Inferential power, formalisms, and scientific models

Ardourel Vincent*, Anouk Barberousse†, Cyrille Imbert§

* IHPST — CNRS, Université Paris 1 Panthéon-Sorbonne
† SND — CNRS, Sorbonne Université
§ Archives Poincaré — CNRS, Université de Lorraine

Abstract
Scientific models need to be investigated if they are to provide valuable information about the systems they represent. Surprisingly, the epistemological question of what enables this investigation has hardly been investigated. Even authors who consider the inferential role of models as central, like Hughes (1997) or Bueno and Colyvan (2011), content themselves with claiming that models contain mathematical resources that provide inferential power. We claim that these notions require further analysis and argue that mathematical formalisms contribute to this inferential role. We characterize formalisms, illustrate how they extend our mathematical resources, and highlight how distinct formalisms offer various inferential affordances.
1. Introduction. When analyzing scientific representations, philosophers of science are keen on mentioning that some models provide scientists with “mathematical resources” and “inferential power”, but they seldom give a detailed analysis of these notions. This paper is devoted to the discussion of what appears to us as major mathematical resources, namely, formalisms. We thus present an analysis of the notion of formalism as well as examples from which we argue that formalisms should be acknowledged as major units of scientific activity.

We proceed as follows. In Section 2, we briefly review what philosophers of science have to say about mathematical resource and inferential power and observe that it is disappointing. In order to fill the gap we have identified, we put forward in Section 3 the three components we identify within the notion of mathematical resource. Section 4 is devoted to one of these components, namely, formalism. At last, in Section 5, we provide the reader with examples of how the choice of a formalism influences the type of knowledge scientists may draw from their representations.

2. Scientific representations and inferences therefrom. At what conditions can scientific models be used to gain information about target systems? First, a suitable semantic relation between the model and the system(s) that it stands for should obtain, so that by investigating the model, we can make legitimate inferences about its target system(s). This cannot be done unless nontrivial inferences about the model itself, as a mathematical object, can be carried out. Models are usually referred to by proper names (like “Ising model” or “Lotka-Volterra” model”) or by expressions that highlight some of their mathematical properties (like “the harmonic oscillator” or “the ideal gas”). There is however more to be learnt about them than their prima facie properties. For example, solving the Ising model reveals more about Ising-like systems than their description as “sets of discrete variables representing magnetic dipole moments of atomic spins that can be in one of two states”; similarly, the mathematical content of an harmonic oscillator goes beyond “being a system that, when displaced from its equilibrium position, experiences a restoring force that is proportional to the displacement”. Philosophers of science are aware of the need to investigate the epistemology of models and how we find out about concealed truths about model systems (Frigg,
2010, 257) but are surprisingly silent about how it is actually performed.¹ They are content with saying that the model is “manipulated” (Morgan and Morrison, 1997, chapter 2, passim) or that we can “play” with it (Hughes, 2010, 49), which are suggestive, but metaphorical characterizations.

Surprisingly, even accounts of applied mathematics and scientific representation that give central stage to their inferential role hardly analyze how it is fulfilled and which elements of the models contribute to it. Let us illustrate this point with Bueno’s and Colyvan’s work. They claim that “the fundamental role of applied mathematics is inferential” (Bueno and Colyvan, 2011, 352) and accordingly propose an “inferential conception” of the application of mathematics that extends Hughes’ three-step DDI account of scientific representation (see below).² First, a “mapping from the empirical set up to a convenient mathematical structure” (ibidem, 353) is established (immersion step); by doing so, it becomes possible “to obtain inferences that would otherwise be extraordinarily hard (if not impossible) to obtain” (ibidem, 352) (derivation step); finally, the mathematical consequences that were obtained are interpreted in terms of the initial empirical set up (ibidem, 353) (interpretation step). Bueno and Colyvan further highlight the importance of the inferential role of mathematics for mathematical unification, novel predictions by mathematical reasoning or mathematical explanations (ibidem, 363). However, the analysis of how this inferential role is carried out shines by its absence. Bueno and Colyvan mostly analyze mathematical resources in a semantic perspective³ and insist on the difference in content and interpretation that these make possible, e.g., when “mathematics provides additional entities to quantify

¹ Frigg, while clearly stating the problem, does not really address it and is content with briefly emphasizing the advantages of his fictional account of model concerning the epistemology of models (Frigg, 2010). As to the epistemological section of Frigg and Hartmann’s review article about scientific models, it merely points at experiments, simulations, thought-experiment as ways of investigating models (Frigg and Hartmann, 2017).

² Suarez’s inferential conception (Suarez, 2004) hardly addresses either the question of how inferences from models are actually carried out. For lack of space, we shall not discuss it here.

³ Their discussion is mostly directed at the shortcomings of Pincock’s “mapping account” of the application of mathematics (Pincock, 2004).
over” (complex numbers), or is “the source of interpretations that are physically meaningful” and provide “novel prediction” about physical systems, like with the case of the interpretation of negative energy solutions to Dirac’s equation (ibidem, 366).

In another paper, Bueno suggests that results are derived “by exploring the mathematical resources of the model” in which features of the empirical set up are immersed (Bueno, 2014, 379, see also 387) and that results emerge “as a feature of the mathematics” (ibidem) or by using “the particular mathematical framework” (ibidem, 385). What this inferential power of mathematics should be specifically ascribed to remains unclear. Bueno and Colyvan (2011, 352) just claim that the “embedding into a mathematical structure makes it is possible to obtain inferences”. They also emphasize how, with the help of appropriate idealizations, “the mathematical model [can] directly [yield] the results” (ibidem, 360, our emphasis). But elsewhere in the paper, consequences are said to be drawn “from the mathematical formalism, using the mathematical structure obtained in the immersion step” (ibidem, 353, our emphasis).

What are we to make of these various claims? A prima facie plausible answer to this question might be that structures and formalisms are the two sides of a same inferential coin. However, this answer is not satisfactory, since, as is well-known, mathematical structures can be presented in different formalisms, which, as we shall see in Section 4, are associated with different inferential possibilities. Another blind spot in Bueno’s and Colyvan’s account is that while the derivation step is claimed to be “the key point of the application process, where consequences from the mathematical formalism are generated” (ibidem, 353), the question of how inferences are drawn with the help of formalisms is left under-discussed.

We draw from this brief analysis of Bueno’s and Colyvan’s views that the notions of mathematical resource and inferential power, which are commonly used when discussing applications of mathematics, are often mere labels in need of further investigation. Coming back to the seminal ideas presented by Hughes and extended by Bueno and Colyvan is of little help because Hughes’ paper lacks precise answers to the following precise questions: What are exactly mathematical resources? What is their inferential power? In his DDI (Denotation, Demonstration, and Interpretation) account of scientific representation, Hughes claims that scientific representations have an “internal dynamic”, whose effects we can examine (1997, 332), and “contain resources which enable us to demonstrate the results we are interested in”. A general notion of resource is appropriate to capture the variety of ways in which demonstrations can be
carried out; however, the claim that the deductive power comes from “the deductive resources of mathematics they employ” (ibidem, 332) is too vague and is left unanalyzed.

3. Components of mathematical resources. How are the notions of inferential power and mathematical resources to be analyzed? Are they linked to structures or to symbolic systems and formalisms? In this section, we claim that formalisms are an important component of the notions of inferential power and mathematical resource and should be analyzed in their own right.

Let us begin by briefly presenting what are, according to us, the three main components of the notions of mathematical resource and associated inferential power. First, mathematical structures, to the extent that they are tractable, are undoubtedly an important part of the mathematical resources that are used in mathematical modeling. As argued by Cartwright, theories are no “vending machines” that “drop out the sought-for representation” (1999, 247); scientific models are no vending machines either and scientists must make the best of the models that they know to be tractable. Accordingly, the content of models often needs to be adapted by means of idealizations, approximations (Redhead 1980), abstractions, by squeezing representations into the straightjacket of a few elementary models (Cartwright, 1981), or by drawing, from the start, on the pool of existing tractable models (Humphreys, 2004, Barberousse and Imbert, 2014).

Second, mathematical knowledge associated with structures is also to be counted as a distinct mathematical resource, which allows for new inferences when it is available. Let us take the well-known example of Koenigsberg’s seven bridges. The impossibility of crossing them once and only once in a single trip can be demonstrated by applying a result from graph theory. Similarly, the explanation of the life-cycle of the Magicicada (Baker 2009, Colyvan 2018) is provided by the application of a number-theoretic property of prime numbers to life-cycles of species.

At last, formal settings or formalisms provide languages in which theories are developed, calculations carried out, and inferences drawn from models. Examples of formalisms are Hamiltonian formalism, path integrals, Fourier representation, cellular automata, etc. We provide a detailed analysis of some of these below. Contrary to mathematical structures, formalisms are partly content neutral (though form and content are often intertwined in scientific representations). As providing a partially stan-
standardized way of making inferences, they are important tools for scientists, which in turn justifies considering them as important units of analysis in the philosophy of science. Other authors have started exploring the idea that format matters in scientific activities. Humphreys gives general arguments to this effect and emphasizes the difference between formats that are appropriate for human-made and format that suit computational inferences (2004). Vorms (2009) also emphasizes the general importance of formats of representation when toying with theories or models. Formalisms are a specifically mathematical type of format whose role needs further investigation. This is what we do in the next section.

**4. What are formalisms?** As briefly stated above, formalisms are mathematical languages that allow one to present mathematical statements or objects and draw inferences about them by means of general inference rules. For example, *Hamiltonian formalism* is one of the formalisms through which scientists may find out means to solve differential equations. *Path integrals* is another formalism of this kind, with the help of which one may also solve (partial) differential equations. Let us illustrate the latter point further: the integral solution of the Schrödinger equation requires using a mathematical object, the *propagator*, whose calculation the path integrals formalism makes easier. *Fourier representation or formalism* enables one to represent mathematical functions as the continuous sum of sine functions (or complex exponential functions), so that harmonic analysis, i.e. the decomposition of a signal in its harmonic frequencies, may be performed. It also provides modelers with a way to express the solutions of some partial differential equations, such as the heat equation. Finally, formalisms like *numerical integrators, cellular automata, lattice Boltzmann methods, and discrete variational integrators*, are indispensable in current computational science.

Formalisms consist in the following elements:

i. elementary symbols;

ii. syntax rules that determine the set of well-formed expressions;

iii. inference rules;

iv. a partly detachable interpretation, both mathematical and physical.

Their use is facilitated by

v. translation rules that indicate how to shift from one formalism to another.

Let us illustrate these elements by discussing in more detail the above examples. In the Hamiltonian formalism, elementary symbols are used for a variable and its conju-
gate momentum: “(q, p)”, or for Poisson brackets “{...}”. Among the syntax rules that are specific to Hamiltonian formalism, some allow one to rewrite Hamilton equations by using the canonical variables. Inferences rules allow the users to use action-angle variables (I, \( \theta \)) and to solve equations by using these coordinates because this change of variables opens the possibility to deal with integrable systems, thus providing a systematic method to solve exactly, i.e., in closed forms, differential systems like the simple pendulum, and more generally, any 1D-conservative system. Indeed, due to this change of variables, one takes full advantage of the existence of conserved quantities in mechanical systems, which are then used as variables (actions) in Hamilton equations. This allows constructing the solution of the equations by “quadrature” (Babelon et al. 2003, chapter 2). An example of a translation rule is the Legendre transform that allows one to shift to Lagrangian formalism. Similarly, in the case of Fourier transforms, an elementary specific symbol is \( \hat{f} \), which corresponds to the Fourier transform of the function \( f \). Scientists use sets of rules that describe the Fourier transforms of some typical functions, such as the constant function, the unit step function, and the sinusoids, but also rules for the convolution product, viz. the Fourier transform of the convolution \( f \circ g \) is the product of Fourier transforms of \( f \) and \( g \): \( (f \circ g)^\wedge = f^\wedge \cdot g^\wedge \), so that solutions of equations may be found within Fourier space. An inverse Fourier transform is also defined, which enables one to move back from the Fourier transform \( f^\wedge \) to the function \( f \) (this is again a translation rule).

As emphasized above, formalisms are (partly) content neutral and thus “exportable”, even though they usually come with a privileged physical interpretation. As a matter of fact, most formalisms have been developed within a peculiar modeling context or are linked to a physical theory. From this origin, the most successful ones may become autonomous and depart from their original, physical interpretation. For example, Hamiltonian formalism was initially developed in the context of classical mechanics but is nowadays autonomous and used in other physical contexts. Path integrals originally come from the study of Brownian motion (Wiener 1923) and quantum mechanics (Feymann 1942) but are currently used in other fields like field theory and financial modeling.

The mathematical interpretation of formalisms may sometimes be detachable. For example, the transition rules associated with cellular automata (see below) do not have any obvious mathematical interpretation. Further, although some formalisms are linked to acknowledged mathematical theories (e.g., the Fourier formalism is linked to
the theory of complex functions), they differ from genuine mathematical theories, as shown by the example of path integrals, in which the formalism is used in the absence of any uncontroversial mathematical theory that could back it up. The definition of a path integral:

$$K(b,a) = \int_a^b e^{i\pi \hbar L} \frac{d\alpha}{\hbar} D\alpha(t)$$

requires using a measure “Dx”, to which no general, rigorous definition can be given yet. This mathematical concern does not prevent physicists from using path integrals anyway, as testified by the following quote: “The question of how the path integral is to be understood in full generality remains open. Given this, one might expect to see the physicists expending great energy trying to clarify the precise mathematical meaning of the path integral. Curiously, we again find that this is not the case” (Davey 2003, 450).

Let us finally emphasize that formalisms also differ from formulations of physical theories and allow philosophers of science to address different philosophical problems. Formulations of theories, in particular axiomatic ones, are explored when questions about conceptual content and metaphysical implications are raised. They pertain to foundational issues. Whether a given formulation involves calculus is a peripheral issue in this context. By contrast, the primary virtue of a formalism is to allow modelers to draw actual inferences from a theory or model. The inferential rules it contains are more important than the mathematical rigor of the language in which it is expressed.

5. Choosing a formalism. So far, we have argued that the inferential power that is required to explore models is partly brought about by formalisms, and we have given examples thereof. Accordingly, formalisms have to be carefully examined by philosophers of science if they are to provide a fine-grained analysis of how scientific knowledge is produced in practice. We now aim to show that there is no unique description of formalism-rooted inferential power since different formalisms allow for different types of inferences and are adapted to different types of inquiries. We do so by providing examples of these differences and of the factors that guide scientists when choosing the formalism that is best suited to the task at hand.

How do scientists decide which formalism to use in a given inquiry? The choice may first depend on the type of models at hand. For example, the path integral formalism is
well adapted to solve systems with many degrees of freedom (Zinn-Justin 2009) and makes “certain numerical calculations in quantum mechanics more tractable” (Davey 2003, 449). Lagrangian formalism offers a well-suited framework to solve equations describing constrained systems (Goldstein 2002, 13, Vorms 2009, 15). Fourier representation allows one to solve, e.g., the differential equations describing the time evolution of electrical quantities in networks. In this case, differential equations are transformed into algebraic equations on variables in Fourier space, which may be easier to solve. Finally, with the change of action-angle variables, Hamiltonian formalism potentially provides exact solutions for integrable systems, which have as many independent conserved quantities as degrees of freedom.

The use of a particular formalism is also guided by epistemic goals. Depending on the chosen formalism, different kinds of properties, general (e.g. periodicity, symmetry) or particular (dynamical), may be inferred from the same model. Let us illustrate this point with the example of prey-predator models in ecology. Among these, some obey Lotka-Volterra (LV) equations and represent transforming populations with a system of two coupled equations. If they are investigated within the Hamilton formalism, general properties of these models can be found without setting initial conditions or numerical values for the involved parameters. The reframed models can indeed be shown to be integrable, like the simple pendulum in classical mechanics. Dutt explicitly emphasizes the advantages of using this formalism for a two-species LV system:

“In dealing with the problems involving periodicity, the Hamilton-Jacobi canonical theory has a distinct advantage over the conventional methods of classical mechanics. In this approach, one introduces action and angle variables through canonical transformations in such a way that the angle variable becomes cyclic. One then obtains the frequency of oscillation by taking the derivative of the Hamiltonian with respect to the action variable. One may thus bypass the difficulty in obtaining the complete solutions of the equations of motion, if these are not required.” (Dutt, 1976, 460, our emphasis)

LV models can also be solved with the help of computers and generic numerical integrators when the aim is to obtain particular dynamics for specific values of parameters and initial conditions. Such numerical solutions of the LV model can also be provided by specific formalisms, such as discrete variational integrators (Krauss 2017, 34; Tyranowski 2014, 149). In that case, discrete equations are derived from a discrete least action principle, which is well-suited to conservative systems, like the LV sys-
tem. Discrete variational integrators allow for the preservation of general properties like the conservation of global quantities, viz. energy, momenta, and symplecticity. This discrete formalism comes with mathematical constraints on the discretization of time since the time step has to be adaptive in order to guarantee the conservation of global quantities (Marsden & West 2001, Section 4.1).

Finally, let us mention that LV models can also be studied by using cellular automata (CA) and associated formalism, with the following advantages:

[a rather general predator-prey model] is formulated in terms of automata networks, which describe more correctly the local character of predation than differential equations. An automata network is a graph with a discrete variable at each vertex which evolves in discrete time steps according to a definite rule involving the values of neighboring vertex variables. (Ermentrout and Edemstein-Keshet 1993, 106)

On the one hand, CA are discrete dynamical systems, but on the other, they are also a nice means to practice science with the help of a computationally simple formalism (in terms of transition rules). They can be extremely powerful. For example, rule 110 is Turing complete and, like lambda-calculus, can emulate any Turing machine and therefore complete any computation. In contrast with the case of Hamilton formalism, CA-based inferences from prey-predator models are carried out for specific values and parameters. As CA are described by local rules, these inferences merely pertain to local variations in the model. However, the simplicity of these rules is a tremendous advantage for modeling and code-writing. For instance, CA allow one to easily add rules for the pursuit and evasion of populations as well as rules for age variation (Boccara et al. 1993, Ermentrout and Edemstein-Keshet 1993, see also Barberousse and Imbert 2013 for an analysis of CA as used in fluid dynamics and compared with Navier-Stokes based methods).

Let us now turn to a different example illustrating how different the epistemological effects of using this or that formalism may be. Crystals are currently modeled as lattices that come under two forms, lattices in real space and lattices in reciprocal space. Each is associated with a specific formalism. Within the real space lattice formalism, crystals are described with a vector R expanded on a vector basis (a_1, a_2, a_3) which corresponds to crystal directions, and alpha, beta, gamma are the corresponding angles. Inferences about symmetry of crystals are usually made within this type of representation since the real space is well adapted to studying discrete translations and rotations.
Crystals can also be described with the help of a vector $R^*$ in a lattice in reciprocal space. There is a clear correspondence between the two spaces since they are dual. Given $R$ in the real space, we can derive $R^*$ in the reciprocal space, and conversely. The two spaces are related by a Fourier transform. However, the reciprocal space can be more convenient because inferences about diffraction and interference patterns are easier to carry out in the Fourier representation. As stressed by Hammond in a textbook of crystallography:

the reciprocal lattice is the basis upon which the geometry of X-ray and electron diffraction patterns can be most easily understood and [...] the electron diffraction patterns observed in the electron microscope, or the X-ray diffraction patterns recorded with a precession camera, are simply sections through the reciprocal lattice of a crystal (Hammond 2009, 165).

This example shows that facilitating inferences may have various epistemological effects. Some are relevant to computational aspects and the predictions or explanations that scientists are able to produce in practice. Others pertain to the way scientists understand and reason about models and their target systems. This example also shows how different epistemic goals (symmetry-oriented vs. interference-oriented investigations of crystals) determine which formalism is chosen.

Overall, the above shows that formalisms not only have an important impact on the amount of results scientists may produce, but also on the types of results that are attainable. The examples we have discussed also highlight that the existence of a variety of formalisms is a source of epistemic richness and enhanced inferential power for scientists because it provides them with multiple ways of investigating the same mathematical structures or structures that are related by suitable morphisms.

6. Conclusion. The above proposals are meant to contribute to the epistemological question of what provides models with inferential power and helps scientists succeeding in their inquiries. We have shown that some of this inferential power is brought about by the formal symbolic tools that scientists use to present and investigate mathematical models. Our second claim is that all formal settings do not enable the same types of inferences nor are suited to all epistemic goals. Accordingly, a fine-grained analysis of the conditions of scientific progress needs, among other things, to focus on formalisms.
Our epistemological analysis is not tied to any particular theory of scientific representation. However, by showing that inferences actually hinge on choice of formalism, it suggests that a theory of scientific representation that is cashed out in terms of structures is too abstract to account for the various ways equations are solved in practice and information extracted from scientific models.

References


