

On the epistemological analysis of modeling and computational error in the mathematical sciences

Nicolas Fillion · Robert M. Corless

Received: 9 April 2013 / Accepted: 22 August 2013 / Published online: 13 September 2013
© Springer Science+Business Media Dordrecht 2013

Abstract Interest in the computational aspects of modeling has been steadily growing in philosophy of science. This paper aims to advance the discussion by articulating the way in which modeling and computational errors are related and by explaining the significance of error management strategies for the rational reconstruction of scientific practice. To this end, we first characterize the role and nature of modeling error in relation to a recipe for model construction known as Euler’s recipe. We then describe a general model that allows us to assess the quality of numerical solutions in terms of measures of computational errors that are completely interpretable in terms of modeling error. Finally, we emphasize that this type of error analysis involves forms of perturbation analysis that go beyond the basic model-theoretical and statistical/probabilistic tools typically used to characterize the scientific method; this demands that we revise and complement our reconstructive toolbox in a way that can affect our normative image of science.

Keywords Rational reconstruction · Mathematical modeling · Modeling error · Computational error · Backward error analysis

N. Fillion (✉)

Department of Statistics and Actuarial Sciences, Joseph L. Rotman Institute of Philosophy,
University of Western Ontario, London, ON, Canada
e-mail: nicolas.fillion@gmail.com

R. M. Corless

Department of Applied Mathematics, Joseph L. Rotman Institute of Philosophy,
University of Western Ontario, London, ON, Canada
e-mail: rcorless@uwo.ca

1 Introduction

One of the preeminent problems in the philosophical tradition concerns the nature of genuine knowledge, as opposed to mere opinion or belief. To that effect, epistemology endeavors to specify which propositions among the ones we believe can commendably be deemed knowledge. In the first known writing dedicated to this question, the *Theaetetus*, Plato suggests that the problem be tackled by supplying conditions that beliefs and opinions have to satisfy in order to count as genuine claims to knowledge, such as being based on perception, being true, having a justification, etc. The suggestion that beliefs need justification has been found to be compelling since cases in which one is right by accident have to be excluded; yet, precisely characterizing what an adequate justification for a belief is has proved to be elusive, even for beliefs agreeing with our best science.

In fact, scientific practice is pervaded with falsehoods, errors (intended and not intended), approximations, and uncertainty (including both known and unknown unknowns). Cases of epistemic serendipity, fortunate mistakes, aesthetic preferences, and personal idiosyncrasies of influential figures are also integral parts of real science. However, not all of those factors play an equally important role in epistemology. It is true that epistemology is descriptive insofar as it has “the task of giving a description of knowledge as it really is” (Reichenbach 1938). However, the point of epistemology is to clarify what knowledge in general, and scientific knowledge in particular, is, to explain its reliability, and to answer questions concerning its scope and limits. Epistemology aims to provide grounds for *evaluating* what knowledge claims are in fact genuine knowledge; in particular, which claims, hypotheses, models, theories, methods should be considered scientifically warranted. Epistemology does not, as a result, take as its objects the actual thought processes of scientists, the actual words used by scientists, or even what scientists take their own activity to be, but rather the rationally compelling presentations they *ought* to have. Thus, to use the term introduced by Carnap (1928), the object of the epistemology of science is a *rational reconstruction* of science.

The dimension of the rational reconstruction practice that generates an object of study suitable for a properly epistemological analysis of knowledge is often presented as an invective to distinguish the *context of discovery* from the *context of justification*. Here, discovery and justification should not be thought of as two temporally distinct processes—first, you discover something and then later you justify it—since the typical development of science involves alternating phases of discovery and justification that inform one another. There might be overlap between the two contexts, as emphasized by Salmon (1970). The distinction between the contexts is one between *processes* of discovery versus *methods* of justifications. The phrase “methods of justification” denotes what satisfactorily establishes knowledge claims, independently of what scientists actually claim.

Clearly, what is to be included in the context of justification is determined by what methods and tools are considered rational; different choices might result in different organizations of what belongs to what context. It is important to emphasize that which methods of justification are rationally admissible is not god-given; there is room for disagreement about justification, which may be discussed philosophically. Moreover,

if we use reconstructive tools that are misguided or insufficiently far-reaching, our assessment of important aspects of science can turn out to be wrong. The role of philosophy of science as a branch of epistemology, from this point of view, is to determine by philosophical analysis what should count as satisfactorily establishing knowledge claims, that is, *what counts as a rational method of justification* in science.

This is the most fundamental level at which formal epistemology and philosophy of science interact. Formal epistemology develops formal methods that assist us in elucidating aspects of scientific knowledge. As such, formal epistemology might be regarded as providing the instruments that fill out our “reconstructive toolbox.” Indeed, many such formal tools have either been designed to account for aspects of scientific methodology, or have done so post hoc. Examples include mathematical theories of formal inferences in logic, modal analyses of epistemic attitudes, Bayesian accounts of probabilities in terms of beliefs, logical theories of belief revision, counterfactual accounts of scientific laws, models of computation and the corresponding characterizations of computational complexity, formal-learning-theoretic accounts of inductive processes, and more recently agent-based models of scientific organizations. Such applications of the methods of formal epistemology have contributed significantly to our understanding of many aspects of science, such as the structure of theories, confirmation, explanation, theory choice, etc.

There is, however, another crucial aspect of science that has not been dealt with in this manner so far, or only marginally so. The main contribution of applied mathematicians to experimental and theoretical sciences consists in constructing mathematical representations of *real* physical systems in contexts that essentially involve uncertainty, measurement error, modelling error, analytical approximations, and other forms of guesses and ignorance. Despite the prima facie epistemologically suspicious character of the ingredients, the model construction recipes often provide extremely accurate representations of systems. On the basis of the commonsensical rule “garbage in, garbage out,” the accuracy of the resulting representations appears to be uncanny. And yet, it surely is the case that the success of the methods of applied mathematicians is not entirely accidental. Thus, there is an epistemological story to tell about the relation between all the kinds of errors that contribute to the construction of mathematical representations and their intrinsic accuracy. In addition to the epistemic and semantic deficiencies in the representational assumptions, there are intrinsic limitations regarding what can be mathematically achieved. As a result one must resort to computer simulations and numerical approximations that introduce an additional dimension to the problem.

In order to formulate proposals aiming to supplement the formal methods already in our reconstructive toolbox to determine the sense and the circumstances in which such applied mathematical recipes are justified, it is necessary to tackle the preliminary problem of characterizing the main strategies of management of modeling and computational error at a sufficient level of generality. To put it more pointedly, it is necessary to have an account of what has to be reconstructed in order to discuss how to reconstruct it. This paper aims to provide an answer to this preliminary question by drawing attention to elements that philosophical discussion often sweep under the rug. It is true that philosophers have already written on the epistemological and semantical aspects of modeling error from many different perspectives, and there has been

a steadily growing interest in philosophy of science about the computational aspects of modeling (e.g., [Barberousse et al. 2009](#); [Kelly 1996](#); [Hartmann 1996](#); [Humphreys 2004, 2009](#); [Morrison 2009](#); [Parker 2010](#); [Thagard 1993](#); [Winsberg 2009](#)). However, in the latter case, much less has been said about the numerical methods on which the simulations are based and the error theory used to justify them (a noteworthy exception is [Wilson 2006](#)). Moreover, philosophers of science have by-and-large not yet acknowledged the conceptual connections between the methods of error analysis arising in modeling contexts and in computational contexts.

Be that as it may, it is crucial to understand the relation between modeling and computational error and the accuracy of mathematical representations to have a proper philosophical understanding of the logic of model construction and model assessment. We will articulate this relation from the point of view of the branch of mathematics known as *numerical analysis*. In Sect. 1, we explain the circumstances that make modeling and computational error intrinsic parts of applied mathematics. In Sect. 2, we provide a classification of the types of error encountered in the context of mathematical modeling and identify at which steps of the construction of a model each type of error occurs. Finally, in Sect. 3, we explain what the relation between modeling and computational error is by explaining how computational error can be interpreted in terms of modeling error. This will lead us to draw conclusions on the type of reconstructive concepts that are required to capture the rational justification of the effective modeling strategies that prove to be so successful in the mathematical sciences.

2 Exact and inexact solutions of models

The construction of a mathematical model is a process that seeks to capture the essential synchronic or diachronic features of a system by deriving equations from modeling assumptions.¹ However, in order to make predictions or to explain phenomena by means of model equations, it is crucial to find their static or dynamical solutions, as the case may be. The process of solving model equations typically involves mathematical operations such as evaluating functions, finding zeros of functions, solving systems of equations, solving difference or differential or integral equations, etc. Different branches of mathematics develop different methods to find solutions of such problems; here, we will focus on numerical analysis. Numerical analysis is a branch of mathematics that develops, studies, and compares *efficient numerical methods* designed to find numerical *approximations* to the solution of mathematical problems arising in applications, while *quantifying* the magnitude of the computational error and *qualifying* the possible resulting misrepresentation of the system ([Fillion 2011](#)). The first question to address to understand the role of numerical analysis in science is: why

¹ Charactering what the essential features of a system are is a delicate problem, and many proposals of very different natures have been made in order to address this intricate question. For the purpose of this paper, it suffices to think of them as contextually determined traits that are relevant to understanding the behavior of interest. Apart from the conceptual and logical approaches to relevance, one can understand this in terms of mathematical methods such as asymptotic analysis. For this latter approach in the philosophical literature, see, e.g., [Batterman \(2002a,b\)](#), [Fillion \(2012\)](#) and [Pincock \(2012\)](#).

would a discipline devote so much effort to *approximate* solutions, instead of developing new methods to find *exact* solutions? After all, is it not the case that exact solutions (often called “analytic”) provide us with the best mathematical answers to our problems? It is important to address this question, since knowing *why* we have to talk about approximations will suggest *how* we should talk about them.

We suggest that there are four reasons. The first reason is a pragmatic one, namely, the exigencies of scientific practice:

The applications of mathematics are everywhere, not just in the traditional sciences of physics and chemistry, but in biology, medicine, agriculture and many more areas. Traditionally, mathematicians tried to give an exact solution to scientific problems or, where this was impossible, to give exact solutions to modified or simplified problems. With the birth of the computer age, the emphasis started to shift towards trying to build exact models but resorting to numerical approximations. (Butcher 2008)

Thus, there are pressing demands from scientists to reliably simulate complex systems with many parameters, which are typically remarkably hard to solve analytically. The second reason is also pragmatic: even if the equations we obtain from our models are exactly formulated, there is always an appeal to experimental data; in this respect, there is a practical necessity to resort to modification, uniformization, compression, and simplification of the data. In addition, since there is always a certain degree of uncertainty in measurements, an understanding of the effects of approximations on the solutions of models is already required.

The third reason is brought about by theoretical necessity. More specifically, mathematicians have produced many impossibility theorems, i.e., they have shown that some types of problems are not solvable, so that there is no computational route that leads to the exact solution. For instance, Abel showed that it is not possible to solve general polynomial equations of degree five or more in radicals (although there is a less-well-known algorithm using elliptic functions for the quintic itself). Liouville showed that many important integrals could not be expressed in terms of elementary functions (and provided a basic theory to decide just when this could in fact be done). Turing has shown that some number-theoretic problems cannot be finitarily decided. With this sort of theoretical limitation in mind, Trefethen (1992) claims that the numerical analysts’ “central mission is to compute quantities that are typically uncomputable, from an analytic point of view, and to do it with lightning speed.”

Finally, the fourth reason is that it is important to look for approximate solutions because exact solutions might be of little value. A typical example of this happens when analytic solutions do not have closed form representations. A famous example of this situation is the global solution of the n -body problem provided by Wang (1990). Another example—we will revisit it in Sect. 4—is the Airy function. In addition, it also happens that even short finite closed form solutions may be of little value.

In such cases, we *have* to resort to approximation in order to use our mathematical models to predict and explain phenomena. Accordingly, the central problem of numerical analysis is an *epistemological* one: *when one cannot know the true solution of a mathematical problem, how should one determine how close to the true solution the*

(presumably) approximate solution is? The similarity with other traditional questions about the adequacy of our knowledge with reality is striking.²

Now, given that both the nature of mathematics in itself and the role of mathematics in science require a perspective on and a theory of numerical approximation, *how* should we talk about computational error? The guiding principle is that numerical methods should be discussed as part of a more general practice of mathematical modeling as found in applied mathematics and engineering. Once mostly absent from texts on numerical methods, this *desideratum* has become an integral part of much of the active research in various fields of numerical analysis. This being said, in order to articulate more precisely what is meant by the claim that “we should evaluate numerical methods in their modeling context,” we need to explain the way in which measures of computational error can be directly interpreted in terms of modeling error. To do so, we discuss the concept of modeling error in more detail in the next section. On this basis, we will then present a formal framework to characterize the relation between computational and modeling error, and the accuracy of mathematical representations.

3 Modeling and computational error

In order to delineate the concept of modeling error, we first distinguish between *theory* and *model*, based on another distinction between two kinds of statements:

1. general principles, often referred to as field equations or conservation laws;
2. constitutive equations, sometimes referred to as specializing relations.

The most important property of the *general principles* is that they are *common to all media*. As is customary in the physics literature, we use the term ‘medium’ to refer to any material, whether real or ideal. Thus, general principles are genuinely universal claims. They determine the general mathematical structure that is used to describe motion, deformation, flow, etc. They are sometimes called “field equations of balance,” but they are best known as *conservation laws*. For example, the axioms of continuum mechanics usually state six conservation principles: conservation of *mass*, *linear momentum*, *moment of momentum*, *energy*, *electric charge*, and *magnetic flux*. Taken together, with a model of space–time, they form the mathematical structure that applies universally to all bodies in any circumstance,³ and they are the proper subject of the branch of mathematics known as *kinematics*. We will refer to this level as the level of *theory*.

Theories are at the level of general principles; they do not by themselves account for phenomena. To account for phenomena, we need to *construct models*. In other words, the general principles, in themselves, are not sufficient to determine the *evolution* (i.e., motion, deformation, etc.) of bodies in a system. As a result, a theory, understood as the

² One such very similar traditional question results from a sceptical worry that lies at the very core of epistemology. In somewhat Kantian terms, it can be formulated as follows: given that the noumenal truths are not accessible, how should one determine the status of such knowledge-claims?

³ A remark is in order. The mathematical structure is universal in the sense that it is *treated* as if it were. No particular constraints on its application is suggested by the theory. However, this is strictly true only insofar as we are dealing with classical (non-quantum) systems, in non-general relativistic space–time.

logical closure of a collection of universal laws, has *no* observational consequences.⁴ In order to formulate a determinate *dynamical* problem, it is required to specify body forces (e.g., universal gravitation, Coulomb’s law, etc.) and the kind of material to which the general principles and the body forces apply. The specification of a material (or of many different materials) is made by means of *constitutive equations*. Despite the fact that these equations are often labeled “laws,” it must be emphasized that the name is somewhat inappropriate, because they cannot be universal laws of nature, or even theoretical principles, since they contradict one another.⁵ Rather, they define ideal materials, and serve as modeling assumptions.

In order to provide a concrete idea of the entire list of ingredients required to obtain a model describing the behaviour of bodies in physical systems through time, we refer to Euler’s recipe⁶:

- (a) Delineate a class of bodies to be studied.
- (b) Determine what specific forces act between these bodies, i.e., what special force laws hold between them.
- (c) Choose Cartesian coordinates and decompose each of the specific forces along the axes of this coordinate system.
- (d) For each body, and for each axis, sum the component forces acting upon this body in the direction of the axis.
- (e) Set this sum of forces equal to $m \frac{d^2 \mathbf{x}}{dt^2}$ (Newton’s Second Law).
- (f) Solve the differential equation, i.e., find $\mathbf{x}(t)$.

We will shortly return to this recipe to identify the sources of errors arising in modeling. However, we must first classify the types of error arising in model construction⁷:

$$\left. \begin{array}{l} \text{systemic error} \\ \text{experimental error} \end{array} \right\} \text{modeling error}$$

$$\left. \begin{array}{l} \text{truncation \& discretization error} \\ \text{roundoff error} \end{array} \right\} \text{computational error}$$

On the one hand, modeling error includes what philosophers of science have called omission, simplification, distortion, idealization, and abstraction. They thus include things such as neglecting air resistance on a projectile, neglecting the gravitational influence of distant stars and not-so-distant celestial bodies, assuming the constancy of parameters that are not constant (e.g., the stiffness of a spring), and treating elastic bodies as being rigid (e.g., a billiard ball collision). But it also includes experimental

⁴ This point is elucidated by Smith (2001, 2002) and Earman et al. (2002). See also Putnam (1991) and Stein (1995) for an illuminating discussion of this fact.

⁵ The sense in which they contradict each other is that they cannot simultaneously apply to the same body, as they can characterize its dynamical properties in mutually exclusive ways.

⁶ We draw this description of the recipe from Wilson (1998) and Smith (2002).

⁷ This classification is an adaptation from Neumann and Goldstine (1947). As always, the difference between error and uncertainty should be borne in mind. An error is simply the difference between a value and the true value, whereas an uncertainty is an interval within which the true value is believed to lie. For more precise definitions, see, e.g., Taylor and Kuyatt (1994) and Joint Committee (2008).



(a) Structure of single-precision floating-point numbers. Each of the 32 numbers is a 1 or a 0, since the format is binary. There is a bit for the sign, 8 bits for the exponent (with base 2), and the remaining 23 bits are for the fractional part. This is to be contrasted with possibly infinitely long decimal expansions of real numbers.



(b) Discreteness of the floating-point “number line”, with variable density.

Fig. 1 Floating-point numbers

errors of various kinds. On the other hand, computational errors are essentially of three types. Firstly, truncation error consists in replacing functions, integrals, differential vector fields, etc., by truncated asymptotic series. Such truncated expressions are computationally important, since we often have no closed form solutions, and it is impossible to add an infinite number of terms in series. Secondly, discretization error consists in replacing continuous flows of the form $\dot{\mathbf{x}} = \mathbf{f}(t, \mathbf{x}(t); \mu)$ by discrete maps of the form $x_{k+1} = \Phi(t_k, x_k, \dots, x_0, h, \mathbf{f})$. This substitution is the basis for most methods of numerical differentiation and integration.⁸ Finally, we typically don’t compute the value of functions using field arithmetic (e.g., the familiar arithmetic of real numbers), since computers cannot handle such entities. Thus, it is replaced with a finite computer arithmetic known as floating-point arithmetic (see Fig. 1).⁹ All of these computational approximations are made because we can only execute finite, discrete operations. Computational error typically arises in steps (c), (d), and (f) of Euler’s recipe.

Now, let us return to Euler’s recipe in order to identify the potential sources of modeling error and their nature. The first step includes the specification of the number and types of bodies (i.e., mass-point particles, rigid bodies, continuously deformable bodies) that are part of the system. Two kinds of modeling error can be introduced here: we can neglect the presence of some bodies altogether, and we can assume that some bodies are simpler than they in fact are (e.g., assuming that a body is rigid, that it is a point particle, or that a fluid is inviscid). It also includes the specification of a number of parameters, such as the kinematical constants (e.g., mass, charge, shear stress, etc.) and the initial values of state variables. The former introduce systemic error and the latter introduce experimental error.

The second step involves a decision about which body-force laws will apply between bodies. For example, one can often suppose that gravitational effects or electromagnetic effects can be neglected. Moreover, this step involves the choice of constitutive equations, as well as the values of the phenomenological parameters they contain. A simple example would be the choice of Hooke’s law $\mathbf{F} = -k\mathbf{x}$ for a spring; there is a source of error in the choice of the parameter k , but also in the fact that springs are not exactly Hookean, since their stiffness is non-constant. At this stage again, we find

⁸ They are thus key for dynamical simulations.

⁹ For more details concerning floating-point number systems, see for example Corless and Fillion (2014, Appendix 1).

both systemic and experimental error. Accordingly, it is steps (a) and (b) of the model construction procedure that we should focus on to understand modeling error.

Note that, to decide whether a model so constructed accounts for some set of phenomena, the step (e) \rightarrow (f) has to be *efficiently computed*, whether exactly or not. In other words, without efficient computation, one cannot decide whether the model accounts for the phenomena, i.e., one cannot determine what the observational consequences are. Moreover, it should be emphasized that, as a result of this requirement of efficient computability, most situations involve a choice between further idealizing the assumptions contributing to the construction of the model and being able to solve the equations exactly, or having less idealized modeling assumptions and being forced to use computational methods that contain an error component.¹⁰ This is why the computational aspects of science cannot be altogether ignored, if one wishes to adequately reconstruct the confirmational and explanatory aspects of science.

These considerations should provide a sufficient clarification of our guiding principle: the role of mathematics in science prescribes that *computational errors should be analyzable in the same terms as modeling and experimental errors*. By that we mean that if truncation, discretization, and roundoff errors are small compared to the modeling and experimental error, then for all we know, our approximate numerical answer can be the right one.

4 Problems and methods in error analysis

In this section, we describe a formal model that will allow us to identify the key problems and methods of error analysis. On this basis, we will explain how computational error can be physically interpreted. It is important to recognize the generality of the method. The analysis extends to many problems in science and engineering, e.g., function evaluation, polynomial equations, series algebra, root finding, numerical linear algebra, numerical quadrature, numerical differentiation, numerical solutions of ordinary differential equations, partial differential equations and many others.

To begin with, we represent a mathematical problem by an operator φ , that has an *input* (data) space \mathcal{I} as its domain and an *output* (result, solution) space \mathcal{O} as its codomain:

$$\varphi : \mathcal{I} \rightarrow \mathcal{O}.$$

Since φ is the problem we are interested with in the first place, we call it the *reference problem*. In many cases, however, we do not have a way to determine the exact solution y to the problem φ at our disposal; this happens in the cases described in Sect. 1. In this very typical case, one can construct a modified problem (using discretization, truncation, and roundoff) for which we can find an exact solution in an efficient way. Accordingly, we introduce the notion of an *engineered problem* $\hat{\varphi}$ (which is by design computable). For some Δy , we obtain this commutative diagram:

¹⁰ This point is articulated more thoroughly by Batterman (2002a). See also Wilson (2006), Pincock (2012, Chap. 11) and Fillion (2012). The point is particularly important to understand the virtues of models at different scales.

$$\begin{array}{ccc}
 x & \xrightarrow{\varphi} & y \\
 & \searrow \hat{\varphi} & \downarrow \Delta y \\
 & & \hat{y}
 \end{array} \tag{1}$$

The Δy is called the *forward error*, and is defined by $\Delta y = \hat{y} - y = \hat{\varphi}(x) - \varphi(x)$. When this is defined, dividing by y gives the *relative forward error*, denoted δy . It represents the difference between the exact and the approximate solution. Accordingly, we can write both $\hat{y} \approx \varphi(x)$ or $\hat{y} = \hat{\varphi}(x)$. In this way, instead of saying that \hat{y} is the *approximate solution to φ* , we can say that it is the *exact solution to $\hat{\varphi}$* . This allows us to emphasize that, instead of focusing on approximate truth, we focus on modified problems; then the investigation is turned into one of characterizing *nearness* of problems. Moreover, *modified problems can be thought of as resulting from model equations derived from slightly modified modeling assumptions*.¹¹

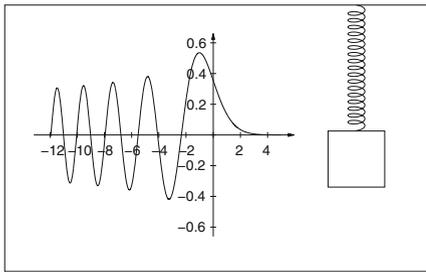
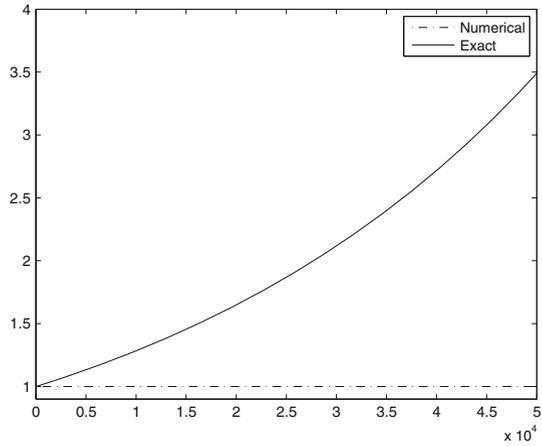
Replacing the reference problem with an engineered problem can lead to surprisingly large forward error. In fact, it is surprising to many that this happens in very simple physical setups. A simple example arises from setups described by a simple homogeneous second-order linear differential equation, say $\ddot{x} - 20,000\dot{x} + x = 0$, which could represent an oscillating mass attached to a Hookean spring immersed in a thick fluid occasioning large damping (here, 20,000 would be the damping coefficient). Then a solution to this differential equation will have the form $x(t) = ce^{\lambda t}$, where λ is a root of the quadratic equation $\xi^2 - 20,000\xi + 1 = 0$ and c is some constant. If we use the quadratic formula to find the roots on a calculator with standard precision, we find that one of the root returned is 0. However, it is not hard to figure out that the true value is 5×10^{-5} . The difference is small (in absolute terms), and yet if we consider the difference between ce^{0t} and $ce^{5 \times 10^{-5}t}$ for large values of t , it can have major repercussions, as we see in Fig. 2. From this we can infer that the problem in question is sensitive to perturbations, since a small variation in the value of the eigenvalue λ can provoke a bifurcation.

This example, however, is not conceptually of much interest, since it is relatively easy to find the exact answer and use it as benchmark. But it is not so for many common problems arising in practice. For example, consider the undamped motion of a weight attached to a spring that becomes linearly stiffer with time (see Fig. 3a). It is described by the differential equation $x''(t) + tx(t) = 0$. A solution to this equation is $\text{Ai}(-t)$, where Ai is the (first) Airy function:

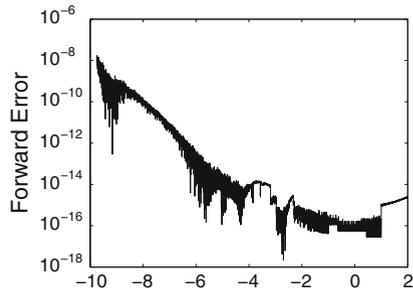
$$\text{Ai}(x) = \frac{1}{\pi} \int_0^{\infty} \cos\left(\frac{1}{3}t^3 + xt\right) dt$$

¹¹ This approach is also central to the so-called method of modified equations (see, e.g., Corless 1994; Corless and Fillion 2014) and, in fluid mechanics, to the so-called method of manufactured equations (see, e.g., Roache 2001; Oberkampf et al. 2004).

Fig. 2 Important qualitative difference resulting from a small change in an eigenvalue



(a) Graph of Ai and of the mass attached to the spring.



(b) Computational error with respect to the variable x .

Fig. 3 A simple physical setup represented by the Airy function

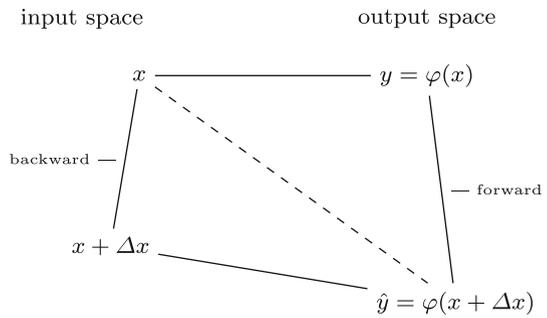
$$= 3^{-\frac{2}{3}} \sum_{n=0}^{\infty} \frac{x^{3n}}{9^n n! \Gamma(n + \frac{2}{3})} - 3^{-\frac{4}{3}} \sum_{n=0}^{\infty} \frac{x^{3n+1}}{9^n n! \Gamma(n + \frac{4}{3})},$$

where Γ is the Gamma function (see [Bender and Orszag 1978](#)). Note that, even if, theoretically, the series converges for all x , it is of almost no practical use. If we use a standard Taylor series computation in standard floating-point arithmetic to compute $f(-12.82)$, near the tenth zero, the absolute error grows very fast as x increases negatively (see [Fig. 3b](#)). Even though the series converges uniformly, the floating-point computation diverges.¹² The same loss of convergence would arise for other finite precision arithmetics, or for computations involving data containing some inaccuracies. This limitation mirrors that of systems of significance arithmetic¹³ used to mathematically analyze experimental data.

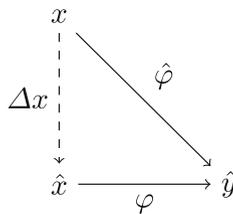
¹² Notice that increasing the floating-point precision will not stop that from happening. Is this really a catastrophe? From the modeling point of view, no. The difficulty stems from radical scale changes, and in this context, it makes sense to consider scale as a fundamental factor in our search for solutions.

¹³ A significance arithmetic is simply a system of calculation rules that takes into consideration the number of significant digits of the operands.

Fig. 4 Backward error analysis:
The general picture



Knowing that the forward error has a certain size, however, is not informative enough. Having a forward error as small as possible is a desideratum, but there remains the question of determining *acceptance criteria*: *when is the forward error small enough to satisfy our modeling needs?* This is why, in applications, it is also important to consider errors in x , the input data of the reference problem φ . This error can have many different sources, e.g., error in preparation of the system, measurement of the data and perturbations of the system. We thus define a quantity $\Delta x = \hat{x} - x$ that corresponds to the size of a modification of x . The smallest such Δx that makes the diagram

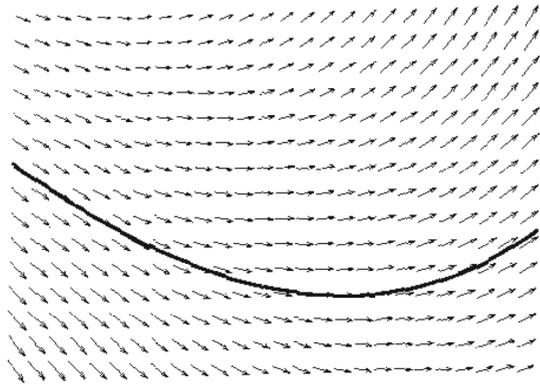


commute is called the *backward error*. As we can see in Fig. 4, we factor the map $\hat{\varphi}$ through \hat{x} instead of through y (as was done in equation 1). This is advantageous since in general we can exactly find or closely estimate Δx , even though we may have no direct information concerning the value of Δy .

Switching our focus from forward error to backward error gives rise to a very general and powerful method called *backward error analysis*.¹⁴ The objective here is to explain the error in the computed solution \hat{y} in terms of errors in the input x . In other words, we ask: *how much error in the input would be required to explain all output error?* Formally, this happens when the diagram in Fig. 4 commutes. Thanks to this change of perspective, the central question is now: When we modified the reference problem φ to get the engineered problem $\hat{\varphi}$, for what set of data have we actually solved the problem φ ? If solving the problem $\hat{\varphi}(x)$ amounts to having solved the problem $\varphi(x + \Delta x)$ for a Δx smaller than the modeling error, then our solution \hat{y} can be considered completely satisfactory.

¹⁴ For a historical account of backward error analysis, see Wilkinson (1971). For a recent exposition and application of this method, see Corless and Fillion (2014), whose afterword contains a brief discussion of potential limitations.

Fig. 5 A vector field with a nearly tangent computed solution



On the basis of the presentation of Sect. 2, this approach can be put in an even more suggestive way: if the computational error accumulated in the steps (c), (d), and (f) of Euler’s recipe corresponds to a backward error smaller than the modeling error accumulated in the steps (a) and (b) of Euler’s recipe, then our computed solution is as satisfactory as the modeling context can demand (no matter how large the forward error is). In such a case, we have successfully extracted the observational consequences from our model and we can use those numerical values to compare with observable phenomena.

The success of this formal model to analyze computational error in terms of modeling error is perhaps best illustrated with the case of initial-value problems. The *standard form of an initial value problem* is

$$\dot{\mathbf{x}}(t) = \mathbf{f}(t, \mathbf{x}(t)), \quad \mathbf{x}(t_0) = \mathbf{x}_0, \quad (2)$$

where $\mathbf{x}(t) : \mathbb{R} \rightarrow \mathbb{C}^n$ is the vector-solution as a function of time, $\mathbf{x}_0 \in \mathbb{C}^n$ is the initial condition, and $\mathbf{f} : \mathbb{R} \times \mathbb{C}^n \rightarrow \mathbb{C}^n$ is the function equal to $\dot{\mathbf{x}}$. For dynamical systems, \mathbf{f} is a velocity vector field (or slope field, or flow) and \mathbf{x} is a curve in phase space that is tangent to the vector field at every point (see Fig. 5). Typically, the solution of this problem will not be directly computable. In this situation, we then resort to some numerical procedure to solve the differential equation. In accord with the formal model proposed, let $\hat{\mathbf{x}}(t)$ be the solution of an engineered problem (say, the map computed by a numerical scheme known as the Runge–Kutta–Fehlberg method) that we would denote $\hat{\varphi}$ here.¹⁵ The backward error turns out to be given by the expression $\Delta(t) = \hat{\mathbf{x}}(t) - \mathbf{f}(t, \hat{\mathbf{x}}(t))$. As a result, we can express the original problem in terms of a modified, or perturbed problem, so that our computed solution is an *exact* solution

¹⁵ It is important for the purpose of applying backward error analysis to numerical solutions of ordinary differential equations that we consider numerical solutions to be \mathcal{C}^1 , i.e., continuously differentiable; otherwise, the backward error would not be globally defined on the interval of integration (see, e.g., Corless and Fillion 2014, part 4).

to this modified problem¹⁶:

$$\mathbf{z} = \mathbf{f}(t, \mathbf{z}) + \Delta(t).$$

From the point of view of dynamical systems, the backward error measures how far from satisfying the differential equation our computed trajectory $\hat{\mathbf{x}}(t)$ is, i.e., how close it is to being tangent to the vector field. In Fig. 5, we see a trajectory that is nearly tangent to the vector field. In an even more suggestive way, we can say that the backward error allows us to find to which perturbed vector field our computed solution is tangent. Thus, as $\Delta(t)$ is a small inhomogenous quantity, we can think of it as a modeling error, say a wind blowing on the system, or a small gravitational attraction from a distant body, or a measurement error on some parameters. As a result, we can directly compare the order of the computational error and the modeling error, and determine whether the consequences obtained by computation are genuinely informative. This is the key point that underlies the claim that the formal model provides measures of computational errors that are directly interpretable in terms of modeling error.

Now, the next question is: *what is the relationship between the forward and the backward error?* The relationship we seek lies in a *problem-specific* coefficient of magnification, i.e., the sensitivity of the solution to perturbations in the data, that is called the *condition of the problem*.¹⁷ The *normwise relative condition number* κ is the supremum of the ratio of the relative change in the solution to the relative change in input, which is expressed by

$$\kappa_{rel} = \sup_x \frac{\|\delta y\|}{\|\delta x\|} = \sup_x \frac{\|\Delta y/y\|}{\|\Delta x/x\|} = \sup_x \frac{\|(\varphi(\hat{x}) - \varphi(x))/\varphi(x)\|}{\|(\hat{x} - x)/x\|}$$

for some norm $\|\cdot\|$. As a consequence, we can show that the relation

$$\|\delta y\| \leq \kappa_{rel} \|\delta x\| \tag{3}$$

holds between the forward and the backward error. We clearly see from this inequality that the condition number acts as a magnifying factor of the error in the data. Knowing the backward error and the condition number thus gives us an upper bound on the forward error. If κ has a moderate size, we say that the problem is *well-conditioned*. Otherwise, we say that the problem is *ill-conditioned*.¹⁸ Thus, if the problem is well-conditioned, i.e., $\kappa \approx 1$, then the error in the solution cannot possibly be much larger

¹⁶ Such an equation in \mathbf{z} can be called a *reverse-engineered* problem. The name is suggestive because we first solve the problem numerically, and then we use the computed solution to determine what perturbed problem we have in fact solved exactly.

¹⁷ Well-conditioning must be distinguished from the concepts of stability of a problem-solving method. There is no unique way of formalizing the notion of *numerical stability*, but its underlying intuitive idea is that an algorithm is numerically stable if it returns results that are about as accurate as the problem and the resources available (typically determined by choosing a system of floating-point arithmetic) allow. Thus, it is similar to the concept of conditioning, but it is a property of methods rather than problems. For rigorous definitions, see, e.g., Higham (2002) and Deuffhard and Hohmann (2003).

¹⁸ Infinitely ill-conditioned problems are known as ill-posed problems in analysis, following Hadamard. See Earman (1986) for a rare discussion in the philosophical literature. Moreover, even if he doesn't specifically

than the error in the data. In such a case, we can conclude that *our strategy provides a solution that is just as good as the exact solution to the reference problem, even if this solution is unknown.*

The condition number, depending on the context, will be given by mathematical quantities such as vector and matrix norms, Lipschitz constants, Gröbner functions, Lyapunov exponents, and other coefficients of sensitivity/stability commonly used in perturbation theory. As a result, not only is the measure of computational error directly interpretable in terms of modeling error, but the analysis of the quality of solutions mirrors the standard methods of perturbation theory for dynamical systems, including systems studied in physics, chemistry, biology, economics, etc.

5 Conclusion

As we have seen, backward error analysis is a form of error analysis that permits us to substitute all sources of variation in the solution of a problem by an equivalent perturbation in the input of the original problem, and this whether or not the exact solution is known. Thus, the computational error is mathematically equivalent to a modeling error in the first sense. Accordingly, the further task of integrating those methods within philosophy of science does not amount to developing a new metaphysics, epistemology, semantics, or methodology of science. Rather, the task is to better delineate the role perturbative methods play in science, and extract insights for the problems of philosophy of science as they are currently construed (along the line of Batterman (2002b) and Wilson (2006)). As suggested above, this can be done by extending the formal methods of epistemology to make explicit the sense in which concepts from perturbation theory complement the more commonly employed concepts of satisfaction (from model theory) and probability.

However, to achieve this goal, rational reconstructions of scientific practice have to be more sensitive to the mathematical difficulties encountered in practice. Analyzing the nature of “in principle” science is a common gambit in philosophy of science. Nonetheless, some adequacy with the practice of model construction has to be preserved. In this respect, we should remind ourselves that “the assumption that as soon as a fact is presented to a mind all consequences of that fact spring into the mind simultaneously with it [...] is a very useful assumption under many circumstances, but one too easily forgets that it is false” (Turing 1950). In real as opposed to “in principle” science, the fact that scientists are swamped with intricate computational complications cannot be disregarded. Without efficient numerical and analytic approximation schemes, there is no explanation and no prediction, no empirical or pragmatic success to boost our confidence in the correctness of the models and of the general laws. In this respect, *in principle* science has no empirical nor theoretical grounds. Hence, our rational reconstruction of theories should include the concepts of computational error analysis.

discuss the concept of well-conditioning, Duhem (1906) has an extended discussion of “les mathématiques de l’à peu près” based on Hadamard’s work.

Moreover, as we have argued, it should not only include them, but it should also explicate the way in which they relate to modelling error. Accordingly, rational reconstructions should focus not only on what theories are, and on what models are, but also on how models are *constructed* within theories by deriving equations from modeling assumptions, and how different modeling assumptions compare with respect to their solutions. The formal methods will be extended adequately only if they can explain the way in which scientific arguments and representations are *effectively* constructed, and how the derived model equations are *effectively* solved. Accordingly, the point of the epistemology of sciences is not to try to counterfactually understand how science would be without errors and uncertainty, but rather the point is to understand how we can live with them. For, as Russell (1954) put it, “[a]lthough this may seem a paradox, all exact science is dominated by the idea of approximation.”

Acknowledgments First and foremost, we would like to thank Robert Batterman. We would also like to thank Erik Curiel, Bill Harper, Robert Moir, Chris Pincock, Bryan Roberts, Chris Smeenk, and two anonymous referees for their useful suggestions.

References

- Barberousse, A., Franceschelli, S., & Imbert, C. (2009). Computer simulations as experiments. *Synthese*, 169(3), 557–574.
- Batterman, R. W. (2002a). Asymptotics and the role of minimal models. *British Journal for the Philosophy of Science*, 53, 21–38.
- Batterman, R. W. (2002b). *The devil in the details: Asymptotic reasoning in explanation, reduction, and emergence*. Oxford: Oxford University Press.
- Bender, C., & Orszag, S. (1978). *Advanced mathematical methods for scientists and engineers: Asymptotic methods and perturbation theory* (Vol. 1). New York: Springer.
- Butcher, J. (2008). Numerical analysis. *Journal of Quality Measurement and Analysis*, 4(1), 1–9.
- Carnap, R. (1928). *The logical structure of the world* (R. A. George, Trans., 1967). Berkeley: University of California Press.
- Corless, R. M. (1994). Error backward. In P. Kloeden & K. Palmer (Eds.), *Proceedings of chaotic numerics*, Geelong, 1993, volume 172 of *AMS contemporary mathematics* (pp. 31–62).
- Corless, R. M., & Fillion, N. (2014). *A graduate introduction to numerical methods, from the viewpoint of backward error analysis*. New York: Springer. <http://www.springer.com/mathematics/computational+science+%26+engineering/book/978-1-4614-8452-3>
- Deuffhard, P., & Hohmann, A. (2003). *Numerical analysis in modern scientific computing: An introduction* (Vol. 43). New York: Springer.
- Duhem, P. (1906). *La Théorie physique: Son objet et sa structure*. Paris: Chevalier & Rivière.
- Earman, J. (1986). *A primer on determinism*, volume 32 of *The University of Western Ontario series in philosophy of science*. Dordrecht: D. Reidel Publishing Company.
- Earman, J., Roberts, J., & Smith, S. R. (2002). Ceteris paribus lost. *Erkenntnis*, 57, 281–301.
- Fillion, N. (2011). *Backward error analysis as a model of computation for numerical methods*. Master’s thesis, The University of Western Ontario, London, ON.
- Fillion, N. (2012). *The reasonable effectiveness of mathematics in the natural sciences*. PhD thesis, The University of Western Ontario, London, ON.
- Hartmann, S. (1996). The world as a process. Simulations in the natural and social sciences. In R. U. M. Hegselmann & K. Troitzsch (Eds.), *Modelling and simulation in the social sciences from the philosophy of science point of view* (pp. 77–100). Dordrecht: Kluwer.
- Higham, N. J. (2002). *Accuracy and stability of numerical algorithms* (2nd ed.). Philadelphia: SIAM.
- Humphreys, P. (2004). *Extending ourselves: Computational science, empiricism, and scientific method*. Oxford: Oxford University Press.
- Humphreys, P. (2009). The philosophical novelty of computer simulation methods. *Synthese*, 169(3), 615–626.

- Joint Committee for Guides in Metrology. (2008). *Evaluation of measurement data—Guide to the expression of uncertainty in measurement*. Technical report JCGM 100:2008, Bureau International des Poids et Mesures. Revised Edition of GUM 1995.
- Kelly, K. (1996). *The logic of reliable inquiry*. Oxford: Oxford University Press.
- Morrison, M. (2009). Models, measurement and computer simulation: The changing face of experimentation. *Philosophical Studies*, 143, 33–57.
- Oberkampf, W., Trucano, T., & Hirsch, C. (2004). Verification, validation, and predictive capability in computational engineering and physics. *Applied Mechanics Review*, 57(5), 345–384.
- Parker, W. S. (2010). An instrument for what? Digital computers, simulation and scientific practice. *Spontaneous Generations: A Journal for the History and Philosophy of Science*, 4(1), 39–44.
- Pincock, C. (2012). *Mathematics and scientific representation*. Oxford studies in the philosophy of science series. Oxford: Oxford University Press.
- Putnam, H. (1991). The “corroboration” of theories. In R. Boyd, P. Gasper, & J. D. Trout (Eds.), *The Philosophy of science* (pp. 121–137). Cambridge, MA: MIT Press.
- Reichenbach, H. (1938). *Experience and prediction: An analysis of the foundations and the structure of knowledge*. Chicago: The University of Chicago Press.
- Roache, P. (2001). Code verification by the method of manufactured solutions. *Journal of Fluids Engineering*, 124(1), 4–10.
- Russell, B. (1954). *The scientific outlook* (2nd ed.). London: George Allen & Unwin Ltd.
- Salmon, W. (1970). Bayes’s theorem and the history of science. In R. Stuewer (Ed.), *Historical and philosophical perspectives of science*, Vol. 5 of *Minnesota studies in the philosophy of science* (pp. 68–86). Minneapolis: University of Minnesota Press.
- Smith, S. R. (2001). Models and the unity of classical physics: Nancy Cartwright’s dappled world. *Philosophy of Science*, 68(4), 456–475.
- Smith, S. R. (2002). Violated laws, ceteris paribus clauses, and capacities. *Synthese*, 130, 235–264.
- Stein, H. (1995). Some reflections on the structure of our knowledge in physics. *Studies in Logic and the Foundations of Mathematics*, 134, 633–655.
- Taylor, B. N., & Kuyatt, C. E. (1994). *Guidelines for evaluating and expressing the uncertainty of NIST measurement results*. Technical report NIST technical note 1297, National Institute of Standards and Technology.
- Thagard, P. (1993). *Computational philosophy of science*. Cambridge, MA: The MIT Press.
- Trefethen, L. N. (1992). The definition of numerical analysis. *SIAM News*, 25, 6–22.
- Turing, A. (1950). Computing machinery and intelligence. *Mind*, 59(236), 433–460.
- Von Neumann, J., & Goldstine, H. (1947). Numerical inverting of matrices of high order. *Bulletin of the American Mathematical Society*, 53(11), 1021–1099.
- Wang, Q. (1990). The global solution of the n -body problem. *Celestial Mechanics and Dynamical Astronomy*, 50(1), 73–88.
- Wilkinson, J. H. (1971). Modern error analysis. *SIAM Review*, 13(4), 548–568.
- Wilson, M. (1998). Mechanics, classical. In E. Craig (Ed.), *Routledge encyclopedia of philosophy*. London: Routledge.
- Wilson, M. (2006). *Wandering significance: An essay on conceptual behaviour*. Oxford: Oxford University Press.
- Winsberg, E. (2009). Computer simulation and the philosophy of science. *Philosophy Compass*, 4, 835–845.