the standard given by  $\Omega'$  (see equation (8)), the  $\varphi'$  that results from  $\bar{\sigma}$  has a wholly different interpretation from the  $\varphi'$  that results from  $\hat{\sigma}$  (despite the fact that both have the same value): the former is a mere redescription of the entire scenario, akin to a global change of coordinates, whereas the latter is generated by a non-trivial charge and thus capable of describing a physical difference in relation to some fixed standard for a charged state.

The resolution of the tension between (a) and (b) can then be put as follows: saying that  $s \odot e$  and  $\bar{\sigma}|_S(s) \odot e' = \sigma(s) \odot e'$  are physically distinct universe states amounts to more than

just giving the values of these respective states, otherwise the final state of the transition  $s \odot e \mapsto \bar{\sigma}|_g(s) \odot e'$  could equally well be expressed as either  $\bar{\sigma}|_g(s) \odot \hat{\sigma}(e) = \sigma(s) \odot e'$  or  $\bar{\sigma}(s \odot e) = \sigma(s) \odot e'$ , and the latter is incompatible with the assumption that the universe states are distinct. The missing information is precisely what is needed to underwrite the physical distinctness of  $e$  and  $e'$  by the standard of  $\Omega'$ , viz. the proviso that the transition  $e \mapsto e'$  is generated by the boundary charge. Once one has included this proviso, one sees that the desired transition  $s \odot e \mapsto \bar{\sigma}|_g(s) \odot e'$  can only be effected by  $\bar{\sigma}|_g \odot \hat{\sigma}$  and not by a universe symmetry.

## 4.2 Non-relationality and Galileo's ship again

We have just argued that GW's notion of a  $U(1)$  Type II scenario is incoherent, because the Core Type II features (which can be coherently combined in a model) are inconsistent with the claim of (Empirical Symmetry). However, it is worth emphasizing that one of GW's objectives in introducing Type II scenarios was to provide an example of *non-relational* empirical significance, and there might be something interesting that GW can say in this regard even if they have to give up the claim that it is the subsystem gauge symmetry that is empirically significant.

To that end, consider the scenario in which GW give up (Empirical Symmetry). They are still within their rights to say that Core Type II scenarios have an empirical symmetry, viz. the boundary symmetry (10) of the boundary matter field  $\varphi$ . Is this an example of non-relational empirical significance, or in other words: is this a case of empirical significance that is not of the Galileo's ship type?

The answer to this question turns in part on what exactly one means by 'relational' in this context. Arguably, GW's intended meaning here is a sense of 'relational' that is elaborated in terms of 'relations between different  $n$ -dimensional spacetime regions that



share a boundary’. So, for example, changes in the value of a given quantity defined for some space-time region relative to the value of that quantity in a *different* space-time region would be a generic example of a relational change. If this what is meant (and if one thinks of one of the spacetime regions as being ‘outside’ the boundary of the subsystem) then it is clear that Core Type II scenarios will be non-relational in the above sense, because—as we saw in Section 3.3—the empirical symmetry belongs to a codimension 1 ‘defect’ that lies at the interface of two  $n$ -dimensional spacetime regions. In other words, the changes in charge associated with the empirical symmetry are all changes within a codimension 1 boundary that is not itself one of the subsystem regions for the relata, in GW’s sense (the boundary’s spatiotemporal role is to provide a means of individuating the relata’s regions).

On the other hand, there is also a sense—very much in the spirit of GW’s discussion of Galileo’s ship cases—in which it makes sense to introduce a relational account of why the boundary symmetry is empirical. It can be elicited by posing the question: Why should we consider  $\varphi$  and  $\varphi' = \varphi + \alpha$  to be physically distinct charged states? Surely GW would want to gloss this in terms of a relational comparison between subsystems (not necessarily individuated spatiotemporally)—either between subsystems in field space, between subsystems within the codimension 1 boundary, or with respect to an externally fixed ‘charge frame’. In this sense, we should not take GW to have fully abandoned Galileo’s ship, even on this fallback position.

## 5 Conclusion

We close by briefly summarizing what we have accomplished in this paper. We began our investigation by contemplating a remarkable and somewhat bewildering claim by GW, namely, that subsystem gauge symmetries can display ‘non-relational’ empirical significance,

and that this possibility is bound up with two elements: (i) the notion of ‘boundary charge’ for a finite (i.e. non-asymptotic) region with boundary, and (ii) a heuristic analogy between the (electrostatic) Faraday cage and its hypothetical dynamical analog within Abelian gauge theory. The image given to us by GW of such ‘Type II scenarios’ (as we called them) is tantalizing but also elusive, because so many features of these scenarios are ‘black-boxed’ in their presentation. We thus undertook the task of spelling out exactly what a Type II scenario amounts to in the context of the simplest—i.e. Abelian—gauge theory.

In the course of thinking through these details, we found a surprising connection between the image given to us by GW and a foundational problematic in gauge theory that physicists have been grappling with recently: How can one describe gauge systems on spacetimes with boundary *without* imposing any field boundary conditions on those boundaries? We then provided a philosophical clarification of this question, and argued that its solution provides an effective method for spelling out the details of a Type II scenario.

Based on these results, we proceeded to argue that although GW’s notion of Type II empirically significant gauge symmetries is incoherent, the natural candidates for Type II models do in fact exhibit an interesting kind of non-relational empirical significance, albeit one that has to do with a rigid symmetry of the boundary theory (and not a gauge symmetry). We can thus conclude that although GW were mistaken in ascribing empirical significance to *gauge symmetry* in this context, much of what they say about the conceptual possibility of ‘non-relational empirical scenarios’ is interesting and realizable—indeed perhaps even prophetic given the development of the recent physics literature! If one is fortunate, bold conceptual conjectures sometimes lead to a serendipitous convergence between the philosophy of physics and the practice of physics—we are pleased to report that this is one such case.

## Acknowledgments

NT and SMR thank Henrique Gomes, Alex Schenkel, Philippe Mathieu, and Laura Wells for helpful discussions; NT also thanks John Norton for the opportunity to present this material as a Pittsburgh HPS colloquium talk. NT and SMR acknowledge the support of NSF grant 1947155, and SMR acknowledges the support of a non-teaching fellowship from Notre Dame's History and Philosophy of Science Program.

*Sebastián Murgueitio Ramírez*  
*History and Philosophy of Science*  
*University of Notre Dame, USA*  
*smurguei@nd.edu*

*Nicholas J. Teh*  
*Department of Philosophy*  
*University of Notre Dame, USA*  
*nteh@nd.edu*

## References

- Brading, K. and Brown, H. R. (2004). Are gauge symmetry transformations observable? *The British Journal for the Philosophy of Science*, 55(4):645–665.
- Crnkovic, C. and Witten, E. (1987). Covariant description of canonical formalism in geometrical theories. *Three hundred years of gravitation*, pages 676–684.
- Donnelly, W. and Freidel, L. (2016). Local subsystems in gauge theory and gravity. *Journal of High Energy Physics*, 2016(9):1–45.

- François, J. (2019). Artificial versus substantial gauge symmetries: A criterion and an application to the electroweak model. *Philosophy of Science*, 86(3):472–496.
- Friederich, S. (2015). Symmetry, empirical equivalence, and identity. *The British Journal for the Philosophy of Science*, 66(3):537–559.
- Gomes, H. (2019a). Gauging the boundary in field-space. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 67:89 – 110.
- Gomes, H. (2019b). Holism as the significance of gauge symmetries. <http://philsci-archive.pitt.edu/16499/>.
- Greaves, H. and Wallace, D. (2014). Empirical consequences of symmetries. *The British Journal for the Philosophy of Science*, 65(1):59–89.
- Healey, R. (2009). Perfect symmetries. *The British Journal for the Philosophy of Science*, 60(4):697–720.
- Khavkine, I. (2014). Covariant phase space, constraints, gauge and the peierls formula. *International Journal of Modern Physics A*, 29(5).
- Kosso, P. (2000). The empirical status of symmetries in physics. *The British Journal for the Philosophy of Science*, 51(1):81–98.
- Mathieu, P., Schenkel, A., Teh, N., and Wells, L. (2019). Homological perspective on edge modes in linear yang-mills theory. <http://philsci-archive.pitt.edu/16271/>.
- Nguyen, J., Teh, N. J., and Wells, L. (2018). Why surplus structure is not superfluous. *The British Journal for the Philosophy of Science*.

Teh, N. J. (2015). Galileo's gauge: Understanding the empirical significance of gauge symmetry. *Philosophy of Science*, 83(1):93–118.

Zangwill, A. (2013). *Modern Electrodynamics*. Modern Electrodynamics. Cambridge University Press.