

# String Theory And The Universe

Edward Witten

IFT-UNESP, São Paulo

February 14, 2011

This will be a talk about string theory,

This will be a talk about string theory, but more precisely, I will tell the story by explaining how string theory has grown beyond anything that anyone has conceived at any step along the way.

This will be a talk about string theory, but more precisely, I will tell the story by explaining how string theory has grown beyond anything that anyone has conceived at any step along the way.

But my purpose is not mainly a historical retrospective.

This will be a talk about string theory, but more precisely, I will tell the story by explaining how string theory has grown beyond anything that anyone has conceived at any step along the way.

But my purpose is not mainly a historical retrospective.

The main point is rather that I believe that this process is far from finished.

This will be a talk about string theory, but more precisely, I will tell the story by explaining how string theory has grown beyond anything that anyone has conceived at any step along the way.

But my purpose is not mainly a historical retrospective.

The main point is rather that I believe that this process is far from finished.

We still can't really conceive the end of the journey.

Actually, I am at best only semi-qualified to offer a historical account for at least one obvious reason:

Actually, I am at best only semi-qualified to offer a historical account for at least one obvious reason:

I didn't get to see, much less participate in, all of the important stages.

Actually, I am at best only semi-qualified to offer a historical account for at least one obvious reason:

I didn't get to see, much less participate in, all of the important stages.

Indeed, few people have, as this process has by now gone on for forty years.

Actually, I am at best only semi-qualified to offer a historical account for at least one obvious reason:

I didn't get to see, much less participate in, all of the important stages.

Indeed, few people have, as this process has by now gone on for forty years.

In my case, I wasn't directly involved until the early 1980's (or the late 1970's if you are generous) by which time a lot of things had happened.

So I can't say what I'd have thought at the time of the idea of inventing a new theory by concocting a scattering amplitude out of whole cloth

So I can't say what I'd have thought at the time of the idea of inventing a new theory by concocting a scattering amplitude out of whole cloth – namely the Veneziano amplitude:

$$A(s, t) = \frac{\Gamma(-s)\Gamma(-t)}{\Gamma(-s-t)}.$$

So I can't say what I'd have thought at the time of the idea of inventing a new theory by concocting a scattering amplitude out of whole cloth – namely the Veneziano amplitude:

$$A(s, t) = \frac{\Gamma(-s)\Gamma(-t)}{\Gamma(-s-t)}.$$

I can only say that looking at it now, I am surprised that trying to guess a formula like that proved to be the starting point for something significant.

The motivation had to do with hadron physics. The proliferation of hadronic resonances had made some physicists despair of describing the strong interactions via the known framework of local quantum field theory.

The motivation had to do with hadron physics. The proliferation of hadronic resonances had made some physicists despair of describing the strong interactions via the known framework of local quantum field theory. But one can't get anywhere with just a negative assumption – no local fields. It takes a positive idea.

The positive idea that led to the Veneziano amplitude was this: as there were so many hadron resonances, it seemed clear that resonant scattering was an important part of hadronic interactions.

The positive idea that led to the Veneziano amplitude was this: as there were so many hadron resonances, it seemed clear that resonant scattering was an important part of hadronic interactions.

This led to the following question: Can a scattering amplitude that is given by a sum of resonant contributions only

$$A(s, t) = \sum_n \frac{P_n(t)}{s - M_n^2}$$

The positive idea that led to the Veneziano amplitude was this: as there were so many hadron resonances, it seemed clear that resonant scattering was an important part of hadronic interactions.

This led to the following question: Can a scattering amplitude that is given by a sum of resonant contributions only

$$A(s, t) = \sum_n \frac{P_n(t)}{s - M_n^2}$$

be the whole story?

The positive idea that led to the Veneziano amplitude was this: as there were so many hadron resonances, it seemed clear that resonant scattering was an important part of hadronic interactions.

This led to the following question: Can a scattering amplitude that is given by a sum of resonant contributions only

$$A(s, t) = \sum_n \frac{P_n(t)}{s - M_n^2}$$

be the whole story? With only finitely many poles, the answer is definitely “no”; can it work with infinitely many?

Not having participated in the physics of the 1960's, I can't really say how well-motivated this question was at the time, but it certainly proved to be a fruitful one.

Not having participated in the physics of the 1960's, I can't really say how well-motivated this question was at the time, but it certainly proved to be a fruitful one.

First of all, it turned out that there is a physical picture behind the Veneziano amplitude.

A meson is a little string with charges at its ends:



The meson resonances – which correspond to the poles of the Veneziano amplitude – are vibrational states of the string.

This physical picture – which first emerged from studies of the Veneziano amplitude – is now believed to be qualitatively correct as a description of strongly interacting particles

This physical picture – which first emerged from studies of the Veneziano amplitude – is now believed to be qualitatively correct as a description of strongly interacting particles or at least mesons

This physical picture – which first emerged from studies of the Veneziano amplitude – is now believed to be qualitatively correct as a description of strongly interacting particles or at least mesons – but (at least in its simple versions) only qualitatively correct.

This physical picture – which first emerged from studies of the Veneziano amplitude – is now believed to be qualitatively correct as a description of strongly interacting particles or at least mesons – but (at least in its simple versions) only qualitatively correct. Quantum Chromodynamics or QCD – which emerged about five years after the Veneziano amplitude – is far more successful in describing the details of strong interactions... so much so that for some years string theory went into eclipse.

Later developments showed that the apparent failure of string theory as a theory of strong interactions wasn't the last word.

Later developments showed that the apparent failure of string theory as a theory of strong interactions wasn't the last word. It has actually taught us a lot about QCD, with more probably still to come.

Later developments showed that the apparent failure of string theory as a theory of strong interactions wasn't the last word. It has actually taught us a lot about QCD, with more probably still to come. The modern view is actually that QCD may be equivalent to a not-yet-discovered string theory.

But the bigger reason that string theory survived was not that it actually has some relevance to strong interactions.

But the bigger reason that string theory survived was not that it actually has some relevance to strong interactions. It was that although the idea of a theory without local fields was wrong for strong interactions, there is another problem in physics for which it is right.

This is the problem of quantum gravity.

This is the problem of quantum gravity. We need a theory of quantum gravity, since quantum mechanics and gravity do exist in the real world.

This is the problem of quantum gravity. We need a theory of quantum gravity, since quantum mechanics and gravity do exist in the real world. And it seems that we cannot understand quantum gravity via the usual algorithm “quantize classical fields,”

This is the problem of quantum gravity. We need a theory of quantum gravity, since quantum mechanics and gravity do exist in the real world. And it seems that we cannot understand quantum gravity via the usual algorithm “quantize classical fields,” since if we try to do this for General Relativity, we run into a maze of infinities.

In the context of gravity, there cannot be a gauge-invariant local field  $\phi(x)$

In the context of gravity, there cannot  
be a gauge-invariant local field  $\phi(x)$   
... since  $x$  itself isn't gauge-invariant!

In the context of gravity, there cannot be a gauge-invariant local field  $\phi(x)$  ... since  $x$  itself isn't gauge-invariant!

This was pointed out by Bryce de Witt in 1968 – coincidentally the same year the Veneziano amplitude was written down, though the two insights were related only much later.

So a theory with gauge-invariant local fields cannot describe quantum gravity.

So a theory with gauge-invariant local fields cannot describe quantum gravity.

Of course, there is no immediate converse to this statement.

The fact that the inventors of string theory developed a theory without gauge-invariant local fields did not in and of itself ensure that the theory was going to describe quantum gravity.

The fact that the inventors of string theory developed a theory without gauge-invariant local fields did not in and of itself ensure that the theory was going to describe quantum gravity.

But that turned out to happen.

From rather early days, it was found that string theory generated massless spin one particles with gauge-like properties (open strings) and massless spin two particles with graviton-like properties (closed strings).

From rather early days, it was found that string theory generated massless spin one particles with gauge-like properties (open strings) and massless spin two particles with graviton-like properties (closed strings). This seemed bad ...

From rather early days, it was found that string theory generated massless spin one particles with gauge-like properties (open strings) and massless spin two particles with graviton-like properties (closed strings). This seemed bad ... since the theory was supposed to describe the strong interactions, which have a “mass gap” – all particles are massive.

From rather early days, it was found that string theory generated massless spin one particles with gauge-like properties (open strings) and massless spin two particles with graviton-like properties (closed strings). This seemed bad ... since the theory was supposed to describe the strong interactions, which have a “mass gap” – all particles are massive. People tried hard to get rid of these massless particles

From rather early days, it was found that string theory generated massless spin one particles with gauge-like properties (open strings) and massless spin two particles with graviton-like properties (closed strings). This seemed bad ... since the theory was supposed to describe the strong interactions, which have a “mass gap” – all particles are massive. People tried hard to get rid of these massless particles ... and failed.

But eventually, some physicists became daring enough to propose that string theory had been misinterpreted

But eventually, some physicists became daring enough to propose that string theory had been misinterpreted ... that the strings were *much* smaller than had been assumed

But eventually, some physicists became daring enough to propose that string theory had been misinterpreted ... that the strings were *much* smaller than had been assumed ... that the theory should be used to describe quantum gravity and other fundamental forces.

But eventually, some physicists became daring enough to propose that string theory had been misinterpreted ... that the strings were *much* smaller than had been assumed ... that the theory should be used to describe quantum gravity and other fundamental forces. (These are still developments that occurred before I was in the field, so I am giving only a second-hand summary.)

So – not for the first or last time – the theory veered in a direction that wasn't foreseen by those who had gotten the ball rolling.

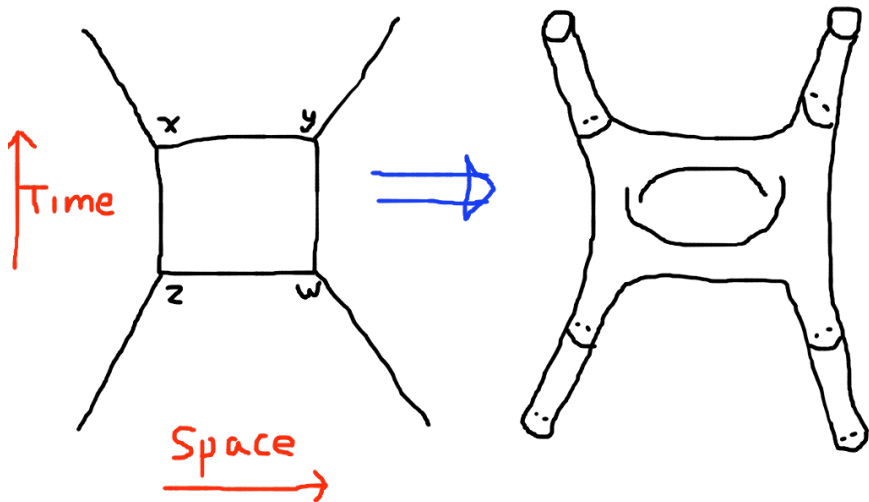
So – not for the first or last time – the theory veered in a direction that wasn't foreseen by those who had gotten the ball rolling.

There is actually a simple picture that helps explain why string theory has a mind of its own.

So – not for the first or last time – the theory veered in a direction that wasn't foreseen by those who had gotten the ball rolling.

There is actually a simple picture that helps explain why string theory has a mind of its own.

It has to do with what happens to Feynman diagrams in going to string theory.



This really has a lot of consequences.  
Here are three:

This really has a lot of consequences.  
Here are three:

- \* The infinities of Feynman diagrams disappear.

This really has a lot of consequences.  
Here are three:

- \* The infinities of Feynman diagrams disappear.

- \* What the interactions are is not up to us – they are fixed once we pick the “string.”

This really has a lot of consequences.  
Here are three:

- \* The infinities of Feynman diagrams disappear.

- \* What the interactions are is not up to us – they are fixed once we pick the “string.”

- \* There is no definite notion of a spacetime event.

So physicists arrived at a theory that

So physicists arrived at a theory that  
(i) is completely finite, since the  
interaction vertices are missing,

So physicists arrived at a theory that  
(i) is completely finite, since the  
interaction vertices are missing,  
(ii) has a massless spin two particle,  
since no one could quantize the string  
without getting one, and

So physicists arrived at a theory that  
(i) is completely finite, since the  
interaction vertices are missing,  
(ii) has a massless spin two particle,  
since no one could quantize the string  
without getting one, and  
(iii) has its own mind concerning  
what the interactions will be.

The interactions that the theory chooses are those of General Relativity (plus short distance corrections that preserve the general structure of Einstein's theory).

But our picture has many other consequences.

But our picture has many other consequences. Here are three more:

- \* There are far fewer theories than in standard quantum field theory – five as counted in the mid-1980's and only one as understood today.

But our picture has many other consequences. Here are three more:

- \* There are far fewer theories than in standard quantum field theory – five as counted in the mid-1980's and only one as understood today.
- \* The theory picks the spacetime dimension.

But our picture has many other consequences. Here are three more:

- \* There are far fewer theories than in standard quantum field theory – five as counted in the mid-1980's and only one as understood today.
- \* The theory picks the spacetime dimension.
- \* The theory picks its own symmetries. We don't pick them.

The spacetime dimension turned out to be 10 – which no one wanted going in.

The spacetime dimension turned out to be 10 – which no one wanted going in. At first it must have sounded like a joke.

The spacetime dimension turned out to be 10 – which no one wanted going in. At first it must have sounded like a joke. But when the theory was reinterpreted as a candidate unified theory of all particles, the extra dimensions gave “room” to derive the complexity of the real world from a simple starting point.

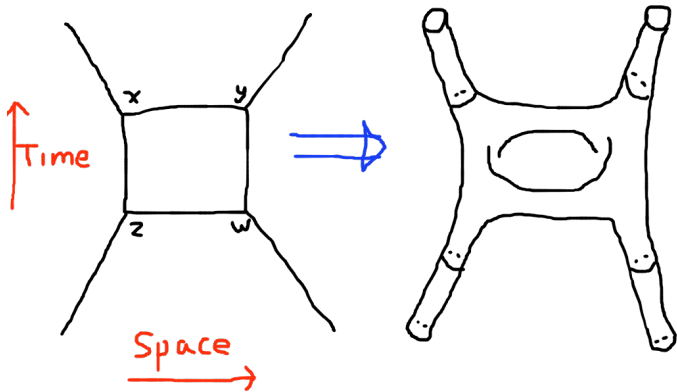
As for the symmetries, in elucidating them, physicists re-discovered general covariance and gauge symmetry (which of course were already known) but also discovered a new type of symmetry – supersymmetry.

Supersymmetry is the main contribution that string theory has made to predicting what is new that might be discovered in particle physics. (Other more exotic possibilities involve “large” extra dimensions and the like.)

Supersymmetry led to the extraordinarily rich subject of supergravity – which is what we get in making a supersymmetric theory that describes gravity.

Supersymmetry and supergravity are actually the tip of a much bigger iceberg – string theory is somehow based on a new kind of geometry that we don't understand.

Certainly we don't understand what kind of new geometry does this to the Feynman diagrams:



We don't even know a world in which this is a sensible question.

We don't even know a world in which this is a sensible question. That is why it has been wisely said (not originally by me) that string theory is part of twenty-first century physics that fell by chance into the twentieth century.

We don't even know a world in which this is a sensible question. That is why it has been wisely said (not originally by me) that string theory is part of twenty-first century physics that fell by chance into the twentieth century. No one had the conception for it, and we still don't today.

Somehow, there is a new kind of geometry in which one isn't allowed to talk about points or lines in spacetime, but one is allowed to talk about (quantum) minimal surfaces.

Somehow, there is a new kind of geometry in which one isn't allowed to talk about points or lines in spacetime, but one is allowed to talk about (quantum) minimal surfaces. Certainly no one at the start foresaw something like this.

Somehow, there is a new kind of geometry in which one isn't allowed to talk about points or lines in spacetime, but one is allowed to talk about (quantum) minimal surfaces.

Certainly no one at the start foresaw something like this.

So again, this is part of what I mean in saying that the theory has kept imposing its will upon us.

These questions really were not very popular until 1984 when a new method of anomaly cancellation (Green and Schwarz) and the heterotic string (Gross, Harvey, Martinec, and Rohm) made it possible to derive from string theory a decent rough draft of the world, with the gauge forces and quarks and leptons plus quantum gravity.

To me the anomaly cancellation was  
a sort of signal from heaven.

To me the anomaly cancellation was a sort of signal from heaven. A lot of miracles had gone into the development of string theory up to this point, but this was the first one that really happened while I was watching closely.

To me the anomaly cancellation was a sort of signal from heaven. A lot of miracles had gone into the development of string theory up to this point, but this was the first one that really happened while I was watching closely. From this point on, string theory and related matters have been my main scientific interest.

The prevailing view of this period was that string theory differs from pre-twentieth century physics by the presence of two parameters, namely  $\hbar$  and  $\alpha'$ .

The prevailing view of this period was that string theory differs from pre-twentieth century physics by the presence of two parameters, namely  $\hbar$  and  $\alpha'$ . Every physicist knows about  $\hbar$ .

The prevailing view of this period was that string theory differs from pre-twentieth century physics by the presence of two parameters, namely  $\hbar$  and  $\alpha'$ . Every physicist knows about  $\hbar$ . In string theory,  $\alpha'$  is a fundamental unit of length, below which ordinary ideas in geometry fail.

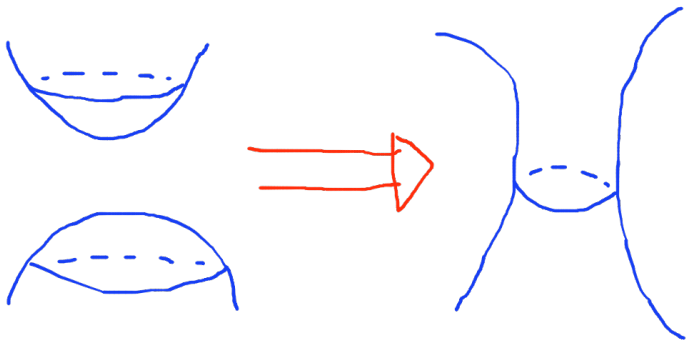
If string theory is right, the deformation by  $\alpha'$  is as fundamental and interesting as the deformation by  $\hbar$ , though quite different.

Most string theorists (including me, for the most part) spent the decade after 1984 studying the  $\alpha'$  deformation.

Most string theorists (including me, for the most part) spent the decade after 1984 studying the  $\alpha'$  deformation. The main tool was two-dimensional conformal field theory, which describes propagation of the string.

Most string theorists (including me, for the most part) spent the decade after 1984 studying the  $\alpha'$  deformation. The main tool was two-dimensional conformal field theory, which describes propagation of the string.

Typical results of the period involved smooth descriptions of processes that in ordinary geometry require passing through a singularity:



The prevailing paradigm held that string theory is based on an unusual (because  $\alpha'$ -deformed) classical theory that then gets quantized (i.e.  $\hbar$ -deformed) in a “normal” way.

Two-dimensional conformal field theory – which describes the motion of a string – was seen as the best tool to study the  $\alpha'$ -deformed classical theory.

Two-dimensional conformal field theory – which describes the motion of a string – was seen as the best tool to study the  $\alpha'$ -deformed classical theory. Thus it seemed like the important branch of quantum field theory for string theorists.

I remember that this frustrated me a bit, since I had spent much of my youth learning about more general varieties of quantum field theory – starting with the four-dimensional case, of course.

I remember that this frustrated me a bit, since I had spent much of my youth learning about more general varieties of quantum field theory – starting with the four-dimensional case, of course. But I don't think I ever channeled this frustration into trying to improve on the then-prevailing paradigm.

A few (largely British) physicists did try to go beyond it, starting with a simple question: Why stop at strings? Why not membranes?



Taken literally, the question “why not membranes?” actually has a good answer:

Taken literally, the question “why not membranes?” actually has a good answer: Strings work better than membranes or other objects of higher dimension because of the unique properties of complex numbers.

Taken literally, the question “why not membranes?” actually has a good answer: Strings work better than membranes or other objects of higher dimension because of the unique properties of complex numbers. A complex number  $z = x + iy$  is a two-dimensional object, like the worldsheet of a string.

But this did not prove to be the whole story.

But this did not prove to be the whole story. Eventually a more subtle version of the membrane idea emerged, and now we know that membranes and higher objects are actually part of string theory – not an alternative theory but part of the same theory.

Meanwhile another idea was developing that also played a role in challenging the established paradigm.

Meanwhile another idea was developing that also played a role in challenging the established paradigm. This was electric-magnetic duality.

Since the nineteenth century, physicists had been fascinated by the symmetry of Maxwell's vacuum equations between electric and magnetic fields:  $E \leftrightarrow B$ .

Since the nineteenth century, physicists had been fascinated by the symmetry of Maxwell's vacuum equations between electric and magnetic fields:  $E \leftrightarrow B$ . This symmetry seems to be spoiled by the fact that in nature we see electric charges and no magnetic ones.

Quantum mechanics seems to tell us that the symmetry is actually impossible since we need to introduce a vector potential in order to write the Schrodinger equation in a magnetic field.

Quantum mechanics seems to tell us that the symmetry is actually impossible since we need to introduce a vector potential in order to write the Schrodinger equation in a magnetic field. We derive the magnetic field from a vector potential and the electric field from a scalar potential.

Dirac was the one who explained how magnetic monopoles can be consistent with quantum mechanics. This required his famous quantization law:

$$eg = 2\pi\hbar c \cdot \text{integer}.$$

Dirac was the one who explained how magnetic monopoles can be consistent with quantum mechanics. This required his famous quantization law:

$$eg = 2\pi\hbar c \cdot \text{integer}.$$

But it wasn't until the 1970's that convincing quantum theories were described that have magnetic monopoles.

Also in the 1970's, QCD was developed and it raised a lot of intractable problems – the most obvious being why quarks are “confined.”

Also in the 1970's, QCD was developed and it raised a lot of intractable problems – the most obvious being why quarks are “confined.” It was suggested (by Polyakov, 't Hooft, Mandelstam, and Nambu, among others) that this could be explained in an alternative description of QCD using “magnetic variables.”

It sounds nice, but there is a catch.

It sounds nice, but there is a catch.  
No one knew what the magnetic  
variables would be in the case of QCD  
...

It sounds nice, but there is a catch.  
No one knew what the magnetic variables would be in the case of QCD ... There wasn't any real progress until the mid-1990's, and even today we only have a partial understanding.

The first relatively clear conjecture for how a nonabelian gauge theory could have a symmetry between electricity and magnetism was made (in steps, starting with Goddard-Nuyts-Olive and Montonen-Olive) in the 1970's.

The first relatively clear conjecture for how a nonabelian gauge theory could have a symmetry between electricity and magnetism was made (in steps, starting with Goddard-Nuyts-Olive and Montonen-Olive) in the 1970's. But it took a long time until these ideas were developed in a way that most physicists found convincing and useful.

By the mid-1990's, a number of clues suggested that electric-magnetic duality (that is, symmetry between electricity and magnetism) is important in the general structure of string theory.

By the mid-1990's, a number of clues suggested that electric-magnetic duality (that is, symmetry between electricity and magnetism) is important in the general structure of string theory. The most direct was the fact that electric-magnetic duality is important in supergravity (which is a semiclassical limit of string theory).

The various strands – such as membranes and electric-magnetic duality – were knitted together in the mid-1990's.

The various strands – such as membranes and electric-magnetic duality – were knitted together in the mid-1990's. But all I will really be able to do today is to just give a few hints about the new paradigm that resulted.

If a symmetry that exchanges electric and magnetic charge is important in the structure of the theory, then because of the  $\hbar$  in Dirac's formula

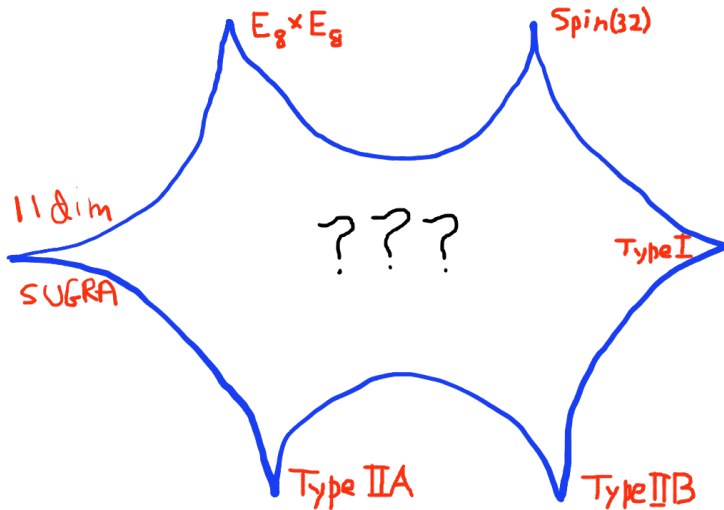
$$eg = 2\pi\hbar c \cdot \text{integer},$$

the basic structure of the theory can only be understood quantum mechanically.

The idea “ $\alpha'$  first,  $\hbar$  second” cannot be correct. They must be treated together.

The idea “ $\alpha'$  first,  $\hbar$  second” cannot be correct. They must be treated together. The theory can't be properly understood as quantization of a classical theory so in some sense, if correct, it must give a new interpretation of what quantum mechanics means.

When  $\alpha'$  and  $\hbar$  are treated together, one gets a nice surprise – it turns out that there is only one string theory, that is only one candidate for superunification of the laws of nature.

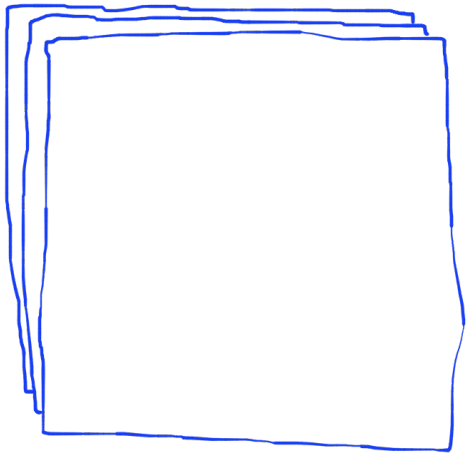


Previous paradigm: String theory is based on two-dimensional conformal field theory.

Previous paradigm: String theory is based on two-dimensional conformal field theory.

New view: All quantum field theory in all dimensions is part of the input for string theory.

For example,  $U(N)$  gauge theory in any dimension describes the behavior of  $N$  parallel branes:



This has numerous applications,  
among them

- \* Quantum mechanics of black holes
- \* Better understanding of four-dimensional gauge theory

For the black holes: Build a black hole from parallel branes. Describe the quantum states of the black hole using gauge theory. Get a microscopic explanation of the Bekenstein-Hawking entropy of the black hole, by counting quantum states, using the methods of gauge theory.

To understand gauge theory better, do the reverse: Map questions in gauge theory into questions about gravity that (when one is fortunate) are more easily answered.

To understand gauge theory better, do the reverse: Map questions in gauge theory into questions about gravity that (when one is fortunate) are more easily answered. For example, in the right setup, it turns out that confinement of quarks is a consequence of the existence of a black hole horizon.

The most spectacular success along these lines has been the modeling of the quark-gluon plasma produced at RHIC in terms of string theory and black holes.

In sum, in the 1990's the prevailing view of what string theory means and how one has to try to understand it was greatly enlarged again.

In sum, in the 1990's the prevailing view of what string theory means and how one has to try to understand it was greatly enlarged again. One may ask, what is next? What might be the next major change in perspective?

Famously, it is hard to make predictions, especially about the future.

Famously, it is hard to make predictions, especially about the future. But here is a case in which it is difficult to know what to say about the recent past.

Famously, it is hard to make predictions, especially about the future. But here is a case in which it is difficult to know what to say about the recent past. There possibly has been another major change in the interpretation of the theory in roughly the last decade ... but it is hard to know.

Since there is only one string theory, it seems reasonable to hope that it might have only one quantum vacuum describing our Universe and leading to definite predictions for the values of all of the dimensionless natural constants.

Since there is only one string theory, it seems reasonable to hope that it might have only one quantum vacuum describing our Universe and leading to definite predictions for the values of all of the dimensionless natural constants. But research in string theory has never supported this hope.

For example, back in the mid-1980's, models of string compactification to four dimensions gave not one possibility but many that were at least roughly compatible with what we observe.

For example, back in the mid-1980's, models of string compactification to four dimensions gave not one possibility but many that were at least roughly compatible with what we observe. I always hoped that this would go away, but later research always had the effect of just generating more possibilities.

I was one of the people working on these constructions.

I was one of the people working on these constructions. I am tempted to say that the proliferation of possibilities was the embarrassment of my youth.

But meanwhile, developments in astronomy have made some of our most distinguished colleagues suspect that the real world may be based on just such a “megaverse” of possibilities.

The two most relevant developments are the apparent observation of a tiny but not zero cosmological constant, and the successes of the theory of cosmic inflation in describing the fluctuations of the cosmic microwave radiation.



The process of learning what string theory means has a long way still to run.

The process of learning what string theory means has a long way still to run. I don't think we are close to seeing the end of it.