

Undecidability: The No-Man's Land between Theory and Experiment

MATTY J HOBAN

Department of Computer Science, University of Oxford, UK

March 3, 2015

Abstract

Theoretical Computer Science was one of the great intellectual developments of the 20th Century and only very recently have its ideas been integrated into basic science. I argue that these ideas and tools could be vital to understanding the relationship between mathematics and physics. In particular, if we are to ask if the utility of mathematics is a trick or otherwise, then we can use methods from the theory of cryptography and computation to pitch theory and experiment up against each other. Given this perspective we can ask if an experimenter can convince us to accept the theorist's version of events. I will outline that this contest is not as simple as it initially seems. There could be a stalemate between theory and experiment and it comes in the form of undecidability: a theorist can neither confirm or deny if Nature is behaving according to our most-trusted physical theories. Surprisingly, undecidability is a very real prospect in our current theories as I will discuss, and it forces us to question the relationship between mathematics and physics.

Introduction

Modern physics is a collaborative effort. The scientific ideal being that theorists model physical systems and predict the outcomes of experiments and the experimentalists confirm or refute these predictions. During this whole process, mathematics makes up a significant part of our toolbox of scientific discovery from the inception of the theoretical idea to analysis of data and errors.

It is now also clear that computers are an indispensable part of modern physics. Not only are they utilised for repetitive and difficult calculations as well as collecting and analysing experimental data, elements of theoretical computer science have entered into the sphere of theoretical physics. For example, the language of computational complexity can indicate how hard a physical problem is to solve.

The fact that our physical theories are phrased mathematically allows us to translate ideas and practices from one field to another. This occurs even if these fields initially had nothing to do with each other. It is also true that the relationship between computer science and physics can be a dialogue where the latter could potentially challenge key concepts in the former. One key example is the emergence of quantum computing which seems to question the long-held convictions summarised in the extended Church-Turing thesis: any physically realizable computation can be efficiently simulated by a probabilistic Turing Machine [1].

This interplay between computer science and physics is an intriguing development in basic science. I want to argue that if we are to understand the relationship between mathematics and physics then ideas in computer science must be incorporated. This young field not only gives us the language in which one can ask about tricks and truths but we can also obtain answers methodically. Just like learning any new language, if physics learns more about computer science then a whole new perspective presents itself.

If we are being tricked by Nature into accepted a particular picture of the world, then the 20th Century gave us cryptographic tools to protect ourselves from such deception. In particular, the interactive proof system (IPS) is the perfect platform for the task of challenging claims made by powerful agents or individuals [2]. In an IPS, there are two parties which we'll call Alan and Kurt as shown in Figure 1. Kurt is insanely intelligent and powerful, so much so that when he says something, it is so beyond our grasp that we cannot immediately be sure that it is true. We wish to ascertain if Kurt's claims are trick or truth. Alan, on the other hand, is cunning and methodical, so he devised a machine that can try to systematically check if Kurt is telling the truth. The computational power lies in not only having access to solutions of very hard problems but also being able to check these solutions.

We can now construct an IPS for science [3]. In this scenario, Kurt is now some experiment producing data and Alan wishes to confirm that the predictions made by his physical theory are consistent with this data. I want to reiterate that the tools developed for IPS's were developed *independently* of physics but in the context of studying information processing. This independence allows us to ask questions about the *process* of doing science and not just the results themselves. For example, following Aharonov and Vazirani in [4] we can ask the question, 'is quantum mechanics falsifiable?' Since quantum theory says that the dimension of a quantum system scales exponentially in the number of its sub-systems it may get extremely hard to make predictions for large systems, or what the authors call the "high complexity limit":

"Moving beyond the practical difficulties in experimentally dealing with large scale quantum systems ... we ask whether testing [Quantum Mechanics] in the high complexity limit is even theoretically possible, or whether there are fundamental obstacles that prevent such testing." [4]

This question is a tantalising one and great theoretical progress has been made in how one could certify quantum behaviour [5]. I want to go beyond the issue of complexity towards the question of computability. In particular I want to address the subject of decidability: given a theory and a description of an experiment, can we always compute in a finite amount of time what will happen in that experiment? The computation could take a second, a week or a century but it should still stop after some amount of time.

Returning to the IPS, imagine that Kurt is asked to perform an experiment and then tells Alan what the experimental outcomes were. Recall that Kurt might want to trick us into accepting a lie as truth. Can we confirm that the sequence of outcomes that he produced are compatible with the experiment we suggested to him? More generally, can we always predict that certain outcomes will occur or not?

Remarkably, given our current theories this is not always possible as demonstrated in very recent research [6]. However, experimentally we could observe such an outcome occurring, but does this mean that our theory is no longer useful if it cannot make particular predictions? Has the traditional backwards and forwards of prediction and confirmation been disrupted by the possibility of a "don't know"? This is now a fascinating third way for the relationship between theory and experiment that has the potential to undermine the hitherto comfortable marriage between maths and physics.

Importantly, the notion of decidability that we will discuss here is explicitly about performing computations and not necessarily about making deductions from a set of logical axioms. The ground-breaking work of Gödel falls into this category of undecidable [?]. The notion of decidability that will concern us is the one that led Turing to develop his now-synonymous machine that abstracts the work of mathematician making calculations [8]. The Turing Machine in time has been generalised to allow non-deterministic computations (such as quantum computations) where a machine does not have a unique, determined calculation but many

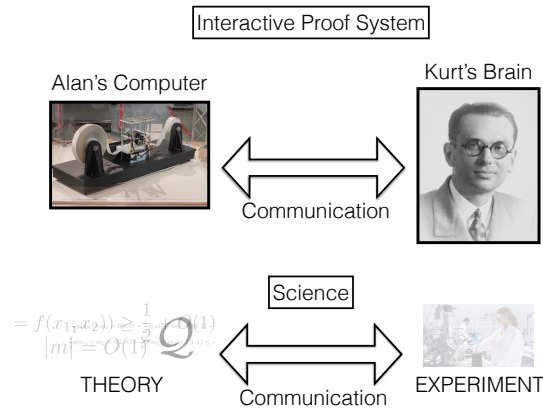


Figure 1: An IPS allows two parties to communicate (we call them Alan and Kurt). Kurt has a very powerful brain and claims to give solutions that are compatible with a computational problem Alan has but Kurt is so powerful that he cannot be trusted to use his power for good. He might want to trick Alan. Alan, on the other hand, has a machine that can verify if the solution is correct and thus tell if he is being tricked. It has very recently been suggested that the scientific process can be viewed as an interactive proof where we want to check if a particular experiment is being implemented. One can imagine that Nature herself is now the all-powerful Kurt and we wish to see if Nature's workings are compatible with our best physical theories. Perhaps Nature could even trick us into accepting a particular (mathematical) view of the world. A fundamental question is if we can always calculate if particular experimental outcomes are compatible with our current theories. [Image of Turing Machine by Rocky Acosta (2012)]

possibilities.

We may obtain the tools to fundamentally probe Nature but by results in computability theory we may also find that some questions just cannot be answered by our theories. To avoid this, could we now demand that decidability be a foundational principle for science? This could create an inconsistent mathematical framework but could it also create a consistent scientific framework?

I will outline these fundamental concepts in computer science and show how they behave when exposed to our current physical theories. I want to show that there can be limitations to the relationship between theory and experiment and if we are to challenge the role of mathematics in physics, there is no better place to start than with these limitations.

The Finite and the Infinite

The casual reader who has tried to download anything will be fully aware that a computer has finite storage capacity. It is common knowledge that information in computing is broken down into its constituent bits. However, our theories are often described in terms of continuous parameters such as time and space. Since these parameters can take continuous values the difference between them can be infinitesimally small. Wouldn't a computer then need an infinite amount of storage space to even write down the value of a particular variable? If we have to store an infinitely large number then no wonder a computer cannot perform certain calculations in a finite amount of time.

Anyone who has done some computer programming with objects such as floating point numbers will be aware that methods exist that *approximate* arbitrary numbers that aren't just integer values. These approximations only store a finite number of bits and thus pose no problem for a computational device. We will assume that the values of parameters in our theories are approximated by some number that has some finite precision. This approximation will often be a rational approximation where we find the closest fraction to a number. If we take the constant π as an example, it can be approximated by the fraction $22/7$ to a precision of 2 decimal places.

As it turns out later for our later discussion on decidability, we don't need to consider arbitrary numbers. It is possible to show that certain events are undecidable even if the values of the parameters in our theory are rational numbers. Therefore, undecidability need not be about our computer having finite precision, it is more fundamental than this.

Another more practical consideration is that, at the end of the day, we are doing physics. Physics requires experimental observation or validate itself as an empirical endeavour. As is taught to schoolchildren performing their first experiments, our experiments have sources of error and our tools have a finite precision. When taught to measure distances and time intervals, we were also taught to calculate the error in our measurements so that these observations are given some finite precision.

Experimental precision improves year upon year but there is no getting around finite precision. If our computer can operate to a precision even better than our experiments then confirmation or refutation by experiment is compatible with our predictions. Turing saw the computational machine he described in his seminal work as a way of automating the process of mathematicians performing calculations. Whilst a mathematician would feel at home contemplating the continuous and the infinitesimal, the machine would struggle with these concepts. I would argue then that the machine is at home with experimental physics despite the fact that it is a mathematical object. This now leads us to discuss this abstract model in all its glory.

Computation and Turing Machines

What does a mathematician need? A pencil, brain and an endless supply of paper. A Turing Machine replaces each of these elements with some mechanical equivalent [8]. First, we imagine an infinitely long roll of tape that will suffice as our endless supply of paper. We then have the machine that is able to read information written on the tape as well as able to delete and right over this information. Finally, it also has a series of rules that it follows telling the machine where to move, what to calculate and what to read and write. This abstract model encompasses a multitude of forms of computation. That has led to the famous Church-Turing thesis which claims informally that every form of computation can be simulated by a computation on a Turing Machine [9].

The rules of computation in the Turing Machine dictate that what is written on the tape is written in terms of a finite collection of symbols (called an alphabet). Also, the machine can be one of a finite number of possible states. An extremely basic Turing Machine such as a power-switch has two possible states: on and off. In addition we have two special states which are the starting and final states; the former informs the machine that it can begin computing where as the latter tells it to stop. The only thing that is infinite about the machine is its endless tape. Thus, in full abstraction, a machine could go on computing forever on its tape without ever ending up in its final state.

A computational process involves giving a Turing Machine an input sequence, or string, of

symbols and asking it to decide whether this string has a particular property. If the machine can decide whether the string indeed has this property, then it ends up in the final state and indicates its answer. We will return to this point later when we discuss decidability.

Interactive Proofs

In Figure 1 we can see that the computational device that Alan uses is a Turing Machine and so we wish to expand our discussion to incorporate IPS's. Here, we allow Kurt to communicate to Alan, and he can copy the messages onto his tape. The first examples of IPS's were developed in the context of computational complexity. For example, NP is a collection, or class, of computational problems where we can check solutions efficiently and deterministically where P is the class of problems that have efficient deterministic algorithms for solving them [10]. Thus all those problems in NP have an IPS where Kurt provides Alan with potential solutions and Alan can verify them in a short amount of time.

This framework has been expanded to include multiple rounds of communication between Kurt and Alan as well probabilistic computation where Alan is allowed some small error in his computation [2]. Now, as previously mentioned, we can adapt this scenario to the problem of falsifying theoretical predictions. In the direction of certifying whether Kurt has a machine behaving according to the rules of quantum mechanics, impressive results have been obtained [5, 11, 12].

An emerging body of work has shown that if Alan has access to small quantum systems that he can communicate to Kurt, then Alan can certify that Kurt is behaving quantum mechanically [11, 12]. A fascinating consequence of this is also that Alan does not need to simulate Kurt's experiment in its entirety as this may be difficult for him. One can even do this by restricting Alan to having only efficient (probabilistic) classical computation without the need for him to have any quantum system [5]. To do this, we require more-than-one Kurt and these Kurts are now allowed to communicate to each other. However, they can share quantum entanglement, a truly non-classical property of quantum theory where we have imperfect knowledge of individual systems but perfect knowledge about the system as whole. This can lead to quantum systems having stronger-than-classical correlations.

These results are only the beginning of an extremely exciting research direction that could push our existing theories to their limits. For as Aharonov and Vazirani indicate, most of our current experiments only look at particular aspects of quantum theory and do not exploit all of the exotic aspects of the theory [4]. One intriguing open question regarding the testing of quantum systems comes from allowing multiple Kurts to share entanglement. If we allow them to share an infinite amount of entanglement so that they have an infinite source of super-classical resources. Is it still possible for Alan to certify Kurt's experiment? This sort of question is not even known to be decidable in the manner discussed earlier. The short digression leads us nicely to our central theme: decidability.

(Un)decidability

So far I hope I have convinced you that concepts in theoretical computer science can provide new insights into old questions. I also hope I have convinced you that the Turing Machine is a fine model for reasoning about how we make calculations. Because Turing aimed to abstract the work of a mathematician, this machine naturally abstracts the work of a theorist making predictions. Also, I have claimed that issues about finite precision in calculations are non-issues and are in fact very naturally motivated in the context of physics: if we are to simulate Nature,

why not simulate it warts and all. Having set the scene, I want to show that there are still unfortunate weaknesses in the model of computation as introduced by Turing.

As indicated earlier, a Turing Machine has the capacity to go on calculating forever. That is, if after some amount of time it is not in its final state, it will just continue applying its rules of calculation. The computer cannot decide on an output and will just keep dithering for all eternity. One could imagine very simple programs that are just designed to make the machine keep going on forever. It is then important that we consider all possible programs for a particular problem and ask if one amongst them can help us solve a problem. The natural question is then: 'are there natural problems for which a Turing Machine cannot decide on the answer?'

As indicated earlier, the answer is yes and there is now a growing repertoire of questions we can ask that are undecidable (Poonen gives a nice review in Reference [13]). The most striking example of an undecidable problem is a very simple one, the HALTING PROBLEM: 'if given a description of a Turing Machine and a particular input, will the machine halt (end up in the final state in some finite time) on that input?'

The reason why this question cannot, in general, be answered is that if we have a program P that decides halting on a particular input, we can use this to design a new program P' that decides if the machine *does not* halt on the same input. We then ask that program P halts on the input if, and only if, the program P' does not halt on the same input. The clever part is now that we could always give this Turing Machine a description of the programs themselves as an input. Recall that our programs P and P' are assumed to halt and decide for us if the Turing Machine does likewise, therefore, they will be have a finite, mechanistic description. Imagine you are playing a game where a person is being described to you, and you have to say whether you know the person or not. Well, the description that you could be given could be of yourself and you should be able to apply the same reasoning for that description as well as to any other person. Going back to the HALTING PROBLEM, since program P applied to the Turing Machine as an input halts if and only if program P' does not halt when applied to this same input, we have a contradiction. Therefore, neither P nor P' could decide the halting or non-halting in the first place.

I personally view this as an incredible result in the history of basic science. It tells us that given a particular mechanism with a particular input, we cannot be sure that we can give an answer. On the other hand, is the example we give not a little bit contrived and suited to the mathematical constructions of Turing Machines? When it comes to doing proper calculations in physics and other sciences, won't we always just have some nice decidable problem? As I suggested at the outset, undecidability is not a mathematical artefact but can appear in our physical theories. I want to turn to my particular research field, that of quantum mechanics and describe an undecidable process in the foundations of quantum theory.

Quantum Indecisiveness

Imagine a simple system that is just a sequence of observations, one after the other. In quantum theory, we are told that a quantum state is prepared prior to measurement and this is the mathematical description of the system. Then a sequence of observations are quantum measurements applied to the state and, in general, they will affect this state in a dramatic fashion. Notions such as the "collapse of the wavefunction" can be found in many textbooks in the subject. Many of these textbooks focus on a particular kind of measurement: the projective measurement. As well as this, they tend to focus on quantum states of maximal knowledge: the pure states (Peres' book in Reference [14] takes great care to introduce these concepts). For the projective measurements, each outcome is associated with the (probabilistic) preparation of a new pure

quantum state.

Now, given the preparation of a pure quantum state and a description of the sequence of projective measurements, we can predict if certain outcomes will occur. For those familiar with linear algebra and the basics of quantum mechanics, this is a relatively simple task. Students across the world can readily code-up something that calculates the answer to this question (again up to some finite precision). The quantum state can be described as a vector of numbers (these numbers are approximated by rational numbers) and then the probability of obtaining certain outcomes is a function of the inner product between two vectors. We can then view the state after measurement as a new pure quantum state or column vector of numbers (as described by the measurements). Therefore, to see if the next outcome in the sequence can occur, we calculate the inner product again and so on. A Turing Machine, given enough tape, in a finite amount of time can calculate (up to some precision) whether a particular outcome will occur or otherwise.

However, this experimental set-up just described is an idealisation of what actually happens. In this set-up, we have maximal knowledge of the system (described as a pure state) and a series of measurements that are the most-precise that can be envisaged within the theory. Of course we will not always have this maximal knowledge but the theory can cope with a messier state-of-affairs. A pure state can now be a mixed state, which is a probabilistic ensemble of pure states; with some probability we are in a particular pure state but *a priori* do not know which one. Measurements are now not projective measurements, but positive-operator valued measures (POVM) that do not always change the quantum state to a new pure state but could take it to a mixed state. A POVM can be viewed as a noisy measurement, therefore being more relevant for the finite precision of the experimental world.

Given this noisier set-up, can we still predict if certain measurement outcomes will occur? Now the answer is a remarkable 'no', as shown by Eisert, Müller and Gogolin, this problem is undecidable [6]. In other words, if we could predict the outcome of experiments then would could solve the HALTING PROBLEM. But what is more, they also showed that the analogue of this scenario for classical systems is always decidable. Obviously, we must consider finite, discrete parameters in our classical system so that the description of the experiment can be input into a Turing Machine in finite time. Furthermore, as indicated earlier, the parameters describing our measurements and state can be written in terms of rational numbers and the undecidability still holds. It truly is a property of quantum theory as is.

Those overly familiar with quantum theory will recall that every noisy process in quantum theory can be turned into a process with pure quantum states and projective measurements by the so-called Stinespring dilation [15]. It is a dilation because this pure process is now in a much larger system with more degrees of freedom. However, after performing this process in the pure world, we discard these extra degrees of freedom. This is all a mathematical process and not necessarily a physical one. Noise can affect our systems in many different ways. It is interesting that the act of discarding these extra degrees of freedom we end up with a decidable problem even though everything appears decidable in the pure case.

After our diversion into the word of decidability and quantum undecidability, I want to return to our model of IPSs. How can we predict the activities of Nature if we cannot even decide if certain events will occur? Does this render an IPS for Nature ultimately useless? The answer is not completely clear to me and this could be precisely for the reasons alluded to immediately above. Perhaps our view of Nature as implementing an undecidable process is precisely because our information about the experiment is limited and if we could move beyond these limitations, we could resume normal service. In the larger system that is mathematically equivalent to our smaller, noisier system, we could decide what will happen.

I want to address these concerns by mentioning another very recent result obtained about

condensed matter systems [16]. Here our question is not about measurement outcomes but about properties of a physical system. The question is if the energy levels of a physical system with a given Hamiltonian (describing the systems interactions) do or do not have a gap. It turns out this problem is also undecidable in general. This occurs even if there are finitely many systems each with a finite number of dimensions. Also, this process is not about noisy processes but about the ideal mathematical model of a physical system.

It is not clear at all if we will ever be able to avoid undecidability in our physical theories. These remarkable results must be taken into account when we try and understand the relationship between theory and experiment. But what is more exciting, is that only very recently have we become equipped with the tools and perspective to explore this relationship. Now is a thrilling time in the development of basic science.

Conclusion

In the title of this essay I wanted to evoke the image of a stalemate between mathematics and physics. I feel that the issue of decidability is the latest twist in the story of the marriage between these two venerable subjects. I also feel the subject is more than just a mathematical quirk but could force us to re-examine the relationship between theory and experiment. Many ask what good is it to have a physical theory if it cannot give us good predictions? Well, how much more baffling is it that a theory cannot give us any predictions *at all*. Given our current difficulties in extracting testable experimental predictions from potential theories of quantum gravity, one could put this down to the difficulty in constructing experiments where quantum gravity will apply. It would be a perverse turn of events that even if we were to construct these experiments, our theories still could not produce the predictions that we crave.

I briefly mentioned the speculation that we might insist on decidability as a principle for future physical theories. We should explore the consequences of this principle and whether its logical consequence is to give us diverging predictions from those of quantum theory. For example, in some of my previous work with collaborators we indicated that there may be reasonable generalisations of quantum theory that give decidable predictions in an IPS [17]. On the other hand, it seems very unlikely in full generality that we will get decidable predictions from quantum theory.

I have aimed to argue that a deep result in the mathematics of computation could have consequences for how we do physics in the future. Undecidability is not a mathematical curiosity confined to the study of extravagantly contrived computational devices, it can emerge in our best physical theories. In asking whether the relationship between maths and physics is a fundamental truth or just trickery, a more interesting and exciting answer might be that it is fundamentally impossible to decide.

References

1. E. Bernstein and U. Vazirani, "Quantum complexity theory", Proc. of STOC '93, pp. 11-20 (1993)
2. S. Goldwasser, S. Micali and C. Rackoff, "The knowledge complexity of interactive proof systems", SIAM Journal on Computing 18 (1): pp. 186-208 (1989)
3. J. Yaari, Ph. D. thesis, "Interactive Proofs as a Theory of Confirmation", The Hebrew University of Jerusalem (2011)

4. D. Aharonov and U. Vazirani, "Is Quantum Mechanics Falsifiable? A computational perspective on the foundations of Quantum Mechanics", in *Computability: Gödel, Turing, Church, and beyond*, Eds: Copeland, Posy, Shagrir, (MIT) 2012
5. B. W. Reichardt, F. Unger and U. Vazirani, "Classical command of quantum systems", *Nature* 496, pp. 456-460 (2013)
6. J. Eisert, M. P. Müller and C. Gogolin, "Quantum measurement occurrence is undecidable", *Phys. Rev. Lett.* 108, 260501 (2012)
7. K. Gödel, "Über formal unentscheidbare Sätze der Principia Mathematica und verwandter System I", *Monatshefte für Math. und Physik* 38, pp. 173-198 (1931)
8. A. M. Turing, "On computable numbers, with an application to the Entscheidungsproblem", *Proc. London Math. Soc.* (2) 42, pp. 230-265 (1936)
9. A. M. Turing, Ph. D. thesis, "Systems of Logic Based on Ordinals", Princeton University (1938)
10. C. Papadimitriou, "Computational Complexity", Addison Wesley (1993)
11. A. Broadbent, J. Fitzsimons and E. Kashefi, "Universal blind quantum computation", *Proc. of FOCS 2009*, pp. 517-526 (2009)
12. D. Aharonov, M. Ben-Or and E. Eban, "Interactive Proofs For Quantum Computations", arXiv:0810.5375 [quant-ph] (2008)
13. B. Poonen, "Undecidable problems: a sampler", in *Interpreting Gödel: Critical essays*, Ed: J. Kennedy, pp. 211-241, Cambridge University Press (2014)
14. A. Peres, "Quantum Theory: Concepts and Methods", Springer (1995)
15. V. Paulsen, "Completely Bounded Maps and Operator Algebras", Cambridge Studies in Advanced Mathematics (2003)
16. T. Cubitt, D. Perez-García and M. M. Wolf, "Undecidability of the Spectral Gap", arXiv:1502.04573 [quant-ph] (2015)
17. M. Navascués, Y. Guryanova, M. J. Hoban and A. Acín, "Almost quantum correlations", *Nature Communications* 6: 6288 (2015)