Models and Fiction

Synthese 172(2), 2010, 251-268

Roman Frigg

London School of Economics

r.p.frigg@lse.ac.uk

Most scientific models are not physical objects, and this raises important questions. What sort of entity are models, what is truth in a model, and how do we learn about models? In this paper I argue that models share important aspects in common with literary fiction, and that therefore theories of fiction can be brought to bear on these questions. In particular, I argue that the pretence theory as developed by Walton (1990) has the resources to answer these questions. I introduce this account, outline the answers that it offers, and develop a general picture of scientific modelling based on it.

1. Introduction

The first step in tackling a scientific problem often is to come up with a suitable model. When studying the orbit of a planet we take both the planet and the sun to be spinning perfect spheres with homogenous mass distributions gravitationally interacting with each other but nothing else in the universe; when investigating the population of fish in the Adriatic Sea we assume that all fish are either predators or prey and that these two groups interact with each other according to a simple law; and when studying the exchange of goods in an economy we consider a situation in which there are only two goods, two perfectly rational agents, no restrictions on available information, no transaction costs, no money, and dealings are done in no time.[1] In other words, what we are presented with is a highly stylised and distorted rendering of the system under investigation. A popular introduction to physics describes the situation as follows:

[1] The first of these is the Newtonian model the solar system that is discussed in most elementary physics textbooks (see e.g. Young and Freedman 2000). The is second the so-called 'Lotka-Volterra model' (Volterra 1926); see Weisberg (2007) for a discussion. The third is what is now called the 'Edgeworth Box' (Edgeworth 1881); see Morgan (2004) for a discussion.

> 'In physics a model is a simplified version of a physical system that would be too complicated to study in detail. […] Suppose we want to analyze the motion of a baseball thrown through the air. How complicated is this problem? The ball is neither perfectly spherical nor perfectly rigid; it has raised seams, and it spins as it moves through the air. Wind and air resistance influence the motion, the earth rotates beneath it, the ball's weight varies a little as its distance from the earth changes, and so on. If we try to include all these things, the analysis gets hopelessly complicated. Instead, we invent a simplified version of the problem. We neglect the size and shape of the ball by representing it as a point object, or particle. We neglect air resistance by making the ball move in the vacuum, we forget about the earth's rotation, and we make the weight a constant. Now we have a problem that is simple enough to deal with.' (Young and Freedman 2000, 3)

And this point is not specific to physics; the same can be said about modelling in important parts of biology, chemistry, economics, psychology, and other disciplines.

When presenting a model, scientists perform two different acts: they present a hypothetical system as object of study, and they claim that this system is a representation of the particular part or aspect of the world that we are interested in, the so-called the target system.[2] In working scientists' presentations of models these two acts are usually not clearly distinguished, and the emphasis is often on the first act rather than on the specification of the representational relation. Nevertheless, from an analytical perspective it is important to keep these two acts separate. To this end, let us introduce some terminology. I refer to the hypothetical system proffered as an object of study as the 'model system'. Model systems can be (and often are) used to represent a target system, but the intrinsic nature of the model system does not depend on whether or not this is the case; model systems are objects of sorts and as such can be studied in themselves. Furthermore, I use the term 'modelling' to refer to the entire practice of devising, describing and using a model system.

This raises two sets of questions. The first is concerned with the nature of model systems. What kind of things are they? What makes statements about them true or false? And how do we learn about them? The second is concerned with how model systems represent their targets. What relation does the model have to bear to the target and what is the role of conscious users when a model system is used to represent something? This paper is concerned with the first set of questions. It puts forward the claim that scientific modelling shares important aspects in common with literary fiction and develops a view of model systems based on the so-called pretence theory of fiction. This choice should not suggest that the

[2] This point has been emphasised by Godfrey-Smith (2006) and Weisberg (2007).

second group of questions is either straightforward or unimportant. On the contrary, it is far from clear how to analyse representation and much will have to be said about this problem at some later point. This paper paves the ground for such a discussion because the issue of scientific representation can be addressed properly only once we understand the nature of model systems.

2. Strictures on Structures

What kind of things are model systems? Some, for instance wood models of a car that we put into a wind tunnel, are physical objects. But most models, among them those mentioned in the introduction, are not. An influential view, originating with Suppes (1960) and now held, among others, by van Fraassen (1997) and French and Ladyman (1997), takes model systems to be structures (in the set-theoretical sense). A structure $S = [U, R]$ consists of a non-empty set U of objects (the domain of the structure), and a non-empty indexed set R of relations on U. This definition does not assume anything about the nature of the objects in U. Likewise, relations are defined purely extensionally (i.e. an n-place relation is defined as a class of ordered n-tuples), and hence have no properties other than formal properties such as transitivity or reflexivity.[3]

This conception of model systems is too narrow. Although structures do play an important role in scientific modelling (I come back to this below), model systems cannot be identified with structures. What is missing in the structuralist conception is an analysis of the physical character of model systems. The view of model systems that I advocate regards them as imagined physical systems, i.e. as hypothetical entities that, as a matter of fact, do not exist spatio-temporally but are nevertheless not purely mathematical or structural in that they would be physical things if they were real.[4] If the Newtonian model system of sun and earth were real, it would consist of two spherical bodies with mass and other concrete properties such as hardness and colour, properties that structures do not have; likewise, the populations in the Lotka-Volterra model would consist of flesh-and-blood animals if they were real, and the agents in Edgeworth's economic model would be rational human beings.

[3] Russell (1919, 60) presents a clear account of this aspect of structures.

[4] This view of models is advocated by Frigg (2003) and Godfrey-Smith (2006).

There are two reasons to prefer this take on model systems over the structuralist account. The first is that scientists often talk about model systems as if they were physical things, which is a natural thing to do if models are imagined physical entities. Young and Freedman, when presenting their model of the baseball (in the above quote), do not say that they present a mathematical structure; rather they describe a hypothetical situation in which a perfectly spherical rigid ball moves without air resistance and in the absence of other confounding factors. This way of thinking about model systems is typical in mechanics as well as many branches of physics. And the same is true in biology. Godfrey-Smith (2006, 736-8) points out that Levins' work on population biology as well as Maynard Smith and Szathmáry's models in evolutionary theory are best understood as describing hypothetical physical populations, and he adds that this way of looking at model systems is crucial for the discovery of novel phenomena, for making sense of the treatment of certain issues (e.g. the discussion of robustness in Levins), and for the communication of results in papers and books.

The second reason has to do with how model systems relate to the world. A structure is not about anything in the world, let alone about a particular target system. Those who take model systems to be structures suggest connecting structures to target systems by setting up a morphism between them (the most common morphism is isomorphism; other suggestions include partial isomorphism, homomorphism, and embedding). But a morphism holds between two structures and not between a structure and a part of the world *per se*. In order to make sense of the notion that there is a morphism between a model system and its target we have to assume that the target exemplifies a particular structure, and this cannot be had without bringing non-structural features into play.

The argument for this latter claim proceeds in two steps. The first consists in realising that structural claims are abstract in the sense that they cannot be true unless some more concrete claims are true as well.[5] This is best illustrated with an example. Consider $S_t = [U=(a, b, c), R=(\langle a, b\rangle, \langle b, c\rangle, \langle a, c\rangle)]$, the structure consisting of a three-object domain endowed with a transitive relation R (where '$\langle\ ,\ \rangle$' denotes an ordered tuple). The claim that a part of the physical world has structure S_t is true only if it is also true that it consists of three iron rods of different length (in which case a, b, and c are the iron rods and R the *shorter than* relation), or if it is also true that it consists of three not equally expensive books (in which case a, b, and c are the books and R the *more expensive than* relation), etc. There are innumerable descriptions

[5] This point is discussed in more detail in Frigg (2006).

that make the structural claim that a part of the physical world exemplifies S_t true, but the claim is never true unless some non-structural claim is true as well.[6]

Hence, in order for it to be true that a target system possesses a particular structure, a more concrete description must be true of the system as well. This by itself would not have to worry the structuralist. The problem, and this is the second step, arises when we realise that the descriptions we choose to ground structural claims are almost never true descriptions of the target system. The examples mentioned in the introduction make this sufficiently clear. The structure on which the formal treatment of solar system is based is not fitted out by a realistic description of the solar system, but by one that takes planets to be ideal spheres with homogenous mass distributions, etc. This is what Downes has in mind when he says that there is no empirical system corresponding to the equation of the ideal pendulum (1992, 145), and what Cartwright (1983, Ch. 7) emphasises when she points out that we have to come up with a 'prepared description' of the system in order to make it amenable to mathematical treatment.

Taken literally, descriptions that ground structural claims (almost always) fail to be descriptions of the intended target system. Instead, they describe a hypothetical system distinct from the target system. This has unfortunate consequences for the structuralist. If the descriptions employed to attribute a structure to a target system were just plain descriptions of that system, then the claim that model systems are just structures would appear at least prima facie plausible. But once we acknowledge that these descriptions describe hypothetical systems rather than real target systems, we also have to acknowledge that hypothetical systems are an important part of the theoretical apparatus we employ, and that they therefore have to be included in our analysis of how scientific modelling works. This can, of course, be done in different ways. My suggestion is that these hypothetical systems in fact *are* the models systems we try to understand, and I therefore I reserve the term 'model system' for the hypothetical physical entities described by the descriptions we use to ground structural claims; I refer to the relevant structures as 'model structures'. This facilitates the analysis in what follows, but ultimately nothing hangs on this choice; one could just as well say that model systems are a composite entities consisting of a hypothetical system and a structural

[6] One might try to avoid this problem by positing that models represent data (or data models) rather than target systems. This seems to be wrong. Data play an important role in confirming a model, but they are not what the model represents. For a discussion of this point see Bogen and Woodward (1988) and Teller (2001).

part. What does matter, however, is that we acknowledge that scientific modelling indeed involves such hypothetical systems.[7]

3. Model Systems and Fiction

The question regarding the nature of model systems is reminiscent of another time-honoured philosophical problem, the nature of fiction. Although hardly at the centre of attention, the parallel between certain aspects of science and literary fiction has not gone unnoticed. It has been mentioned by Maxwell (see Cat (2001) for a discussion), and occupied centre stage in Vaihinger's (1911) philosophy of the 'as if'. In more recent years, the parallel has also been drawn specifically between models and fiction. Cartwright observes that 'a model is a work of fiction' (1983, 153) or an 'intellectual construction' (*ibid*, 144), which motivates her move to view physics as theatre (*ibid*., 139) and to later suggest an analysis of models as fables (1999, Ch. 2). Fine notes that modelling natural phenomena in every area of science involves fictions (1993, 16), and Elgin (1996, Ch. 6) argues that science shares important epistemic practices with artistic fiction. Hartmann (1999) and Morgan (2001) emphasise that stories and narratives play an important role in models, and Morgan (2004) stresses the importance of imagination in model building. Sugden (2000) points out that economic models describe 'counterfactual worlds' that are constructed by the modeller, and McCloskey (1990) regards economists as 'tellers of stories and makers of poems'.[8] Finally, in a recent paper Godfrey-Smith has explicitly put forward the view that models are similar to the fictional entities we encounter in novels:

> 'In making this argument, I take at face value the fact that modelers often *take* themselves to be describing imaginary biological populations, imaginary neural networks, or imaginary economies. […]

[7] One could try to avoid the commitment to hypothetical systems by renouncing a literal understanding of the relevant descriptions and arguing that it does not follow from the fact that descriptions are poor or highly idealised that they are not descriptions of the target at all; it just means that they are idealised descriptions. This move is of no avail. Being an idealised description is not primitive concept and it calls for analysis. On the most plausible analysis, *D* is an approximate description of object *O* iff what *D* literally describes is in some relevant sense an idealisation of *O*. But what *D* literally describes is a hypothetical system, and so we back to where we started.

[8] Giere (1988, Ch. 3) argues that models are 'abstract entities', which could be also interpreted as a fiction based view of models. In personal communication he pointed out to me that this is not what he intended.

> Although these imagined entities are puzzling, I suggest that at least much of the time they might be treated as similar to something that we are all familiar with, the imagined objects of literary fiction. Here I have in mind entities like Sherlock Holmes' London, and Tolkein's Middle Earth. These are imaginary things that we can, somehow, talk about in a fairly constrained and often communal way. On the view I am developing, the model systems of science often work similarly to these familiar fictions. The model systems of science will often be described in mathematical terms (we could do the same to Middle Earth), but they are not just mathematical objects.' (Godfrey-Smith 2006, 735)

I support a point of view that aims to understand model systems in analogy with entities that occur in literary fiction. What we have to recognise, though, is that this analogy is only a starting point. If put forward without further qualifications, explaining model systems in terms of fictional characters amounts to explaining the unclear by the obscure. And Godfrey-Smith agrees. In the paragraph immediately following the above passage he observes that '[a]t the end of the day, of course, some general account must be given of the imaginary objects of both ordinary fiction and scientific modeling'. To present such an account is the aim of this paper.

Before delving into theories of fiction, more needs to be said about the questions we expect an account of scientific models to answer, and about why the analogy between model systems and fictions is fruitful.

Every tenable account of model systems has to address the following issues.

(I1) *Identity conditions*. Model systems are often presented by different authors in different ways. Nevertheless, many different descriptions are meant to describe the same model system. When are the model systems specified by different descriptions identical?

(I2) *Attribution of properties*. In the previous sections I have argued that model systems have physical properties. How is this possible if model systems do not exist in space and time? What sense can we make of statements like 'the ball is charged' or 'the population is isolated from its environment' if there are no actual balls and populations? In fact, it has been claimed that such statements are outright contradictory because abstract objects like the ideal pendulum cannot have the same properties as concrete physical systems (Hughes 1997, 330).

(I3) *Comparative statements*. Comparing a model and its target system is essential to many aspects of modelling. We customarily say things like 'real agents do not behave like the agents in the model' and 'the surface of the real sun is unlike the surface of the model sun'. How can we compare something that does not exist with something that does? Likewise, how are we to analyse statements that compare features of two model systems with each other like 'the agents in the first model are more rational than the agents in the second model'?

(I4) *Truth in model systems*. There is right and wrong in a discourse about model systems. But on what basis are claims about a model system qualified as true or false, in particular if the claims concern issues about which the description of the system remains silent? What we need is an account of truth in model systems, which, first, explains what it means for a claim about a model system to be true or false and which, second, draws the line between true and false statements at the right place (for instance, an account on which all statements about a model systems come out false would be unacceptable).

(I5) *Epistemology*. We do investigate model systems and find out about them; truths about the model system are not forever concealed from us. How do we find out about these truths and how do we justify our claims?

(I6) *Metaphysical commitments*. The metaphysics of fictional entities is an issue fraught with controversy (see Friend (2007) for a review). For this reason we need to know what kind of commitments we incur when we understand model systems along the lines of fiction, and how these commitments, if any, can be justified.

It is the contention of this paper that the pretence theory of fiction fits the bill.[9] The next section provides a brief introduction to this theory, and Section 5 outlines the responses that we get from this theory (I1) – (I6).[10]

---

[9] Strictly speaking, Walton (1990) restricts the use of 'pretence' to verbal (or more generally behavioural) participation, which does not include the activity of someone reading on his own. However, it has become customary to use 'pretence' as synonymous with 'make-believe' and I stick to this wider use in what follows.

[10] For want of space I cannot discuss competing approaches. In a nutshell, their problems seem to be the following. The paraphrase account (Russell 1905) does not offer a workable theory of truth in fiction (Crittenden 1991, Ch. 1). The neo-Meinongean view (Parsons 1980) runs into difficulties with incompleteness (Howell 1979, Sec. 1) and as a consequence does not offer a satisfactory answer to (I5). Finally, Lewis' (1978) account is too permissive about what counts as true in a fictional context (Currie 1990, Sec. 2.3; Lamarque and Olsen 1994, Ch. 4).

Before delving into a discussion of fiction, let me give reasons for believing that thinking about model systems as akin to the objects of literary fiction helps us answering these questions. First, characteristically there is nothing in the world of which essential passages of the text of a novel are a true description, and the names of fictional persons and objects characteristically do not denote real persons or objects. Competent readers are fully aware of this and do not mistakenly believe that they read a description of fact when they engage with the text. The situation is the same in science. As Thomson-Jones (2007) points out, scientific textbooks and journal articles abound with passages that are meaningful plain descriptions of physical systems from the domain of enquiry of the scientific discipline in question, but which do not describe *actual* systems and which would not be taken to do so by any competent practitioner in the field. Frictionless planes, spherical planets, infinitely extended condenser plates, infinitely high potential wells, massless strings, populations living in isolation from their environment, animals reproducing at constant rate, perfectly rational agents, markets without transaction costs, and immediate adjustment to shocks are but some objects or features that figure prominently in many model systems and yet fail to have counterparts in the real world.

Second, we can truly say that in David Lodge's 'Changing Places' Morris Zapp is a professor of English literature at the State University of Euphoria. We can also truly say that in the novel he has a heart and a liver, but we cannot truly say that he is a ballet dancer or a violin player. Only the first of these claims is part of the explicit content of the novel, yet there is a matter of the fact about what is the case 'in the world of the story' even when claims go beyond what is explicitly stated. The situation with model systems is the same. The description of a model system only specifies a handful of essential properties, but it is understood that the system has properties other than the ones mentioned in the description. Model systems are interesting exactly because more is the true of them than what the initial description specifies; no one would spend time studying model systems if all there was to know about them was the explicit content of the initial description. It is true that the Newtonian model solar system is stable and that the planets move in elliptical orbits, but none of this is part of the explicit content of the model system's original specification.

Third, a story not only has content that goes beyond what is explicitly stated, we also have the means to learn about this 'extra content' by using certain (usually implicit) rules of inference. The same goes for model systems. Finding out what is true in a model system beyond what is explicitly specified in the relevant description is a crucial aspect of our engagement with the system – in fact the bulk of the work that is done with a model system is usually expended on establishing whether or not certain claims about it hold true.

Fourth, although we sometimes read just for pleasure, when we read serious literature we often engage in comparisons between situations in the fiction and real circumstances, and in doing so we learn about the world.[11] Again, this has parallels in the context of modelling. At what point exactly comparisons between model system and target become relevant depends on one's theory of representation; but on every account of representation one has to compare features of the model system with features of the target at some point, even if only to assess how good an approximation the former is of the latter.

4. A Primer on Pretence Theory

Kendall Walton's (1990) point of departure is the capacity of humans to imagine things.[12] Sometimes we imagine something without a particular reason. But there are cases in which our imagining something is prompted by the presence of a particular object, in which case this object is referred to as a 'prop'. 'Object' has to be understood in the widest sense possible; anything capable of affecting our senses can serve as a prop. An object becomes a prop due to the imposition of a rule or 'principle of generation' (p. 38), prescribing what is to be imagined as a function of the presence of the object. If someone imagines something because he is encouraged to do so by the presence of a prop he is engaged in a game of make-believe. Someone who is involved in a game of make-believe is pretending; so 'pretence' is just a shorthand way of describing participation in such a game (p. 391) and has (in this context) nothing to do with deception (p. 392). The simplest examples of games of make-believe are children's plays (p. 11). In one such play stumps may be regarded as bears and a rope put around the stump may mean that the bear has been lassoed; or pointing the index finger at someone and saying 'bang' may mean that the person has been shot.

[11] There is, however, controversy over how and what we learn from fiction; see Kivy (2006) for a discussion.

[12] I here discuss pretence theory as it is presented by Walton (1990); Currie (1990) and Evans (1982, Ch. 10) develop different versions. Parenthetical references in the text of this and the following section are to Walton's book.

Pretence theory considers a vast variety of different props ranging from novels to movies, from paintings to plays, and from music to children's games. In the present context I only discuss the case of literature. Works of literary fiction are, on the current account, regarded as props as they prompt the reader to imagine certain things. By doing so a fiction generates its own game of make-believe. This game can be played by a single player when reading the work, or by a group when someone tells the story to the others.

Some rules of generation are *ad hoc*, for instance when a group of children spontaneously imposes the rule that stumps are bears and play the game 'catch the bear'. Other rules are publicly agreed on and hence (at least relatively) relatively stable. Games based on public rules are 'authorized'; games involving *ad hoc* rules are 'unauthorized'. By definition, a prop is a representation if it is a prop in an authorised game. On this view, then, stumps are not representations of bears because the rule to regard stumps as bears is an *ad hoc* rule; *Hamlet* is a representation because everybody who understands English is invited to imagine its content, and this has been so since the work came into existence. Within pretence theory 'representation' is used as a technical term. Representations are not, as is customary, explained in terms of their relation (e.g. resemblance or denotation) to something beyond themselves; representations are things that possess the social function of serving as props in authorised games of make-believe.

Props generate fictional truths by virtue of their features and principles of generation. Fictional truths can be generated directly or indirectly; directly generated truths are 'primary' and indirectly generated truths are 'implied' (p. 140). Derivatively, one can call the principles of generation responsible for the generation of primary truths 'principles of direct generation' and those responsible for implied truths 'principles of indirect generation'. The leading idea is that primary truths follow immediately from the prop, while implied ones result from the application of some rules of inference. The reader of *Changing Places* reads that Zapp 'embarked […] on an ambitious critical project: a series of commentaries on Jane Austen which would work through the whole canon, one novel at a time, saying absolutely everything that could possibly be said about them' and is thereby invited to imagine the direct truth that Morris Zapp is working on such a project. The reader is also invited to imagine that Zapp is overconfident, arrogant in an amusing way, and pursues a project that is impossible to complete. None of this is explicitly stated in the novel. These are inferred truths, which the reader deduces from common knowledge about academic projects and the psyche of people pursuing them.[13] What rules can legitimately be used to reach conclusions of this sort is a difficult issue fraught with controversy. I will return briefly to it below; for the time being all that matters is that there are such rules, no matter what they are.

One further issue deserves brief mention, the question of what distinguishes a fictional text from texts of other sorts. A common view associates fiction with falsity and lack of reference: a fictional text is one that talks about persons who don't exist and events that have never occurred. Although it often is the case that fictions are literally false (a fact that I have used in previous sections to motivate the view under discussion), it is important to recognise that a fictional text cannot be *defined* in terms of contrast to truth or fact. A news report can be wrong, but this does not turn it into fiction; and a novel can refer to real persons and talk about actual events without thereby becoming historical treatise (it can do so either by pure chance, or because the author intends this to be the case). On the current account, to be a text of fiction is to possess the function to serve as a prop in a game of make-believe (p. 102).[14] Hence, the essential difference between a fictional and non-fictional text lies in what we are supposed to do with it: a text of fiction invites us to imagine certain things while a report of fact leads us to believe what it says. We can imagine both what is and what is not the case and hence fictional truth is compatible, and may in fact coincide, with actual truth; but this does not render a fictional text non-fictional, just as the fact that a news report is a lie does not turn it into fiction.

5. Models and Imagination

Pretence theory, with some minor qualifications, has the resources to respond to the issues discussed in Section 3. Model systems usually are presented to us by way of descriptions, and these descriptions should be understood as props in games of make-believe. Characteristically, model system descriptions begin with 'consider' or 'assume' and thereby make it clear that they are not descriptions of fact, but an invitation to ponder – in the present

[13] The distinction between primary and inferred truths is not always easy to draw, in particular when dealing with complex literary fiction. Walton also guards against simply associating primary truth with what is explicitly stated in the text and inferred ones with what follows from them (see Walton (1990, Ch. 4) for a discussion). For the purpose of the present discussion these subtleties are inconsequential.

[14] For details see Walton (1990, Ch. 2), Currie (1990, Ch. 1) and Lamarque and Olsen (1994, Chs. 2 and 3).

idiom: imagine – a particular situation. Although it is often understood that this situation is such that it does not occur anywhere in reality, this is not a prerequisite; models, like literary fictions, are not *defined* in contrast to truth. In elementary particle physics, for instance, a scenario is often proposed simply as a suggestion worth considering and only later, when all the details are worked out, the question is asked whether this scenario bears an interesting relation to what happens in nature, and if so what the relation is.[15]

The 'working out' of the details usually consists in deriving conclusions from the primary assumptions of the model and some general principles or laws that are taken for granted. For instance, we derive that the planets move in elliptical orbits from the basic assumptions of the Newtonian model and the laws of classical mechanics. This is explained naturally in the idiom of pretence theory. What is explicitly stated in a model description (that the model-planets are spherical, etc.) are the primary truths of the model, and what follows from them via laws or general principles are the implied truths; the principles of direct generation are the linguistic conventions that allow us to understand the relevant description, and the principles of indirect generation are the laws that are used to derive further results from the primary truths.

We can now address the questions raised in Section 3. The attribution of certain concrete properties to models (I2) is explained as it being fictional that the model system possesses these properties. To say that the model-population is isolated from its environment is just like saying that Zapp drives a convertible. Both claims follow from a prop together with rules of generation. In other words, saying that a hypothetical entity possesses certain properties involves nothing over and above saying that within a certain game of make-believe we are entitled to imagine the entity as having these properties. For this reason there is nothing mysterious about ascribing concrete properties to nonexistent entities, nor is it a category mistake to do so.

Let us now discuss the issue of truth in model systems (I4), which will also provide us with solutions to the other open questions. There is close connection between truth in fiction and truth in a model system, and hence an explication of the former doubles as an explication of the latter. So let us start with truth in fiction. The question is: what exactly do we assert when we qualify 'Zapp drives a convertible' as true in the fiction while 'Zapp drives a Mini

[15] For an accessible account of particle physics that makes this aspect explicit see Smolin (2007), in particular Ch. 5

Cooper' as false?[16] To begin with, it is crucial to realise that there are three different kinds of statement in connection with fiction in science, and that these require a different treatment when it comes to the questions of truth; I refer to these as intrafictional, metafictional, and transfictional statements.[17] For someone sitting in an armchair reading *Changing Places* 'Morris jumped into the paternoster on the downside' is an intrafictional statement because the reader is involved in playing the game defined by the novel and imagines that the sentence's content is the case. Someone who read the novel a while ago and asserts in discussion that Zapp jumped into a paternoster makes a metafictional statement because he is talking about the fiction. If he then also asserts that Zapp, his quirks notwithstanding, is more likeable than any literature teacher he ever had or that Zapp is smarter than Candide, he makes transfictional statements as he is comparing Zapp to real persons and characters in other fictions.

Intraficational propositions are made within the fiction and we are not meant to believe them, nor are we meant to take them as reports of fact; we are meant to imagine them. Although some statements are true in the fiction as well as true *tout court* ('1968 was the year of student revolts' is true and true in *Changing Places*), we often qualify false statements as true in the fiction ('Zapp is a literary theorist' is false because there is no Zapp) and true statements as false in the fiction ('white light is composed of light of other colours' is false in Goethe's *Faust*). So truth and truth in fiction are distinct; in fact, truth in fiction is not a species of truth at all (p. 41). For this reason it has become customary when talking about what is the case in a fiction to replace locutions like 'true in the fiction' or 'true in a fictional world' by the term of art 'being fictional'; henceforth '$F_w(p)$' is used as an abbreviation for 'it is fictional in work w that p', where p is a placeholder for an intrafictional proposition like 'Zapp pursues an impossible project'.[18]

[16] There is controversy over this issue even within pretence theory. It is beyond the scope of this paper to discuss the different proposals and compare them to one another. In what follows I develop an account of truth in fiction that is based on elements from different theories and that is tailored towards the needs of a theory of model systems.

[17] All theories of fiction acknowledge this distinction. My terminology is adapted from Currie (1990, Ch. 4) who speaks about the 'fictive', 'metafictive' and 'transfictive' use of fictional names.

[18] I here follow Currie (1990, Ch. 2) and assume that sentences like 'Zapp drives a convertible' express propositions, something that Walton denies (p. 391). This assumption greatly simplifies the statement of truth conditions for fictional statements, but nothing in the present paper hangs on it. Essentially the same results can be reached only using sentences and pretence (see pp. 400-105).

The question now becomes: when is *p* fictional in *w*? Let the *w*-game of make-believe be the game of make-believe based on work *w*, and similarly for '*w*-prop' and '*w*-principles of generation'. Then, *p* is fictional in *w* iff *p* is to be imagined in the *w*-game of make-believe (p. 39). In more detail:

> *p* is fictional in work *w* iff the *w*-prop together with the *w*-principles of generation prescribes *p* to be imagined

From this it becomes immediately clear how truth and fictionality are connected: *p* is fictional in work *w* iff $F_w(p)$ is true; the fact that $F_w(p)$ is true is referred to as a 'fictional truth' (p. 35). For a proposition to be fictional in work *w* it is not necessary that it is actually imagined by anyone: fictional propositions are ones for which there is a prescription to the effect that they *have to be imagined* (p. 39), and whether a proposition is to be imagined is determined by the prop and the rules of generation. Hence, props, via the rules of generation, make propositions fictional – and thereby create fictional truths – independently of people's actual imaginings (p. 38), and for this reason there can be fictional truths that no one knows of. With this in place we can now also render concept of a 'fictional world' or 'world of a fiction' precise: the world of work *w* is the set of all propositions that are fictional in *w*.[19]

This analysis of truth in fiction carries over to model systems one to one simply by replacing *p* by a claim about the model, *w* by the description of the model system, and *w*-principles of generation by the laws and principles assumed be at work in the model. For instance, 'the solar system is stable' is true in the Newtonian model of the solar system systems iff the description of the system together with the laws and principles assumed to hold in the system (the laws of classical mechanics, the law of gravity, and some general assumptions about physical objects) imply that this is the case. This gives us a straightforward answer to the question about identity conditions (I1). Two models are identical iff the worlds of the two models – the set of all propositions that are fictional in the two models – are identical.[20]

[19] Fictional worlds thus defined are rather different from possible worlds as used in modal logic, the most significant difference being that the former are incomplete while the latter are not. See Currie (1990, 53-70) for a discussion of possible worlds and fiction.

[20] An interesting consequence of this identity condition is that not all models with the same prop are identical, because they can operate with different rules of indirect generation. This is the case, for instance, when the 'same

Metaficational propositions make genuine claims that can be true or false in the same way in which claims about chairs and tables can be true or false. But how can such statements be true if the singular terms that occur in them have no referents? A solution emerges when we realise that statements like 'Zapp is a professor' are ellipses for 'in *Changing Places*, Zapp is a professor'. So when we metafictionally assert *p*, what we really assert is 'in work *w*, *p*' (p. 397). Asserting that something is the case in a work of fiction is tantamount to asserting that it is fictional in that work. Hence asserting 'in work *w*, *p*' amounts to asserting '*p* is fictional in work *w*', which in turn in is equivalent to 'it is fictional in work *w* that *p*'. The last sentence is, of course, just $F_w(p)$. Hence metafictionally asserting *p* amounts to asserting $F_w(p)$. The truth condition for this assertion follows from what has been said above:

> $F_w(p)$ is true iff *p* is fictional in *w*, which in turn is the case iff the *w*-prop and together with the *w*-principles of generation prescribes *p* to be imagined.

Derivatively, *p*, when uttered as a metafictional claim, is true iff *p* is fictional when uttered as an intrafictional claim.[21] Again, this analysis translates to scientific statements without further ado.

Transfictional propositions pose a particular problem because they – apparently – involve comparing something with a nonexistent object, which does not seem to make sense: we cannot compare someone with Zapp if there is no Zapp. Different authors have offered very different solutions to this problem.[22] Fortunately we need not deal with the problem of transfictional statements in its full generality because the transfictional statements that are relevant in connection with model systems are of a particular kind: they compare features of the model systems with features of the target system. For this reason, transfictional statements about models should be read as prefixed with a clause stating what the relevant respects of the comparison are, and this allows us to rephrase comparative sentences as comparisons between

model' is treated first classically and then quantum mechanically; on the current view, the classical and the quantum model are not identical.

[21] In some places Walton ties the truth of such statements to *authorised* games (e.g., p. 397-8). This restriction seems unnecessary as the analysis works just as well for unauthorized games.

[22] Lamarque and Olsen (1994, Ch. 4), for instance, solve the problem by introducing characters. Walton, by contrast, renounces the commitment to characters and instead analyses transfictional statements in terms of unauthorized games (pp. 405-416).

properties rather than objects, which makes the original puzzle go away. Hence, truth conditions for transfictional statements (in the context of scientific modelling) come down to truth conditions for comparative statements between properties, which are unproblematic in the current context (that is, the problems that attach to them have nothing to do with issues surrounding fictional discourse). For instance, when I say 'my friend Peter is just like Zapp' I am not comparing my friend to a nonexistent person. What I am asserting is that both my Peter and Zapp possess certain relevant properties (Zapp possesses properties in the sense explained above) and that these properties are similar in relevant ways. Likewise, when I say that the population of rabbits in a certain ecosystem behaves very much like the population in the predator-prey model, what I assert is that these populations possess certain relevant properties which are similar in relevant respects. What these relevant properties are and what counts as being similar in relevant respects may well depend on the context. But this is not a problem. All that matters from a semantic point of view is that the apparent comparison with a nonexistent object eventually comes down the unproblematic comparison of properties, and the statement making this comparison is true iff the statement comparing the properties with each other is true. Obviously, statements comparing two nonexistent objects are analysed in exactly the same way.

This is the sought after answer to (I3) and (I4). And what is more, this take on truth also provides us with an answer to the question about the epistemology of models (I5): we investigate a model by finding out what follows from the primary truths of the model and the rules of indirect generation. This seems to be both plausible and in line with scientific practice because a good deal of the work that scientists do with models can accurately be described as studying consequences of the basic assumptions of the model.

What metaphysical commitments do we incur by understanding models in this way? The answer is: none. Walton's theory is antirealist in that it renounces the postulation of fictional or abstract entities, and hence a theory of scientific modelling based on this account is also free of ontological commitments. This, of course, is not a refutation of metaphysically less parsimonious views such as Meinong's, and there may be reasons to eventually prefer such a view over an antirealist one. The point to emphasise here is that whatever these reasons may be, the needs of science are not one among them.

This concludes the discussion of (I1) – (I6). Before integrating the various insights gained so far into a consistent picture of scientific modelling, I would like to address a potential misunderstanding. In common parlance, 'imagination' has subjective overtones, and this might suggest to some that an understanding of models as imagined entities makes them subjective because every person imagines something different. This is not so. In pretence theory, imaginations in an authorised game of make-believe are sanctioned by the prop itself and the rules of generation, both of which are public and shared by the relevant community. Therefore, someone's imaginings are governed by intersubjective rules, which guarantee that, as long as the rules are respected, everybody involved in the game has the same imaginings. For this reason models are indeed the same for everybody.

The framework of pretence theory needs a minor amendment at one particular point to be of use for the analysis of scientific modelling. A representation, by definition, is a prop in an authorised game of make-believe. On this view, the text of a novel and the description of a model system are representations. But in science the term 'representation' is also used in a different way, namely to denote a relation between the model system and its target (and, depending on one's views about representation, also other relata like users and their intentions). These two senses of 'representation' need to be clearly distinguished, and for this reason I call the former 'p-representation' ('p' for 'prop') and the latter 't-representation' ('t' for target). Using this idiom, the two acts mentioned in the introduction can be described as, first, introducing a p-representation specifying a hypothetical object and, second, claiming that this imagined object t-represents the relevant target system.[23]

[23] Notice that this distinction also offers a straightforward answer to the alleged puzzle of representations without target: these objects are p-represented but not t-represented – a Bacchus picture p-represents Bacchus, but it does not (and never can) t-represent him.

6. Interlude: The Use of Mathematics

Mathematics plays an important role in scientific modelling, and view of scientific modelling is acceptable only if can explain how mathematics comes into the picture. How does the view developed so far, which places emphasis on the non-formal character of model systems, account for the use of mathematics in modelling? Fortunately this is a variant of a well-known and much discussed problem, namely the problem of the applicability of mathematics. Since Wigner observed that 'the enormous usefulness of mathematics in the natural sciences is something bordering on the mysterious and that there is no explanation for it' (1960, 2), a sizeable body of literature has grown that is devoted to dispelling the air mystery surrounding the use of mathematics in the sciences.[24] The main question addressed in this literature is how mathematics hooks onto something non-mathematical, and various answers have been given. The point to emphasise here is that all these answers carry over to the current context without further ado. The problem of how a piece of mathematics latches onto an imagined object is the same as the problem of how a piece of mathematics latches onto a part of the material world. If we aim to come up with a mathematical description of a population of bees, say, it does not matter whether this population is real or imagined. So when asked to explain how mathematics enters into the picture of modelling canvassed so far, we can simply appeal to whatever account of the application of mathematics we prefer.

In what follows I assume that a structuralist account as outlined in Shapiro (1983) is a compelling option (but nothing hinges on this; the points I make in what follows could be made from any other point of view). The leading idea of this account is that mathematics is the study of abstract structures, and that mathematical expressions like equations should be understood as describing structures. If a particular bit of mathematics, e.g. a particular equation, applies to a non-mathematical object, this is so because the structure it describes is instantiated in that object. Structures themselves are assumed to be Platonic entities in that they exist independently of human minds.

[24] For surveys see Brown (1999, Ch. 4) and Shapiro (1997, Ch. 8).

7. The Elements of Scientific Modelling

The elements introduced in the previous sections together add up to a general picture of what is involved in scientific modelling. This picture is schematically illustrated in Figure 1.[25] An immediate reaction to this picture might be to ask: but where is the model? There is no single answer to this question. With the exception of the target system, every part of the schema legitimately may be, and sometimes is, referred to as 'model'. Once it is acknowledged that scientific modelling involves all the elements that appear in Figure 1, it becomes a purely terminological question which one of these we call 'the model', and this choice largely depends on the context.

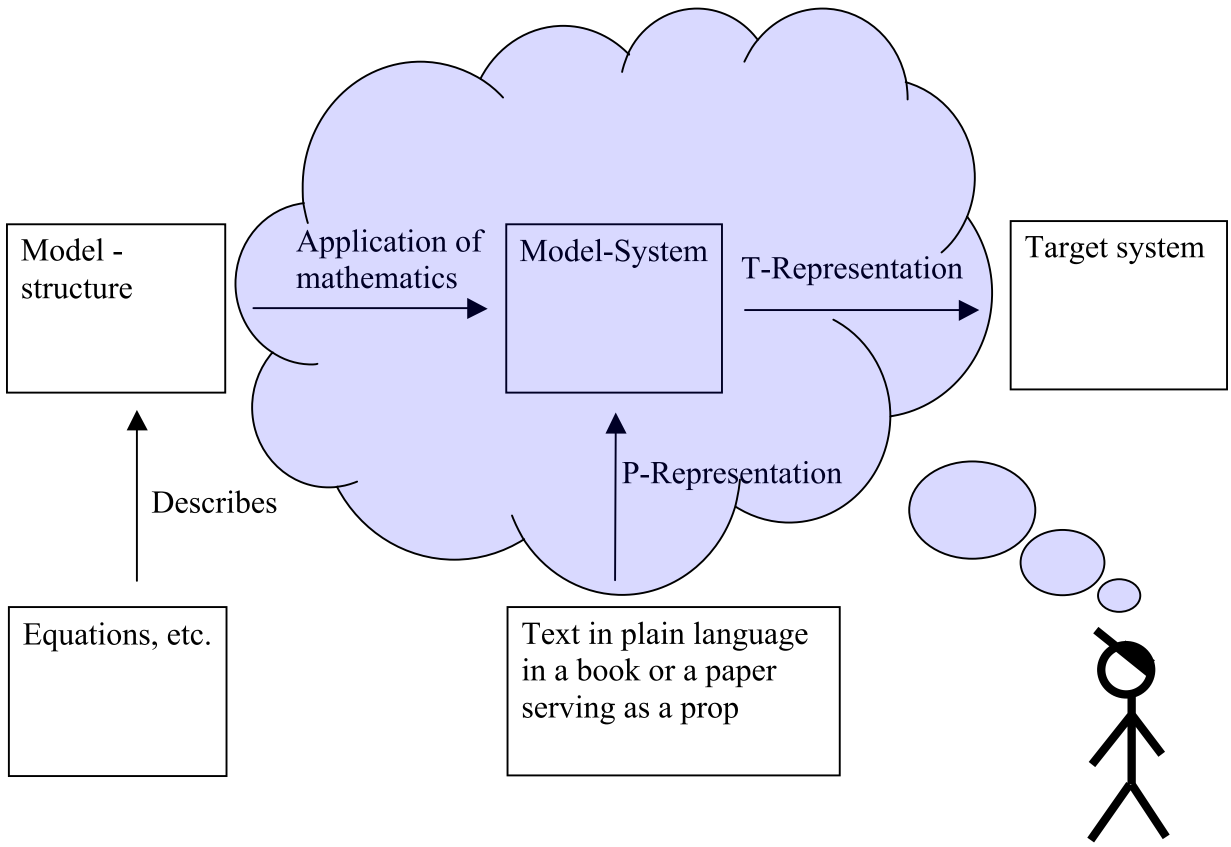

Figure 1 – The elements of scientific modelling.

[25] Parts of this diagram are motivated by the figure in Giere (1988, 83).

8. Conclusion and Outlook

I have argued that scientific modelling shares important aspects in common with literary fiction, and that for this reason theories of fiction can be brought to bear on issues in connection with modelling. I have identified six such issues and suggested that pretence theory offers satisfactory responses to them. From this discussion emerges a general picture of scientific modelling, which views scientific modelling as a complex activity involving the elements shown in Figure 1.

Before pointing to some problems and open questions, let me briefly mention a further advantage of this view. Although there does not seem to be a clear distinction between modelling and thought-experimenting in scientific practice, there has been little interaction between the respective philosophical debates. This is lamentable because it seems to be important to understand how models and thought experiments relate to each other. In a recent paper Davies (2007) argues that there are important parallels between fictional narratives and thought experiments, and that exploring these sheds light on many aspects of thought experiments. This take on thought experiments is congenial to the view on models presented in this paper and suggests an avenue of investigation that understands modelling and thought-experimenting as intrinsically related.

Needless to say, pretence theory is not without problems.[26] Two of them are particularly relevant to understanding scientific modelling. The first is that 'imagination' has different meanings, some of which are unsuited for the purposes of science (see Currie (2004) for a discussion of various notions of imagination). Hence, more needs to be said about what exactly imagining amounts to in science and about how it differs from imagining in other contexts, as well as how it differs from other activities like considering, pondering, and entertaining.

The second problem is that although the general idea of rules of generation is intuitively clear, it turns out to be difficult to give an account of these rules. The two most important rules in the context of literary fiction, the reality principle and the mutual belief principle, not only suffer from intrinsic problems (Walton 1990, Ch. 4), it can also be shown that they give

[26] For critical discussion see, among others, Lamarque (1991), Budd (1992), and the contributions to the symposium on Walton's book in *Philosophy and Phenomenological Research* 51 (1991).

wrong results when put to work in science. So what are the rules of generation in scientific fictions? This is a substantial question that needs to be addressed, but we should not expect a single unified answer. On the contrary, different disciplines have different rules, and understanding what these rules are will shed light on how modelling in these disciplines works.

Finally, we need to address the second question introduced in Section 1, namely how model systems t-represent their respective targets. This is a formidable task. At this point I can only indicate that an account based on the notion of exemplification as discussed by Elgin (1996, Ch. 6) seems to me to be the most promising route to follow.

Acknowledgements

A special thanks goes to the members of the Logos fictionalism reading group, in particular Manuel García-Carpintero and José Díez, without whom this paper would not have been written. Furthermore I would like to thank Nancy Cartwright, Anjan Chakravartty, Stacie Friend, Ron Giere, Arnon Levy, Genoveva Martí, Margaret Moore, Isabelle Peschard, Julian Reiss, Susan Sterrett, and Michael Weisberg for helpful discussions and/or comments on earlier drafts; thanks also the participants of workshops in Madrid, Barcelona and Paris for their critical feedback. Finally I would like to acknowledge financial support from two project grants of the Spanish Ministry of Science and Education (SB2005-0167 and HUM2005-04369).

Bibliography

Bogen, James, and James Woodward (1988), "Saving the phenomena", *Philosophical Review* 97:303-352.

Brown, James Robert (1999), *Philosophy of mathematics. An introduction to the world of proofs and pictures*. London: Routledge.

Budd, Malcolm (1992), "Review of 'Mimesis as Make-Believe'", *Mind* 101:195-198.

Cartwright, Nancy (1983), *How the laws of physics lie*. Oxford: Oxford University Press.

——— (1999), *The dappled world. A study of the boundaries of science*. Cambridge Cambridge University Press.

Cat, Jordi (2001), "On understanding: Maxwell on the Methods of illustration and scientific metaphor", *Studies in History and Philosophy of Science* 32 (3):295-441.

Crittenden, Charles (1991), *Unreality. The metaphysics of fictional objects*. Ithaca and London: Cornell University Press.

Currie, Gregory (1990), *The nature of fiction*. Cambridge: Cambridge University Press.

——— (2004), "Imagination and make-believe", in Berys Gaut and Dominic McIver Lopes (eds.), *The Routledge Companion to Aesthetics*, London: Routledge, 355-346.

Davies, David (2007), "Thought experiments and fictional narratives", *Croatian Journal of Philosophy* 7 (19):29-45.

Edgeworth, Francis Ysidro (1881), *Mathematical psychics: An essay on the application of mathematics to the moral sciences*. London: Kegan Paul.

Elgin, Catherine Z. (1996), *Considerate judgement*. Princeton: Princeton University Press.

Evans, Gareth (1982), *The varieties of reference.* Edited by John McDowell. Oxford: Oxford University Press.

Fine, Arthur (1993), "Fictionalism", *Midwest Studies in Philosophy* 18:1-18.

French, Steven, and James Ladyman (1997), "Reinflating the semantic approach", *International Studies in the Philosophy of Science* 13:103-121.

Friend, Stacie (2007), "Fictional Characters", *Philosophy Compass* 2 (2):141–156.

Frigg, Roman (2003), *Re-presenting scientific representation*, PhD Thesis. London: University of London.

——— (2006), "Scientific representation and the semantic view of theories", *Theoria* 55:49-65.

Giere, Ronald N. (1988), *Explaining science. A cognitive approach*. Chicago: Chicago University Press.

Godfrey-Smith, Peter (2006), "The strategy of model-based science", *Biology and Philosophy* 21:725-740.

Hartmann, Stephan (1999), "Models and stories in hadron physics", in Mary Morgan and Margaret Morrison (eds.), *Models as Mediators. Perspectives on natural and social science*, Cambridge: Cambridge University Press, 326-346.

Howell, Robert (1979), "Fictional objects: How they are and how they aren't", *Poetics* 8:129-177.

Hughes, R. I. G. (1997), "Models and Representation", *Philosophy of Science (Proceedings)* 64:325-336.

Kivy, Peter (2006), *Performance of reading: An essay in the philosophy of literature*. Oxford: Blackwell.

Lamarque, Peter (1991), "Essay review of 'Mimesis as make-believe: On the foundations of the representational arts' by Kendall Walton", *Journal of Aesthetics and Art Criticism* 49 (2):161-166.

Lamarque, Peter, and Stein Haugom Olsen (1994), *Truth, fiction, and literature*. Oxford: Clarendon Press.

Lewis, David (1978), "Truth in fiction", in David Lewis (ed.), *Philosophical Papers, Volume I*, Oxford: Oxford University Press 1983, 261-280.

McCloskey, Donald N. (1990), "Storytelling in economics", in Christopher Nash (ed.), *Narrartive in culture. The uses of Storytelling in the sciences, philosophy, and literature*, London: Routledge, 5-22.

Morgan, Mary (2001), "Models, stories and the economic world", *Journal of Economic Methodology* 8 (3):361-384.

——— (2004), "Imagination and Imaging in Model Building", *Philosophy of Science* 71 (4):753–766.

Parsons, Terrence (1980), *Nonexistent objects*. New Haven: Yale University Press.

Russell, Bertrand (1919), *Introduction to mathematical philosophy*. London and New York Routledge 1993.

Russell, Bertrand (1905), "On denoting", in, *Logic and knowledge*, London Routledge 1956, 39-56.

Shapiro, Stewart (1983), "Mathematics and Reality", *Philosophy of Science* 50:523-548.

——— (1997), *Philosophy of mathematics. Structure and ontology*. New York: Oxford University Press.

Smolin, Lee (2007), *The trouble with physics: The rise of string theory, the fall of a science and what comes next*. London: Allen Lane.

Sugden, Robert (2000), "Credible worlds: The status of theoretical models in economics", *Journal of Economic Methodology* 7 (1):1–31.

Suppes, Patrick (1960), "A comparison of the meaning and uses of models in mathematics and the empirical sciences", in Patrick Suppes (ed.), *Studies in the methodology and foundations of science. Selected papers from 1951 to 1969*, Dordrecht Reidel 1969, 10-23.

Teller, Paul (2001), "Whither Constructive Empiricism", *Philosophical Studies* 106:123-150.

Thomson-Jones, Martin (2007), "Missing systems and the face value practice", *Synthese* forthcoming.

Vaihinger, Hans (1911), *The philosophy of 'as if'*. English Translation London: Kegan Paul 1924.

van Fraasen, Bas C. (1997), "Structure and perspective: Philosophical perplexity and paradox", in Marisa L. Dalla Chiara (ed.), *Logic and Scientific Methods*, Dordrecht: Kluwer, 511-530.

Volterra, Vito (1926), "Fluctuations in the abundance of a species considered mathematically", *Nature* 118:558-560.

Walton, Kendal L. (1990), *Mimesis as make-believe: On the foundations of the representational arts*. Cambridge/MA: Harvard University Press.

Weisberg, Michael (2007), "Who is a Modeler?" *British Journal for Philosophy of Science* 58:207-233.

Wigner, Eugene (1960), "The unreasonable effectiveness of mathematics in the natural sciences", *Communications on Pure and Applied Mathematics* 13:1-14.

Young, Hugh D., and Roger Freedman (2000), *University physics. With modern physics*. 10th ed. San Francisco and Reading (MA).