On Gauge, Symmetry and Duality

S. De Haro and J. Butterfield
sd696@cam.ac.uk ; jb56@cam.ac.uk

University of Amsterdam, NL; and Trinity College, Cambridge UK
1707.06681, 1806.01505, 1801.09095, 1707.06581

100 Years of Gauge Theory, Bad Honnef, 31 July 2018

Introduction

The idea of gauge as descriptive redundancy prompts the question:
Do a state and its symmetry-transform represent the same physical state of affairs?

Certainly not always, on usual definitions of 'symmetry'.
For example: as a unitary commuting with the Hamiltonian. We should not say any two states in the same energy eigenspace represent the same physical state of affairs!

So: *Under what conditions is the answer 'Yes'?*

Counting possibilities is a favourite topic for philosophers, since e.g. the Leibniz-Clarke debate.

We approach this question using our account of duality ...

A duality is 'a symmetry of a whole theory'. Indeed: dual theories are, in general, not equivalent.
We propose conditions for equivalence to hold: in short, that the dual theories are internally interpreted and are unextendable.

Outline

Symmetry introduced:
We write $\langle Q, s\rangle$ for the value of the quantity Q in state s: for a classical state (point in phase space), a possessed value; for a quantum state, an expectation value (and similarly for matrix elements $\langle s|Q|s'\rangle$).

It is usual to take a symmetry as a map on states. So a symmetry a (we write a for 'automorphism') carries a state s in a state space $\mathcal{S}$ to another state $a(s)$, where s and $a(s)$ assign the same values to all the quantities Q in some salient, usually large, set of quantities $\mathcal{Q}$: $\langle Q, a(s)\rangle = \langle Q, s\rangle$.

Or dually, with the symmetry as a map on quantities, that preserves the value on a given state. Recall that given any map $a : \mathcal{S} \to \mathcal{S}$, we define its *dual map* on quantities, $a^* : \mathcal{Q} \to \mathcal{Q}$, by requiring:

for $s \in \mathcal{S}$ and $Q \in \mathcal{Q}$: $\langle a^*(Q), s\rangle := \langle Q, a(s)\rangle$.

Thus we can similarly define a symmetry as a map on quantities that preserves values.

Some cases involve, not equality of values, i.e. *invariance*, but some suitable *covariance*. Typically, there are two naturally related maps on states and quantities, say $a : \mathcal{S} \to \mathcal{S}$ and $\alpha : \mathcal{Q} \to \mathcal{Q}$, with:

$$\langle \alpha(Q), a(s) \rangle = \langle Q, s \rangle.$$

Of course, there is usually a group G of symmetries.
So our opening question becomes:
Under what conditions are a state and its transform, s and $a(s)$, with $a \in G$, 'gauge-equivalent'?

The recent philosophical literature (Belot, Caulton, Dewar, Read and Møller-Nielsen) advises caution ...

Duality introduced:
We take a duality between two theories to be a matter of:

(a): the two theories sharing a common core; (itself a theory, the *bare theory*); and

(b): the two given theories being isomorphic models of this common core: here, 'model' means a homomorphic copy (i.e. representation in the sense of representation theory).

We call the two dual i.e. isomorphic theories, *model triples*, the 'triple' referring to the fact that the theory consists of three items: a state-space, a set of quantities, and a dynamics: $\langle \mathcal{S}, \mathcal{Q}, \mathcal{D} \rangle$.

So: the bare theory can be realized (we say: modelled) in various ways: like the different representations of an abstract group or algebra. These models are in general *not* isomorphic, and they differ from one another in their *specific structure*: like inequivalent representations of a group. But we say:

When the models are isomorphic, there is a duality.

Our usage:
Beware: the word ‘model’, as contrasted with ‘theory’, often connotes:
(i): a specific solution for the physical system concerned, whereas the ‘theory’ encompasses all solutions—and in many cases, for a whole class of systems;
(ii): an approximation whereas the ‘theory’ deals with exact solutions;
(iii): being part of the physical world (especially: being empirical, and-or observable) that gives the interpretation, whereas the ‘theory’ is not part of the world, and so stands in need of interpretation.

Our use of ‘model’ rejects all three connotations.

Duality as surprising:
We usually discover a duality in the context of studying, not a bare theory, but rather: two interpreted models of such a theory.

Usually, we do not initially believe them to be isomorphic in any relevant sense. Or even: to be models of any single relevant theory (even of a bare one).

The surprise is to discover that they are such models—indeed are isomorphic ones. And the surprise is greater, the more detailed is the common structure (like '10-dimensional semisimple Lie group', as against 'group').

Duality defined:

The idea: a duality is an isomorphism between two *model triples*. The model M has *specific structure* $\bar{M}$; but the model triple is separated from $\bar{M}$, and expresses only the model's representing the bare theory. So $M = \langle \mathcal{S}_M, \mathcal{Q}_M, \mathcal{D}_M, \bar{M} \rangle =: \langle m, \bar{M} \rangle$, where $m := \langle \mathcal{S}_M, \mathcal{Q}_M, \mathcal{D}_M \rangle$ is the model triple (or *model root*).

A **duality** between $m_1 = \langle \mathcal{S}_{M_1}, \mathcal{Q}_{M_1}, \mathcal{D}_{M_1} \rangle$ and $m_2 = \langle \mathcal{S}_{M_2}, \mathcal{Q}_{M_2}, \mathcal{D}_{M_2} \rangle$ requires:
an isomorphism between Hilbert spaces (for classical theories: manifolds):

$$d_s : \mathcal{S}_{M_1} \to \mathcal{S}_{M_2} \ \text{ using } d \text{ for 'duality' ;} \tag{1}$$

and an isomorphism between the sets (almost always: algebras) of quantities

$$d_q : \mathcal{Q}_{M_1} \to \mathcal{Q}_{M_2} \ \text{ using } d \text{ for 'duality' ;} \tag{2}$$

such that:

(i) the values of quantities match:

$$\langle Q_1, s_1 \rangle_1 = \langle d_q(Q_1), d_s(s_1) \rangle_2 \ , \ \ \forall Q_1 \in \mathcal{Q}_{M_1}, s_1 \in \mathcal{S}_{M_1} . \tag{3}$$

(Similarly for matrix elements such as $\langle s|Q|s'\rangle$, for quantum theories).

(ii) d_s is equivariant for the two triples' dynamics, $D_{S:1}, D_{S:2}$, in the Schrödinger picture; and d_q is equivariant for the two triples' dynamics, $D_{H:1}, D_{H:2}$, in the Heisenberg picture: see Figure 1.

$$\begin{tikzcd}
\mathcal{S}_{M_1} \arrow[r, "d_s"] \arrow[d, "D_{S:1}"] & \mathcal{S}_{M_2} \arrow[d, "D_{S:2}"] \\
\mathcal{S}_{M_1} \arrow[r, "d_s"] & \mathcal{S}_{M_2}
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathcal{Q}_{M_1} \arrow[r, "d_q"] \arrow[d, "D_{H:1}"] & \mathcal{Q}_{M_2} \arrow[d, "D_{H:2}"] \\
\mathcal{Q}_{M_1} \arrow[r, "d_q"] & \mathcal{Q}_{M_2}
\end{tikzcd}$$

Figure : Equivariance of duality and dynamics, for states and quantities.

Comparing dualities and symmetries:

Broadly speaking, a duality reduces to a symmetry, in the case of 'self-duality': i.e. where there is just one model, and one model triple, at issue, i.e. $M_1 = M_2$ and $m_1 = m_2$.

Thus consider a symmetry as a map, $s_1 \mapsto a(s_1)$, on states $s_1 \in \mathcal{S}_{M_1}$, that preserves values of quantities Q_1: $\langle Q_1, a(s_1)\rangle_1 = \langle Q_1, s_1\rangle_1$.
Defining the duality map d_s on states to be a, and the duality map d_q on quantities to be the identity map $Q_1 \mapsto Q_1$, yields a duality, in the sense defined. For the requirement that a duality match values, eq. 3, reduces to the symmetry requirement that values are equal.

Similarly if we consider a symmetry as a map, $Q_1 \mapsto \alpha(Q_1)$, on quantities $Q_1 \in \mathcal{Q}_{M_1}$, that preserves values on states s_1: $\langle \alpha(Q_1), s_1\rangle_1 = \langle Q_1, s_1\rangle_1$.

A duality preserves any symmetry of its model triples.
(1): *There is a commuting square diagram of isomorphisms.*
The duality maps $d_s : \mathcal{S}_{M_1} \to \mathcal{S}_{M_2}$, and $d_q : \mathcal{Q}_{M_1} \to \mathcal{Q}_{M_2}$ are isomorphisms; and 'is isomorphic to' is symmetric and transitive. So the diagram, with a any automorphism of $\mathcal{S}_{M_1}$, commutes:

$$\begin{array}{ccc} \mathcal{S}_{M_1} & \xrightarrow{a} & \mathcal{S}_{M_1} \\ \downarrow d_s & & \downarrow d_s \\ \mathcal{S}_{M_2} & \longrightarrow & \mathcal{S}_{M_2} \end{array}$$

Figure : Commutativity of duality and symmetry for states.

Similarly for quantities, as against states. Since d_q is required to be an isomorphism of quantities, this diagram, with α any automorphism of $\mathcal{Q}_{M_1}$, must commute:

$$\begin{tikzcd} \mathcal{Q}_{M_1} \arrow[r, "\alpha"] \arrow[d, "d_q"] & \mathcal{Q}_{M_1} \arrow[d, "d_q"] \\ \mathcal{Q}_{M_2} \arrow[r] & \mathcal{Q}_{M_2} \end{tikzcd}$$

Figure : Commutativity of duality and symmetry for quantities.

These two diagrams are just what we mean by saying a duality d preserves an automorphism of the states/quantities in its domain model triple, and preserves the structure of $\mathcal{S}_{M_1}$ / $\mathcal{Q}_{M_1}$.

(2): *The value of a quantity on a given state equals the value of the dual-quantity on the dual-state.*
A symmetry involves more than an automorphism of the state-space, and of the set of quantities. The values of quantities in states, $\langle Q, s \rangle$, must (for a large and salient set of quantities, though usually not *all* quantities) be preserved under the symmetry.

For duality, the corresponding condition is (i) in our definition of duality: that values are equal between states and quantities that correspond by the duality. Recall Eq. (3):

$$\langle Q_1, s_1 \rangle_1 = \langle d_q(Q_1), d_s(s_1) \rangle_2 \,, \quad \forall Q_1 \in \mathcal{Q}_{M_1}, s_1 \in \mathcal{S}_{M_1} . \tag{4}$$

(Similarly for matrix elements such as $\langle s|Q|s' \rangle$, for quantum theories).

(3): *The same verdict—that a duality preserves any symmetry of its model triples—applies to dynamical symmetries.*
We need to compose, for each side of the duality, the equivariance condition that states a dynamical symmetry, with the duality map.
So with the Schrödinger picture of dynamics, and D_{t,t_0} the dynamics on $\mathcal{S}_{M_1}$: the dynamics on $\mathcal{S}_{M_2}$ is given by (reading down): $d_s \circ D_{t,t_0} \circ d_s^{-1}$:

$$
\begin{tikzcd}
\mathcal{S}_{M_2} \arrow[r] \arrow[d, "d_s^{-1}"] & \mathcal{S}_{M_2} \arrow[d, "d_s^{-1}"] \\
\mathcal{S}_{M_1} \arrow[r, "a"] \arrow[d, "D_{t,t_0}"] & \mathcal{S}_{M_1} \arrow[d, "D_{t,t_0}"] \\
\mathcal{S}_{M_1} \arrow[r, "a"] \arrow[d, "d_s"] & \mathcal{S}_{M_1} \arrow[d, "d_s"] \\
\mathcal{S}_{M_2} \arrow[r] & \mathcal{S}_{M_2}
\end{tikzcd}
$$

Figure : Commutativity of duality, symmetry, and dynamics.

Duals are, in general, physically inequivalent:
Duality is a *formal* relation. So once duals are *interpreted*, they can disagree, i.e. make different claims about the world—despite the formal equality of values.

They can disagree: either by

contradicting each other about a single subject-matter: (Contr);

or by

making assertions about different subject-matters: (Diff).

We call such disagreements *physical inequivalence*.

Duals can also *agree*—be *physically equivalent*—while yet the duality is surprising.
Examples: (1) Position-momentum duality in quantum mechanics gives surprisingly different descriptions of the same system/degrees of freedom—but without disagreement; (2) Bosonization.

We introduce the ideas of interpretation and subject-matters. Then we give some examples of duals that disagree, and duals that agree, in classical and quantum physics. Then we give, more tentatively, examples in string theory.

Interpreting physical theories:
We endorse the framework of *intensional semantics*, in the style of Frege, Carnap and Lewis. Words and sentences are assigned intensions: maps from the set of worlds W to extensions.

This framework has the great merit of respecting the meanings of words! That may seem an obviously mandatory feature for any endeavour calling itself ‘semantics’. But the ‘semantics’ in books of logic and model theory investigate the mathematical consequences of assigning arbitrary meanings (specifically, extensions) to words...

It also makes precise the notion of a *subject-matter*, as a partition of the set W of worlds. In a cell of the partition, any two worlds match as regards the subject-matter. Thus a proposition is *entirely about* a subject-matter if the set of worlds at which it is true is a union of cells of the subject-matter.

Subject-matters: (Contr) and (Diff):

We now make precise how dual theories can disagree: either by
(Contr): contradicting each other about a subject-matter; or by
(Diff): describing different (though 'isomorphic') subject-matters.

(Contr): Each of two dual model triples is interpreted as wholly true (its conjunctive proposition is wholly true) at a union of cells of a common subject-matter. But these two unions are disjoint: for the propositions contradict each other.

(Diff): Two dual model triples are interpreted as wholly true (the conjunctive proposition of each is wholly true) at distinct sets of worlds. Each set is a union of cells of the triple's subject-matter, i.e. partition. But the partitions are different, and so are the sets. The sets need not be disjoint: both the model triples could be, both of them, wholly true. But the sets are distinct.

Disagreeing duals in classical physics:

(1): Newtonian mechanics with different absolute rests: (Contr)

Two formulations of Newtonian point-particle mechanics (say N particles with gravitation), that disagree in what they identify as absolute space: what inertial timelike congruence is 'truly at rest'.
Thus the spacetime is $\mathbb{R}^4$ and the bare theory is a neo-Newtonian (Galilean) formulation of point-particle mechanics.

Nowadays, we usually take the contrary specifications of absolute rest as 'gauge-fixings', 'a distinction without a difference', 'a sign that we should move to a neo-Newtonian formulation', in which the 'surplus structure of absolute rest is eliminated'. In terms of our Schema: the inertial congruence is part of the specific structure, which is not "mapped across".

But this is with the benefit of hindsight: of our now knowing the neo-Newtonian formulation. Returning to 1700: the views of Newton and Clarke are tenable—and then the duality illustrates (Contr). In terms of our Schema: the inertial congruence is "mapped across"—but interpreted differently.

(2): Kramers-Wannier duality in classical statistical mechanics: (Diff)

The bare theory is the equilibrium theory of a two-dimensional square lattice with the Ising Hamiltonian: i.e. the canonical ensemble with probabilistic weights for a configuration s given by $\exp(-\beta H[s])$, where H is the Ising Hamiltonian and $\beta \equiv 1/kT$ is the inverse temperature.

The duals are approximations to the partition function $Z \equiv \Sigma_s \exp(-\beta H[s])$ that are valid at low and high temperatures T, respectively: say, low on the 'left' and high on the 'right'.

We write the partition function using the dimensionless inverse temperature $\nu := J/kT$, and we define ν^* by $\tanh \nu^* := \exp(-2\nu)$. So $\nu^* = 0/\infty$ iff $\nu = \infty/0$.

The left dual is the expansions parameterized by T being in some low range $[T_1, T_2] \subset \mathbb{R}$, i.e. by large $\nu := J/kT$ in the range $[J/kT_2, J/kT_1]$. The right dual is the expansions parameterized by small ν^* in the range $[\tanh^{-1}(\exp(-2J/kT_1)), \tanh^{-1}(\exp(-2J/kT_2))]$.

The duality map d_s is

$$d_s : Z(\nu) \mapsto Z(\nu^*) := \frac{Z(\nu)}{2^{1-N}(2\sinh 2\nu)^N} \ . \qquad (5)$$

This is a case of (Diff). The low temperature regime and high temperature regime are different though isomorphic subject-matters.

This duality has various generalizations and modifications: lattices with different geometries, different Hamiltonians and couplings to external sources etc. Cf. the later discussion of unextendability.

Agreeing duals in quantum physics:

(3): Bosonization: the free, massless case
The duality, in two dimensions, between:

(1): the free, massless bosonic scalar field, ϕ; and
(2): the free, massless Dirac fermion, ψ.

There is an isomorphism between the quantised fields:

$$\begin{array}{rcl} \partial\phi(z) & \leftrightarrow & :\psi^\dagger(z)\,\psi(z): \\ :e^{i\,\phi(z)}: & \leftrightarrow & \psi(z)\,. \end{array} \tag{6}$$

This isomorphism is ensured by the two models' sharing the same enveloping Virasoro algebra (the Virasoro algebra with central charge $c = 1$, coupled to an abelian affine algebra at level $k = 1$).

The unitary, irreducible representations of this algebra are unique up to unitary equivalence. So, in this case, duality is a consequence of the *unitary equivalence of the two representations (models).*

Disagreeing duals in string theory:

Tentatively! ... Gauge/gravity duality, and T-duality, provide examples of (Contr): we discuss the first ...

(4): Gauge-gravity duality: (Contr)
'Gauge/gravity duality' is the umbrella term for dualities between a string theory (hence including a description of gravity) on a $(d+1)$-dimensional spacetime (the 'bulk') and a quantum field theory (a gauge theory, with no description of gravity) on a d-dimensional space or spacetime that forms the bulk's boundary.

The common core, i.e. bare theory, of which the bulk and boundary theories are models (in our representation-theory sense) has as its spacetime: the d-dimensional boundary manifold equipped—not with a metric, but merely—with an equivalence class of them, under local conformal transformations. Details below.

Suppose the bulk theory (the 'left dual') says spacetime is five-dimensional ($d + 1 = 5$); so the boundary theory, the right dual, says it is four-dimensional.

But both theories are putative 'theories of everything', 'toy cosmologies'. They are both about a single topic, the cosmos; in philosophers' jargon, the actual world. So the theories make contrary assertions about that single topic, the actual world, namely about the dimension of its spacetime. So this is a case of (Contr).

In terms of the simple logic—or rhetoric!—of the situation, we have come full circle, back to Example (1).

We agree that there is a temptation to say: 'the real truth lies in what is in common, or what is behind, the two duals'. That is: either

(i) to formulate the duals' common core/bare theory (if we have no formulation or a defective one), and-or

(ii) to formulate another theory 'behind the duals', of which they are approximations, not representations.

Recall the analogous temptation for Example (1): either

(i) to move to Galilean (neoNewtonian) spacetime, or

(ii) to move to geometrized gravity, such as in general relativity.

Of course, this temptation is scientifically, heuristically, invaluable—and rightly stressed by physicists. (We call (ii) the 'heuristic function of dualities', i.e. helping us find "new physics" beyond the common core.)

But setting aside the future development of our theories, to concentrate on interpreting them *as now formulated*:— here is a case of (Contr).

Kinds of interpretation: external and internal

The previous discussion relied on defining relevant subject-matters before the duality is interpreted (e.g. by assuming possible worlds independent of theory).

For example, in the example of Newtonian mechanics, Newton and Clarke assumed that there is a subject matter 'absolute space', and the two duals disagreed about the correct standard for it—thus a case of (Contr).

This is what we call an **external interpretation:** i.e. one that also interprets the specific structure which we use to build the models, but which does not 'take part' in the isomorphism. Namely, an absolute specification of rest is not left invariant by a Galilean transformation.

The internal interpretation

But what if we are "Leibnizians", and we do not assume we can always define the subject-matter beforehand? Then we would be cautious about interpreting the specific structure that we use to build our models, since it is not mapped across by the duality. That is: we would try to use the theory, and the duality between the models, to define the subject-matter.

In the example of Newtonian mechanics: Rather than including a standard of 'absolute rest' in our ontology, we would look for frame-independent states and quantities.

We thus come to the idea of an:

Internal interpretation: an interpretation that maps all of and only the model root, regardless of the specific structure of the model. So in the case of dualities, the internal interpretation interprets only the bare theory.

The internal interpretation

Since the internal interpretation only interprets the common core that is isomorphic between the two models and not the specific structure, one way to obtain it is to start with the external interpretations of the two models, and see where they agree: "cross out" the interpretations of the specific structures, and interpret the model triples in the same way.

This is what we do if we are Leibnizians about Newtonian mechanics: we have concepts in our theory that Clarke also has, but we simply drop the interpretation of the notions associated with 'absolute rest'.

This does not mean that the specific structure has no place in interpretation. It can correspond to some facts in the world, once a *stipulation* is made. In this sense, external and internal interpretations can complement each other.

Gauge-gravity duality

The schema can be illustrated in **gauge-gravity dualities**: they relate $(d+1)$-dimensional string theories (models) to d-dimensional quantum field theories (QFT models).

We **do not have an exact (or non-perturbative) definition** of the models, or of the duality. Having a rigorous definition of it would almost be like 'proving' the duality.

But we can get important insights about the theory, and thus illustrate the schema, by considering the **semi-classical limit**
(technically: strong 't Hooft coupling).

Suppose two such models are dual. What is their **common core**?

Gauge-gravity duality (pure gravity: no matter)

The gravity theory is defined under two **boundary conditions** for the metric and the stress-energy tensor, defined at spacelike infinity:

(i) A boundary condition for the **metric** at infinity, which is defined up to conformal transformations.

We have a d-dimensional conformal manifold, $\mathcal{M}$, at infinity, with a conformal class of metrics, $[g]$. This is identified with the manifold on which the CFT, with its conformal class of metrics, is defined.

Thus the pair $(\mathcal{M}, [g])$ is part of the **common core.**

Gauge-gravity duality

The asymptotic symmetry algebra associated with the gravity model is the d-dimensional **conformal algebra** (for a flat boundary, $\mathrm{o}(2, d-1)$), and the representations of this algebra form the set of admissible states belonging to the Hilbert space of admissible states, $|s\rangle_{\mathcal{M},[g]} \in \mathcal{H}$.

(ii) A boundary condition is also required for the asymptotic value of the **canonical momentum**, Π_g, conjugate to the metric on the boundary, evaluated on all the states, $\langle s|\Pi_g|s\rangle$. This choice further restricts to a subspace of the previous Hilbert space $\mathcal{H}$: it determines a subset of states of the conformal algebra.

The simplest (and usual) choice, $\langle s|\Pi_g|s\rangle = 0$, preserves the full conformal symmetry.

Gauge-gravity duality

In the case of interest (pure matter on the gravity side), the only operator turned on in the CFT is the **stress-energy tensor**, T_{ij}.

The duality dictionary tells us that: $\Pi_g \leftrightarrow T_{ij}$.

Thus, the two models share the d-dimensional conformal manifold $\mathcal{M}$ with its conformal class of metrics $[g]$ and the operators Π_g, T_{ij} and their algebra.

These determine the values of the infinite set of **correlation functions**:

$$_{\mathcal{M},[g]}\langle s|\ T_{i_1j_1}(x_1)\cdots T_{i_nj_n}(x_n)\ |s\rangle_{\mathcal{M},[g]}\ , \tag{7}$$

which agree between the two models, with the approximations made (and for the appropriate subset of states).

This discussion can be generalised to include other states and operators.

Physical (In-)Equivalence: Two Interpretative Cases

Since the conception of duality is formal, it allows the idea of '**distinct but isomorphic sectors of reality**'—namely as the ranges of the interpretation maps i on the two sides of the duality. On an external interpretation, the models are **physically inequivalent**.

But this of course does not forbid the other sort of case: where the two models are **physically equivalent**, i.e. describe '**the same sector of reality**'. This is modelled by the internal interpretation maps i having the same images: e.g. for quantities in the world, written as $\mathcal{Q}_{\text{world}}$:

$$\begin{tikzcd} \mathcal{Q}_{M_1} \arrow[rr, "d_q"] \arrow[dr, "i_1"'] & & \mathcal{Q}_{M_2} \arrow[dl, "i_2"] \\ & \mathcal{Q}_{\text{world}} & \end{tikzcd}$$

Figure : The two sides of the duality describe 'the same sector of reality'.

When does duality lead to physical equivalence?

Thus we get the following **formulation for physical equivalence**:

$$i^1_{\text{int}} = i^2_{\text{int}} \circ d \ . \tag{8}$$

We will say that '**interpretation commutes with duality**'
(in the sense that the three maps form a commuting diagram).

This formalises what we mean by the phrase 'physical equivalence', as sameness of interpretation of two models.

What are sufficient conditions for duality to lead to physical equivalence under all suitable interpretations? Roughly, we need two conditions:

(i) The internal interpretation can be adopted because each dual is exact and describes the whole world.
(ii) When we say that the duals have the "same domains" of application, we need an agreed philosophical conception of what is a domain of application.

Two conditions for duality to lead to physical equivalence

(i) **Unextendability:** Each dual 'is about the whole world', i.e. cannot be extended beyond its domain of application.
Symmetry principles often help here, in that they ensure that we have the most general theory with the given symmetries and field content, which cannot be further extended without breaking that symmetry.

(There is a possible weakening of this condition: that the theory can be extended beyond its domain, but in doing so the interpretation does not change, i.e. it is "stable" or "robust" against such extensions.)

(ii) **Defining domains:** While realists and empiricists can construct the same interpretations of a theory ('the picture of the world drawn by the theory', van Fraassen (1980)), they have different degrees of belief in the entities that it postulates. If we want to claim that two models, with their interpretations, are physically equivalent, then we have to make these ontological commitments explicit.

The need for unextendable theories

Extendable theories provide examples of Read and Møller-Nielsen's (2018) observation that an explication of the shared ontology of two models need not exist. If our model is extendable through its having some inconsistencies (e.g. the force between two masses in Newtonian mechanics diverges when the two masses are on top of each other) then the model is mathematically inconsistent, and we should expect its ontology to be of limited value. We cannot tell whether duality leads to physical equivalence under all suitable interpretations.

Thus, to fully discuss physical equivalence (through real examples, rather than toy models), it is important that we deal with mathematically well-defined theories—unextendable ones.

If the theory is mathematically inconsistent, then there is no possible world associated with it!

The challenge posed by dualities of unextendable theories

This gives an interesting distinction in the types of dualities:

(A) For **extendable theories**, the question of physical equivalence is limited by the fact that possible extensions of the theory may lead to radical modifications of the interpretation. For example: Newtonian mechanics, Kramers-Wannier duality.

(B) For **unextendable theories**, there is an interesting question about whether the models are physically equivalent under all suitable interpretations. For example: some QFTs (including bosonization), possibly quantum gravity dualities.

Summary

1. We have illustrated two ways in which dual pairs can disagree: either by contradicting each other about a single subject-matter, or by making assertions about different subject-matters.

2. So just as we should be cautious about physically identifying a state s and its symmetry-transform $a(s)$, so also about dual theories.

3. There are internal interpretations, which take the duality as the starting point to interpret the theory.

4. On an internal interpretation, we can take dual pairs to be physically equivalent, provided the models are unextendable, and one has an agreed philosophical conception of the domain of application.

Vielen Dank!

References

Butterfield, J. (2017). "On Dualities and Equivalences Between Physical Theories", in: *Philosophy Beyond Spacetime*, B. Le Bihan, et al. (Eds.) (OUP), to appear: 1806.01505.

De Haro, S. and Butterfield, J. (2017). "A Schema for Duality, Illustrated by Bosonization", in: *Foundations of Mathematics and Physics One Century After Hilbert*, Kouneiher, J. (Ed.), Springer. 1707.06681.

De Haro, S. (2016). "Spacetime and Physical Equivalence". In *Space and Time after Quantum Gravity*, Huggett, N., Wüthrich, C. (Eds.) (CUP), to appear. http://philsci-archive.pitt.edu/13243/. 1707.06581

De Haro, S. (2018). "The heuristic function of dualities". *Synthese*, https://doi.org/10.1007/s11229- 018-1708-9. 1801.09095

Dieks, D., van Dongen, J., and de Haro, S. (2015). "Emergence in holographic scenarios for gravity". *Studies in History and Philosophy of Modern Physics*, **52**, 203-216.

Notation for theories and models

A notation for a model M that exhibits how M augments the structure of the theory T with specific structure, $\bar{M}$ say, of its own:—

Do **not** write $M = \langle T, \bar{M} \rangle$, since M uses $\bar{M}$ to build a representation of T's structure. Better to write: $M = \langle T_M, \bar{M} \rangle$. So the subscript M on T reflects that the specific structure $\bar{M}$ is used to build the representation of T. We call T_M, the 'part' of M that represents T, the **model root**.

Thus for a theory as a triple, $T = \langle \mathcal{S}, \mathcal{Q}, \mathcal{D} \rangle$: we write a model as a quadruple:

$$M = \langle \mathcal{S}_M, \mathcal{Q}_M, \mathcal{D}_M, \bar{M} \rangle =: \langle m, \bar{M} \rangle \ , \tag{9}$$

where $m := T_M := \langle \mathcal{S}_M, \mathcal{Q}_M, \mathcal{D}_M \rangle$ is called the **model triple**, as well as *model root*.

(5): T-duality: (Contr)

'T-duality' is the umbrella term for two dualities between two pairs of string theories (as currently formulated). Both dualities involve inverting the radius of one of the compact ('curled up' like a circle) dimensions of space. Thus a type IIA theory postulating that a certain dimension of space has radius R is dual to a type IIB theory where the dimension is $1/R$.

Objection! If one theory, say a type IIA theory, postulates a radius R so small that it could not be empirically detected, $1/R$ may well be so large that it *could* be detected—if it was real.
Reply! Measuring the radius of a putative compact dimension—say by sending off a particle and timing how long it takes to return to you—can be naturally accommodated by *both* dual theories.
For what one dual describes as a journey through physical space, is described by the other dual as a journey through an internal space.

We interpret the duals as both being about a single topic: the cosmos, the actual world. They make contrary assertions about this topic. So they disagree: a case of (Contr).

Just like Example 4: except that disagreement over a spatial radius replaces disagreement over spatial dimensionality.

Again: we here set aside the heuristic function of dualities.

We agree that maybe one could treat the two string theories, not as theories of everything (TOEs, 'toy cosmologies'), but as both true in a single cosmos/possible world with, say a 10-dimensional space.

Namely: the type IIA describes one compact dimension as radius R, and the type IIB describes another compact dimension as radius $1/R$.

This turns the duals' disagreement into a case of (Diff), not (Contr).

Forgive us, O guru from Illinois...

Huggett (2017) takes the two duals in T-duality to *agree*. He writes:

> *[He concedes that it] would not be a logical fallacy, nor [contravene] unavoidable semantic or ontological principles, [to deny that] the duals describe the same physical possibility. [But ...] from a practical scientific point of view, it makes sense to treat those differences as non-physical . . . long established well-motivated scientific reasoning should lead us to think that dual total theories represent the same physical situation (2017: 86).*

He goes on to address the resulting question: how can we make sense of the 'appearance' that the dual theories contradict each other about the radius of space?

He distinguishes two answers, called 'interpretation one' (p. 84) and 'interpretation two' (p. 85), and argues in favour of the second ...