Undecidable problems in quantum field theory

Yuji Tachikawa (Kavli IPMU)

Talk at ICTS, May 25, 2022

[2203.16689]

Life is unpredictable.

Theoretical physics is no exception:

- Probability inherent in quantum mechanics.
- Chaotic behavior of deterministic systems.
- **a third type of unpredictability.**

Theoretical physicists working on a topic $\boldsymbol{X}$
often spend their entire lives on answering the following question:

Given a system $\boldsymbol{x}$ in this topic $\boldsymbol{X}$, does it have property $\boldsymbol{P}$?

- $\boldsymbol{X}$: 4d gauge theories
  $\boldsymbol{P}$: $\boldsymbol{x}$ confines in the infrared limit
- $\boldsymbol{X}$: 1d quantum spin chains
  $\boldsymbol{P}$: $\boldsymbol{x}$ is gapless in the continuum limit
- $\boldsymbol{X}$: supersymmetric theories
  $\boldsymbol{P}$: $\boldsymbol{x}$ spontaneously breaks supersymmetry
- $\boldsymbol{X}$: various materials
  $\boldsymbol{P}$: $\boldsymbol{x}$ is superconducting at room temperature and normal pressure
- the list continues…

Theoretical physicists working on a topic $\boldsymbol{X}$
often spend their entire lives on answering the following question:

Given a system $\boldsymbol{x}$ in this topic $\boldsymbol{X}$, does it have property $\boldsymbol{P}$?

- $\boldsymbol{X}$: 4d gauge theories
  $\boldsymbol{P}$: $\boldsymbol{x}$ confines in the infrared limit
- $\boldsymbol{X}$: 1d quantum spin chains
  $\boldsymbol{P}$: $\boldsymbol{x}$ is gapless in the continuum limit
- $\boldsymbol{X}$: supersymmetric theories
  $\boldsymbol{P}$: $\boldsymbol{x}$ spontaneously breaks supersymmetry
- $\boldsymbol{X}$: various materials
  $\boldsymbol{P}$: $\boldsymbol{x}$ is superconducting at room temperature and normal pressure
- the list continues…

Theoretical physicists working on a topic $\boldsymbol{X}$
often spend their entire lives on answering the following question:

Given a system $\boldsymbol{x}$ in this topic $\boldsymbol{X}$, does it have property $\boldsymbol{P}$?

- $\boldsymbol{X}$: 4d gauge theories
  $\boldsymbol{P}$: $\boldsymbol{x}$ confines in the infrared limit
- $\boldsymbol{X}$: 1d quantum spin chains
  $\boldsymbol{P}$: $\boldsymbol{x}$ is gapless in the continuum limit
- $\boldsymbol{X}$: supersymmetric theories
  $\boldsymbol{P}$: $\boldsymbol{x}$ spontaneously breaks supersymmetry
- $\boldsymbol{X}$: various materials
  $\boldsymbol{P}$: $\boldsymbol{x}$ is superconducting at room temperature and normal pressure
- the list continues…

Theoretical physicists working on a topic $\boldsymbol{X}$
often spend their entire lives on answering the following question:

Given a system $\boldsymbol{x}$ in this topic $\boldsymbol{X}$, does it have property $\boldsymbol{P}$?

- $\boldsymbol{X}$: 4d gauge theories
  $\boldsymbol{P}$: $\boldsymbol{x}$ confines in the infrared limit
- $\boldsymbol{X}$: 1d quantum spin chains
  $\boldsymbol{P}$: $\boldsymbol{x}$ is gapless in the continuum limit
- $\boldsymbol{X}$: supersymmetric theories
  $\boldsymbol{P}$: $\boldsymbol{x}$ spontaneously breaks supersymmetry
- $\boldsymbol{X}$: various materials
  $\boldsymbol{P}$: $\boldsymbol{x}$ is superconducting at room temperature and normal pressure
- the list continues…

For example, last year, **I spent 50% of my research time to answer if the following 1d spin chain is a CFT**: The local Hilbert space have states

$$|1\rangle\,, |\alpha\rangle\,, |\alpha^2\rangle\,, |\rho\rangle\,, |\alpha\rho\rangle\,, |\alpha^2\rho\rangle\,,$$

with some nearest-neighbor constraints, and the Hamiltonian is

$$H = -\sum_i P_\rho^{(i)},$$

where

$$P_c^{(i)}\,|a_{i-1}a_ia_{i+1}\rangle = \sum_{a_i'} (F_{a_{i+1}}^{a_{i-1}\rho\rho})_{a_ic}\overline{(F_{a_{i+1}}^{a_{i-1}\rho\rho})}_{a_i'c}\,|a_{i-1}a_i'a_{i+1}\rangle\ ,$$

for a certain set of constants F_d^{abc}.

Details are not important for today's purpose.
It is just that this particular model is of some interest.

No exact solutions is known. So we needed to resort to numerical simulations on clusters for months.

It seems it leads to a CFT of $c \sim 2.0$ or 2.1.

[Huang-Lin-Ohmori-YT-Tezuka 2110.03008]
[Vanhove-Lootens-Van Damme
-Wolf-Osborne-Haegeman-Verstraete 2110.03532]
[Lin-Ryu-Zhu 2203.14992]

Why do we have to work this hard?

Shouldn't there be an algorithm, at least in the future when the technique is sufficiently developed, which allows us to determine whether a given 1d spin chain Hamiltonian gives a CFT or not?

Then I learned a shocking fact around last May:

It can happen that, for certain choices of (X, P),

> ***there is no uniform algorithm deciding***
> ***whether $P(x)$ holds for an arbitrary given $x \in X$.***

and

> ***there is a specific $x_0 \in X$ such that***
> ***whether $P(x_0)$ holds or not is unprovable.***

Let us say that such (X, P) is undecidable.

The existence of such undecidable questions $(\boldsymbol{X}, \boldsymbol{P})$ in physics was first explicitly pointed out in [Cubitt, Perez-Garcia, Wolf 1502.04135] to my knowledge.

There, the undecidability of the following was shown:

$\boldsymbol{X}$: 2d nearest-neighbor quantum spin systems

$\boldsymbol{P}$: $\boldsymbol{x}$ is gapless in the continuum limit

This result was further strengthened to 1d spin chains in [Bausch, Cubitt, Lucia, Perez-Garcia 1810.01858].

So, there cannot be any such algorithm solving the question for us.

In this talk, I would like to give a brief review of this result.

I will also show that the following question is also undecidable:

$\boldsymbol{X}$: 2d $\boldsymbol{\mathcal{N}=(2,2)}$ supersymmetric quantum field theory

$\boldsymbol{P}$: $\boldsymbol{x}$ has supersymmetric vacua.

All we need is some basic materials from

- **Theoretical computer science,**
- **Gödel's incompleteness theorems,**
- **Theory of Diophantine equations.**

I stress that what we need to use are all extremely fundamental and basic results in each of the subjects, discovered long time ago,
at least half a century and often even more.

They might be a bit exotic for standard mathematical physicists, though.

So my main aim is to give a very rough overview
of these standard materials.

I am not an expert myself.
If there is one in the audience,
please correct me if I say something wrong!

Halting problem

Let us start with a fundamental result
in theoretical computer science:

***There is no algorithm which tells
whether a program halts in finite time or not.***

Let us recall why this is the case.

We need to define what an algorithm is.
For us, a rough discussion suffices.

It is a finite-length program, written in a text file,
excecuted by a CPU deterministically,
depending on an input, which is given in a text file.

In your first computer science course, you learn how to program.

It often runs into an infinite loop and does not stop/halt.
You need to debug it, painstakingly.

We all wish a program to solve the halting problem: we want a program to see if another program runs into an infinite loop or not.

But this is impossible.

To see this, assume otherwise.

Then there is a program **h** which tells whether a program **p**, when given a text file $\boldsymbol{t}$ as an input, halts in finite time.

We now write the following program **c** which accepts an input text file $\boldsymbol{t}$:

- Regard the content of $\boldsymbol{t}$ as a program **t**
- Use **h** to see if **t** halts in finite time given $\boldsymbol{t}$ as input
- if **h** says it halts, go to an infinite loop.
- if **h** says it does not halt, halt.

(You are feeding the program text file **t** itself into the program $\boldsymbol{t}$.
The program $\boldsymbol{t}$ usually does not expect this.
But we do this anyway to stress-test the program.)

We now write this program into a text file $\boldsymbol{c}$ and
consider how **c** behaves with $\boldsymbol{c}$ as the input.

c with c as the input does the following:

- Regard the content of c as a program **c**
- Use **h** to see if **c** halts in finite time given c as input
- if **h** says it halts, go to an infinite loop.
- if **h** says it does not halt, halt.

We see that it halts if it doesn't , and it doesn't halt if it does.
Contradiction!

So there cannot be any program **h** which determines the halting problem.

Gödel's incompleteness theorems

Next, we give a brief review of Gödel's second incompleteness theorem, which says

> ***Whether the standard set theory is consistent or not cannot be proved within the standard set theory.***

To get some sense of what this means, you need some preparations.

A model of a given set of axioms is a mathematical structure satisfying the said set of axioms.

For example, let $\mathcal{G}$ be the set of axioms of a group.

A model of $\mathcal{G}$ is a group.

Whether a statement is true or false depends on a model.

For example, "$ab = ba$ for all a and b" is true for Abelian groups, but not for non-Abelian groups.

If a statement P can be proved from the said set of axioms, P is true in all models.

If a statement P cannot be proved from the said set of axioms, the validity of P depends on the chosen model.

So, there should be no problem in accepting that the questions

- **whether a statement can be proved or not**
- **whether a statement is true or false**

are a priori different for a given set of axioms. There can be unprovable statements, and its validity can depend on the model considered.

It is slightly weird that this state of affairs is applicable to the standard set theory and therefore the entire mathematics, but that's called life.

Therefore there are mathematical statements whose truth value depends on the model of set theory.

Gödel's second incomplete theorem says that the consistency of the set theory itself is one such statement.

What does it mean?

A set of axioms is called **consistent**
if you cannot prove a contradiction (such as $\mathbf{0} = \mathbf{1}$) from that.

Next, **a proof** of a theorem is a series of logical deductions starting from a set of axioms.

Finding a proof requires ingenuity.

In contrast, checking whether a proof is valid can be done mechanically in principle.

It is actually a thriving field of study, known as automatic proof verifiers.

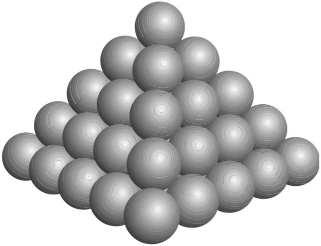
For example, the proof of Gödel's second incompleteness theorem has been verified [Paulson, 2104.13792, 2104.14260].

(The preprint versions are from last year, but the papers were published somewhat earlier in 2014.)

An aside

As another complicated proof which has been successfully verified ...

The statement that this standard packing



is the densest in 3d is known as Kepler's conjecture.

An aside

Thomas Hales announced his proof of it in [math.MG/9811078].

After a 7-year referring process,
it was accepted in [Annals of Mathematics 162(2005)1065].

But the referee did not say that the proof was perfectly correct,
since it used a lot of case-by-case analysis done by computer.

Hales was not satisfied, so decided to go into the proof verifier business
and verify the proof himself.

Started the [flyspeck project] to do this,
which took him and many other collaborators until 2014.

Which was then published in [Forum of Mathematics, 2017].

Coming back to our main topic:

The set of axioms we want to start from
can be written in a text file $\boldsymbol{a}$.

The statement of the theorem we want to prove
can be written in a text file $\boldsymbol{t}$.

The proof of the theorem from the axioms
can also be written in a text file $\boldsymbol{p}$.

A text file is just a series of 0s and 1s.
So it can be regarded as a (gigantic) integer.

So,

The set of axioms we want to start from
can be encoded in an integer $\boldsymbol{a}$.

The statement of the theorem we want to prove
can be encoded in an integer $\boldsymbol{t}$.

The proof of the theorem from the axioms
can be encoded in an integer $\boldsymbol{p}$.

Whether $\boldsymbol{p}$ is a valid proof of $\boldsymbol{t}$ from $\boldsymbol{a}$ can be algorithmically determined,
involving only arithmetic.

Therefore, the consistency of a set of axioms $\boldsymbol{a}$ can be phrased
arithmetically:

there is no integer $\boldsymbol{p}$ *encoding a proof of* $\mathbf{0 = 1}$.

So, the consistency of a set of axioms a is an arithmetical statement.

Gödel's second incompleteness theorem says that:

> ***for any consistent set of axioms a***
> ***which is rich enough to discribe arithmetic in it,***
> ***the consistency of a, phrased in this arithmetical manner,***
> ***cannot be proved.***

The standard set theory is an example of such a.

I don't have any ability to explain here how this theorem is proved.
Please read any standard textbook.

The statement "there is no integer $\boldsymbol{p}$ encoding the proof of $\mathbf{0 = 1}$" cannot be proved.

Therefore, **there are models in which this statement is true**, and **there are models in which this statement is false**.

In the latter case, the statement "there is an integer $\boldsymbol{p}$ encoding the proof of $\mathbf{0 = 1}$" is true.

How can that be, if the set theory is consistent (which is part of our assumption)?

Well, the set of integers "$\mathbb{N}$" in such models is **larger** than our $\mathbb{N}$, so there is an "integer" in "$\mathbb{N}$" which encodes the proof of $\mathbf{0 = 1}$ there.

It is somewhat like the fact that $\boldsymbol{x}^{\mathbf{2}} + \mathbf{1} = \mathbf{0}$ can be solved in $\mathbb{C}$ but not in $\mathbb{R}$, although what is going on in this case is subtler.

We now consider a program **x**:

- Enumerate all possible text files one by one.
- Check if it is the description of a proof of $\mathbf{0 = 1}$.

Deciding whether this program halts in finite time is equivalent to proving whether the standard set theory is consistent or not.

So, this is impossible.

We have thus established that the question

$\boldsymbol{\Xi}$: computer programs

$\boldsymbol{\Pi}$: $\boldsymbol{\xi}$ halts in finite time

is undecidable in the sense that

> *there is no uniform algorithm deciding*
> *whether $\boldsymbol{\Pi}(\boldsymbol{\xi})$ holds for an arbitrary given $\boldsymbol{\xi} \in \boldsymbol{\Xi}$.*

and

> *there is a specific $\boldsymbol{\xi_0} \in \boldsymbol{\Xi}$ such that*
> *whether $\boldsymbol{\Pi}(\boldsymbol{\xi_0})$ holds or not is unprovable.*

The first is the halting problem,
and the second is Gödel's second incompleteness theorem.

Diophantine equations

Next let us talk about Diophantine equations.

Hilbert, in his address to ICM Paris in 1900, listed his famous 23 problems. **The tenth** asks :

> *Is there an algorithmic method to decide whether a Diophantine equation is solvable, i.e. if a polynomial equation with integer coefficients,*
>
> $$P(x_1, \ldots, x_k) = 0,$$
>
> *has a solution in integers $x_i \in \mathbb{Z}$?*

This was answered negatively in 1970 by Matiyasevich, following the works of Davis, Putnam, Robinson.

The point was that, any program ξ can be translated (or compiled) into a Diophantine equation

$$P_\xi(x_1, \ldots, x_k) = 0$$

such that ξ **halts in finite time if and only if there is a solution** $x_i \in \mathbb{Z}$.

When ξ is the program looking for the proof of $0 = 1$ in the set theory, **proving that** $P_\xi(x_1, \ldots, x_k)$ **has no solution** is equivalent to **proving the consistency of the set theory**, and therefore impossible.

In a certain encoding, whether a program v halts on the input x is equivalent to whether the following Diophantine equation has a solution or not [Jones, Journal of Symbolic Logic 47 (1982) 549]:

$$
\begin{aligned}
0 = {} & (((zuy)^2 + u)^2 + y - v)^2 + (elg^2 + \alpha - (b - xy)q^2)^2 + (q - b^{5^{60}})^2 \\
& + (\lambda + q^4 - 1 - \lambda b^5)^2 + (\theta + 2z - b^5)^2 \\
& + (l - u - t\theta)^2 + (e - y - m\theta)^2 + (n - q^{16})^2 \\
& + (r - (g + eq^3 + lq^5 + (2(e - z\lambda)(1 + xb^5 + g)^4 + \lambda b^5 \\
& \qquad + \lambda b^5 q^4)q^4)(n^2 - n) - (q^3 - bl + l + \theta\lambda q^3 + (b^5 - 2)q^5)(n^2 - 1))^2 \\
& + (p - 2ws^2r^2n^2)^2 + (p^2k^2 - k^2 + 1 - \tau^2)^2 + (4(c - ksn^2)^2 + \eta - k^2)^2 \\
& + (k - (r + 1 + hp - h))^2 + (a - (wn^2 + 1)rsn^2)^2 + (c - 2r - 1 - \phi)^2 \\
& + (d - (bw + ca - 2c + 4a\gamma - 5\gamma))^2 + (d^2 - (a^2 - 1)c^2 - 1)^2 \\
& + (f^2 - (a^2 - 1)i^2c^4 - 1)^2 \\
& + ((d + of)^2 - ((a + f^2(d^2 - a))^2 - 1)(2r + 1 + jc)^2 - 1)^2.
\end{aligned}
$$

So there is no way to decide algorithmically whether this equation has an integer solution or not uniformly for all v and x.

Also, let us take $\boldsymbol{\xi}$ be the program looking for a proof of $\mathbf{0 = 1}$. Then

$$\boldsymbol{P_{\xi}(x_1, \ldots, x_k) = 0}$$

has an integer solution **if and only if the standard set theory is inconsistent**.

Therefore you can never prove that it has no solution,
although many people believe that it has no solution.

Cond-mat realization

In [Cubitt, Perez-Garcia, Wolf 1502.04135],
the undecidability of the problem

$\boldsymbol{X}$: 2d nearest-neighbor quantum spin systems

$\boldsymbol{P}$: $\boldsymbol{x}$ becomes CFT in the continuum limit

was shown by constructing a spin system $\boldsymbol{x_\xi}$ depending on a computer program $\boldsymbol{\xi}$ such that $\boldsymbol{x_\xi}$ is gapped if and only if $\boldsymbol{\xi}$ halts in finite time.

Very briefly, **they represented** the running of a computer program in the ground state of a spin system such that **the running time is encoded spatially.**

In slightly more detail, given a program ξ, they construct a lattice spin system H_ξ with the following property:

$$\begin{cases} E_0 < -1 & (L < L_\xi), \\ E_0 > +1 & (L > L_\xi). \end{cases}$$

Here, L is the system size, L_ξ is a constant related to the total running time of the program ξ, and E_0 is the ground state energy.

In the large size limit, $E_0 > +1$ or $E_0 < -1$ depending on whether ξ halts or not.

Let us say that H_ξ acts on $\bigotimes_{\text{site } i} V_i = \bigotimes_{\text{site } i} \mathbb{C}^n$.

Choose two of your favorite spin systems $H_{A,B}$ acting on

$$\bigotimes_{\text{site } i} A_i = \bigotimes_{\text{site } i} \mathbb{C}^a, \qquad \bigotimes_{\text{site } i} B_i = \bigotimes_{\text{site } i} \mathbb{C}^b.$$

Normalize so that the ground states of H_A and H_B have zero energy.

Now consider a combined spin system acting on

$$\bigotimes_i \left[(V_i \otimes A_i) \oplus B_i \right]$$

with the Hamiltonian

$$H_{\text{combined}} = H_\xi \otimes H_A + H_B + \text{(penalty term)}$$

where the 'penalty term' is to disfavor wavefunctions on both $V \otimes A$ and B.

So the ground state of H_{combined} is either

$$\text{(ground state of } H_\xi) \otimes \text{(ground state of } H_A)$$

or

$$\text{(ground state of } H_B),$$

depending on which has lower energy.

By construction,

$$(\text{ground state of } \boldsymbol{H_\xi}) \otimes (\text{ground state of } \boldsymbol{H_A})$$

has energy

$$\begin{cases} \boldsymbol{E > +1} & \boldsymbol{(L > L_\xi)}, \\ \boldsymbol{E < -1} & \boldsymbol{(L < L_\xi)} \end{cases}$$

while

$$(\text{ground state of } \boldsymbol{H_B}),$$

has energy $\boldsymbol{E = 0}$.

Therefore, the ground state of the combined system is

$$(\text{ground state of } \boldsymbol{H_\xi}) \otimes (\text{ground state of } \boldsymbol{H_A})$$

when $\boldsymbol{L < L_\xi}$ and is

$$(\text{ground state of } \boldsymbol{H_B}),$$

when $\boldsymbol{L > L_\xi}$.

Let $\boldsymbol{H_A}$ **be a known massive theory** and $\boldsymbol{H_B}$ **be a known CFT**.

Then **the combined system is CFT if and only if $\boldsymbol{\xi}$ halts**.

Let $\boldsymbol{H_A}$ **be a known normal theory** and
$\boldsymbol{H_B}$ **to be a known superconducting theory**.

Then **the combined system is superconducting if and only if** $\boldsymbol{\xi}$ **halts**.

Let us further take $\boldsymbol{\xi}$ to be the program looking for a proof of $\mathbf{0 = 1}$.

Then **the combined system is superconducting**
if and only if the standard set theory is inconsistent.

This might actually be a way to check if the standard axioms of math is consistent, if such a material can be synthesized.

We can look for proofs of $\mathbf{0 = 1}$ whose lengths approaches Avogadro's number, which is considerably longer than any proofs published in any journal!

2d SUSY realization

The only new result in my talk today is that the question

$\boldsymbol{X}$: 2d $\boldsymbol{\mathcal{N}=(2,2)}$ supersymmetric quantum field theories

$\boldsymbol{P}$: $\boldsymbol{x}$ has supersymmetric vacua

is undecidable.

Again the strategy is the same:

the reduction to the undecidability of the halting problem.

I use an intermediate step, namely

the undecidability of solvability of Diophantine equations.

Recall that a Diophantine equation is a polynomial equation of many variables

$$P(x_1, x_2, \ldots, x_n) = 0$$

where the coefficients are all in $\mathbb{Z}$ and the unknowns x_i are also all in $\mathbb{Z}$.

Recall also that for any program ξ there is a Diophantine equation P_ξ such that

P_ξ *has a solution if and only if the program* ξ *halts in finite time.*

Given this, it is easy to translate it to the undecidability of the existence of supersymmetric vacua of 2d $\mathcal{N}=(2,2)$ theories.

We promote $x_1, \ldots, x_n$ to chiral superfields $X_1, \ldots, X_n$.

We introduce superfields $Z_1, \ldots, Z_n$ and finally Y.

Consider the superpotential

$$W_\xi = Y P_\xi(X_1, \ldots, X_n)^2 + \sum_a Z_a (\sin 2\pi i X_a)^2.$$

Consider the superpotential

$$W_\xi = Y P_\xi(X_1, \ldots, X_n)^2 + \sum_a Z_a (\sin 2\pi i X_a)^2.$$

- $\dfrac{\partial W_\xi}{\partial Z_a} = 0$ imposes $\sin 2\pi i X_a = 0$, i.e. $X_a \in \mathbb{Z}$.
- $\dfrac{\partial W_\xi}{\partial Y} = 0$ imposes $P_\xi(X_1, \ldots, X_n) = 0$.
- $\dfrac{\partial W_\xi}{\partial X_a} = 0$ does not impose any further conditions.

This system has supersymmetric vacua (parameterized by Z_a and Y) if and only if the Diophantine equation P_ξ has a solution.

Done.

So, **there cannot be an algorithm deciding whether a given 2d $\mathcal{N}=(2,2)$ Wess-Zumino model has SUSY vacua or not.**

Also, **taking ξ to be the program looking for the proof of $0=1$, the corresponding 2d $\mathcal{N}=(2,2)$ theory has SUSY vacua if and only if the standard set of axioms of math is inconsistent!**

Summary

There are cases that, for certain choices
of theoretical physics questions (X, P),

> ***there is no uniform algorithm deciding whether $P(x)$ holds for an arbitrary given $x \in X$,***

and

> ***there is a specific $x_0 \in X$ such that whether $P(x_0)$ holds or not is unprovable.***