The Logic in Computer Science Column

by

Yuri Gurevich

Microsoft Research
One Microsoft Way, Redmond WA 98052, USA
gurevich@microsoft.com

WHY SOME PHYSICISTS ARE EXCITED ABOUT THE UNDECIDABILITY OF THE SPECTRAL GAP PROBLEM AND WHY SHOULD WE

Vladik Kreinovich
Department of Computer Science
University of Texas at El Paso
El Paso, TX 79968, USA
vladik@utep.edu

Abstract

Since Turing's time, many problems have been proven undecidable. It is interesting though that, arguably, none of the working physicist problems had been ever proven undecidable – until T. Cubitt, D. Perez-Garcia and M. M. Wolf proved recently that, for a physically reasonable class of systems, no algorithm can decide whether a given system has a spectral gap. We explain the spectral gap problem, its importance for physics and possible consequences of this exciting new result.

1 First physically meaningful un-computability result

Since the pioneering 1936 result of Alan Turing, many decision problems have been shown to be undecidable.

At first glance, we should expect physicists to be excited about these results. After all, physicists deal with mathematical models of physical phenomena. Sometimes, they discover mathematical models of new phenomena and update the existing models if experiments show the need for such an update. However, most of the time they deal with the existing models, be they Maxwell's equation for electrodynamics, Einstein's general relativity equations for gravity, Schrödinger's equations for quantum physics, or whatever.

From this viewpoint, the fact that some mathematical problems are undecidable should be of interest to physicists. And indeed, on the philosophical level,

many physicists appreciate these results. However, in their view, these results are removed from their day-by-day activities.

Yes, it is interesting whether a Turing machine will eventually stop. However, from the practical viewpoint, the question usually is: will a machine stop in a year? in hundred years? OK, if it is related to something long-term, like storing used radioactive materials, the question becomes: will the machine continue working million years form now? These are all reasonable questions, and for all these questions there are algorithms for solving them. The only time when we have an algorithmically undecidable problem is when we ask what happens "eventually" – but eventually, the Sun will stop shining, and in practice, the machine will break anyway, no matter how small the probability is that it will break in any given year.

Physicists might have thought: yes, philosophical and mathematical problems are tough – and some of them may be undecidable – but our practice-related problems, they are all decidable; for each of them, we will eventually know the solution.

Meanwhile, computer scientists became interested in decidability of physical problems. While physicists were, in general, not interested in analyzing theoretical computability of physical theories, many logicians and computer scientists have formulated and analyzed how computable predictions of physical theories are. Probably the best known was the result of Georg Kreisel, who, in his 1974 paper [6], clearly explained the motivation for analyzing these problems: if some (deterministic) physical events are not computationally predictable on the usual computers — or at least if the computational prediction of a future event is possible only long after the event has happened — then we can speed up computations by utilizing this phenomenon in our computational devices.

This speed-up is easy to explain. What we want to predict is the (rational-valued) result r of measuring an appropriate physical quantity in the after-event state of the corresponding physical system. For a well-defined deterministic theory, this result can be described as the only rational number r that satisfies a well-defined property $P(r)$ (the property describing both the general physical theory and the corresponding specific event). If computing such r on the usual computers takes longer than the duration of the event, then, instead of computing r on usual computers, we can, at least in principle, launch the corresponding physical process and measure the result. This will enable us to get r faster than the usual computers – and thus, speed up the corresponding computations.

This is similar to the fact that in the ancient times it was probably faster to obtain a good approximation to $\sqrt{2}$ by accurately drawing a right square and measuring its diagonal than by computing – and to the fact that in the 1930s, for solving some systems of differential equations, analog devices were faster than numerical computations.

In the four decades after the 1974 paper, the analysis of computability of physi-

cal phenomena has not uncovered any undecidable physical problems. (It is worth mentioning, however, that the related analysis of whether we can speed up computations by using different physical processes resulted in a great triumph of *quantum computing* [7], where an explicit use of quantum phenomena in computations can indeed lead to a drastic speed-up.)

Until recently, when mathematicians and computer scientists analyzed the computability and decidability of physical problems (see, e.g., [2] and references therein), their interesting (to us) research had little effect on working physicists.

Suddenly, a new result [3], [4] by Toby Cubitt, David Perez-Garcia, and Michael M. Wolf seems to have changes all this. Cubitt, Perez-Garcia, and Wolf took a real-life physical problem, a problem of the type that working physicists are usually working on – namely, the problem of existence of a spectral (energy) gap – and showed that this problem is algorithmically undecidable. This may be the first time when a practical physical problem has been shown to be algorithmically undecidable.

Physicists seem to be excited – otherwise, the paper containing this result would not have been published in *Nature*, the journal publishing major breakthroughs. But why are they excited? This is what we will try to explain – what spectral gaps are, why these gaps are important, what exactly the authors did, and how it is all related to the fundamental questions of physics and computing. We hope that after reading this essay, we will all join physicists in their excitement.

2 What are spectral (energy) gaps and why spectral gaps are important: an informal introduction

A short answer to the question why spectral gaps are important is: because quantum physics is *all* about gaps, and especially spectral gaps. I realize that this does not sound convincing to a non-physicist. To explain, we need to go back to the origin of quantum physics.

Many popular expositions of quantum physics emphasize that quantum physics is counter-intuitive and weird but confirmed by experiments. One may get an impression that experiments are the only problem with classical physics.

This picture is oversimplified. Yes, many phenomena of quantum physics sound weird to us, but largely, what motivated quantum physics was not the new experiments: classical (pre-quantum) physics turned out to be inconsistent and incapable to explain simple physical phenomena like light or atoms. To see this, let us go back to the very origin of quantum physics – the 1900 work by Max Planck (after whom numerous Max Planck Institutes are named).

The problem that Planck worked upon was related to the generation of light;

see, e.g., [5], Vol. 3.

When it is dark, we need light. How can we make light? Thousands years ago, our ancestors already knew the answer: light a fire. Primitive people lit pieces of wood, medieval people lit candles, 20 century folks lit Tungsten inside an Edison lamp, but the process is always the same: we heat something, and that something starts to glow, starts to emit electromagnetic waves – in particular, light.

It is also a well-known empirical fact that the observed color (or, more generally, wavelength) of this light depend on the temperature. When we heat an object a little bit, it does not yet emit visible light. As the temperature increases, we see the light, first red, then yellow, etc.

For each temperature, this light has a certain typical wavelength – and there is also radiation on nearby wavelengths. How can we describe this phenomenon in precise terms?

Of course, there are many different sources of heat, many different materials, many different processes may be going on: for Edison lamp, it is simply heating; for a candle, it is burning, etc. What is characteristic of all these processes is that if there is a small deviation, the process quickly goes back to an equilibrium. If a wind blows, for a moment, the flames will change, but once the wind stops, we see the same process as before. An area of physics called thermodynamics describes such equilibria in quantitative terms. It seems reasonable to apply the corresponding equilibria formulas to light emission.

According to thermodynamics, particles in a heated gas bump into each other and exchange energy. As a result, their velocity distribution reaches an equilibrium. Similarly, light rays of different frequencies can interact with each other until they reach an equilibrium. However, for light, physicists encountered a problem. In principle, interactions can produce electromagnetic waves of arbitrarily large frequencies. According to thermodynamics formulas, in the equilibrium, we have *the same* energy for each frequency. In practice, however, for sufficiently large frequencies, the radiation energy *decreases* with frequency.

From the mathematical viewpoint, this decrease makes sense: otherwise, the overall energy of the radiation-in-equilibrium would be infinite. So, we need to modify classical physics to avoid this infinity – and to explain observations.

Classical physics assumes that, for each frequency, energies of the corresponding electromagnetic waves can take any real value. Since we do not see light at high frequencies, something prevents lower-frequency electromagnetic waves to pass energy to higher-frequency ones. To explain this, Planck conjectured that there is the smallest amount of energy $E(\omega)$ that an electromagnetic wave of frequency ω can carry, and that the wave's energy is always a multiple of $E(\omega)$. This smallest amount increases with frequency; as a result, for sufficiently high frequencies, there is not enough energy to generate corresponding waves.

In other words, Planck conjectured that for electromagnetic waves of a given

frequency ω, not all energy values are possible: there is an *energy gap* between energy 0 and the smallest possible energy $E(\omega)$.

How does this smallest energy $E(\omega)$ depend on the frequency ω? Planck started by trying the simplest possible – linear – dependence $E(\omega) = \hbar \cdot \omega$, for come constant $\hbar$, and he got a perfect fit with the observations! (This constant $\hbar$ is the famous Planck's constant, which is present in most equations of quantum physics.)

This idea of the smallest possible energy (*quantum*) turned out to be very productive[1].

Quanta helped explain another phenomenon in which classical physics had difficulty: the existence of atoms.

According to Maxwell's equations of electrodynamics, an accelerating charged particle emits radiation. This is the reason why heated bodies emit radiation: heat means that the particles move around, and as charged electrons and nuclei move around, they emit radiation.

An electron going around the atom's nucleus has a non-zero acceleration – otherwise, it would have followed a straight line. Thus, according to classical physics, the electron should emit radiation – and hence, lose energy. As it loses energy, it will fall closer and closer to the nucleus, just like a satellite's orbit eventually get closer and closer to the Earth until the satellite burns in the atmosphere.

A satellite burns down after several thousand rotations; similarly, a classical-physics electron would fall down after thousands of rotations. The difference is that for a satellite, this takes decades, while for an electron, with its small size and fast rotation, it would occur in microseconds.

So, according to the classical physics, atoms cannot exist: if an atom existed, it would disappear in a few microseconds. How can we reconcile physics with the fact that atoms do exist? This problem was worked upon by Niels Bohr, another eventual Nobelist. According to Bohr's 1913 idea, since electrons do not go into lower-energy states, this means that not all energy values are possible, that there is an energy gap within which energies are not physically possible.

The above two examples describe well-solved problems. Since 1913, energy gaps appeared in many interesting physical results and in many challenging open problems.

One of the remaining open problems is the *mass-gap* problem. It is one of the seven Millennium Prize Problems selected in 2000 by the Clay Mathematics Institute, with a million dollar prize for solving each of them. This problem is formulated in somewhat complicated precise mathematical terms, but its physical

[1] It is worth mentioning that Einstein received his Nobel prize *not* for relativity theory, but for a series of papers in which he used Planck's quantum (gap) ideas to quantitatively explain *photoelectric effect*, when light generates electricity (as in solar panels).

meaning is simple: to explain stability of matter.

Atoms consist of electrons, protons, and neutrons. There are also photons, particles that form light (and electromagnetic radiation in general). There are hundreds of other particles, most of them unstable. For example, a stand-alone neutron transforms itself into a proton, and electron, and an anti-neutrino, with an average decay time of about 15 minutes. Such a decay is possible, because the overall rest mass of the three products is smaller than the rest mass of the neutron, and the overall electric charge of the decay products is equal to the 0, the neutron's charge[2].

Physicists usually assume that any transformation that does not violate any conservation laws is possible. From this viewpoint, the fact that, e.g., electrons are stable seems to indicate that there are no charged particles of smaller mass – i.e., that for all possible electrically charged elementary particles, there is a mass (energy) *gap* between 0 and the rest mass of the electron. The challenge is to explain this gap.

Energy gaps are useful not only for foundations; they are also important for practical applications, e.g., for computer design.

To design a computer, we need to design cells for storing and processing zeros and ones. In principle, it is possible to select two values of a physical quantity, and identify one of these values with 0 and another with 1. The problem is that in the absence of a gap, small perturbations will eventually "slide" the cell from the 0 state to the 1 state (or vice versa). To make computer operations more stable, it is desirable to select quantities for which there is a *gap* between the 0 and 1 states.

The need for such a gap is one of the reasons *semiconductors* are ubiquitous in computer design: while in a conductor, an electron can have any possible value of energy, in a semi-conductor, there is a *gap* between states in which electrons practically do not move and states in which they form a current.

3 How to describe energy gaps: from informal idea to equations

How can energy-gap ideas be translated into equations? To explain this transition to a computer science audience, let us start with the ideas that many computer scientists know from popular expositions of quantum computing; see, e.g., [7].

In the quantum version of a bit (known as *qubit*), in addition to the usual state 0 and 1 (which are usually denoted by $|0\rangle$ and $|1\rangle$), we can also have linear combinations ("superpositions") of these states, i.e., states of the type $\psi \stackrel{\text{def}}{=} c_0|0\rangle + c_1|1\rangle$, where c_0 and c_1 are complex numbers for which $|c_0|^2 + |c_1|^2 = 1$. In such a

[2]neutron is electrically neutral, as its name indicates

state ψ, the probability to observe 0 is equal to $|c_0|^2$, and the probability to observe 1 is equal to $|c_1|^2$ (and the fact that we always observe either 0 or 1 explains the requirement $|c_0|^2 + |c_1|^2 = 1$).

The only information we can infer from the experiments – with some fidelity – are the absolute values $|c_i|$ of the complex numbers c_i (and the absolute values of their linear combinations). These absolute values we can infer by statistical sampling.

From the geometric viewpoint, the absolute value is the *length* of the 2-D vector (x, y) representing a complex number $x + \mathrm{i}y$. Lengths do not change if we rotate the vectors. In geometric terms, rotation corresponds to multiplication by $\exp(\mathrm{i}\alpha)$ for a real α. Thus, the states ψ and $\exp(\mathrm{i}\alpha) \cdot \psi$ are indistinguishable – and are therefore considered to be the same state.

How to describe a particle? If we have finitely many spatial locations $x^{(1)}, \ldots, x^{(n)}$, then, similarly, a general quantum state can be described a linear combination of particles in these locations: $\psi = \psi(x^{(1)})|x^{(1)}\rangle + \ldots + \psi(x^{(n)})|x^{(n)}\rangle$. Such a state is characterized by a mapping from spatial locations $x^{(i)}$ into complex numbers $\psi(x^{(i)})$. Here, $|\psi(x^{(i)})|^2$ is the probability to observe the particle in the location $x^{(i)}$. For each set $S \subseteq \{x^{(1)}, \ldots, x^{(n)}\}$, the probability to find a particle in one of the locations from this set is thus equal to the sum $\sum\limits_{x \in S} |\psi(x)|^2$.

To describe a general particle, we can consider more and more locations and tend to a limit. In the limit, we get a function ψ that assigns, to each location x, a complex number $\psi(x)$ so that for every spatial region Ω, the probability of finding the particle in this region is equal to the *integral* $\int\limits_{\Omega} |\psi(x)|^2\, dx$. This function is known as a *wave function*; it describes the *state* of a quantum particle.

How do states change with time? In quantum physics, superposition (or, in mathematical terms, linear combination) is a fundamental idea. Any change should preserve linear combinations and should thus be linear. So, in a differential equation $\dfrac{\partial \psi}{\partial t} = f(\psi)$ describing this change, the right-hand side $f(\psi)$ must be linear in ψ, i.e., the equation must have the form

$$\frac{\partial \psi}{\partial t} = H_0 \psi \tag{1}$$

for some linear operator H_0[3]. This equation is known as the *Schrödinger equation* after the Nobelist who discovered it.

The Schödinger equation describes dynamics on the micro-level. However, most of our observations are of macro-objects. From the viewpoint of these ob-

[3]The operator H_0 must also guarantee that the integral $\int |\psi(x)|^2\, dx$ representing full probability does not change with time.

servations, fast micro-level changes have already occurred. So, we are mostly interested in the stabilized ("equilibrium") states.

Let us recall that, in quantum physics, we identify states ψ and $\exp(\mathrm{i}\alpha)\cdot\psi$. Thus, the requirement that the state be stationary, i.e., that at different moments of time t and t_0 the state be the same, means that for some $\alpha(t)$ we have $\psi(x,t) = \exp(\mathrm{i}\cdot\alpha(t))\cdot\psi(x)$, where $\psi(x) = \psi(x,t_0)$. Substituting this expression into the equation (1), we conclude that

$$\exp(\mathrm{i}\alpha(t))\cdot\mathrm{i}\cdot\frac{d\alpha}{dt}\cdot\psi(x) = \exp(\mathrm{i}\alpha(t))\cdot H_0\psi.$$

Dividing both sides by $\exp(\mathrm{i}\alpha(t))$, we get $\mathrm{i}\cdot\dfrac{d\alpha}{dt}\cdot\psi(x) = H_0\psi$, i.e., $\mathrm{i}\cdot\dfrac{d\alpha}{dt} = \dfrac{H_0\psi}{\psi(x)}$. The right-hand side of this equality does not depend on time, so the left-hand side does not depend on time either, and is, thus, a constant. This constant is usually denoted by ω. So, $\alpha(t) = \omega\cdot t$ for some constant ω, hence $\psi(x,t) = \exp(\mathrm{i}\cdot\omega\cdot t)\cdot\psi(x)$, and the stationary Schrödinger equation takes the form $\mathrm{i}\cdot\omega\cdot\psi(x) = H_0\psi$.

In mathematical terms, the values $\mathrm{i}\cdot\omega$ are eigenvalues of the linear operator H_0, while the corresponding functions $\psi(x)$ are eigenfunctions of this operator.

There is an additional reason why stationary solutions are important: if we know *stationary* solutions of the Schrödinger equation (1), then we can find its general (time-dependent) solutions. Indeed, any vector can be represented as a linear combination of eigenvectors. In particular, the function $\psi(x,t_0)$ can be represented as a linear combination of the corresponding eigenfunctions $\psi_\omega(x)$: $\psi(x,t_0) = \sum\limits_{\omega} c_\omega\cdot\psi_\omega(x)$ for some coefficients c_ω. We know how eigenfunctions change in time, so we can therefore get an explicit expression for $\psi(x,t)$: $\psi(x,t) = \sum\limits_{\omega} c_\omega\cdot\exp(\mathrm{i}\cdot\omega\cdot t)\cdot\psi_\omega(x)$.

In physical terms, a stationary solution represents oscillations with frequency ω. In these terms, a general solution is a linear combinations of stationary solutions corresponding to different frequencies.

For light, the representation as such a linear combination was first discovered in the Newton's prism experiment: a light passing through a prism gets visually separated into different colors, i.e., into different wavelengths. The resulting set of lines is known as the *spectrum*.

For example, by observing the spectrum of a star, we can learn what elements are present there. Because of this example, in the general case, the set of all possible eigenvalues ω is also known as the *spectrum*.

If we use Planck's formula $E = \hbar\cdot\omega$ to express frequency ω in terms of energy E, as $\omega = \dfrac{E}{\hbar}$, then we conclude that $\mathrm{i}\cdot\dfrac{E}{\hbar}\cdot\psi = H_0\psi$, i.e., equivalently, that

$$H\psi = E\cdot\psi, \tag{2}$$

where we denoted $H \stackrel{\text{def}}{=} -\mathrm{i}\hbar \cdot H_0$.

So, the operator H describes how the energy of the system depends on its state. This operator is thus known as the *energy operator*, or *Hamiltonian.* In terms of the energy operator, the general dynamic Schrödinger equation has the form $\mathrm{i} \cdot \hbar \cdot \dfrac{\partial \psi}{\partial t} = H\psi$.

The apparent importance of energy sounds somewhat unexpected: in classical physics, energy was just one of the many physical characteristics. The above equation shows that in quantum physics, energy takes central stage. To describe the dynamics of a quantum system, it is sufficient to know how energy depends on the state.

This may sound counter-intuitive, but even in classical physics, many solutions can be obtained solely from the expression for energy. For example, a general way to explicitly solve 1-D Newton's differential equations $m \cdot \dfrac{d^2x}{dt^2} = -\dfrac{dV(x)}{dx}$ for a particle in a potential field $V(x)$ is to use the expression for total energy as the sum of kinetic and potential energies:

$$E = E_{\text{kin}} + E_{\text{pot}} = \frac{1}{2} \cdot m \left(\frac{dx}{dt} \right)^2 + V(x).$$

From this expression, we conclude that

$$\frac{dx}{dt} = \sqrt{\frac{E - V(x)}{m/2}}$$

hence, separating the variables,

$$\frac{dx}{\sqrt{E - V(x)}} = \sqrt{\frac{m}{2}} \cdot dt.$$

Integrating, we get an explicit solution:

$$\int \frac{dx}{\sqrt{E - V(x)}} = \sqrt{\frac{m}{2}} \cdot t + C_0.$$

For a particle in an electromagnetic field, we know the expression of energy in terms of location and velocity, so we can write the corresponding quantum equation. This is exactly how Schrödinger came up with his equations.

It is worth mentioning that, according to the late physicist Yuri Rumer – who worked closely with many founders of quantum physics – Schrödinger was motivated by Einstein, who, not believing in the probabilistic ideas of quantum physics, wanted Schrödinger to formalize the then-vague ideas about particles-as-waves into precise equations. Einstein's hope was that by solving the resulting

equations, we would clearly see that their solutions are not consistent with the known observations, and thus, we would be able to reject the probabilistic ideas.

When Schrödinger performed this formalization, to his surprise, he got a perfect explanation of the observed atomic spectra! According to Rumer, Schrödinger offered Einstein – who suggested the original idea – to be a co-author of his ground-breaking paper. However, Einstein was so against the probabilistic ideas of quantum physics that not only he refused this offer, he even discouraged Schrödinger from even mentioning that this paper was based on Einstein's idea.

Meanwhile, Einstein continued trying to disprove quantum ideas. Next, he came up with the now-well-known Einstein-Podolsky-Rosen "paradox". Again, it turned out that this is not really a paradox – that the corresponding (somewhat counterintuitive) *entanglement* is an observable feature of quantum physics – a feature which underlies many modern quantum computing algorithms.

Again and again, great Einstein tried to disprove quantum physics – and every time he actually helped it a lot (just like the biblical prophet Balaam who, when trying to curse the Israelites, ended up twice blessing them).

4 From elementary particles to more complex (e.g., solid-body) physics

For a single particle – e.g., for the electron in a hydrogen atom – the Schrödinger equation can be solved analytically.

For larger atoms, the situation is more complicated. For m particles, a basic *classical* state can be described by the locations x_i of all these particles. and thus, has the form $|x_1, \ldots, x_m\rangle$. So, a general *quantum* state of an m-particle system can be described by the (complex-valued) coefficients $\psi(x_1, \ldots, x_m)$ at all these classical states – i.e., taking into account that each spatial location x_i is characterized by three coordinates, by a wave function depending on $3m$ real-valued variables.

For atoms, $m \le 100$, so we can solve the corresponding equation numerically and get good approximations to the observed spectrum. However, for a piece of solid body, with $m \approx 10^{23}$ atoms, there is no easy way to solve a differential equation describing a function of $3 \cdot 10^{23}$ variable. What can we do?

We know that any real-life piece of solid body is bounded. However, from the viewpoint of an individual atom, the size of this piece is so much larger than the atom's size, that for most practical purposes, we can view this piece as infinite. This is especially true since far-away atoms practically do not interact with each other. As a result, to describe the next state of each atom, it is sufficient to know the current state of this atom and of a few of its neighbors.

This is similar to the fact that to describe the trajectory of a near-Earth satellite,

it is sufficient to know how the Earth and the Sun (and maybe the Moon) affect its trajectory; the effect of all other planets can be safely ignored. This means that the expression for the satellite's energy includes terms depending on the the satellite's distance from the Earth, but there are no terms depending on the distance between the satellite and, say, Jupiter.

For solid bodies, this idea works well: usually, as we increase the number of atoms, the system's behavior very fast falls into the same pattern, which we can then easily extrapolate to the case of 10^{23} particles.

Sometimes, one needs to consider a rather large number of interacting particle before we get to the limit behavior, but in most cases, the corresponding number is still tractable.

Based on the numerous successes of this approach, physicists thought that in general, analyzing the properties of such large well-organized groups of particles should be an algorithmically solvable problem: just study big enough sets, and you will see the pattern.

Interestingly, the new undecidability-of-spectral gap result shattered this belief. The new result shows that, contrary to the physicists' belief, even for simple versions of the problem, a general algorithm is *not* possible.

Specifically, the authors consider a 2-D lattice in which each particle can be in finitely many (d) different classical states – so that the quantum state of each particle can be described by a (complex-valued) d-dimensional vector. It is assumed that this lattice is *translation-invariant* in the sense that each particle interacts with its neighbors in exactly the same way. It is also assumed that only nearest neighbors interact with each other, i.e., that the overall Hamiltonian (that describes the energy of the system) is a sum of expressions which depend only on the states of four direct neighbors.

For each size L of the corresponding $L \times L$ 2-D lattice, we get a finite-dimensional system, with finitely many eigenvalues. Let $\lambda_0(L)$ denote the smallest of these eigenvalues – i.e., the smallest possible energy of the system. The corresponding lowest-energy state is called the *ground state*. Let $\lambda_1(L)$ denote the second smallest eigenvalue. According to the traditional physicists' belief, by observing the difference $\lambda_1(L) - \lambda_0(L)$ for a sufficiently large L, we should be able to tell whether there is an energy gap that makes the ground state stable:

- for some systems, the difference $\lambda_1(L) - \lambda_0(L)$ tends to a positive number $\gamma > 0$; we therefore conclude that for such systems, we have a semiconductor-type energy gap;
- for other systems, the difference $\lambda_1(L) - \lambda_0(L)$ tends to 0; for such systems, we have a conductor-type gapless behavior.

The authors prove that distinguishing between these two cases is, in general, *not*

algorithmically possible. Namely, they construct a family of Hamiltonians for which, for each Hamiltonian:

- either there is a positive number $\gamma > 0$ for which $\lambda_1(L) - \lambda_0(L) > \gamma$ for all sufficiently large sizes L,
- or there exists a positive number $c > 0$ such that, for every $\varepsilon > 0$ there exists an L_0 for which, for all sizes $L \geq L_0$, every number in the interval $[\lambda_0(L), \lambda_0(L) + c]$ is ε-close to some eigenvalue of the corresponding Hamiltonian.

For this family of Hamiltonians, they prove that no algorithm can decide whether we are in a first (gap) case or in the second (gapless) case. In other words, contrary to what physicists believed, *it is not possible to tell whether a given physical system has a spectral (energy) gap or not*.

Readers familiar with constructive mathematics may suspect that this result could be caused by considering general computable real numbers as coefficients of the corresponding Hamiltonians. Indeed, it is known that if we consider general computable real numbers, then even checking equality of real numbers is not algorithmically decidable; see, e.g., [1], [8]. The authors are well aware of this possible suspicion, so they made sure that all the coefficients that relate the Hamiltonian to the local states are either rational numbers, or simple algebraic numbers of the type $r + r' \cdot \sqrt{2}$, where r and r' are rational. For such numbers, equality is algorithmically decidable.

5 How the undecidability of the spectral gap problem is proved: a very brief description

The main objective of this essay is to explain why the undecidability-of-spectral-gap result is interesting.

The proof is very interesting but very complex (takes more than 100 pages). It is impossible to describe all its nuances in a short essay; so we will only give a very brief description.

In a nutshell, the undecidability of the spectral-gap problem is proven in a rather straightforward way: by reducing the halting problem for Turing machines to the spectral-gap problem, i.e., by proving that if we had an algorithm for solving the spectral-gap problem, then we would be able to algorithmically solve the halting problem.

However, instead of using *traditional* Turing machines, the authors use the halting problem for *quantum* Turing machines.

At first glance, the use of quantum Turing machines should not make the description more complex: aren't traditional Turing machines a particular case of the quantum ones?

Well, not exactly. We can indeed *emulate* classical (non-quantum) computations – in particular, Turing machines – by quantum ones, but this emulation is *not* trivial. The reason for this non-triviality is that quantum processes are *reversible*: given a future state, we can uniquely reconstruct the past state. In contrast, traditional Turing machines are *not* reversible: a machine computing a function $f(n)$ transforms n into $f(n)$; so, if $f(n) = f(m)$ for some $n \le m$, then, based on the future state $f(n) = f(m)$, we cannot tell whether the original state was n or m.

Schrödinger equations are, as one can easily see, reversible: just like we can integrate these equations forward in time to predict the future, we can also integrate these equations back in time to reconstruct the past[4].

This reversibility is a known feature of quantum computing. As a result, in quantum computing, even a simple classically irreversible bit operation $f(x_1, \ldots, x_n)$, like the "and" operation $f(x_1, x_2) = x_1 \,\&\, x_2$, is represented in a rather complicated form, as a mapping from an $(n+1)$-element tuple $|x_1, \ldots, x_n, y\rangle$ to $|x_1, \ldots, x_n, y \oplus f(x_1, \ldots, x_n)\rangle$, where $\oplus$ is addition modulo 2 (= exclusive or).

For Turing machines, the reduction to reversible ones is even more complicated. So, to develop the proof, the authors had to develop a whole theory of quantum Turing machines and of their reversible classical counterparts.

6 Instead of a conclusion: what are the possible consequences of this result?

The above algorithmic result implies that if we have an axiomatic mathematical theory (like the usual Zermelo-Fraenkel axiomatic set theory ZF), then there exist physical systems for which, from these axioms, we cannot deduce neither the existence not the absence of the energy gap. In other words, we have Gödel-type statements which have direct physical meaning.

This conclusion is easy to make: by applying all possible rules to different axioms, we can algorithmically generate all possible results provable in a given axiomatic theory. Thus, if for each system, we would be able to prove either the existence or the absence of the proof, then we should be able to algorithmically

[4]This, by the way, is not a specific feature of quantum physics: the usual Newton's equations $m \cdot \dfrac{d^2x}{dt^2} = F$ are reversible as well. This is counter-intuitive: I can easily break a cup, but no natural physical process can combine the resulting pieces back into a whole cup. To explain this seeming paradox, it is necessary to go so deeply into physics – and into the corresponding ideas of algorithmic randomness and Kolmogorov complexity – that this would require another essay.

decide whether a given system has a gap – by simply waiting for a corresponding positive or negative statement to appear in the sequence of all provable results.

What are the possible computational consequences of these results? Let us recall one of the main reasons why computer scientists are interested in these physics-related problems: if we find a physical phenomenon whose computational prediction takes longer than simple waiting for this phenomenon to occur, then we could use this phenomenon to speed up computations.

For quantum computing, we did not find an undecidable problem, but we found phenomena which turned out to be useful for computing – and this led to the development of quantum computing.

For spectral gap, we do seem to have an undecidable problem, but this problem is only undecidable in the limit. For each specific size of the lattice, the problem is finite-dimensional and therefore, algorithmically decidable. However, the new result seems to indicate that for large real-life (finite) systems, it may be difficult to algorithmically predict the presence of a spectral gap. In this case, we may be able to speed up computations by designing the corresponding lattice and measuring the corresponding eigenvalues.

In other words, it may turn out that this spectral-gap result is not only interesting from the physical and philosophical viewpoints, it may be a first step towards yet another breakthrough in quantum computing.

Let's work to make this happen!

References

[1] E. Bishop, *Foundations of Constructive Analysis*, McGraw-Hill, New York, 1967.

[2] D. Bridges, "Can constructive mathematics be applied in physics?", *Journal of Philosophical Logic*, 1999, Vol. 28, No. 5, pp. 439–453.

[3] T. S. Cubitt, D. Perez-Garcia, and M. M. Wolf, "Undecidability of the spectral gap", *Nature*, 2015, Vol. 528, pp. 207–211.

[4] T. S. Cubitt, D. Perez-Garcia, and M. M. Wolf, *Undecidability of the spectral gap*, http://arxiv.org/abs/1502.04573, 2015.

[5] R. Feynman, R. Leighton, and M. Sands, *The Feynman Lectures on Physics*, Addison Wesley, Boston, Massachusetts, 2005.

[6] G. Kreisel, "A notion of mechanistic theory", *Synthese*, 1974, Vol. 29, pp. 11–26.

[7] M. A. Nielsen, I. L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, Cambridge, U.K., 2000.

[8] K. Weihrauch, *Computable Analysis*, Springer Verlag, Berlin, 2000.