THEORIES AND MODELS IN SCIENTIFIC PROCESSES

William E. Herfel
Władysław Krajewski
Ilkka Niiniluoto and
Ryszard Wójcicki

Rodopi

Poznań Studies in the Philosophy of the Sciences and the Humanities
1995, Vol. 44, pp. 137 - 149

Nancy Cartwright
Towfic Shomar
Mauricio Suárez

THE TOOL BOX OF SCIENCE

Tools for the Building of Models with a Superconductivity Example

Abstract. We call for a new philosophical conception of models in physics. Some standard conceptions take models to be useful approximations to theorems, that are the chief means to test theories. Hence the heuristics of model building is dictated by the requirements and the practice of theory-testing.

In this paper we argue that a theory-driven view of models can not account for common procedures used by scientists to model phenomena. We illustrate this thesis with a case study: the construction of one of the first comprehensive model of superconductivity by the London brothers in 1934. Instead of a theory-driven view of models, we suggest a phenomenologically-driven one.

1. Introduction and Acknowledgements

The following paper is divided into two parts. The first part comprises section 2 and is a statement of views advocated by Nancy Cartwright. The second part includes sections 3 and 4. This part is the work of Mauricio Suárez and Towfic Shomar. We see both parts as complementary. The second part illustrates the first and the first part provides the framework for the second. The paper was produced in the context of the research project in modelling in Physics and Economics at the Centre for the Philosophy of Natural and Social Sciences, LSE. We are grateful to all the participants in the weekly Wednesday meetings for their suggestions and comments, both on the issues raised in this paper and on issues related to modelling in general. For their comments on sections 3 and 4, we would like to thank especially Daniel Hausman, Martin Jones, Mary Morgan, Margaret Morrison, Marco Del Seta, Max Steuer and Andrew Warwick. Of course none of the views expressed here should be impugned to them. Marco Del Seta also helped with the typing of the paper. Finally we would like to acknowledge the financial support of the research project at the Centre for the Philosophy of Natural and Social Sciences towards the writing of this paper. Mauricio Suárez also would like to acknowledge the financial support of a *La Caixa-British Council* scholarship.

2. *The Tool Box of Science*

A decade and a half ago from a variety of very different points of view, philosophers, historians and sociologists of science began a revolution against what they labelled the "theory-dominated" view of science. In the study of physics for example there was the work of Ian Hacking, Peter Galison and Allan Franklin on experiment, Kristin Schreder-Frechette on big machines and big applications, Steve Shapin and Simon Schaeffer on the intertweaving of science, politics, instrumentation and manners, Crosbie Smith and Norton Wise in their study of Kelvin on the industrial and technological sites of advancing scientific knowledge, and Harry Collins' stress on the importance of implicit versus explicit knowledge in the design and reproduction of lasers.

Under the new regime philosophy of science could no longer be viewed as the philosophy of scientific theory. Even when the discussion is restricted to the philosophy of scientific knowledge, that can not be identified with the philosophy of scientific theory either. Sociologists made clear their claims that the repudation of theory did not mean that they were unconcerned with understanding and cognition by naming their field, the "Sociology of Scientific Knowledge".

I use the image of the tool-box of science to describe a kind of instrumentalism that I defend as a part of this movement to undermine the domination of theory. For much of post-war philosophy of science, instrumentalism has been read narrowly as one of two ways of understanding *theory.* There was the instrumentalist interpretation and there was the realist interpretation. I want to urge instead an instrumentalist account of science, with theory as one small component. Our scientific understanding and its corresponding image of the world is encoded as much in our instruments, our mathematical techniques, our methods of approximation, the shape of our laboratories, and the pattern of industrial developments as in our scientific theories. My claim is that these bits of understanding so encoded should not be viewed as claims about the nature and structure of reality which ought to have a proper propositional expression that is a candidate for truth or falsehood. Rather they should be viewed as adaptable tools in a common scientific tool box. But what are the tools for? What are we meant to be doing with them?

Ian Hacking (Hacking, 1983) has sketched two broad scientific endeavours, representing and intervening. The most familiar versions of instrumentalism focus on the second activity: science is a tool for intervening, manipulating and changing the natural order. In this paper we focus on the first. Physics does aim to represent the world, but it represents it not in its theories but in its models. My chief object of attack here is the covering-law account of the relation of theory to model. This account gives us a kind of homunculus image of model creation: Theories have a belly-full of tiny already-formed models buried within them. It takes only the midwife of deduction to bring them forth. On the semantic view, theories are just collections of models; this view offers then a modern japanese-style automated version of the covering-law account that does away even with the midwife.

The covering-law account was accepted for a very long time by instrumentalists equally with realists. Consider just for one example the Duhem-Quine problem, which has been one of the chief weapons in the instrumentalist arsenal. Theories are underdetermined by any possible amount of data about the real world. Hence it becomes questionable what sense can be made for claiming that one rather than another of them is true of the world. In the simplest version the Duhem-Quine thesis asserts that for any data d there will always be a number of alternative contradictory pairs of scientific theories and auxiliaries that "account for" d. That is, for any d, there exist incompatible $< T, A >$, $< T', A' >$ such that

$$T + A \longrightarrow d$$

and also

$$T' + A' \longrightarrow d$$

I used to argue (for example in *How the Laws of Physics Lie* (Cartwright, 1983)) that the implications of theory, read literally, were almost universally false. I too was taken in by the covering-law account. Nowadays I want to argue that it does not make any sense to talk of "reading the theory" literally, nor to talk of what the theory "implies". It is the implication in $T + A \longrightarrow d$ that I want to challenge. Theories and auxiliaries do not imply data –or better following Matthias Kaiser's advice in this volume, "phenomena"– even in principle. Representations of phenomena must be constructed and theory is one of the many tools we use for the construction.

Consider the alternative: it is not just the models but theories themselves that represent; consider it in the context of modern physics. What can the abstract "laws" of mathematical physics represent? They are certainly not inductive generalizations of the behaviour of real systems. That lesson of *How the Laws of Physics Lie* I am still prepared to stand by. The alternatives? Edicts written by God in his great "Book of Nature"? Or, the fundamental structure of the universe? I want to urge that fundamental theory represents nothing and there is nothing for it to represent. There are only real things and the real ways they behave. And these are represented by models, models constructed with the aid of all the knowledge and technique and tricks and devices we have. Theory plays its own small important role here. But it is a tool like any other; and you can not build a house with a hammer alone.

I defend this view from a number of different angles. One of these arguments is an excursion into metaphysics. Where does the behaviour we record in our models come from? It does not, I claim, drop out as a special case –a 'concretization' of covering laws (see Cartwright, forthcoming). Rather it is a consequence of the repeated and smooth running operation of nomological machines. Another kind of argument depends on looking at how we do arrive at models in physics –not by deduction from basic theory it turns out. This is the kind of argument put forward in the following sections of this paper.

3. *The Theory-Driven View of Models*

3.1 *Models as Approximations*

Let us begin by explaining what we call the *theory-driven view of models.* This view is often articulated around the notion of "*approximation*". A nice example is Michael Redhead's discussion of models as approximations in his 1980 paper 'Models in Physics' . In Redhead's example we are entitled to treat the model as an approximation because the mathematical form of the equations guarantees that the solution of the equation in the model will not diverge greatly from the exact solution of the equation in the theory. Consider a theory that contains the following equation

$$\frac{dy}{dx} - \lambda y = 0 \qquad (1)$$

which has solutions of the form

$$y_n = 1 + \lambda x + \frac{\lambda^2 x^2}{2!} + \ldots + \frac{\lambda^{n-1} x^{n-1}}{(n-1)!} + \ldots$$

Then the n-th order approximation to the solution would be

$$y_n = 1 + \lambda x + \frac{\lambda^2 x^2}{2!} + \ldots + \frac{\lambda^{n-1} x^{n-1}}{(n-1)!}$$

which is an exact solution of the following equation:

$$\frac{dy}{dx} - \lambda y + \frac{\lambda^n x^{n-1}}{(n-1)!} = 0 \qquad (2)$$

According to Redhead, equation (2) is an *impoverishment* model, in the sense that it approximates the equation in the theory (1) for small λ. So far this limit is defined only mathematically. If we want the model to represent some physical system then the relevant question is how to justify the limit on physical grounds. There are a large number of possible limiting equations to the theory, many of which would converge at the same rate. Only one gives the simplified solution above.

But suppose that the theory is true in its domain and that, as is often the case, we find out that the model is a reasonably accurate representation of the behaviour of the system. Then it seems surprising that the very approximation that has a solution is precisely the one that has descriptive power. The fact that this is the case calls for a physical explanation. Something must be true of the phenomena that we are attempting to model in virtue of which the mathematical limiting equation gives a correct description.

3.2 *Models as Idealizations*

A simple case that shows this kind of physical explanation is the construction of the damped linear oscillator from the simple harmonic oscillator. The equation of the simple harmonic oscillator is:

$$m\frac{d^2x}{dt^2} = -\left(\frac{mg}{l}\right)x, \qquad (3)$$

while the equation that describes a damped harmonic oscillator is:

$$m\frac{d^2x}{dt^2} = -\left(\frac{mg}{l}\right)x + bv. \qquad (4)$$

The extra term bv '*models*' the friction the oscillator is subject to. In classical mechanics friction is often modelled by a linear function of velocity, on a number of physically plausible grounds[1].

Equation (4) tends to equation (3) in the limit $b \longrightarrow 0$, as required for an approximation. The crucial issue is that this is not an arbitrarily chosen mathematical limit (several are possible). On the contrary, the precise

[1] For a discussion of modelling friction see eg. Herbert Goldstein [1980], p. 24.

form of the limit is justified by the putative relations between the objects that the models represent. Equation (3) is satisfied by a linear oscillator with no friction; equation (4) is satisfied by a linear oscillator subject to friction. There are well-established techniques in the theory to represent this difference formally.

The *theory-driven view* of models states that as more correction terms are introduced, the model becomes a more realistic representation of the system. The presumption is that, in the ideal limit of introducing correction terms, a final representation will be achieved. Complementary to this presumption for the *theory-driven view* is the thesis that the role of scientists in the construction of models is limited to introducing physically well motivated correction terms into the theoretical, more idealized, models.

We want to argue against the *theory-driven view* on the following grounds: *it is rarely the case that models of the phenomena are arrived at as de-idealizations of theoretical models.* Rather than trying to convey a general picture of modelling in physics, we would like to concentrate upon a specific historical case of phenomenological model building, the construction of the first comprehensive model of superconductivity by the London brothers in 1934. However specific, this case is also paradigmatic. It gives a good description of what is an extended, and common, practice of model building in physics. We choose this example for various reasons. First, it was constructed by two very well known and highly regarded physicists, one of them a theoretician (Fritz London), the other an experimentalist (Heinz London). Second, it was perceived at the time, by the physics community, as a great achievement and one that increased radically our understanding of superconductivity. Finally, it happens to be a case that is to a degree explicit about *not* being a case of de-idealizing a theoretical model.

4. The Construction of Phenomenological Models: Superconductivity

4.1 The Hallmarks of Superconductivity

The model that Fritz and Heinz London (London & London, 1934) proposed for superconductors in 1934 greatly influenced the development of theoretical treatments of superconductivity for very many years afterwards. Theories of superconductivity such as the one of Ginsburg and Landau owe a great deal to the formalism developed by the London brothers and to the profound intuitions contained in their work. Perhaps more importantly, the theory of Bardeen, Cooper, Schrieffer (*BCS* theory) which arises as a development of the Ginsburg-Landau theory is based upon the insight of the formation of Cooper pairs, an insight already anticipated in a speculative manner at the end of the paper by Fritz and Heinz London.

The London equations were indeed a major achievement: for the first time a model was proposed that could account for the fundamental features of superconductivity phenomena. Superconductors are materials that exhibit extraordinary conducting behaviour under specific circumstances. The hallmarks of superconducting behaviour are the following two well established phenomenological findings: resistanceless conductivity and the Meissner effect.

4.2 Resistanceless conductivity

Kamerlingh Onnes (Onnes, 1913) found in 1911 that when mercury is cooled below $4.2K^o$ its electrical resistance falls to near zero. In 1914 he discovered that the effect does not take place in the presence of an intense magnetic field. This is the first phenomenological trait of superconductivity: under a certain critical transition temperature (T_C) and in the absence of strong magnetic fields, a superconductor will exhibit almost perfect resistanceless conductivity. Almost perfect resistanceless conductivity is confirmed by the presence of a stationary current through, say, the surface of a superconducting ring. The current flows at virtually the same constant rate and does not die off.

The relation between the transition temperature (T_C) and the critical magnetic field (B_C) was explored experimentally by Onnes himself. He found that the following relation held with an accuracy of a few percent:

$$B_C = B_0\{1 - \left(\frac{T}{T_c}\right)^2\}. \tag{5}$$

This equation defines the domain of superconductivity. Figure 1 is the graph of equation (5); we indicate on it the region where superconductivity occurs. The graph clearly shows that there are two different ways to approach the superconducting regime. One way is to bring down the ambient temperature to T_C while mantaining constant the weak external magnetic field (weaker than B_C). The other way is to decrease the magnetic field below B_C and to mantain constant the temperature at some $T < T_C$. Both strategies will work. Superconducting behaviour is suddenly exhibited when the critical phase transition takes place.

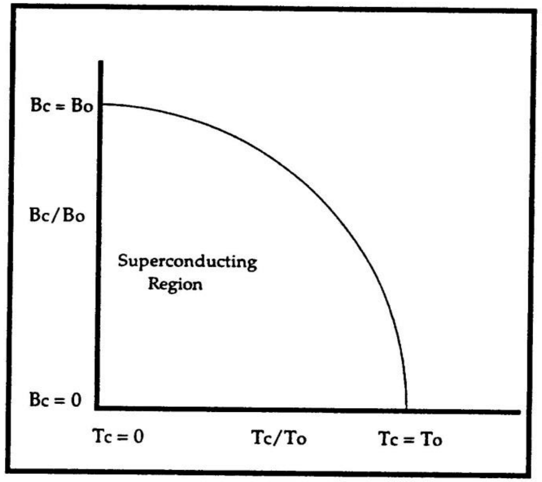

Figure 1: The Domain of Superconductivity

This equation defines the domain of superconductivity. Figure 1 is the graph of equation (5); we indicate on it the region where superconductivity occurs. The graph clearly shows that there are two different ways to approach the superconducting regime. One way is to bring down the ambient temperature to T_C while mantaining constant the weak external magnetic field (weaker than B_C). The other way is to decrease the magnetic field below B_C and to mantain constant the temperature at some $T < T_C$. Both strategies will work. Superconducting behaviour is suddenly exhibited when the critical phase transition takes place.

4.3 The Meissner Effect

The second, equally important, trait of superconductivity was found in 1933 by Meissner and Ochsenfeld (Meissner, Ochsenfeld, 1933). The *Meissner effect* is the sudden expulsion of magnetic flux from a superconductor when cooled below its transition temperature. As such, the Meissner effect defines a fundamental property of superconductors, namely their diamagnetism. A distinction is usually made between Type I and Type II superconductors. Type I superconductors exhibit total diamagnetism: i.e. *all* magnetic flux is expelled in the phase transition. In the case of Type II superconductors, the

diamagnetism is less than perfect. In this paper we will be concerned only with Type I superconductivity. The construction of Type II superconductors took place much later, and the distinction played no role in the historical example that we wish to discuss. So in what follows, by "superconductors" we will mean Type I superconductors. Type I superconductors are thin films made out from metals like zinc (Zn), aluminium (Al), mercury (Hg), lead (Pb).

4.4 The London Equations

In between 1911 and 1933, that is in between the discovery of resistanceless conductivity and the discovery of the Meissner effect, a model was found for superconductors very much along the lines of the *theory-driven view*. It was based upon an 'acceleration equation':

$$\Lambda \frac{d\vec{J}}{dt} = \vec{E} \qquad (6)$$

where $\Lambda = \frac{m}{ne^2}$ (a constant that depends upon the mass m, charge e and number density of electrons n). The 'acceleration equation' can account for a stationary current flowing at a constant rate in the absence of electric fields. For in that case $\vec{E} = 0$ and it follows that

$$\frac{d\vec{J}}{dt} = 0$$

which implies a constant current density $\vec{J}$. The 'acceleration equation' is derived in classical electromagnetic theory by taking the acceleration of the electrons to be caused by a Lorentz force (see, for instance, Bleaney and Bleaney).

London and London were quick to realize that this 'acceleration equation' can not account for the Meissner effect. In fact the equation contradicts the Meissner effect. They prove this in their 1934 paper by deriving from equation (6) a constraint on the magnetic field ($\vec{H}$) *inside* the superconductor[2]. The constraint is given by the equation

$$\Lambda c^2 \nabla^2 \frac{d\vec{H}}{dt} = \frac{d\vec{H}}{dt}. \qquad (7)$$

[2] The derivation involves only the standard electromagnetic identifications: 1) $curl\vec{E} = \frac{1}{c}\frac{d\vec{H}}{dt}$ and 2) $\frac{1}{c}\vec{J} = curl\vec{H}$, together with a number of algebraic manipulations and some vector calculus. The interested reader can follow the detailed derivation in London and London .

Integrating with respect to time one finds the following nonhomogeneous equation for $\vec{H}$:

$$\Lambda c^2 \nabla^2 (\vec{H} - \vec{H}_0) = \vec{H} - \vec{H}_0 \quad (8)$$

$\vec{H}_0$ denotes the magnetic field at the time t=0 (i.e. possibly before the transition phase has occured!).

One solves a nonhomogeneous equation by solving the corresponding homogeneous equation first and then superposing this solution to any particular solution of the nonhomogeneous equation. The corresponding homogeneous equation is:

$$\Lambda c^2 \nabla^2 \vec{H} = \vec{H} \quad (9)$$

The solutions to equation (9) are exponentials $e^{\sqrt{\Lambda}cx}$ that decrease very quickly with distance x from the surface of the superconductor. A particular solution to equation (8) is simply $\vec{H}_0$. The general solution to equation (8) is then given by a superposition of $\vec{H}_O$ and the previously mentioned exponentials. Hence all possible solutions to equation (7) involve the initial field in the superconductor (i.e. the field before the transition phase to the domain of superconductivity). It follows, *a fortiori*, that any solution will necessarily contradict the Meissner effect. As London and London write:

> The general solution means, therefore, that practically the original field persists for ever in the supraconductor. The field $\vec{H}_0$ is to be regarded as 'frozen in' and represents a permanent memory of the field which existed when the metal was last cooled below the transition temperature.

A useful way to visualize this result is by plotting the paths that would be followed in the graph for the phase transition (see figure 2 – which is taken from [Bleaney, Bleaney 1976], p. 399.). Suppose that we begin in the position designated by P in the graph. There is no initial field in the superconductor and hence no flux. In that case the prediction of the 'acceleration equation' model is as in the Meissner experiment: after the transition to the superconducting domain we expect to end up in the position given by S. In the final state there is no flux inside the bulk of the superconductor. But consider the other case when *there is* some initial flux in the superconductor. For instance, suppose that the material is put in the presence of some magnetic field before the transition takes place, such as is designated by position

Q in the graph. In this case the prediction of the 'acceleration equation' model contradicts the Meissner effect. The equation implies that after the transition we end up in R; the Meissner effect entails that we end up in S nevertheless.

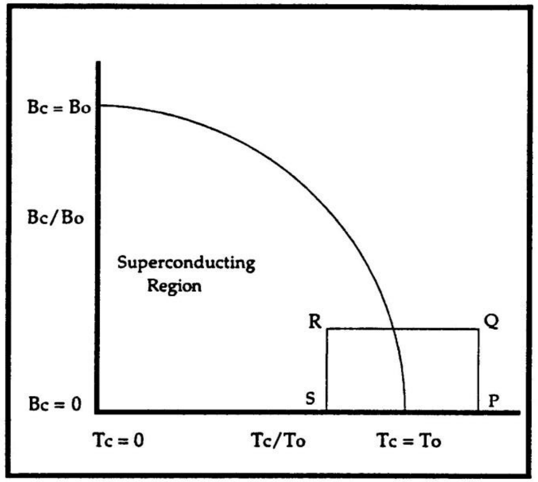

Figure 2: The Meissner effect

A new model had to be constructed to account for the Meissner effect. According to the *theory-driven view of models* there are two possible ways to do so. Either we introduce some correction factors well motivated from the point of view of physical theory into the 'acceleration equation' model or its equivalent (equation (7)). Or we go through the derivation that took us from Maxwell's equations to the 'acceleration equation' model and revise some of the physical assumptions along the way (such as, for instance, the assumption that what causes the electrons to accelerate is a Lorenz force).

Look at what Fritz and Heinz London do in fact. They inspect the behaviour of the homogeneous equation (9). They then realize that the solutions to equation (9) are exponentials that decrease very quickly with distance x from the surface of the superconductor. That simply means that there is no field inside the bulk of the superconductor. But that is precisely what the Meissner experiment provides evidence for!

They abandon the old 'acceleration equation' and instead propose equation (9) as the '*fundamental law*' of superconductivity. Notice what this means for the *theory-driven view*. The transition from the general non-homogeneous equation $\Lambda c^2 \nabla^2 \frac{d\vec{H}}{dt} = \frac{d\vec{H}}{dt}$ to the corresponding homogeneous equation $\Lambda c^2 \nabla^2 \vec{H} = \vec{H}$ is not warranted by either of the two available techniques for the *theory-driven view*. In particular the new equation is clearly *not* found by introducing correction terms into the old equation. Nor is it found by motivating the restriction on theoretical grounds. It is not derived from basic electromagnetic equations with a set of different physical assumptions. In fact, from the point of view of the *theory-driven view* the restriction to the homogeneous equation just looks ad hoc. It seems to have been contrived only to account for the Meissner effect.

Indeed we feel that we have stumbled upon an example of phenomenological model building about which the *theory-driven view* has little to say. What is needed is the recognition of the independence from theory, in methods and aims, of the scientific activity we have come to call phenomenological model building. Anything less than that will fail to make sense of the justification offered by the London brothers for their proposal of a new model for superconductivity (London & London, 1934):

> Until recently the existence of 'frozen in' magnetic fields in supraconductors was believed to be proved theoretically and experimentally. By Meissner's experiment, however, it has been shown that this point of view cannot be mantained.

We believe this is good enough justification.

REFERENCES

Bleaney B., Bleaney B., 1976, *Electricity and Magnetism*, Oxford University Press.

Cartwright N., Forthcoming in "The Monist", *Ceteris paribus laws and socio-economic machines.*

Cartwright N., 1983, *How the Laws of Physics Lie*, Oxford: Oxford University Press.

Goldstein H., 1980, *Classical Mechanics*, Addison-Wesley.

Hacking I., 1983, *Representing and Intervening*, Cambridge University Press.

London F., London H., 1934, *The electromagnetic equations of the supraconductor*, Proceedings of the Royal Society, **A149**.

Meissner W., Ochsenfeld, 1933, Naturwissenschaften, **11**, p. 787.

Onnes K., 1913, *Commun. kamerlingh onnes lab.*, Tech. Rep. **34b**, Univ. Leiden.

Redhead M., 1980, *Models in physics*, British Journal for the Philosophy of Science, **31**.

Department of Philosophy
Logic and Scientific Method
London School of Economics London, ENGLAND