Some remarks on Rovelli's 'Why Gauge?'

Nicholas J. Teh*

DAMTP and Department of History and Philosophy of Science, University of Cambridge

October 16, 2013

Contents

1 Introduction

In an insightful recent paper, Rovelli (2013) has argued against the orthodox claim that *gauge-dependent* quantities have no physical significance.

Let $\mathcal{S}$ be a physical system with gauge-invariant quantities $\{O_i\}$, i.e. each O_i is invariant under the action of a gauge transformation. A gauge-dependent quantity A is one that is *not* invariant under the action of a gauge transformation. The orthodoxy identifies the gauge-invariant quantities with the *physical* quantities of $\mathcal{S}$; gauge-dependent quantities, on the other hand, are characterized as 'redundancies' or 'descriptive fluff'.

There are various ways in which the orthodoxy can be challenged and refined. For instance, various authors have argued that the asymptotic gauge symmetries of a subsystem have physical significance (see e.g. Teh (2013); Greaves and Wallace (2013)), and arguably, the gauge transformations (relating gauge-dependent quantities) of a classical system contain physical information pertaining to the system's quantization.[1] Rovelli's particular challenge draws on the following key idea:

*njywt2@cam.ac.uk

[1]This latter point can be fully appreciated by considering the role that such transformations play in determining the existence of gauge anomalies.

(Couple) The gauge-dependent quantities of a system $\mathcal{S}$ contain information about how $\mathcal{S}$ *could* couple to other physical systems.

Thus gauge-dependent quantities contain physical information, albeit of a modal and relational type. Indeed, he takes (Couple) to explain (i) what it means to say that the world is described by a gauge theory, and relatedly, (ii) why our world is so well-described by gauge theories.[2]

Rovelli's strategy is as follows. He first tells a parable involving an elementary mechanical model in order to illustrate (Couple). He then argues that despite being rather more mathematically sophisticated, typical examples of theories with local symmetry – such as General Relativity and Yang-Mills-type theories – are analogous to the parable insofar as they illustrate (Couple).

The goal of this paper is to argue that, despite the obvious power of Rovelli's insight, his analogy is limited: Yang-Mills type theories illustrate (Couple) in a different and rather more subtle way than his parable.

In Section 2, we review Rovelli's parable and his subsequent claims about Yang-Mills-type theories. Section 3 discusses the disanalogies between Yang-Mills-type theories and his Parable. Finally, we return to claims (i) and (ii) the Section 4.

2 The parable of the spaceships

Rovelli illustrates (Couple) by telling the following parable. Let $\mathcal{S}_1$ be a system of N spaceships in Euclidean space, modeled as point particles $\{x_n(t)\}$, where $n = 1, \ldots, N$. The dynamics is specified by the Lagrangian

$$L_1 = \frac{1}{2}\sum_{n=1}^{N-1}(\dot{x}_{n+1} - \dot{x}_n)^2, \tag{1}$$

and the equations of motion are invariant under the gauge transformation

$$x_n(t) \mapsto x_n(t) + \lambda(t). \tag{2}$$

The gauge-invariant quantities of $\mathcal{S}_1$ are the relative distances $a_n := x_{n+1} - x_n$, where $n = 1, \ldots N-1$.

Now let $\mathcal{S}_2$ be a second system like the first, albeit with M spaceships $\{y_n(t)\}$, $n = 1, \ldots, M$. The Lagrangian L_2 and the gauge transformations are the same as (1) and (2), but with y_n coordinates instead. Similarly, the gauge-invariant quantities of $\mathcal{S}_2$ are the relative distances $b_n := y_{n+1} - y_n$, where $n = 1, \ldots, M-1$.

The parable then goes as follows. Considered as isolated and non-interacting systems (i.e. suppose the two fleets are very far apart) the fleets of $\mathcal{S}_1$ and $\mathcal{S}_2$ have $N-1$ and $M-1$ degrees of freedom respectively, which are captured by the gauge-invariant quantities a_n and b_n. For such systems, one has no need to invoke the gauge-dependent quantities x_n and y_n when describing the physical degrees of freedom.

But now consider what happens when the two fleets approach each other and develop an interaction term L_{int} in their composite Lagrangian $L_{12} = L_1 + L_2 + L_{int}$, where

$$L_{int} = \frac{1}{2}(\dot{y}_1 - \dot{x}_N)^2. \tag{3}$$

[2] Here Rovelli has in mind *local* gauge theories.

Evidently, the interacting system $\mathcal{S}_{12}$ now has $N + M - 1$ relative distances (or gauge-invariant quantities).

Rovelli then asks us to think about the following equation for counting the difference between the number of degrees of freedom (#) of the interacting composite system $\mathcal{S}_{12}$, on the one hand, and its isolated subsystems $\mathcal{S}_1$ and $\mathcal{S}_2$, on the other hand:

$$\#\mathcal{S}_{12} - (\#\mathcal{S}_1 + \#\mathcal{S}_2) = 1. \tag{4}$$

Evidently, $\mathcal{S}_{12}$ has an extra degree of freedom which comes from the gauge-invariant quantity $(y_1 - x_N)$, which is in turn formed by combining two gauge-dependent quantities: one from $\mathcal{S}_1$ and $\mathcal{S}_2$ respectively, i.e. x_N and y_1. No wonder then that the gauge-invariant quantities of $\mathcal{S}_1$ and $\mathcal{S}_2$ alone do not suffice to capture the full system's degrees of freedom!

Rovelli draws the following morals from this parable and seeks to apply them to the general case of theories with local symmetries, e.g. Yang-Mills theories and General Relativity.

(A) The parable illustrates the connection between 'gauge' and relationism about physical systems: the observables of a system are not its gauge-dependent quantities (e.g. the positions of spaceships or material objects), but rather the relations between such quantities (e.g. the relative distances between spaceships). These relational observables are precisely the gauge-invariant quantities.

(B) As illustrated by (4), there is a holism inherent in the system formed by combining two interacting subsystems: its observables do not reduce to just those of its subsystems. The additional observables come from interactions between the subsystems, i.e. relations between the gauge-dependent quantities of the subsystems. This Rovelli takes to vindicate (Couple): gauge-dependent quantities contain essential physical information about how such relations can be formed between subsystems.

(C) This insight furnishes us with a relational notion of what it means to 'measure' a gauge-dependent quantity x_n of a system $\mathcal{S}_1$. In order to make sure a measurement, one needs to let a measuring device system $\mathcal{S}_2$ interact with $\mathcal{S}_1$ in such a way that one forms a gauge-invariant quantity $c(x_n, y_n)$ of the composite system $\mathcal{S}_{12}$, where y_n is a gauge-dependent quantity of $\mathcal{S}_2$. E.g., in the parable, one fleet of spaceships served as the system being measured, and the other fleet served as a 'probe' or measuring device.

As we will see in the next section, none of these morals transfer straightforwardly to the case of Yang-Mills-type theories. Nonetheless, we will see that (Couple) remains a genuine insight, and that there is still a point to speaking of the relationship between gauge-invariance and relational observables, and how this in turn impinges upon the measurement of guage-dependent quantities, as Rovelli himself notes.

3 Yang-Mills-type theory and three disanalogies

Instead of discussing an $SU(2)$ gauge field coupled to fermions (as Rovelli does), we will opt to study a simpler theory, viz. scalar electrodynamics (SE). In this theory, the composite system $\mathcal{S}_{12}$ contains a $U(1)$ gauge field coupled to a scalar field, and its Lagrangian is

$$L_{12} = \frac{1}{2}(D_\mu\phi)^* D^\mu\phi - U(\phi^*\phi) - \frac{1}{2}F_{\mu\nu}F^{\mu\nu}. \tag{5}$$

In order to obtain an analogy with Rovelli's parable, we will have to start farther back than he does, viz. by asking the question: what are the analogs of the interacting subsystems $\mathcal{S}_1$ and $\mathcal{S}_2$ that appear in his parable? Only then will we be able to evaluate whether his morals (A), (B), and (C) carry over to SE.

We immediately notice two important differences. First, in the parable, the subsystems are of the same type – they have the same physical quantities. By contrast, in our example, $\mathcal{S}_1$ is a theory of $U(1)$ gauge field and $\mathcal{S}_2$ is a scalar field theory. Second, $\mathcal{S}_1$ an $\mathcal{S}_2$ can be defined independently in the parable. By contrast (as we will discuss in (B) below), a realistic scalar field theory with local symmetry cannot be defined independently of a gauge field.

Be that as it may, $\mathcal{S}_1$, a free $U(1)$ gauge theory, can indeed be defined independently – can we at least describe it along the lines suggested by (A)? The answer to this question turns in part on how one understands what 'relationism' is for fields. Recall that the classic relationist position – as in the case of (A) – holds that the true set of physical properties is nothing other than relations between material bodies; thus the absolute positions of materials bodies are 'gauge-dependent' and the relative distances are 'gauge-invariant'.

One way of setting up the analogy between $U(1)$ gauge theory and (A) is thus to say that the gauge field A is just a collection of assignments of properties to spacetime points $\{x \in M\}$, and that the gauge-dependent relata are the properties $\{A(x)\}$. The field strength values $F = dA$ at each point can then be described as local differences (i.e. relations) between these relata, just as in the case of (A). On the other hand, in order to obtain the full set of gauge-invariant observables, we will need to supplement these local relations with non-local relations (obtained by integrating the gauge field over loops), as we know from studying the Aharonov-Bohm effect or gauge theory on non-simply-connected spacetimes.

Nonetheless, it is arguable that a closer analogy to (A) is to think of the field itself as a unified and extended material substance – and thus take the field function (and not the values that it assigns to spacetime points) to play the role of a relata. This can indeed be done if one appeals to the descent-theoretic formulation of gauge theory. We will not need the full details of the formalism; the important point is that in this picture, the kinematic structure of gauge theory is defined as a set of relations between gauge fields defined on patches of spacetime. More precisely, the gauge fields $\{A_\alpha : U_\alpha \to \mathbb{C}\}$ correspond to the elements of a 'good' cover $\{U_\alpha\}$ of spacetime – one thus individuates fields according to the patches of spacetime on which they are defined. Relations between the fields are then specified on each intersection $U_\alpha \cap U_\beta$, i.e. $A_\alpha - A_\beta = d \log g_{\alpha\beta}$, where $g_{\alpha\beta} : U_\alpha \cap U_\beta \to U(1)$ is a transition function. The totality of these relations between gauge fields on different patches of spacetime – in conjunction with a suitable 'gluing procedure' – suffices to capture the full gauge-invariant content of the theory, i.e. it defines a principal $U(1)$-bundle, which is often taken as a starting point for the formalism of $U(1)$ gauge theory. This conception of 'relations between gauge fields' brings under a single rubric both local and non-local gauge-invariant physical quantities.

Let us take stock: while it is indeed possible to develop an analog of (A) from the gauge theory perspective, several new features also appear. First, there arises an interpretive question about how exactly to understand the notion of relata (especially in connection with material bodies) whose relations we are to interpret as gauge-invariant observables. And second, these relations must include non-local quantities, which are intimately related to the global topological features of spacetime.

When we proceed to contemplate the analogy with (B), we can no longer ignore the question of how to define $\mathcal{S}_2$. Unlike the case of the parable, it does not make sense to try to count the (gauge-invariant) degrees of freedom of $\mathcal{S}_1$ and $\mathcal{S}_2$ separately, and then show that their sum is less than the degrees of freedom of the composite system $\mathcal{S}_{12}$. This is because the geometry of a local scalar field theory is standardly constructed as a vector bundle that is associated to the principal bundle of a gauge theory – in other words, $\mathcal{S}_2$ cannot be defined independently of $\mathcal{S}_1$.

(A caveat must here be added to the claim that one cannot define a scalar field theory with local symmetry in a manner that is independent of gauge fields. First, this applies to *realistic* theories – one can of course write down trivially invariant Lagrangians with local symmetry such as $L = \partial_\mu(\phi^*\phi)\partial^\mu(\phi^*\phi) - U(\phi^*\phi)$.[3] Second and more importantly, physics generally requires that we have a way to compare the internal degrees of freedom at different spacetime points, and the mathematical device that allows us to do this, viz. a connection on a bundle, is precisely what gives rise to local gauge fields.)

Nonetheless, one can still try to obtain an analogy with (B) by adopting a slightly different strategy: indeed, doing so provides an explanation for why it is reasonable to move from a field theory with global symmetry to one with local symmetry (and is thus a partial response to the complaint at the end of Section B of Rovelli (2013), viz. that it is hard to understand the usual story about why global symmetries should be 'gauged').

The idea behind this strategy is to start with a scalar field theory $\mathcal{S}'_2$that can be constructed independently, viz. a scalar field ϕ with a *global* $U(1)$ symmetry, and ask how it might be possible to couple this theory to the free gauge theory $\mathcal{S}_1$.

In order to make progress with this story, we will now need to pursue the analogy with (A) in the case of $\mathcal{S}'_2$. It is in fact not hard to find. The gauge-dependent variables here are the phase values of the scalar field at each point in spacetime, and the gauge transformations are given by the action of the global $U(1)$ symmetry on these quantities. Gauge-invariant relational observables can then be obtained in situations where two different regions of spacetime have different phase values respectively – they are the relative phases between the fields on these different regions. For instance, one can set up up a relative phase difference between the field ϕ on some compact region R (with boundary ∂R) of the spacetime M, and the field ϕ on the complement $M - R$ (where we require a compatibility condition such as $\phi|_{\partial R} = 0$ to hold on the boundary separating these regions).[4] The analogy with (A) is thus restored for $\mathcal{S}'_2$.

We can now ask ourselves what it would take in order to couple the gauge-dependent variables of $\mathcal{S}_1$ with those of $\mathcal{S}'_2$. If we take Rovelli's tack, i.e. introduce couplings by means of possible invariant terms in a Lagrangian, we see that terms like $(D_\mu\phi)^* D^\mu\phi$ need to be invariant. But this can only be the case if the $U(1)$ symmetry of ϕ is modified to become a local symmetry – it is the need for coupling that explains this passage. Nonetheless, the Lagrangian and its terms are local objects – they are defined point-wise – and do not contain information about the global geometry of the local scalar field theory. It is for this reason that the story that we have told is not quite correct: the full picture can only be obtained by starting with the gauge theory $\mathcal{S}_1$ and using its geometry to construct that of a local scalar field.

We can thus see the partial analogy with (B) and its limitations. We cannot independently construct one of the subsystems of interest, but only a closely-related version – thinking about how to couple the two systems and making the necessary modification (to local symmetry) in order to do so then allows us to construct gauge-invariant terms in the Lagrangian of the composite system. (Couple) is still vindicated, but in a rather roundabout way.

Finally, let us consider the analogy with (C). In Rovelli's parable, the gauge-dependent quantities

[3] I thank Carlo Rovelli for this example.

[4] This can be done in a manner akin to Galileo's Ship, see Teh (2013).

of $\mathcal{S}_2$ (the analog of the scalar field theory) were used as a reference point in order to measure the gauge-dependent quantities of $\mathcal{S}_1$ (the analog of the gauge field theory). By contrast, in a Yang-Mills-type theory that is coupled to matter (scalar fields or fermions), the natural way of thinking about measurement – as Rovelli himself notes – is by means of a standard matter field ϕ' (representing the measuring device) in relation to the matter field ϕ that one wishes to measure; the relation is specified by the gauge field A. This is indeed a rather different picture from (C), although it is of course true that the gauge-invariant quantities – Wilson loops or lines – combine gauge-dependent quantities such as A and the scalar field ϕ.

4 Discussion

As we have just seen, the parable's resemblance to Yang-Mills-type theories (and also General Relativity) has its limits, but Rovelli's main insight, i.e. (Couple), still stands.

Why is a gauge-dependent quantity like the gauge field A so good at capturing information about how it couples to other fields? One easy answer is that, from the Lagrangian perspective, coupling terms are often introduced by contracting gauge-dependent quantities in order to form an invariant, in particular a gauge-invariant. Furthermore, such couplings are local, whereas the gauge-invariant quantities of a Yang-Mills-type theory are typically non-local. Nonetheless, this answer is not complete. In order to prove a strong version of Rovelli's point, one would have to give a general argument to show that couplings can *only* be introduced by combining gauge-dependent variables (or at least, that this was the case for some class of theories). We leave the consideration of this question to future work.

One might naively think that (Couple) could be shown to be superfluous by directly coupling matter fields to gauge-invariant quantities such as Wilson loops or lines. However, this sort of coupling is typically accomplished by adding an additional scalar field term to the integral of the gauge field A in the Wilson loop (see e.g. Giombi (2009)), so one still ends up invoking gauge-dependent quantities. Perhaps a true advance in this direction can only be achieved, as (Zee, 2010, p. 456,457) suggests, by finding a manifestly gauge-invariant way of formulating the theory and its observables, i.e. one that does not involve integrating A over a curve.

One possibility for pursuing the line of thought suggested by Zee is by appealing to dualities such as AdS/CFT, according to which gauge symmetry of one theory is 'completely invisible' from the perspective of its dual theory. As explained by Horowitz and Polchinski (2009), an illuminating way of understanding this idea is by analogy with certain condensed matter systems with a field variable $e(x)$ that separates into $e(x) = b(x)f^\dagger(x)$ in certain phases, where $b(x)$ and $f(x)$ are gauge-dependent fields which transform under the action of $\exp(i\lambda)$, and $e(x)$ remains gauge-invariant under this combination of transformations. Nonetheless, this example of a mapping from $b(x), f(x)$ in one theory to $e(x)$ in the dual theory does not show that one can dispense with gauge-dependent quantities for defining interactions, for one does not see such interactions in the dual theory. A much more thorough investigation is needed into the relevance of dualities for our understanding of (Couple). For now, we only wish to remark that the standard 'dictionary' for AdS/CFT, i.e. the GKPW formula (see e.g. Witten (1998)), is constructed precisely by coupling the gauge-dependent quantities of one theory to those of its dual. Thus it may well turn out that gauge-dependent quantities are not just important for encoding information about how theories couple, but indeed for understanding how two *prima facie* different theories can be mapped onto each other.

Acknowledgements The author thanks Carlo Rovelli and Jeremy Butterfield for helpful discussions about this article.

References

Giombi, S. (2009). Supersymmetric wilson loops from weak to strong coupling. *Notes from Simons Workshop*, insti.physics.sunysb.edu/conf/simonswork7/talks/Giombi.pdf?

Greaves, H. and Wallace, D. (2013). Empirical consequences of symmetries. *British Journal for the Philosophy of Science*.

Horowitz and Polchinski (2009). Gauge/gravity duality. In Oriti, D., editor, *Approaches to Quantum Gravity: Toward a New Understanding of Space, Time and Matter*. CUP.

Rovelli, C. (2013). Why gauge?

Teh, N. (2013). Galileo's gauge. *In Submission*.

Witten, E. (1998). Anti de sitter space and holography. *Adv. Theor. Math. Phys.*, 2:253.

Zee, T. (2010). *QFT in a Nutshell*. PUP.