The Categories of Physical Systems and Theories

Erik Curiel

Munich Center For Mathematical Philosophy
Ludwig-Maximilians-Universität

Black Hole Initiative
Harvard University

Smithsonian Astrophysical Observatory
Radio and Geoastronomy Division

email: `erik@strangebeautiful.com`

Outline

It needs a lot of intuitive physical sense to know when to expect actual things to behave like the idealized models we make of them.

J. L. Synge
Relativity: The Special Theory
(ch. I, §18, p. 32)

I find myself dragged in contrary directions. On the one hand, I am chary of formal methods in philosophy of science, as they too often seem to be used as ends in themselves; and this seems to me true to an unhealthy degree in the study of the structure and semantics of scientific theories. On the other, I believe that, when used with caution and when supplemented by substantive contact with and constraint by the real empirical content of science, they can be useful, even fruitful.

I will attempt in this paper to submit to the second impulse. I will propose a way of representing a physical theory as a category, which is nothing new. I will also propose, however, a way of representing as a category the physical systems appropriately and adequately represented by a theory, and correlatively a way of representing as a category the family of measuring and observational practices used to bring the systems and the theory into substantive contact with each other.

Various constructions on these categories, such as functors between them, will then be used to capture the idea that the theory represents those systems with propriety and adequacy, when the values of their physical quantities are determined by members of the given family of measuring and observational practices. This will yield a natural necessary and sufficient condition for the equivalence of two theories in a physically substantive sense.

This all may seem to defeat the purpose of formalization, which is often conceived of as a complete abstraction from messy details; not all messy details can be abstracted from, however, on pain of formal vacuity; the goal is to orivde tools to help determine the minimum amount of messy detail one needs to include in order to endow one's formal machinery with empirical content, for the purposes of one's philosophical investigations.

In this case, the messy detail is unavoidable: the whole point is that two theories are equivalent if and only if they can appropriately and adequately represent the same family of physical systems, and positing two formalisms along with a few interpretational principles cannot suffice to show this.

Weatherall (2017, p. 330):

> *[M]y goal is to review some ways in which thinking of a physical theory as a category of models bears fruit for issues of antecedent philosophical interest. The role the category theory ends up playing is to regiment the discussion, providing the mathematical apparatus needed to make questions of theoretical structure and equivalence precise enough to settle.*

The whole issue, though, is how much detail one does need to go into on the representational and empirical side for the formalization to allow one to settle such issues as equivalence.

I do not like a formulation of the question "what is the theory" that admits as an appropriate answer a purely formal structure, along with perhaps the fixing of something like a Tarskian semantics or a sketch of representational capacities by way of, say, the laying down of a few interpretive principles matching parts of the formal structure to some particularly simple family of phenomena the theory purports to treat. A formulation of that kind assumes, with no argument ever given, that it is possible to make a clean separation between, on the one hand, one part of the scientific knowledge the theory embodies, viz., that encoded in the pure formalism and, on the other, the remainder of that knowledge.

The remainder of the theory's knowledge includes at a minimum what is encoded in the practice of modeling particular systems, of performing experiments, of bringing the results of theory and experiment into mutually fruitful contact—including all the *in*exact mathematical techniques that cannot be formalized, approximative and heuristic techniques motivated by loose physical arguments and principles not part of the formalism, and justified only by practical success—including, in a word, real application of the theory in actual scientific practice.

We should not assume such a clean separation is possible without an argument. Although I have many arguments against the position, many of which I think are strong and compelling, I do not claim to have a sockdolager from which there is no recovery (though I suspect that taken in sum my arguments effectively serve as one). For the purposes of this paper, nonetheless, I will make something very close to that assumption, to see how far it may get us.

This will be a programmatic talk, laying out a picture, with no pretense of arguing for its correctness or adequacy, and no attempt to prove any results of interest about the constructed system. The goal is only to show what such a thing may look like by arguing that it latches on to something important—something we should not so cavalierly ignore as is usual—in the philosophical study of theories, so as to elucidate, by way of comparison with the standard treatments, what those standard treatments lack.

example of my discontents: Newtonian gravity

a standard picture of standard Newtonian gravitational theory

1. a family of models

$$(M,\, t_a,\, h^{ab},\, \nabla_a,\, \phi,\, \rho)$$

2. M is a 4-dimensional manifold (usually $\mathbb{R}^4$), t_a and h_{ab} are spatial and temporal “metrics”, and ∇_a is a flat affine connection
3. ρ is matter density, ϕ is gravitational potential, and $\nabla^n \nabla_n \phi = 4\pi\rho$
4. test particles accelerate according to

$$\xi^n \nabla_n \xi^a = -\nabla^a \phi$$

where ξ^a is the particle’s 4-velocity

a standard picture of geometrized Newtonian gravity

1. a family of models

$$(M,\, t_a,\, h^{ab},\, \tilde{\nabla}_a,\, \rho)$$

2. M is a 4-dimensional manifold (usually $\mathbb{R}^4$), t_a and h_{ab} are spatial and temporal "metrics", and ∇_a is a (generally) curved affine connection
3. ρ is matter density, and $R_{ab} = 4\pi\rho t_a t_b$
4. test particles follow geodesics of $\tilde{\nabla}_a$

claim (Weatherall 2017): it is reasonable to assume they have the same representational capacities, and so are equivalent, because:

1. the interpretation of ρ as matter density
2. the relation between

   $$\nabla^n \nabla_n \phi = 4\pi\rho$$

   and

   $$R_{ab} = 4\pi\rho t_a t_b$$

3. the motions of test particles
4. AND each model in each theory has what manifestly appears to be an empirically equivalent one in the other, based on these interpretive principles

It is not at all clear that the criteria used to establish “empirical equivalence” of models in this case—what paths test-particles traverse, and the relationship of this to the distribution of matter—exhausts the empirical content of either theory so as to justify a claim of equivalence.

1. Gravitational force will not necessarily “couple” with radiative fields (*e.g.*, Maxwell) in the same way as curvature will, especially as regards approximations, *e.g.*, with regard to the relative size of characteristic wave-lengths and characteristic curvature scales as opposed to gravitational-field intensity—the “geometrical optics” approximation may well be different in each theory.
2. Well posed initial-value formulations for same types of physical fields? Seems plausible not: affine/metric structure in the one is “dynamical”, with geometrical constraints on evolution, and so its initial-value formulation faces many of the same technical and conceptual difficulties as that of GR.

3. Curvature and gravitational force suggest different physical interpretations, respectively, of what one would otherwise take to be equivalent phenomena: what it means to "extract energy from the gravitational field to heat up my cup of coffee", *e.g.* This will be modeled differently in each theory, with no "obvious translation" between the two. ("Quasi-local" gravitational energy flux always well defined in GNG, not in NGT without asymptotic boundary conditions.)
4. Curvature allows for possibility of introducing measures of gravitational entropy akin to those in GR that "gravitational force" with a flat affine structure does not.

Credo, and Animadversions on the Standard Approach to Theories, Empirical Content, and Equivalence

The Category of Theories

The Category of Physical Systems

The Category of Measurements

The Category of Data Models

How Theory and Experiment Make Contact

A Crude Category-Theoretic Model of Empirical Content

A Crude Proposal for Some Inter-Theory Relations

frameworks

A framework is a system that allows one to formulate propositions and affirm them in principled ways based on evidence gathered according to good principles and applied as evidence based on good principles. So far, that does not differentiate it from a theory. It is so differentiated by the fact that some of the propositions it allows one to formulate are themselves theories.

a framework is a category of theories, **FW**

- Newtonian mechanics
- non-relativistic quantum mechanics
- quantum field theory
- special relativity (if you're Harvey Brown)
- general relativity
- perhaps Lagrangian mechanics, Hamiltonian mechanics, statistical mechanics

objects

1. **theories:** instantiation of framework's abstract equations of motion, a *genus* (Newtonian gravitational theory instantiating Newton's Second Law)
2. **species:** an instantiation of a genus by fixing values for kinematic quantities (number of particles, all masses, in Newtonian gravitational theory)
3. **state:** state of a species (fixing positions and velocities of definite number of particles with given masses in Newtonian gravitational theory); collection forms "space of states"
4. **quantity:** value of a physical quantity accruing to type of system treated by genus

arrows

1. **theory to theory:**
    - approximation, idealization, limiting case (pendulum with θ versus $\sin(\theta)$; friction versus no friction)
    - theorems or constructions capturing counterfactual or global relations (Newton's Precession Theorem, topology)
2. **species to theory:**
    - subobjects (in category-theoretical sense)
3. **species to species:**
    - "accuracy truncation" (ignoring couplings, *e.g.*, Earth-Moon-Sun treating Moon as test body versus as gravitating body)
    - "precision truncation" (dropping significant digits in values of kinematical quantities, *e.g.*, mass)
    - theorems or constructions capturing counterfactual or global relations (measures, topology)

arrows, cont.

1. **state to species:**
   - subobjects (in category-theoretical sense)
2. **state to state:**
   - dynamical evolution, indexed by time and couplings with environment ("externally imposed forces"), starting from a given state: an *individual model*; collectively forming a (sub-)group under composition
   - "precision truncation" (dropping significant digits in values of quantities, *e.g.*, ignoring perturbations in Jupiter's orbit)
   - global structure on space of states (measures, topology)
3. **quantity to state:**
   - value of quantity in that state (should appropriately "commute" with dynamical-evolution arrows)
4. **quantity to quantity:**
   - precision truncation (should appropriately "commute" with precision-truncation arrows among states)
   - value-change during dynamical evolution (induced by dynamical-evolution arrows)

other possibilities, sophistications

1. 2-category for **FW** (Hans? other category-theory wizards?)
2. separate category for each theory **THEOR**, with embedding functors into **FW**, and functors (global structure, approximations, . . .) between theory categories
3. separate category for each species **SPEC** with embedding functors into **THEOR**, and functors between species categories (global structure, truncations, . . .)
4. separate category of individual models for each species: objects include dynamical evolutions (arrows between states in **SPEC**); arrows represent changes in external forces that "push" the system from one to another dynamical evolution $\Rightarrow$ algebraic, topological, *etc.*, structures on families of dynamical evolutions
5. and so on

the category **SYS**:

regime of adequacy

the class of physical systems a theory appropriately and adequately treats

Geroch: "a limiting circumstance in which the effects included within that system [by the theory] remain prominent while the effects not included become vanishingly small"

objects

1. **system:** an individual in the regime (the Solar System, with respect to Newtonian gravitational theory)
2. **state:** state of an individual, a subobject (the values of the positions and velocities of all the bodies in the Solar System at a given time)
3. **quantity:** value of physical quantity in a state, a subobject

arrows

1. **system to system:** external intervention changing system from one species to another (a comet enters the Solar System)
2. **state to system:** identification of state of system, indexed by time
3. **state to state:** dynamical evolution, indexed by time and coupling with environment starting from each state
4. **quantity to state:** identification of value of a quantity of a state, indexed by time (subobjects in category-theoretical sense); should appropriately "commute" with dynamical evolution

Credo, and Animadversions on the Standard Approach to Theories, Empirical Content, and Equivalence

The Category of Theories

The Category of Physical Systems

The Category of Measurements

The Category of Data Models

How Theory and Experiment Make Contact

A Crude Category-Theoretic Model of Empirical Content

A Crude Proposal for Some Inter-Theory Relations

the category **MEAS**:

1. **objects:** a set of measured results for a given experimental arrangement (generally finite and discrete); $\{\mathtt{result}_\lambda\}_{\lambda\in\Lambda}$ for some indexing set Λ, which we will usually suppress reference to

2. **arrows:** the relation of $\mathtt{result}_1$ to $\mathtt{result}_2$ when the two are made on the same physical system:
   - using different experimental arrangements
   - or using the same but with less precision or accuracy
   - or accounting for fewer quantities or relations
   - and so on

so as to accord with the idealized picture of Hertz (discussed soon), this implicitly assumes that measurements are:

- effectively “instantaneous”
- “non-invasive” (don’t destroy the system)

These will not always be adequate idealizations. Then further sophistications must be implemented.

Credo, and Animadversions on the Standard Approach to Theories, Empirical Content, and Equivalence

The Category of Theories

The Category of Physical Systems

The Category of Measurements

The Category of Data Models

How Theory and Experiment Make Contact

A Crude Category-Theoretic Model of Empirical Content

A Crude Proposal for Some Inter-Theory Relations

the category **DATA**

1. **objects:** a highly structured mathematical object `dm`, constructed by analytic, idealizing, interpolative, extrapolative, statistical, and other forms of reasoning applied to an individual measurement result or a set of individual measurement results; includes the special object `null` to be explained later

2. **arrows:** the relation of $\mathtt{dm}_1$ to $\mathtt{dm}_2$ when the two are constructed from measurement results from the same physical system:
    - using different experimental arrangements
    - or using the same but with less precision or accuracy
    - or accounting for fewer quantities or relations
    - or one is a substructure of the other (*e.g.*, representing a state during an entire dynamical evolution)
    - and so on

why I needed quantities as objects: because some data models don't fix all possible dynamic quantities of a system

Credo, and Animadversions on the Standard Approach to Theories, Empirical Content, and Equivalence

The Category of Theories

The Category of Physical Systems

The Category of Measurements

The Category of Data Models

How Theory and Experiment Make Contact

A Crude Category-Theoretic Model of Empirical Content

A Crude Proposal for Some Inter-Theory Relations

A good theory will be *appropriate* and *adequate* for representing and reasoning about the genus of physical systems the theory purports to treat.

appropriate	one can identify in the relevant sense states in the theory's space of states with substructures of the data models of physical systems in the genus treated by the theory
adequate	one can then further identify entire individual models (dynamical evolutions) with entire data models; and then continue on to engage in substantive, successful reasoning about those physical systems based on that identification, and, moreover, one has good reason to believe that such identifications can be carried out for a much broader range of relevantly similar systems than the ones already treated

these identifications need have nothing to do with isomorphism or homomorphism or similarity between data-model structures and theoretical structures, or anything else of the kind

they may merely be, *e.g.*, brute force (approximate) equivalence of values of quantities

example: LIGO

1. deploy instrument so as to be amenable kinematically to appropriate coupling according to the models of one's theories—here, stress-strains: one needs only to know how to measure geodesic deviation (EFE not needed)
2. discriminate "noise" from "signal" (throwing out "bad measurements", signal pollution from other sources, mistakes, instrument malfunctions, *inter alia*)
3. numerical interpolation, extrapolation, manipulation (best-fit statistical analysis, *etc.*)
4. NOW it's a data-model ready possibly to become evidence by further application of physical principles of general relativity to allow interpretation of stress-strain as "distortions in spacetime" and so identify it with a part of an individual model in general relativity; this is where EFE enters

In the opening sentences of his posthumously published essay "Thought", Frege remarks that, as the concept "good" shows the way in ethics and "beautiful" in æsthetics, so must "true" in logic. I am not entirely sure what he meant, but I think it was something related to this: we must look, in deductive logic, for forms of reasoning that preserve truth in moving from antecedents to consequents. In the same vein, I say that "propriety" shows the way for empirical content of scientific theories: in order to explicate the notion of meaning, we must find forms of reasoning that preserve propriety from initial propositions in a representation to derived propositions in the representation, and, more generally, that preserve propriety from initial representations to other representations derived from them.

We form for ourselves images or symbols of external objects; and the form which we give them is such that the necessary consequents of the images in thought are always the images of the necessary consequents in nature of the things pictured.

H. Hertz
<u>The Principles of Mechanics Presented in a New Form</u>
(Intro., p. 1)

Appropriate, adequate models, in other words, represent physical systems, albeit with this peculiar proviso: the construction of representations of the system in the theory's terms at a moment *commutes* with the physical evolution of the system as determined by experimental measurement at that and at later moments.

A dynamically appropriate and adequate theoretical representation is a functor from **THEOR** to **SYS**, constructed from intermediate functors from **THEOR** to **SPEC** to **DATA** to **MEAS** to **SYS**, that makes theoretical states and dynamical evolutions in **THEOR** appropriately commute with physical states and evolutions in **SYS**.

[PICTURE]

example: SPEC → DATA

a few properties:

- many-to-one on objects and arrows: the same states and dynamical evolutions can be appropriately identified with many data models
- states and dynamical evolutions not realizable by any actual physical system (spatiotemporal scales too small, energy scales too large, *etc.*) map to the data model `null`
- it is essentially surjective (every data model is realized by some state or dynamical evolution, even if only `null`)
- it may or may not be full: all *local* relations—arrows—among states and dynamical evolutions can be captured by data models, but global ones?
- it is not faithful: different uses of data models, reasonings and constructions based on them for the purposes of identification with theoretical structures, can capture the same things

N.b.: this picture may not capture everything we do with theories, *e.g.*, calculating the density of lead using solid-state quantum theory—not clear to me yet

Credo, and Animadversions on the Standard Approach to Theories, Empirical Content, and Equivalence

The Category of Theories

The Category of Physical Systems

The Category of Measurements

The Category of Data Models

How Theory and Experiment Make Contact

A Crude Category-Theoretic Model of Empirical Content

A Crude Proposal for Some Inter-Theory Relations

limiting theories

- compare $\mathbf{DATA}_{\mathbf{GR}}$ to $\mathbf{DATA}_{\mathbf{NGT}}$ to determine what sort of functor one needs between them (forgets structure? properties? stuff?)
- determine whether there is a physically significant functor of the same sort between $\mathbf{THEOR}_{\mathbf{GR}}$ and $\mathbf{THEOR}_{\mathbf{NGT}}$ that appropriately commutes with the functor between each theory and its category of data models, and between the two categories of data models, so as to appropriately commute with physical and states and evolutions of the relevant physical systems

complicated evidential relations

- we think of general relativity as "better, deeper" theory than Newtonian gravitational theory
- *but* most data models we use in astrophysics are constructed based on Newtonian gravitational theory
- *and* we use them as an evidentiary basis for general relativity!

solution

- construct data models appropriate for GR from those of NGT, automatically implements functor from **DATA**$_{\mathbf{NGT}}$ to **DATA**$_{\mathbf{GR}}$ (“forgets the right things”, if any)
- define relevant functors from **THEOR**$_{\mathbf{NGT}}$ to **THEOR**$_{\mathbf{GR}}$ (perhaps those defined by the “inverse” of the limiting functors previously constructed)
- determine whether all the functors appropriately commute, and in so doing “forget the right things” (if any)

equivalence

[PICTURE]

we want the diagram to commute—but do any need to be categorical equivalences except between the data-model categories?

goal: an ideal based on which one can construct simplified category-theoretic models, and thereby get a grip on exactly what it is one is ignoring in one's idealized picture of theories and their structure and semantics, so as to be able to argue in a principled way that it is reasonable to ignore what one is

probably impossible to use in real philosophical argument and investigation

but I really like the Hertz stuff