PHYSICS, PHILOSOPHY AND QUANTUM TECHNOLOGY

DAVID DEUTSCH
Oxford University, Oxford, United Kingdom
E-mail: david.deutsch@qubit.org

Quantum theory and the classical theory of computation were perfected in the 1930s, and fifty years later they were unified to form the quantum theory of computation. Here I want to tell you about a speculation — I can't call it more than a "speculation" even though I know it's true — about the kind of theory that might, in another fifty years' time, supersede or transcend the quantum theory of computation.

There are branches of science — in fact most of them are branches of physics — that we expect, by their nature, to have philosophical implications. An obvious example is cosmology. There are other sciences, such as, say, aerodynamics, in which, no matter how startling or important our discoveries may become, we do not expect fundamental philosophical implications. So, various sciences fall at different places on a scale (Fig. 1) ranging from the most fundamental on the left to the least fundamental, the most derivative, on the right.

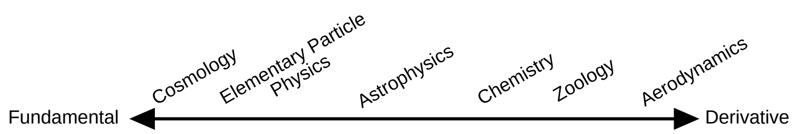

Figure 1. Placing sciences on a scale from fundamental to derivative.

The same holds for mathematics. There are branches of mathematics, such as logic, that we expect by their nature to be relevant to philosophical issues. Then there are other branches, for instance Fourier analysis, which, although they might be very useful both in mathematics itself and in practical applications, are not expected to be philosophically interesting. The vertical axis of Fig. 2(a) measures how mathematical a field is. Halfway up is the line of demarcation between science (which has empirical content) and mathematics (which is purely abstract).

Now, where does the theory of computation lie on that diagram? Traditionally the answer would have been that it is a branch of mathematics, but not a fundamental one: Kronecker said that God made the integers and all the rest is the work of man; more modern formulations give set theory pride of place. So the theory of computation might have been placed at about position 1 on Fig. 2(b). But in the twentieth century, Gödel and Turing argued that the theory of computation ought to be isomorphic to proof theory,

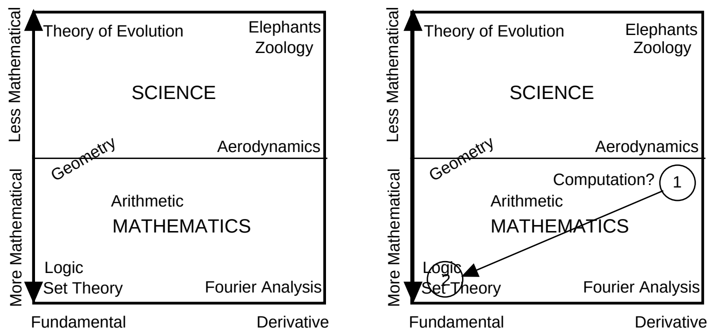

Figure 2. Left panel: (a) assessing the branches of mathematics. Right panel: (b) placing computation in the diagram.

which (it was usually assumed) is at the foundations of mathematics. Such assessments would put it far away at position 2.

What about the theories of information and communication? They were studied under the heading of physics, and would also have been placed well to the right, on the derivative side of the diagram, for all our results in these fields are obtained by taking fundamental theories like the superposition principle and quantum electrodynamics, and deriving conclusions from them. We do not add any new laws.

So we regard our knowledge of the physical world as being structured in the manner of Fig. 3(a): we have laws that apply to absolutely everything, and we apply them to special cases. There's cosmology, and one branch of cosmology would be the physics of galaxies. Within that, there's the physics of solar systems, and within solar systems, planets. On planets, we have quantum computers, elephants, and so on. Alternatively, we can think of elementary particle physics which describes the behavior of all particles in full generality; solid state physics is a special case, and so on up to elephants, humans and quantum computers. Either way, the theory of computers, even quantum computers, would seem to be a highly specialized subfield of a subfield of fundamental physics. After all, it's quite hard to make quantum computers. As far as we know, there are nowhere near as many of them in the universe as there are stars or even elephants. As far as we know there aren't any at all, yet. And in any case, it seems that when we're studying quantum computers we're studying how matter and energy behave under extremely unusual and contrived circumstances — something that may possibly be important to us for practical reasons, but of no fundamental significance. Yet we know that that is the wrong conclusion. We know it because of the existence of computational universality: the laws of physics allow for a machine — a universal quantum computer — with the property that its possible motions correspond

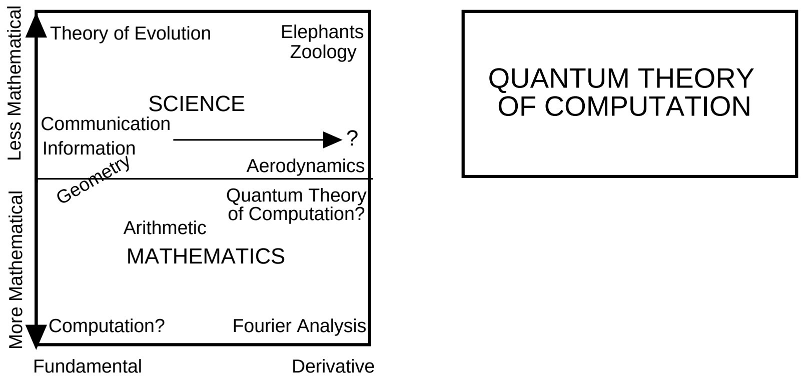

Figure 3. Left panel: (a) placing information and communication in the diagram. Right panel: (b) the encompassing role of quantum computation: does it include all of physics?

in a suitable sense to all possible motions of all possible physical objects. Therefore the whole of physics and more — the study of all possible physical objects — is just isomorphic to the study of all programs that could run on a universal quantum computer.

So the world view portrayed in Fig. 3(a) is wrong, and we're led to a different picture of the relationship between the various sciences, namely Fig. 3(b), in which physics is the quantum theory of computation, and the study of particular physical systems is the study of particular classes of computations. So again galaxies stars and elephants will appear somewhere in this diagram as special cases, but the study of quantum computers is represented by the whole diagram. [As Charles Bennett has pointed out — when I presented this paper at the conference — the quantum theory of computation still does appear as a small region of the diagram as well, but that region has the same structure as the diagram as a whole.]

Yet, even though Fig. 3(b) is our best way of conceiving of the character of physical law, most of our world-views still incorporate something more like the obsolete Fig. 3(a), in which a universal computer is just another physical system. The reason is partly inertia. It is also partly the way in which quantum physics is taught: a course that started with qubits and quantum computational networks instead of the Schrödinger equation and the square well would not only be far closer to the physical foundations of the theory but also inherently simpler. This would have the further merit of allowing quantum theory to be taught earlier in a physics curriculum. But I digress. There is also a good reason for not accepting Fig. 3(b), the computational view of nature, our best available view, as the whole story either.

Admittedly it is very natural, when one has understood the centrality of computation in physics, to make certain speculations about what may be:

maybe the universe that we see — or presumably the multiverse — is really a computer program running on a giant computer. That's an intriguing idea and a rich source of science-fiction plots, but as physics, it is a fundamentally flawed idea. It is a retrograde step in understanding what quantum computation and universality are telling us, and that is for two main reasons. The first is that any cosmology of that type entails giving up on explanations. For if what we see as the laws of physics are actually just attributes of some software, then by the very definition of computational universality, we then have no means of understanding the hardware on which that software is running. So we have no way of understanding the real physics of reality. Thus for this reason, because the properties of this supposed outer-level hardware would never figure in any of our explanations of anything, we have no more reason for postulating that it's there than we have for postulating that there are fairies at the bottom of the garden.

That's a philosophical, or methodological reason. The second reason why I think this idea would be a non-starter is a more technical one, but it too is a killer. We see around us a computable universe; that is to say, of all possible mathematical objects and relationships, only an infinitesimal proportion are ever instantiated in the relationships of physical objects and physical processes. (These are essentially the computable functions.) Now it might seem that one approach to explaining that amazing fact, is to say "the reason why physical processes conform to this very small part of mathematics, 'computable mathematics,' is that physical processes really are computations running on a computer external to what we think of as physical reality." But that relies on the assumption that the set of computable functions — the Turing computable functions, or the set of quantum computable operations — is somehow inherently privileged within mathematics. So that even a computer implemented in unknown physics (the supposed computer that we're all simulations on) would be expected to conform to those same notions of computability, to use those same functions that mathematics designates as computable. But in fact, the only thing that privileges the set of all computational operations that we see in nature, is that they are instantiated by the laws of physics. It is only through our knowledge of the physical world that we know of the difference between computable and not computable. So it's only through our laws of physics that the nature of computation can be understood. It can never be vice versa.

Another of these very natural speculations is that maybe there isn't just one giant computer but an infinite number of them, all running different programs. Or else, forget the computers (the argument I've just given tells us that we have to). We could speculate that there are just many universes all running different laws of physics, as, for instance, in the evolutionary universe idea of Lee Smolin, or the ideas of Seth Lloyd or Max Tegmark. Now I have no objection to the idea that multiple laws of physics in different universes will be implicated in future explanations of the physical world, but it cannot

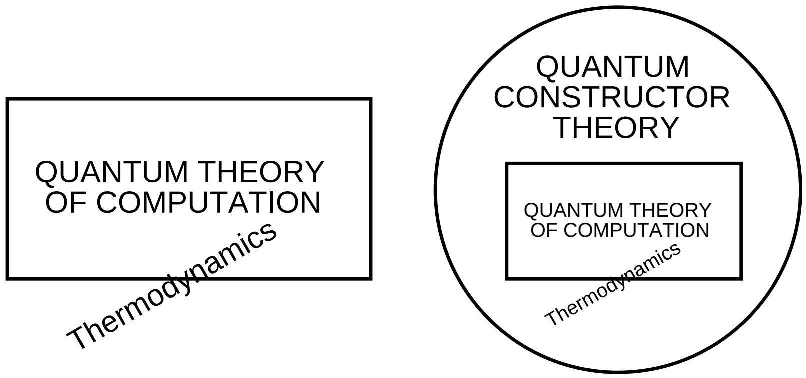

Figure 4. Left panel: (a) thermodynamics is not fully included within the quantum theory of computation. Right panel: (b) quantum constructor theory.

be the whole explanation. It cannot even be the gist of the explanation — for essentially the same reason as the universe-is-a-simulation idea: there is no notion of computation prior to the laws of physics, indeed, there's no notion of laws at all prior to a definition of what is or isn't computable.

Thus, the very power of the principle of the universality of computation, which gives us such a unified view of what physics is, also limits the extent to which physical reality can be regarded as consisting of computations. Because of universality, the nature of computation and the laws governing it are independent of the underlying hardware. Therefore those laws, and that theory, can't explain hardware. Explaining hardware, however, is obviously part of science. Hence there must be something to physics beyond the quantum theory of computation. I think we have to conceive of the quantum theory of computation as a special case of a bigger theory: quantum constructor theory, which is the theory of what physical objects can be constructed, using what resources. Here I don't mean abstract resources, like the number of computational steps or the amount of memory, but physical resources like atoms and energy and entropy and so on.

We do already have some hints about constructor theory: it's not quite true that all of existing physics can be understood as being various aspects of the quantum theory of computation. For instance, we know that it is not possible to build a perpetual motion machine of the second kind. We have a law of physics that tells us so. Yet as far as the quantum theory of computation goes, the time reverse of any physical state is just another physical state. Universality says that the evolution of that state can be simulated by the universal computer with arbitrary accuracy. So the quantum theory of computation does not have the wherewithal to accommodate the whole of, say, quantum statistical mechanics within it [Fig. 4(a)].

Another hint of quantum constructor theory is in the quantum theory of communication. What makes communication different from computation? As in thermodynamics, it is the fact that we have to take into account constraints on computation that are imposed by the actual nature of the physical world. For instance, we may consider two spatially-separated computers, and we have to model what it means for them to 'communicate.' We have to say that certain physical operations, which are perfectly good *computations*, are not allowed because no hardware that is available in nature can implement them because, for instance, that would involve moving information faster than the speed of light. The quantum theory of computation knows nothing of distance; one day, perhaps, distance will be defined in quantum constructor theory, as a certain category of constraints on communication, just as atoms or elementary particles may be defined as certain constraints on the construction of smooth objects. So we have some hints of bits of a future theory that will, in a unified way, address real resources such as energy and volume and time, rather than formal resources such as memory and computational steps and number of computational gates.

The full quantum constructor theory will incorporate the particle physicists' 'theory of everything,' including quantum gravity, as well as the quantum theory of computation and thermodynamics. We may hope that it would be able to answer exotic questions like: can we build a black hole and spin it up until it becomes a time machine? Can we collapse a black hole and have it form new universes which we can design, and if so what are the constraints on that? Before that, we might expect it to resolve more down-to-earth controversies like: can we ever build a controlled fusion reactor? Note that as far as the quantum-theory-of-computation conception of physics goes, the answer would be yes. Building such a device is simply a matter of preparing a certain observable in a fairly sharp state, and in quantum computation that is never a fundamental problem. But that's because the quantum theory of computation doesn't quite capture the whole of physics.

Another such question is of course: can we build a quantum computer? That's a topical issue here, and at the present state of science, still a controversial one. Yet we have to admit that at present, the controversy is not really part of science. It's a bit like glove puppets arguing 'oh yes we can,' 'oh no you can't,' for neither side can appeal to laws of physics. At best they are expressing their intuitions, which may or may not contain wisdom but at the moment are outside physics and outside science. One day the issue will be brought within science, by quantum constructor theory — although I suspect that by then it will have been resolved empirically, by the building of quantum computers.

There will presumably be a constructor generalization of the universal quantum computer, namely the universal quantum constructor, a machine that can be programmed to construct any quantum object that can be constructed, or any quantum object with any achievable properties.

Quantum constructor theory will give us a veritable meta-physics — not metaphysics, the branch of philosophy, but a subject that stands in the same relation to physics as metamathematics does to mathematics. It will be the theory of what can be done, and what can be made, in physical reality. There are some awesome questions waiting for us there in that direction. One of them is: is the human race a universal constructor? — a question with tremendous reverberations for things like epistemology and cosmology, and actual metaphysics.

Our state of knowledge of this field is still very primitive. We are nowhere near answering any such question yet, or even many elementary questions of constructor theory. The subject is still so much in its infancy that even within the quantum theory of computation proper, we don't really have a good idea of what quantum information is, or how entanglement should be quantified and so on. A few years ago the journalist John Horgan wrote a book *The End of Science*, in which he said that all fundamental theories are already known and that from now on fundamental science will all be just dotting the i's and crossing the t's. Horgan concedes that in the year 1900 people were saying much the same things: that physics in the future would be about evaluating the sixth decimal place and so on, and they were proved wrong within a single generation by a series of scientific revolutions. Horgan rightly points out that the fact that those pundits were embarrassingly wrong does not prove that he isn't right. But I think that in 1900 there was a lot more excuse for thinking that physics was almost at an end than there is now. For today's best theories contain within themselves the implication that they can't be complete. There is the well known conflict between the general theory of relativity and quantum theory. There's also the fact I'm stressing, that the principle of universality tells us that there must be a larger unifying theory outside the quantum theory of computation.

So we know, in our field at least, that Horgan is wrong. We know that there are huge gaps in our understanding, and fundamental mismatches between our best theories and the reality they are supposed to explain. We know that if we are at the end of some era then it's only because we are at the beginning of a new one which offers the prospect not only of some quite fundamental discovery but of new kinds of fundamental discovery.