Varieties of Limits to Scientific Knowledge

Piet Hut
Institute for Advanced Study
Princeton, NJ 08540, U.S.A.
email: piet@ias.edu

David Ruelle
Institut des Hautes Etudes Scientifiques
91440 Bures sur Yvette, France
email: ruelle@ihes.fr

Joseph Traub[†]
Computer Science Department
Columbia University
New York, NY 10027, U.S.A.
email: traub@cs.columbia.edu

Abstract

Any area of knowledge is structured by an intricate interplay of limits, intrinsic to that area. Limits have two opposite functions: setting apart and joining. An apparently restrictive limit may actually reveal itself as liberating. We discuss the role that limits play in the real world, in our mathematical idealizations, and in the mappings between them. We propose a classification of limits, and suggest how and when limits to knowledge appear as challenges that can advance knowledge.

[†]) External Professor, Santa Fe Institute

1. Introduction

Studying the limits of knowledge is an effective way to increase knowledge. Charting the boundary between what we do and do not know can enrich science.

A recent burst of activity, centered around the notion of limits to scientific knowledge, has led to a series of workshops, papers, and books (1-3). In this paper, we take a critical look at the notion of limits, and the role they play in scientific research.

1.1. Intrinsic Limits

At any given time, our body of scientific knowledge has limitations that reflect the temporary endpoint of investigations so far. In general, we say that scientific research is conducted at the 'frontiers' of what is known. Expanding the frontiers is indeed an apt metaphor for the growth of scientific knowledge.

In contrast to those ever-expanding outer limits to scientific knowledge, temporary markers for work-in-progress, we may wonder if there are not more permanent frontiers, or holes that in principle cannot be filled in. In other words, we are interested in the question of whether there are intrinsic limits to scientific knowledge.

What can we say more generally about such intrinsic limits? Can we construct a framework within which to discuss these types of limits? The present paper presents an initial attempt to come to grips with these questions, and to point out some possible steps towards a classification of limits.

1.2. Boundaries as Bridges

Scientific progress was considered to be virtually unlimited, even unstoppable, a hundred years ago. Such a positive outlook has given way, in this century, to a more sober reflection on various factors that may limit further growth of scientific knowledge. Economic considerations as well as concern for the environment are two important factors, and we may be reaching technological limits as well.

While it is notoriously difficult to predict technological breakthroughs (scientists, by nature conservative, do not have a good track record in this respect), there seem to be some fundamental limitations that stand in the way of further progress. One example is the finite speed of light, that is likely to provide an increasingly severe barrier for further developments in computer building, and thus in our ability to simulate complex systems.

However, a barrier blocking progress in one direction can also be seen as an invitation to explore other directions. Thus, quantum computation (4) may bypass the limits posed by classical physics. And totally different avenues, such as DNA computing (5) may also offer new possibilities. Historically, attempts to prove that something is impossible, as in the no-go theorems in physics (6), have blocked progress in certain areas for a considerable time by discouraging researchers to even enter those areas, until someone discovered a fatal flaw in some of the unstated assumptions.

In general, any attempt to prove that something is impossible is at the same time an invitation to look for loopholes, hidden assumptions in such a proof, in order to find ways around the difficulties. As we will argue in more detail below, the limits inherent in a body of knowledge are intimately interwoven with the internal structure of that body of knowledge (7), and a deeper understanding of limits therefore leads to a deeper understanding of the matter under consideration.

2. Transcending Limits

2.1. Examples from Mathematics

A spectacular form of intrinsic limit to knowledge was the shocking incompleteness theorem announced by Gödel in 1931: in a formal axiomatic description of a part of mathematics (if not so simple as to be uninteresting, and if free from contradictions) there are statements that can neither be proved nor disproved (8). This applies in particular to formalizations of arithmetics.

Gödel's discovery thwarted Hilbert's program of reducing all of mathematics to mechanical procedures, and seemed to place severe limits on what can be known in mathematics. However, now that we are getting used to incompleteness results, it is not clear whether we really should interpret them as an indication of limits, as something that walls off areas of knowledge to the human investigator.

Instead of trying to answer this question immediately, let us travel back in time, by a couple thousand years, to see whether a more removed historical perspective may be of some help. Let us take, instead of Hilbert's program, Pythagoras' program of understanding the world in terms of properties and interrelations of the natural numbers (we refer to the usual tale; historical accuracy is difficult to assess, and not relevant for our example).

For a while, Pythagoras' program seemed to be sailing along quite nicely. He and his followers also built up an impressive body of mathematical knowledge, and they were able to explain quite a lot in applied areas such as music. But then they discovered that the square root of two 'did not exist as a number' (i.e. is not a rational number), a deeply shocking result. With such a simple number as the length of the diagonal of a square with sides of unit length being 'off limits', mathematics itself seemed to be unexpectedly limited.

Of course, nowadays nobody would consider this result to be in any way a limit. Rather, it is part of the structure of mathematics that $\sqrt{2}$ is irrational, just as $\sqrt{-1}$ is an imaginary number (9).

2.2. Examples from Physics

Let us take a typical example of what strikes us as a severe limit in a physical theory: the fact that the finite speed of light poses an upper limit on the speed of any material (as well as informational) object. Starting from Newton's classical mechanics, in which velocities of any magnitude are allowed, the existence of an upper bound on possible velocities comes as a surprise, and strikes one as a limitation.

However, if we had first learned to think in terms of general relativity theory, and only afterwards tried to understand Newtonian gravity and mechanics in terms of differential geometry, the natural language of general relativity, an altogether different picture of what is limited would have emerged. In this approach, Newtonian gravity is full of limits, in the form of seemingly arbitrary and complicated restrictions, compared to the far simpler way that general relativity can be formulated in geometric terms (10).

Similarly, quantum mechanics seems strikingly incomplete, from the view point of classical mechanics. The uncertainly relation, for example, puts unexpected limits on what can be measured simultaneously. But if we turn the tables, and start with a formalism that is more naturally suited to a quantum mechanical description, it is classical mechanics that is glaringly 'incomplete' and thereby limited, in that it simply omits processes such as quantum mechanical interference (11).

In addition to the fact that some limits are thus seen to be in the eye of the beholder, it is quite possible that some local limits may be circumvented globally. Wormholes in spacetime are an example, through which it might be possible to communicate nearly instantaneously over (seemingly) large distances, without violating the speed limit imposed locally by special relativity (12).

2.3 Boundaries in Mathematics leading to Bridges to new Physics

When physical theories lead to solutions with singularities (for instance infinities), this generally means that the theory breaks down and that new phenomena appear, requiring a new theory. Thus the phenomenon of *breaking of waves* occurs when a simple theory of wave propagation leads to singular solutions.

3. Mathematics, Physics, and Reality

A natural classification of limits to scientific knowledge offers itself: a distinction between limits that show up in our models of the world, versus limits that apply to the world itself, as we know it through observation and experimentation. Within the physical sciences, the models are generally of a mathematical nature. In other areas, such as archaeology, coherent explanations for observational data, while having their own logical structure, are typically not cast in mathematical language.

A more careful consideration, however, throws doubt upon the validity of the model/world distinction. Let us look in some detail at the modeling process that is at the heart of any form of scientific activity.

Science proposes theoretical schemes to describe the 'real world'. The connection with the real world involves an operational description of how observed facts fit into a theoretical scheme. In other words, the challenge is to build increasingly more accurate (usually mathematical) models of the world, together with recipes to connect those models with experiments and observation.

Suppose for instance that we are interested in electric circuits: their mathematical theory is fairly simple, but has to be supplemented by recipes for measuring voltages, resistances, capacitances, *etc.* in the lab. Similarly, land surveying has a simple theory (basically Euclidean geometry), which has to be supplemented by the use of geodesy instruments. Incidentally, these examples call for the remark that one needs the theory of electric circuits (resp. Euclidean geometry) to understand how voltmeters, *etc.* (resp. geodesy instruments) function. What we call recipes may thus be modified or reformulated by use of the associated theory. One can however never completely remove the operational connection between physical world and mathematical model. Because this operational connection, which we call recipes, has no purely mathematical formulation, it tends to be a bit messy, and to be swept under the rug in many formulations of scientific theories (13).

Let us come back to limits, and try to ascribe them to three different types: those in nature, i.e., the real world; those in the models we have built; and those limits that are inherent in our translation recipes. For definiteness, we can take the example of physics, for which these three areas are indicated in figure 1.

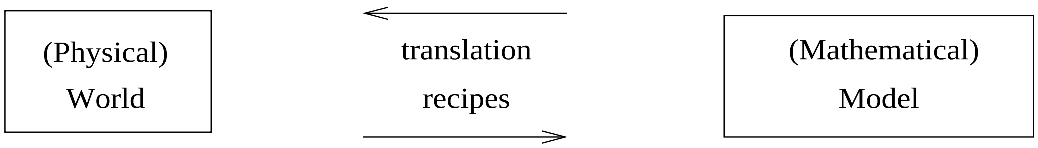

Fig. 1. The three realms of physics.

What makes fundamental progress in physics hard is often not so much the inherent difficulties that are encountered within specific mathematical models, but rather the challenge of finding the proper choice of the pair {translation recipe, model}.

The question of the status of the translation recipes is something that is rarely addressed specifically in academic science courses. Their very existence is often glossed over, and it is not even clear whether they are part of physics of mathematics. If we start from the side of physics, and ask which part of fig. 1 applies to mathematics and which part to physics, we might draw a picture as given in fig. 2a.

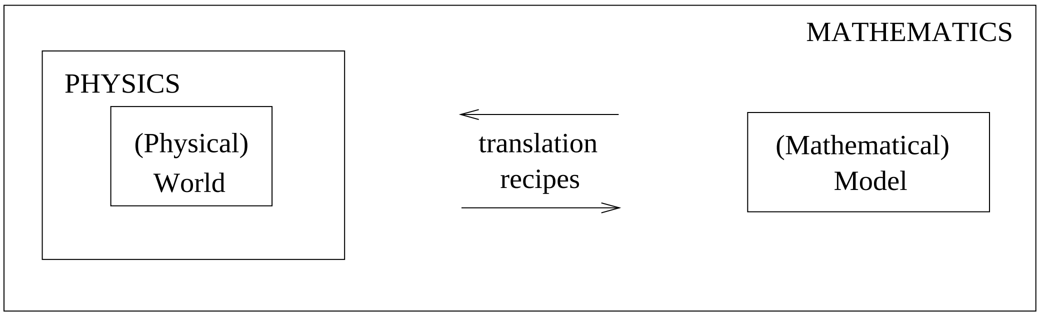

Fig. 2a. A view from physics.

In this figure, the natural world around us is assigned to the domain of physics. Clearly, translation recipes are not part of this natural world. Rather, they are designed by us in order to fit with our models. One might therefore argue that they should be grouped under the general category of mathematical tools.

But as soon as we start from the other side of fig. 1, from the side of mathematics, it is clear that the mathematical models and frameworks are completely self-contained, and do not even leave room for the existence of translation recipes. For example, the notions of ruler and compass may not be part of the fields being measured by a surveyor, but at least they are made from physical material. They are certainly never to be found within Euclidean geometry: they are neither axioms, nor theorems, nor any other form of mathematical entity. This would suggest a classification as given in fig. 2b.

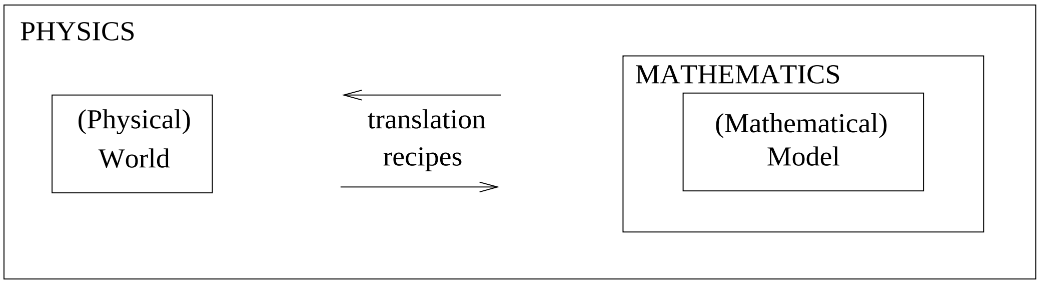

Fig. 2b. A view from mathematics.

Faced with these alternatives, depicted in figs. 2a and 2b, neither of them feels quite right. But perhaps the real problem lies elsewhere. Maybe it is not just the uneasy status of the translation recipes that resists classification. Where else can we look? The mathematical models as such seem to be least problematic. But what about the status of the physical world?

One possibility, in a variation on Kant, would be to declare the physical world, as the realm of the 'things in themselves', off-limits to direct scrutiny. The phenomena we deal with are accessible to us only through the particular ways we choose to investigate them. So we might even argue that we have to shift the legitimate terrain of physics towards a middle position, as illustrated in fig. 3. This may seem a bit extreme, but as we will see below, it can gives us an interesting handle on the status of limits.

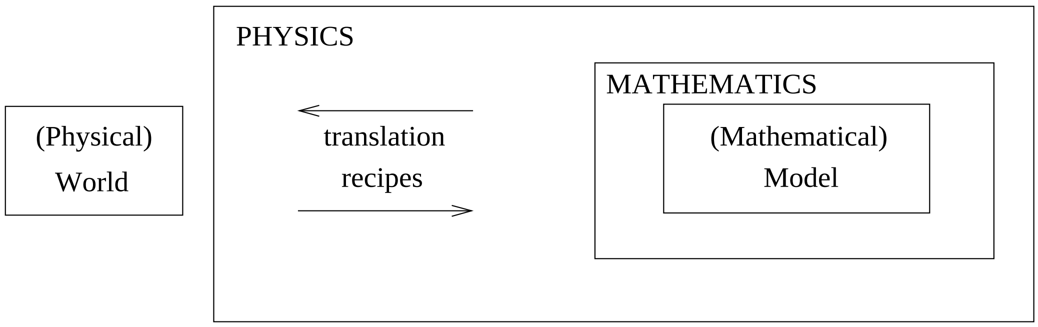

Fig. 3. From ontology to epistemology

We are now in a position to return to the question that led us to this classification: how to distinguish between limits that show up in our models of the world, versus limits that apply to the world itself. In other words, in contrast to scrutinizing models and recipes, what can we say about the existence of limits in the real world?

Questions about the real world do not come equipped with a mathematical model. Examples of questions about the real world include:

- Will there be major global climate changes due to human activities?
- Will the Universe stop expanding?
- How did life originate on Earth?

We get to choose the mathematical model. To rigorously establish a scientific limit we should show that *every* mathematical model that captures the essence of a scientific question is undecidable or intractable (2).

This may be a possible attack in principle, but it is far from evident that it could actually be carried out (Note, however, that in establishing the computational complexity of a mathematical model we do permit *all* possible algorithms to compete).

In scientific practice, all that we encounter are measurements and interpretations, together with applications that in turn are inspected through further measurements. Therefore, if we critically ask what physics deals with, we are really forced to the conclusion depicted in fig. 3.

Does this leave the real world outside of physics? Actually, what is left outside of physics are ontological questions about the real world (14). But the real world is there somehow, and limits the physically acceptable pairs {recipe, model}. Unacceptable pairs are falsified by experiments.

4. A Classification of Limits

If one starts listing limits to knowledge, one soon finds that the items are heterogeneous, but can be fit in categories that we now discuss (we leave out psychological considerations, such as the question to what extent human intelligence may be a limiting factor). The categories overlap somewhat, i.e. some items fit in several categories.

1) The Curse of the Exponential

There are problems that can be solved in principle, but are in fact intolerably hard.

Chaos is an example, where the time evolution of a physical system is assumed to have a good mathematical model, but the actual calculation of the time evolution is limited by the exponential growth of errors (important special cases are weather prediction and the astronomy of the solar system). More generally the problem of solving with suitable accuracy the equations proposed by physical theories may often be practically insuperable.

In computational complexity theory one finds that the running time of any algorithm for solving certain problems increases unacceptably fast with the length of the input (exponentially or worse). For discrete problems, the conjecture $P \neq NP$ suggests a class of intractable problems. On the other hand, many continuous problems have been proven exponentially hard in the number of variables (Note, however, that the conjecture and theorems concern the worst case setting. For some problems intractability can be broken by switching to a stochastic assurance). (15)

2) Asking the Wrong Questions

There are cases where a question one may want to ask can be shown to have no answer in the framework of a given theory (wrong question). In other words, there are structural limitations to the questions one may ask.

In quantum mechanics for instance one should not ask to specify simultaneously the position and momentum of a particle.

The work of Gödel and Turing has shown that it is a bad idea to seek an algorithm deciding whether an arbitrary statement is true or false.

There is a curious relation between category 1) and category 2): the fact that the halting problem for a Turing machine is undecidable can also be expressed by saying that the halting time is not a constructible function. In other words it grows very, very fast with the length of the input: faster than exponential, exponential of exponential, *etc.*

What we now consider to be wrong questions were not always such in the past. It was once reasonable to ask how $\sqrt{2}$ could be written as a rational fraction p/q, or how some algorithm could solve all mathematical problems (Hilbert's dream). The understanding of structural limitations that make such questions unanswerable is scientific progress.

3) Questions of Questionable Status

In physics one often does not have a completely reliable theory, and it is thus unclear if a natural question that appears hard to answer corresponds to a fundamental limitation. A seemingly hard question (such as 'what is the nature of phlogiston?') may turn out to have questionable status, and can eventually evaporate. To investigate whether a limitation is fundamental may thus lead to important theoretical progress. Currently in this category are problems related to cosmology and quantum gravity.

Another example concerns the consistency of arithmetic. Gödel proved that: 1) arithmetic cannot be proved internally to be consistent; and 2) that it is either inconsistent or incomplete. Actually, arithmetic could be inconsistent, but a proof of contradiction would then probably be extremely long, perhaps so long that it could not be implemented in our physical universe (16). Thus even a very fundamental and natural question such as the one concerning the consistency of arithmetic appears upon reflection to be of more questionable nature than one might have liked.

4) Emergent Properties

Another class of very difficult questions, which seem to lead to fundamental limitations of knowledge arise in the study of emergent properties. One can construct hierarchies of systems where each level can in principle be understood in terms of the level below, such as:

??? $<$ quantum field theory $<$ atoms $<$ molecules $<$ life.

In practice, however, we are very limited in understanding the higher levels in terms of the lower levels. In addition, emergent properties in general are fuzzy. This in turn creates difficulties in classification. An example of problematic aspects of this fuzziness can be found in discussions of the anthropic principle, which tend to be inconclusive, simply because we know very little about life forms unrelated to us.

Surprisingly many systems with large numbers of degrees of freedom allow descriptions in far simpler terms, with properties that are not at all obvious from the underlying system. These emergent properties point towards a higher level of organization, exhibiting less complexity than could be anticipated. (Coarse graining is one example of a simplifying process that can uncover emergent properties). But higher levels are not guaranteed to have simpler behavior (and not much has been found, for example, in fully developed turbulence).

5) Limited Access to Data

In some areas of science, the absence of sufficient data leads to severe limitations to knowledge. Historical limitations are in this category (3). We may be able to infer many things about the origin of life, but it appears that we shall never have a detailed story of what really happened. Similar statements may be made about the origin of stars, of mankind, of languages, *etc.*

6) Sample Size of One

Cosmology, the study of the origin of the Universe and its overall structure, is an example of a discipline that deals with a sample size of one. Comparing theory and observation thus becomes more difficult, since a typical theory for the large-scale structure of the Universe predicts numbers that have a statistical uncertainty attached to them (17).

No matter how well we will ever be able to measure the particular realization we live in, this may not enable us to check whether our theory of Universe formation is correct. Certainly any Popperian notion of falsifiability does not apply, as long as we cannot create other Universes to check our theories with. But then again, it is always risky to declare something to be 'off limits', and recently various ideas have been developed concerning 'multiverses' (18).

7) Technological Limitations

They are practically important, but one should be wary of treating as fundamental some limits that can perhaps be overcome by new ideas. For instance, physics at the Planck energy appears very much out of present technological reach, but this does not mean that we shall never understand what is going on in this energy region.

5. Discussion

Limits combine two opposite functions: setting apart and joining. They partition the world (in fact, all that appears, in any form) into separate areas, in intricate and overlapping hierarchies. But at the same time they structure the interrelationships and communication channels between the pieces into which they seem to have carved up the world.

Examples:

- a river or mountain separates and connects two areas;
- cell walls in biology allow the build-up of complex hierarchical structures;
- $\sqrt{2}$ was first seen as not-a-number, later as a different type of number, where rationals and irrationals were joined as numbers but separated by their (ir)rationality property;
- the uncertainty principle: the very fact that position and momentum, for example, are not simultaneously measurable shows the unified nature of quantum reality.

Each limit at first presents itself as a barrier, as something that sets apart.

Further inspection then shows the connecting, bridging, aspect of that limit. Thus each limits provides us with a challenge: to look past its obstacle character, in order to discover its key role in disclosing new areas of knowledge.

Notes

[1] J. L. Casti & J. F. Traub, Eds., *On Limits*, Report, Santa Fe Institute, Santa Fe, NM (1994); J. Casti & A. Karlqvist, Eds., *Boundaries and Barriers* (Addison-Wesley, Reading, MA, 1996); J. B. Hartle, P. Hut & J. F. Traub, Eds., Proceedings of Workshop on *Fundamental Sources of Unpredictability*, to appear in *Complexity* (1997).

[2] J. F. Traub, What is scientifically knowable?, in *Twenty-fifth Anniversary Symposium*, School of Computer Science, Carnegie-Mellon University (Addison-Wesley, Reading, MA, 1991), pp. 489-503.

[3] R. Gomory, *Scient. Amer.*, June 1995, 120.

[4] D. P. DiVincenzo, *Science* **270** (1995), 255.

[5] L. M. Adleman, *Science* **266** (1994), 1021.

[6] M. Kaku, *Quantum Field Theory* (Oxford: Oxford Univ. Pr., 1993), p. 57.

[7] P. Hut, in *Boundaries and Barriers*, J. Casti & A. Karlqvist, Eds. (Addison-Wesley, Reading, MA, 1996), pp. 148-187.

[8] K. Gödel, Monatsh. Math. Phys. **38** (1931), 173.

[9] R. Rosen, in *Boundaries and Barriers*, J. Casti & A. Karlqvist, Eds. (Addison-Wesley, Reading, MA, 1996), pp. 199-214.

[10] C. W. Misner, K. S. Thorne & J. A. Wheeler *Gravitation* (W. H. Freeman, San Francisco, 1973), Ch. 12 and especially Section 12.5.

[11] D. Finkelstein, *Quantum Relativity* (Springer, New York, 1996), Chapter 3 and especially Section 2.4.6.

[10] C. W. Misner, K. S. Thorne & J. A. Wheeler *Gravitation* (W. H. Freeman, San Francisco, 1973), p. 1200.

[13] This point is discussed in D. Ruelle, *Chance and Chaos* (Princeton Univ. Pr., Princeton, 1991), Chapter 2.

[14] D. Finkelstein, Quantum Relativity (New York: Springer, 1996).

[15] J. F. Traub and H. Woźniakowski, Breaking Intractability, *Scientific American*, Jan. 1994, pp. 102-107.

[16] P. Cartier drew our attention to this possibility.

[17] M. White, D. Scott & J. Silk, *Ann. Rev. Astron. Astroph.* **32**, 319 (1994), especially p. 354.

[18] M. Rees, *Before the Beginning: Our Universe and Others* (Addison-Wesley, Reading, MA, 1997).