Physics and Formal Methods

Michael R. Douglas

[1] Simons Center / Stony Brook University

ICMS, May 28, 2019

Rapid progress in computer science and AI has motivated us to meet and discuss the state of the art and prospects for applying formal methods in mathematics: technologies such as automated theorem verification and proof, new formal languages, and new foundations.

In this talk we will review a few problems coming from theoretical physics and initiate an analogous discussion. Here is a brief outline:

1. Broad overview of the physics point of view.
2. Can one base formalization on something other than mathematical foundations?
3. An example from particle physics: compute $g-2$ of the muon.
4. An example from string theory: find Calabi-Yau manifolds with desired properties.
5. Communication in the mathematical sciences: search, integration, formal wiki.

See also my talk at AITP 2018, http://aitp-conference.org/2018/slides/MD.pdf

Rapid progress in computer science and AI has motivated us to meet and discuss the state of the art and prospects for applying formal methods in mathematics: technologies such as automated theorem verification and proof, new formal languages, and new foundations.

In this talk we will review a few problems coming from theoretical physics and initiate an analogous discussion. Here is a brief outline:

1. Broad overview of the physics point of view.
2. Can one base formalization on something other than mathematical foundations?
3. An example from particle physics: compute $g-2$ of the muon.
4. An example from string theory: find Calabi-Yau manifolds with desired properties.
5. Communication in the mathematical sciences: search, integration, formal wiki.

See also my talk at AITP 2018, `http://aitp-conference.org/2018/slides/MD.pdf`

Relations between physics and mathematics

Despite their common origins, and their many common interests, the communities of theoretical physicists and of pure and applied mathematicians work in rather different ways, each with advantages and disadvantages. Let us try to summarize the physics worldview.

The starting point for almost all work in physics is a "theory" (or model), a mathematical framework which describes some aspects of the real world. Examples include

- **CM** the classical mechanics of a finite number of particles, as used in celestial mechanics.
- **QM** the quantum mechanics of a finite number of particles, as used in atomic physics.
- **QED** quantum electrodynamics, a quantum field theory.
- **YM** Yang-Mills theory, the basis of the Standard Model.
- **ST** String/M theory, a candidate fundamental theory which unifies Yang-Mills and quantum gravity.

Relations between physics and mathematics

Despite their common origins, and their many common interests, the communities of theoretical physicists and of pure and applied mathematicians work in rather different ways, each with advantages and disadvantages. Let us try to summarize the physics worldview.

The starting point for almost all work in physics is a “theory” (or model), a mathematical framework which describes some aspects of the real world. Examples include

- **CM** the classical mechanics of a finite number of particles, as used in celestial mechanics.
- **QM** the quantum mechanics of a finite number of particles, as used in atomic physics.
- **QED** quantum electrodynamics, a quantum field theory.
- **YM** Yang-Mills theory, the basis of the Standard Model.
- **ST** String/M theory, a candidate fundamental theory which unifies Yang-Mills and quantum gravity.

The basic equations:

CM State at time t: a list of positions $\vec{x}_i(t)$.
Equation of motion:

$$m_i \frac{\partial^2 \vec{x}_i}{\partial t^2} = -\sum_{j \neq i} \frac{G m_i m_j (\vec{x}_i - \vec{x}_j)}{|\vec{x}_i - \vec{x}_j|^3}.$$

QM State: a ray $\Psi \in L^2(\mathbb{R}^{3N}, \mathbb{C})$.
Equation of motion:

$$i\hbar \frac{\partial}{\partial t} \Psi = \left[-\sum_i \frac{\hbar^2}{2m_i} \frac{\partial^2}{\partial x_i^2} + V(x_i) \right] \Psi.$$

QFT

$$Z = \int [D\phi] \exp - \int d^D x \, (\partial \phi)^2 + \frac{m^2}{2} \phi^2 + \frac{g}{24} \phi^4.$$

Considered as mathematics, the status of these theories varies:

- QM has a satisfactory mathematically rigorous definition, based on the Schrödinger equation and functional analysis.
- CM can be treated rigorously, but it is not fundamental, which leads to problems – singularities, chaos and precision issues.
- QED can be defined rigorously in perturbation theory, meaning as an expansion in the fine structure constant $\alpha \sim 1/137$. This requires renormalization theory. A recent rigorous treatment of perturbative QFT is Kevin Costello's book.
- YM and other QFTs have a rigorous definition – the Wightman axioms supplemented by conditions selecting out a particular case. However, very few results have been derived from it, much less proven. Working rigorously with YM is a central challenge of mathematical physics, *e.g.* see the Clay Millennium prize problem.
- ST has no complete (not to mention rigorous) definition yet, but some simplified versions of it do, such as topological string theory. The concepts needed are not known except in some limiting cases.

Considered as mathematics, the status of these theories varies:

- QM has a satisfactory mathematically rigorous definition, based on the Schrödinger equation and functional analysis.
- CM can be treated rigorously, but it is not fundamental, which leads to problems – singularities, chaos and precision issues.
- QED can be defined rigorously in perturbation theory, meaning as an expansion in the fine structure constant $\alpha \sim 1/137$. This requires renormalization theory. A recent rigorous treatment of perturbative QFT is Kevin Costello's book.
- YM and other QFTs have a rigorous definition – the Wightman axioms supplemented by conditions selecting out a particular case. However, very few results have been derived from it, much less proven. Working rigorously with YM is a central challenge of mathematical physics, *e.g.* see the Clay Millennium prize problem.
- ST has no complete (not to mention rigorous) definition yet, but some simplified versions of it do, such as topological string theory. The concepts needed are not known except in some limiting cases.

Considered as mathematics, the status of these theories varies:

- QM has a satisfactory mathematically rigorous definition, based on the Schrödinger equation and functional analysis.
- CM can be treated rigorously, but it is not fundamental, which leads to problems – singularities, chaos and precision issues.
- QED can be defined rigorously in perturbation theory, meaning as an expansion in the fine structure constant $\alpha \sim 1/137$. This requires renormalization theory. A recent rigorous treatment of perturbative QFT is Kevin Costello's book.
- YM and other QFTs have a rigorous definition – the Wightman axioms supplemented by conditions selecting out a particular case. However, very few results have been derived from it, much less proven. Working rigorously with YM is a central challenge of mathematical physics, *e.g.* see the Clay Millennium prize problem.
- ST has no complete (not to mention rigorous) definition yet, but some simplified versions of it do, such as topological string theory. The concepts needed are not known except in some limiting cases.

Considered as mathematics, the status of these theories varies:

- QM has a satisfactory mathematically rigorous definition, based on the Schrödinger equation and functional analysis.
- CM can be treated rigorously, but it is not fundamental, which leads to problems – singularities, chaos and precision issues.
- QED can be defined rigorously in perturbation theory, meaning as an expansion in the fine structure constant $\alpha \sim 1/137$. This requires renormalization theory. A recent rigorous treatment of perturbative QFT is Kevin Costello's book.
- YM and other QFTs have a rigorous definition – the Wightman axioms supplemented by conditions selecting out a particular case. However, very few results have been derived from it, much less proven. Working rigorously with YM is a central challenge of mathematical physics, *e.g.* see the Clay Millennium prize problem.
- ST has no complete (not to mention rigorous) definition yet, but some simplified versions of it do, such as topological string theory. The concepts needed are not known except in some limiting cases.

Considered as mathematics, the status of these theories varies:

- QM has a satisfactory mathematically rigorous definition, based on the Schrödinger equation and functional analysis.
- CM can be treated rigorously, but it is not fundamental, which leads to problems – singularities, chaos and precision issues.
- QED can be defined rigorously in perturbation theory, meaning as an expansion in the fine structure constant $\alpha \sim 1/137$. This requires renormalization theory. A recent rigorous treatment of perturbative QFT is Kevin Costello's book.
- YM and other QFTs have a rigorous definition – the Wightman axioms supplemented by conditions selecting out a particular case. However, very few results have been derived from it, much less proven. Working rigorously with YM is a central challenge of mathematical physics, *e.g.* see the Clay Millennium prize problem.
- ST has no complete (not to mention rigorous) definition yet, but some simplified versions of it do, such as topological string theory. The concepts needed are not known except in some limiting cases.

Even if their theory of interest has a rigorous definition, physicists and other physical scientists rarely make this definition their starting point. Rather, their goal is to describe the real world and make contact with empirical results (experiment and observation), and this can be done without a mathematically rigorous definition.

Contact between a model and the real world is based on an **interpretation**. One restricts attention to a class of phenomena in the real world and a set of operations which change the state of the world: measurements, symmetry transformations, *etc.*. Each physical operation then corresponds to some mathematical operation in the model. This will be familiar to many of you for QM, where the state is a ray in a Hilbert space $\mathcal{H}$ (say, complex-valued functions on $\mathbb{R}^{3N}$), and operations correspond to projections and unitary operators on $\mathcal{H}$.

In practice one tells a much longer story to make contact with real world operations such as performing scattering experiments at a particle accelerator, measuring physical properties of a chemical compound, or interpreting the data from an orbiting observatory.

Even if their theory of interest has a rigorous definition, physicists and other physical scientists rarely make this definition their starting point. Rather, their goal is to describe the real world and make contact with empirical results (experiment and observation), and this can be done without a mathematically rigorous definition.

Contact between a model and the real world is based on an **interpretation.** One restricts attention to a class of phenomena in the real world and a set of operations which change the state of the world: measurements, symmetry transformations, *etc.*. Each physical operation then corresponds to some mathematical operation in the model. This will be familiar to many of you for QM, where the state is a ray in a Hilbert space $\mathcal{H}$ (say, complex-valued functions on $\mathbb{R}^{3N}$), and operations correspond to projections and unitary operators on $\mathcal{H}$.

In practice one tells a much longer story to make contact with real world operations such as performing scattering experiments at a particle accelerator, measuring physical properties of a chemical compound, or interpreting the data from an orbiting observatory.

Even if their theory of interest has a rigorous definition, physicists and other physical scientists rarely make this definition their starting point. Rather, their goal is to describe the real world and make contact with empirical results (experiment and observation), and this can be done without a mathematically rigorous definition.

Contact between a model and the real world is based on an **interpretation**. One restricts attention to a class of phenomena in the real world and a set of operations which change the state of the world: measurements, symmetry transformations, *etc.*. Each physical operation then corresponds to some mathematical operation in the model. This will be familiar to many of you for QM, where the state is a ray in a Hilbert space $\mathcal{H}$ (say, complex-valued functions on $\mathbb{R}^{3N}$), and operations correspond to projections and unitary operators on $\mathcal{H}$.

In practice one tells a much longer story to make contact with real world operations such as performing scattering experiments at a particle accelerator, measuring physical properties of a chemical compound, or interpreting the data from an orbiting observatory.

The ability to formalize and verify arguments using the interpretation of a physical theory is at least as important as rigor in the mathematics, and arguably more so. Thus, an atomic physicist might ask, can we ground all of the discussion needed to interpret a laboratory experiment in terms of operations in QM, in such a precise way that one could verify that a proposed theoretical analysis is appropriate.

Thus the task we will pose for formal methods is not “can we verify that a specific physical calculation is correct?” Nor is it “can we prove some physically motivated conjecture?” Rather it is:

Can we formalize the interpretation, such that the question “does this set of measurements fit my model, or contradict my model,” could be answered by purely formal reasoning.

Shortly we will discuss the muon $g-2$ measurement as an example. But to get the idea across, it is better to discuss a case for which the underlying physics is well understood, and we all have some familiarity with it. Thus we return to classical physics.

The ability to formalize and verify arguments using the interpretation of a physical theory is at least as important as rigor in the mathematics, and arguably more so. Thus, an atomic physicist might ask, can we ground all of the discussion needed to interpret a laboratory experiment in terms of operations in QM, in such a precise way that one could verify that a proposed theoretical analysis is appropriate.

Thus the task we will pose for formal methods is not “can we verify that a specific physical calculation is correct?” Nor is it “can we prove some physically motivated conjecture?” Rather it is:

Can we formalize the interpretation, such that the question “does this set of measurements fit my model, or contradict my model,” could be answered by purely formal reasoning.

Shortly we will discuss the muon $g-2$ measurement as an example. But to get the idea across, it is better to discuss a case for which the underlying physics is well understood, and we all have some familiarity with it. Thus we return to classical physics.

The ability to formalize and verify arguments using the interpretation of a physical theory is at least as important as rigor in the mathematics, and arguably more so. Thus, an atomic physicist might ask, can we ground all of the discussion needed to interpret a laboratory experiment in terms of operations in QM, in such a precise way that one could verify that a proposed theoretical analysis is appropriate.

Thus the task we will pose for formal methods is not "can we verify that a specific physical calculation is correct?" Nor is it "can we prove some physically motivated conjecture?" Rather it is:

Can we formalize the interpretation, such that the question "does this set of measurements fit my model, or contradict my model," could be answered by purely formal reasoning.

Shortly we will discuss the muon $g-2$ measurement as an example. But to get the idea across, it is better to discuss a case for which the underlying physics is well understood, and we all have some familiarity with it. Thus we return to classical physics.

A formalizable physics problem

Suppose we have a clock, we have observers who can move about on the earth's surface, we grant some fixed landmarks on the surface, and we can make timed observations of celestial bodies using a sextant to determine elevation and the angles to the landmarks, all to some precision. Given a set of observations, show rigorously that they are compatible with Newton's theory, or are not compatible.

This is a somewhat idealized version of the astronomical tests of Newton's theory of celestial mechanics. Of course, we expect this theory to work up to some level of precision, but eventually it should fail. For example, Newton's theory does not correctly describe the perihelion precession of Mercury.

A formalizable physics problem

Suppose we have a clock, we have observers who can move about on the earth's surface, we grant some fixed landmarks on the surface, and we can make timed observations of celestial bodies using a sextant to determine elevation and the angles to the landmarks, all to some precision. Given a set of observations, show rigorously that they are compatible with Newton's theory, or are not compatible.

This is a somewhat idealized version of the astronomical tests of Newton's theory of celestial mechanics. Of course, we expect this theory to work up to some level of precision, but eventually it should fail. For example, Newton's theory does not correctly describe the perihelion precession of Mercury.

What does our “formalized physics” problem involve? We can formalize the mathematics of celestial mechanics as needed, and we are given the observations. The remaining task is to formalize the interpretation to the point where we can make contact between them. This involves:

- Basic notions of astronomy, such as: the observers sit on one of the planets, the earth, whose trajectory is not predetermined but is also to be obtained from the dynamics.
- A model of the observers, say that they each follow some continuous space-time trajectory on the surface of the earth.
- A model of the surface of the earth, say that it is a rotating sphere.
- A model of an observation, which involves the space-time locations of the observer and the target bodies, an assumed behavior of light, and an error model.

Each of these could be made mathematically precise and formalized.

Given these various ingredients of the interpretation, the analog of the "proof" which we would try to formalize and verify is the task of mapping the observations into definite constraints between the positions of the planets and sun at specific times, and finding the solutions of Newton's equations of motion which satisfy these constraints, or else proving that they do not exist.

This task would naturally break down into various parts:

- Identify the interesting objects, whose motion can be usefully measured given the available precision.
- Map measurements into four-vectors in an earth-centered frame.
- Map the earth-centered frame to an inertial frame, centered on the barycenter of the solar system.
- Fit the planetary masses and orbital elements and test the theory.

Of course the N-body problem is not solvable analytically, so we would also need supplementary physical simplifications, for example treating the planet-planet interactions as perturbations.

Given these various ingredients of the interpretation, the analog of the "proof" which we would try to formalize and verify is the task of mapping the observations into definite constraints between the positions of the planets and sun at specific times, and finding the solutions of Newton's equations of motion which satisfy these constraints, or else proving that they do not exist.

This task would naturally break down into various parts:

- Identify the interesting objects, whose motion can be usefully measured given the available precision.
- Map measurements into four-vectors in an earth-centered frame.
- Map the earth-centered frame to an inertial frame, centered on the barycenter of the solar system.
- Fit the planetary masses and orbital elements and test the theory.

Of course the N-body problem is not solvable analytically, so we would also need supplementary physical simplifications, for example treating the planet-planet interactions as perturbations.

While there is no doubt that the astronomers are doing celestial mechanics correctly, formalizing this process would be an amusing and perhaps illuminating example of a formal treatment of a problem for which rigor is not part of the original problem statement.

One might object to this claim – don't we need a rigorous definition of Newton's equations to do this? Yes, but it does not have to be any of the standard mathematical ones. It could be much weaker, because we are using it in such a restricted way. This point might also be interesting to understand in more depth.

The same general idea could then be applied to cutting-edge physics about which we are less certain, and indeed where mistakes are often made. As always in science, most progress in physics comes from disagreement between theory and experiment. The most celebrated examples involve a fundamental disagreement, which contradicts an established theory and may point the way to its successor. But usually the disagreement is caused by an error, in the theoretical calculations, in the experiment, or in the process of comparing the two.

While there is no doubt that the astronomers are doing celestial mechanics correctly, formalizing this process would be an amusing and perhaps illuminating example of a formal treatment of a problem for which rigor is not part of the original problem statement.

One might object to this claim – don't we need a rigorous definition of Newton's equations to do this? Yes, but it does not have to be any of the standard mathematical ones. It could be much weaker, because we are using it in such a restricted way. This point might also be interesting to understand in more depth.

The same general idea could then be applied to cutting-edge physics about which we are less certain, and indeed where mistakes are often made. As always in science, most progress in physics comes from disagreement between theory and experiment. The most celebrated examples involve a fundamental disagreement, which contradicts an established theory and may point the way to its successor. But usually the disagreement is caused by an error, in the theoretical calculations, in the experiment, or in the process of comparing the two.

While there is no doubt that the astronomers are doing celestial mechanics correctly, formalizing this process would be an amusing and perhaps illuminating example of a formal treatment of a problem for which rigor is not part of the original problem statement.

One might object to this claim – don't we need a rigorous definition of Newton's equations to do this? Yes, but it does not have to be any of the standard mathematical ones. It could be much weaker, because we are using it in such a restricted way. This point might also be interesting to understand in more depth.

The same general idea could then be applied to cutting-edge physics about which we are less certain, and indeed where mistakes are often made. As always in science, most progress in physics comes from disagreement between theory and experiment. The most celebrated examples involve a fundamental disagreement, which contradicts an established theory and may point the way to its successor. But usually the disagreement is caused by an error, in the theoretical calculations, in the experiment, or in the process of comparing the two.

$g-2$ of the muon

An example which is somewhat infamous in particle physics is the calculation of $g-2$ for the muon, in other words its magnetic moment. A muon has a lifetime of about 2 microseconds, long enough to do precision measurements, but not to store. Thus the muons must be created in a particle accelerator and this makes the measurement challenging. Nevertheless there has been a major effort to do it and $g-2$ has been measured to better than a part in a billion.

The great interest in $g-2$ is because it can be calculated to high precision in the Standard Model. Its value depends not just on the known particles, but on hypothetical new particles which present-day accelerators cannot produce directly. At present the theoretical and experimental numbers actually differ by more than 3 standard deviations, and a new experiment has begun at Fermilab to improve the precision to 5 standard deviations. If the difference is real, there must be new physics beyond the Standard Model.

$g-2$ of the muon

An example which is somewhat infamous in particle physics is the calculation of $g-2$ for the muon, in other words its magnetic moment. A muon has a lifetime of about 2 microseconds, long enough to do precision measurements, but not to store. Thus the muons must be created in a particle accelerator and this makes the measurement challenging. Nevertheless there has been a major effort to do it and $g-2$ has been measured to better than a part in a billion.

The great interest in $g-2$ is because it can be calculated to high precision in the Standard Model. Its value depends not just on the known particles, but on hypothetical new particles which present-day accelerators cannot produce directly. At present the theoretical and experimental numbers actually differ by more than 3 standard deviations, and a new experiment has begun at Fermilab to improve the precision to 5 standard deviations. If the difference is real, there must be new physics beyond the Standard Model.

Clearly comparing experiment and theory in the ninth decimal place is a tricky business. And this is not the first time $g-2$ has disagreed between theory and experiment. For several years (1996-2001) there was also a 3 sigma disagreement, inspiring many speculations about new physics.

This disagreement turned out to be due to a theoretical mistake. The $g-2$ calculation is particularly tricky, not so much because it involves thousands of Feynman diagrams (this was systematized long ago) but because it requires combining effects from several different sources: QED, the weak interactions, and effects of virtual hadrons, in particular an effect called "hadronic light by light scattering." This last effect cannot be measured, nor can it be calculated from first principles; it is calculated using phenomenological models of hadrons.

Being something of a weak link, this part of the calculation was scrutinized particularly carefully, with several groups each using their own preferred models. Thus it was a surprise when a mistake was discovered.

Clearly comparing experiment and theory in the ninth decimal place is a tricky business. And this is not the first time $g-2$ has disagreed between theory and experiment. For several years (1996-2001) there was also a 3 sigma disagreement, inspiring many speculations about new physics.

This disagreement turned out to be due to a theoretical mistake. The $g-2$ calculation is particularly tricky, not so much because it involves thousands of Feynman diagrams (this was systematized long ago) but because it requires combining effects from several different sources: QED, the weak interactions, and effects of virtual hadrons, in particular an effect called "hadronic light by light scattering." This last effect cannot be measured, nor can it be calculated from first principles; it is calculated using phenomenological models of hadrons.

Being something of a weak link, this part of the calculation was scrutinized particularly carefully, with several groups each using their own preferred models. Thus it was a surprise when a mistake was discovered.

Clearly comparing experiment and theory in the ninth decimal place is a tricky business. And this is not the first time $g-2$ has disagreed between theory and experiment. For several years (1996-2001) there was also a 3 sigma disagreement, inspiring many speculations about new physics.

This disagreement turned out to be due to a theoretical mistake. The $g-2$ calculation is particularly tricky, not so much because it involves thousands of Feynman diagrams (this was systematized long ago) but because it requires combining effects from several different sources: QED, the weak interactions, and effects of virtual hadrons, in particular an effect called "hadronic light by light scattering." This last effect cannot be measured, nor can it be calculated from first principles; it is calculated using phenomenological models of hadrons.

Being something of a weak link, this part of the calculation was scrutinized particularly carefully, with several groups each using their own preferred models. Thus it was a surprise when a mistake was discovered.

arXiv.org > hep-ph > arXiv:hep-ph/0112102

Search or Arti
(Help | Advanced

High Energy Physics – Phenomenology

Comment on the sign of the pseudoscalar pole contribution to the muon g-2

Masashi Hayakawa, Toichiro Kinoshita

(Submitted on 6 Dec 2001 (v1), last revised 16 Dec 2001 (this version, v2))

We correct the error in the sign of the pseudoscalar pole contribution to the muon g-2, which dominates the O(alpha^3) hadronic light-by-light scattering effect. The error originates from our oversight of a feature of the algebraic manipulation program FORM which defines the epsilon-tensor in such a way that it satisfies the relation epsilon_{mu_1 mu_2 mu_3 mu_4} epsilon_{nu_1 nu_2 nu_3 nu_4} eta^{mu_1 nu_1} eta^{mu_2 nu_2} eta^{mu_3 nu_3} eta^{mu_4 nu_4} = 24, irrespective of space-time metric. To circumvent this problem, we replaced the product epsilon_{mu_1 mu_2 mu_3 mu_4} epsilon_{nu_1 nu_2 nu_3 nu_4} by - eta_{mu_1 nu_1} eta_{mu_2 nu_2} eta_{mu_3 nu_3} eta_{mu_4 nu_4} \pm cdots in the FORM-formatted program, and obtained a positive value for the pseudoscalar pole contribution, in agreement with the recent result obtained by Knecht {\it et al}.

Comments:	7 pages, LaTeX 2epsilon
Subjects:	**High Energy Physics – Phenomenology (hep-ph)**
Report number:	KEK-TH-793
Cite as:	**arXiv:hep-ph/0112102**
	(or **arXiv:hep-ph/0112102v2** for this version)

Submission history

From: Masashi Hayakawa [view email]
[v1] Thu, 6 Dec 2001 20:19:12 UTC (8 KB)
[v2] Sun, 16 Dec 2001 12:13:36 UTC (8 KB)

Which authors of this paper are endorsers? | *Disable MathJax* (*What is MathJax?*)

Such mistakes are of course not infrequent, even in the published literature. This one was exceptional only in the long time it took to discover it. Could this mistake, or the inevitable future mistakes, have been caught by any sort of formal verification?

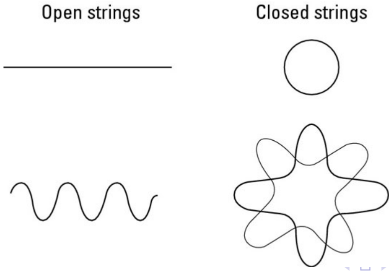
It's not clear to me, but perhaps the answer is yes. But the challenge is that the mistake did not come from any single theoretical assumption or calculation, rather it came at the step of integrating various theoretical subresults derived using different physical subtheories and approximations, and not noticing a difference in conventions in one of the results. So one suspects that it would only be caught by a "large scale" formalization that covered the problem as a whole, not the individual subtheories. And while the subtheories can to some extent be formulated in a mathematically rigorous way, the overall problem cannot.

Such mistakes are of course not infrequent, even in the published literature. This one was exceptional only in the long time it took to discover it. Could this mistake, or the inevitable future mistakes, have been caught by any sort of formal verification?

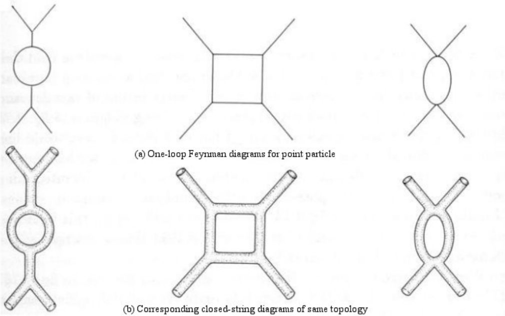
It's not clear to me, but perhaps the answer is yes. But the challenge is that the mistake did not come from any single theoretical assumption or calculation, rather it came at the step of integrating various theoretical subresults derived using different physical subtheories and approximations, and not noticing a difference in conventions in one of the results. So one suspects that it would only be caught by a "large scale" formalization that covered the problem as a whole, not the individual subtheories. And while the subtheories can to some extent be formulated in a mathematically rigorous way, the overall problem cannot.

String theory

Particles (electrons, quarks, photons, etc.) are really small loops of string. Different modes of vibration $\rightarrow$ different particles. Open string $\rightarrow$ one direction of vibration $\rightarrow$ polarization of photon.
Closed strings naturally vibrate in two directions (left and right movers). Spin two particle $\rightarrow$ graviton, so string theory naturally contains gravity (general relativity).
In fact string theory is a quantum unified theory of gravity and Yang-Mills theory coupled to matter, a candidate fundamental theory.



The unification of general relativity and quantum mechanics has been studied intensively for over 60 years and it is a hard problem. Scaling arguments (theory of renormalizability) tell us that in D space-time dimensions, the strength of gravity grows with decreasing length L as L^{2-D}. In $D > 2$ gravity becomes strong at the Planck scale (10^{-33} cm) meaning that the metric is a strongly fluctuating variable. However the observed space-time metric is almost flat, it is not strongly fluctuating.



(a) One-loop Feynman diagrams for point particle

(b) Corresponding closed-string diagrams of same topology

In string theory, there is another preferred scale, the string length. Quantum effects are "cut off" at shorter distances, eliminating the strong metric fluctuations so that quantum gravity is consistent with the observed properties of space-time.

String theory cannot be modified – it is a single unified structure (though with many limits which superficially look different). Thus, it is either right or wrong as a candidate fundamental theory.

And, string theory has many other surprising and important properties:

- Maximal symmetry and supersymmetry
- Exceptional structures such as E_8
- Dualities: strong $\leftrightarrow$ weak, gauge $\leftrightarrow$ gravity.

Many of the alternative approaches to quantum gravity turned out to be particular cases of “string/M theory.”

String theory has explained some (not yet all) of the mysteries of quantum gravity, such as the entropy of black holes. The study of string theory has also led to breakthroughs on many other physical questions: the origin of quark confinement, symmetry breaking, phase diagrams, etc.

String theory cannot be modified – it is a single unified structure (though with many limits which superficially look different). Thus, it is either right or wrong as a candidate fundamental theory.

And, string theory has many other surprising and important properties:

- Maximal symmetry and supersymmetry
- Exceptional structures such as E_8
- Dualities: strong $\leftrightarrow$ weak, gauge $\leftrightarrow$ gravity.

Many of the alternative approaches to quantum gravity turned out to be particular cases of “string/M theory.”

String theory has explained some (not yet all) of the mysteries of quantum gravity, such as the entropy of black holes. The study of string theory has also led to breakthroughs on many other physical questions: the origin of quark confinement, symmetry breaking, phase diagrams, etc.

String theory cannot be modified – it is a single unified structure (though with many limits which superficially look different). Thus, it is either right or wrong as a candidate fundamental theory.

And, string theory has many other surprising and important properties:

- Maximal symmetry and supersymmetry
- Exceptional structures such as E_8
- Dualities: strong $\leftrightarrow$ weak, gauge $\leftrightarrow$ gravity.

Many of the alternative approaches to quantum gravity turned out to be particular cases of “string/M theory.”

String theory has explained some (not yet all) of the mysteries of quantum gravity, such as the entropy of black holes. The study of string theory has also led to breakthroughs on many other physical questions: the origin of quark confinement, symmetry breaking, phase diagrams, etc.

To test string/M theory, we must first show that we can use it to derive the particles and forces we know about, described by the Standard Model. If this is not possible, we will know that string theory does not describe our universe.

How do we do this? The starting point is to realize that in string theory, spacetime actually has ten dimensions. This is not in contradiction with experience but only if we postulate that six of these dimensions form a compact manifold M of diameter much less than a micron (the scale at which Newton's inverse square law of gravity has been tested).

One then needs to work out the theory of strings vibrating on M. The topology and geometry of M translate into properties of the matter and forces we observe – the spectrum and masses of particles, the fine structure constant and other couplings, etc..

To test string/M theory, we must first show that we can use it to derive the particles and forces we know about, described by the Standard Model. If this is not possible, we will know that string theory does not describe our universe.

How do we do this? The starting point is to realize that in string theory, spacetime actually has ten dimensions. This is not in contradiction with experience but only if we postulate that six of these dimensions form a compact manifold M of diameter much less than a micron (the scale at which Newton's inverse square law of gravity has been tested).

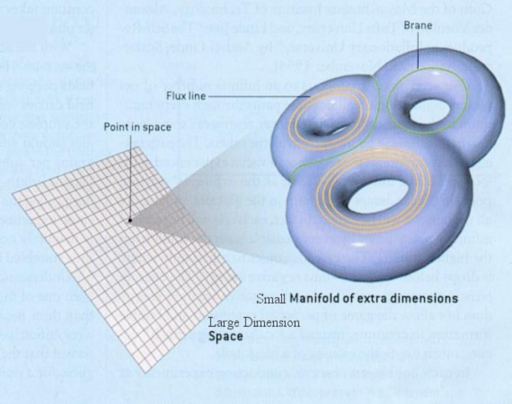
One then needs to work out the theory of strings vibrating on M. The topology and geometry of M translate into properties of the matter and forces we observe – the spectrum and masses of particles, the fine structure constant and other couplings, etc..

To test string/M theory, we must first show that we can use it to derive the particles and forces we know about, described by the Standard Model. If this is not possible, we will know that string theory does not describe our universe.

How do we do this? The starting point is to realize that in string theory, spacetime actually has ten dimensions. This is not in contradiction with experience but only if we postulate that six of these dimensions form a compact manifold M of diameter much less than a micron (the scale at which Newton's inverse square law of gravity has been tested).

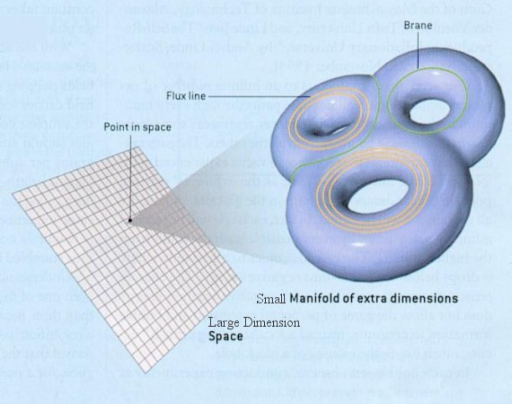
One then needs to work out the theory of strings vibrating on M. The topology and geometry of M translate into properties of the matter and forces we observe – the spectrum and masses of particles, the fine structure constant and other couplings, etc..

There are many possible choices for M – it might be a six-torus, a six-torus quotiented by a discrete group (an “orbifold”), a Calabi-Yau manifold, or many other possibilities. In addition there is additional data (branes, fluxes, *etc.*) to be chosen on M, call such a choice V (we will be a little bit more concrete below).



Each (M, V) can lead to different predictions for observed physics. We do not know *a priori* which of these choices describes our universe. Thus, we need to work out the predictions for many such choices.

There are many possible choices for M – it might be a six-torus, a six-torus quotiented by a discrete group (an “orbifold”), a Calabi-Yau manifold, or many other possibilities. In addition there is additional data (branes, fluxes, *etc.*) to be chosen on M, call such a choice V (we will be a little bit more concrete below).



Each (M, V) can lead to different predictions for observed physics. We do not know *a priori* which of these choices describes our universe. Thus, we need to work out the predictions for many such choices.

The details are lengthy, but the main point for purposes of this talk is,

> *String theory provides a procedure by which we can construct and classify objects with purely mathematical definitions - manifolds with structure* (M, V) *- and for each one derive a candidate theory of the observable physics in our universe.*

Thus the central problem of theoretical physics – what are the possible fundamental laws of nature – is reduced to pure mathematics.

In the next part of the talk, we will describe one of the mathematical structures we need for this project. Many of them are familiar and used throughout the mathematical sciences, such as finite groups. Others are more specialized, but are still of interest to mathematicians as well as physicists.

Many of these structures are combinatorial and their classification is intricate. These are prime candidates for computational mathematics.

The details are lengthy, but the main point for purposes of this talk is,

> *String theory provides a procedure by which we can construct and classify objects with purely mathematical definitions - manifolds with structure* (M, V) *- and for each one derive a candidate theory of the observable physics in our universe.*

Thus the central problem of theoretical physics – what are the possible fundamental laws of nature – is reduced to pure mathematics.

In the next part of the talk, we will describe one of the mathematical structures we need for this project. Many of them are familiar and used throughout the mathematical sciences, such as finite groups. Others are more specialized, but are still of interest to mathematicians as well as physicists.

Many of these structures are combinatorial and their classification is intricate. These are prime candidates for computational mathematics.

The details are lengthy, but the main point for purposes of this talk is,

> *String theory provides a procedure by which we can construct and classify objects with purely mathematical definitions - manifolds with structure* (M, V) *- and for each one derive a candidate theory of the observable physics in our universe.*

Thus the central problem of theoretical physics – what are the possible fundamental laws of nature – is reduced to pure mathematics.

In the next part of the talk, we will describe one of the mathematical structures we need for this project. Many of them are familiar and used throughout the mathematical sciences, such as finite groups. Others are more specialized, but are still of interest to mathematicians as well as physicists.

Many of these structures are combinatorial and their classification is intricate. These are prime candidates for computational mathematics.

Kreuzer-Skarke database

Perhaps the best introduction to computational string theory is to describe the first work of lasting value. This is the Kreuzer-Skarke database of reflexive polytopes, used to classify and work with a particular type of manifold M, the toric Calabi-Yau hypersurfaces.

Let us first give the definition, and then say a few words about where it comes from. A lattice polytope Δ is the convex hull in $\mathbb{R}^{d+1}$ of a finite set of integral points $\nu^{(i)} \in \mathbb{Z}^{d+1}$. Its **dual polytope** is the set

$$\Delta^* \equiv \{y \in \mathbb{R}^{d+1} : x \cdot y \geq -1 \ \forall x \in \Delta\}. \quad (1)$$

A lattice polytope is **reflexive** if its dual is also a lattice polytope.

For fixed d, the set of reflexive lattice polytopes is finite. The Kreuzer-Skarke database lists the 473,800,776 instances for $d = 3$. (See `http://hep.itp.tuwien.ac.at/~kreuzer/CY/`)

Kreuzer-Skarke database

Perhaps the best introduction to computational string theory is to describe the first work of lasting value. This is the Kreuzer-Skarke database of reflexive polytopes, used to classify and work with a particular type of manifold M, the toric Calabi-Yau hypersurfaces.

Let us first give the definition, and then say a few words about where it comes from. A lattice polytope Δ is the convex hull in $\mathbb{R}^{d+1}$ of a finite set of integral points $\nu^{(i)} \in \mathbb{Z}^{d+1}$. Its **dual polytope** is the set

$$\Delta^* \equiv \{y \in \mathbb{R}^{d+1} : x \cdot y \geq -1 \ \forall x \in \Delta\}. \quad (1)$$

A lattice polytope is **reflexive** if its dual is also a lattice polytope.

For fixed d, the set of reflexive lattice polytopes is finite. The Kreuzer-Skarke database lists the 473,800,776 instances for $d = 3$. (See `http://hep.itp.tuwien.ac.at/~kreuzer/CY/`)

Kreuzer-Skarke database

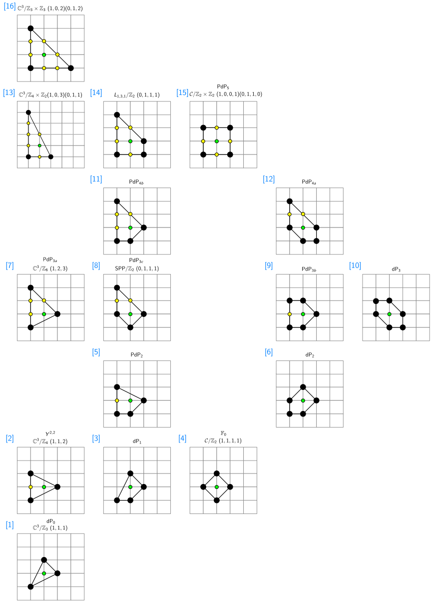
Perhaps the best introduction to computational string theory is to describe the first work of lasting value. This is the Kreuzer-Skarke database of reflexive polytopes, used to classify and work with a particular type of manifold M, the toric Calabi-Yau hypersurfaces.

Let us first give the definition, and then say a few words about where it comes from. A lattice polytope Δ is the convex hull in $\mathbb{R}^{d+1}$ of a finite set of integral points $\nu^{(i)} \in \mathbb{Z}^{d+1}$. Its **dual polytope** is the set

$$\Delta^* \equiv \{y \in \mathbb{R}^{d+1} : x \cdot y \geq -1 \ \forall x \in \Delta\}. \quad (1)$$

A lattice polytope is **reflexive** if its dual is also a lattice polytope.

For fixed d, the set of reflexive lattice polytopes is finite. The Kreuzer-Skarke database lists the 473,800,776 instances for $d = 3$. (See `http://hep.itp.tuwien.ac.at/~kreuzer/CY/`)



From Y.-H. He, R. K. Seong and S.-T. Yau, `arXiv:1704.03462`. The symmetry under reflection across the x-axis relates pairs of dual polytopes.

Why do we care about reflexive polytopes? First, the extra dimensions (in realistic compactifications) must be a six dimensional manifold which approximately satisfies Einstein's equation $R_{ij} = 0$ (Ricci flat). Furthermore we assume there is $N = 1$ supersymmetry in the resulting four dimensional theory. This requires a covariantly constant spinor on M which implies that M is complex and Kähler. A large subset of such M can be obtained as the solution set of a single equation $f(Z) = 0$ in a toric manifold, constructed by gluing charts $(\mathbb{C}^*)^d$ according to a prescription whose data depends on Δ. Then, by Yau's theorem such an M will have a Ricci flat metric iff $c_1(M) = 0$. This will be the case for suitably chosen f if Δ is reflexive.

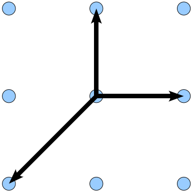
The condition that Δ^* is also a lattice polytope means that the CY associated to Δ has a second "mirror" CY associated to Δ^*. This "mirror symmetry" was noticed by the physicist Philip Candelas and then explained in these terms by Victor Batyrev.

Why do we care about reflexive polytopes? First, the extra dimensions (in realistic compactifications) must be a six dimensional manifold which approximately satisfies Einstein's equation $R_{ij} = 0$ (Ricci flat). Furthermore we assume there is $N = 1$ supersymmetry in the resulting four dimensional theory. This requires a covariantly constant spinor on M which implies that M is complex and Kähler. A large subset of such M can be obtained as the solution set of a single equation $f(Z) = 0$ in a toric manifold, constructed by gluing charts $(\mathbb{C}^*)^d$ according to a prescription whose data depends on Δ. Then, by Yau's theorem such an M will have a Ricci flat metric iff $c_1(M) = 0$. This will be the case for suitably chosen f if Δ is reflexive.

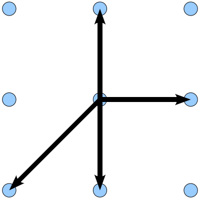
The condition that Δ^* is also a lattice polytope means that the CY associated to Δ has a second "mirror" CY associated to Δ^*. This "mirror symmetry" was noticed by the physicist Philip Candelas and then explained in these terms by Victor Batyrev.

Toric manifolds

A good way to get more CY3's is to keep the single defining equation (a hypersurface), but generalize the ambient space $\mathbb{CP}^4$. A rather general construction of analogous spaces is provided by toric geometry. This combines many of the ideas on the previous slide.



In concrete terms, a toric space is specified by a fan. Each ray of the fan corresponds to a coordinate Z^i, while linear relations between rays correspond to symmetries. Thus, the top figure represents $\mathbb{CP}^2$, and the relation $v_0 + v_1 + v_2 = 0$ corresponds to the action $Z^i \to \lambda Z^i$. The bottom figure has an extra ray and an extra relation. It represents the del Pezzo surface dP_1.



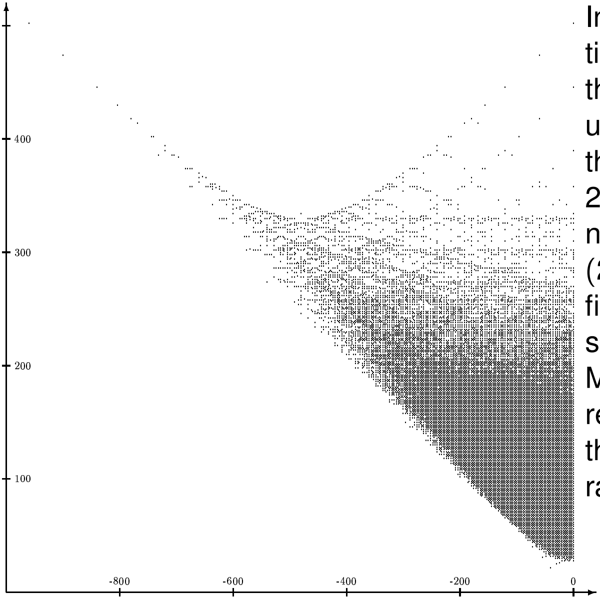


Fig. 1: $h_{11} + h_{12}$ vs. Euler number $\chi = 2(h_{11} - h_{12})$ for all pairs (h_{11}, h_{12}) with $h_{11} \le h_{12}$.

In this figure, the vertical axis is $b^{1,1} + b^{2,1}$, the number of CY moduli. The horizontal axis is the Euler character $\chi = 2b^{1,1} - 2b^{2,1}$, twice the number of generations in $(2,2)$ heterotic compactification. Each dot represents one or more CY_3's. Mirror symmetry is the reflection symmetry with the (omitted) $\chi > 0$ quadrant.

While Kreuzer and Skarke were inspired by mirror symmetry, their dataset displayed a more subtle pattern, the simple shape of its boundary. This is usually called the “shield” after the shape of the full boundary with the mirror symmetry.

While the finiteness of the set was known, the shield was unexpected and was a computer-aided discovery. Recently a conjecture has been made to explain this shape – it is that all CY threefolds with large Betti numbers are elliptic fibrations (Taylor and Wang). The finiteness of the elliptically fibered threefolds was proven long ago by Mark Gross, and four and five-folds more recently by Di Cerbo and Svaldi arXiv:1608.02997. The boundary of this subset can be understood in terms of blow-up transitions of the base. But it is not clear why this argument applies to all threefolds.

In string theory, one asks different questions – for example, what is the prevalence of singularities with certain Kodaira types, or can one find a holomorphic vector bundle V with $c_3(V) = 6$. At present these are largely questions of computational mathematics.

While Kreuzer and Skarke were inspired by mirror symmetry, their dataset displayed a more subtle pattern, the simple shape of its boundary. This is usually called the “shield” after the shape of the full boundary with the mirror symmetry.

While the finiteness of the set was known, the shield was unexpected and was a computer-aided discovery. Recently a conjecture has been made to explain this shape – it is that all CY threefolds with large Betti numbers are elliptic fibrations (Taylor and Wang). The finiteness of the elliptically fibered threefolds was proven long ago by Mark Gross, and four and five-folds more recently by Di Cerbo and Svaldi `arXiv:1608.02997`. The boundary of this subset can be understood in terms of blow-up transitions of the base. But it is not clear why this argument applies to all threefolds.

In string theory, one asks different questions – for example, what is the prevalence of singularities with certain Kodaira types, or can one find a holomorphic vector bundle V with $c_3(V) = 6$. At present these are largely questions of computational mathematics.

While Kreuzer and Skarke were inspired by mirror symmetry, their dataset displayed a more subtle pattern, the simple shape of its boundary. This is usually called the “shield” after the shape of the full boundary with the mirror symmetry.

While the finiteness of the set was known, the shield was unexpected and was a computer-aided discovery. Recently a conjecture has been made to explain this shape – it is that all CY threefolds with large Betti numbers are elliptic fibrations (Taylor and Wang). The finiteness of the elliptically fibered threefolds was proven long ago by Mark Gross, and four and five-folds more recently by Di Cerbo and Svaldi `arXiv:1608.02997`. The boundary of this subset can be understood in terms of blow-up transitions of the base. But it is not clear why this argument applies to all threefolds.

In string theory, one asks different questions – for example, what is the prevalence of singularities with certain Kodaira types, or can one find a holomorphic vector bundle V with $c_3(V) = 6$. At present these are largely questions of computational mathematics.

Platonic datasets

The set of four-dimensional reflexive polytopes is too large to work with by hand or even to do much in-depth work with individual examples. Furthermore we are often interested in questions of existence or statistical questions, *i.e.* what fraction of the cases have a property X of interest, or is there a correlation between the presence of properties X and Y in a given example. In this sense it is a dataset much like the datasets we deal with in computer science and statistics.

Usually one deals with datasets which arise from the physical world: images, waveforms, other measurements – or from the social world: texts, networks of relationships, *etc.*. These are empirical datasets.

The Kreuzer-Skarke dataset, as do other datasets of mathematical objects, has an *a priori* definition. It is a **platonic dataset.** Platonic datasets are a main focus of interest in computational mathematics.

Platonic datasets

The set of four-dimensional reflexive polytopes is too large to work with by hand or even to do much in-depth work with individual examples. Furthermore we are often interested in questions of existence or statistical questions, *i.e.* what fraction of the cases have a property X of interest, or is there a correlation between the presence of properties X and Y in a given example. In this sense it is a dataset much like the datasets we deal with in computer science and statistics.

Usually one deals with datasets which arise from the physical world: images, waveforms, other measurements – or from the social world: texts, networks of relationships, *etc.*. These are empirical datasets.

The Kreuzer-Skarke dataset, as do other datasets of mathematical objects, has an *a priori* definition. It is a **platonic dataset.** Platonic datasets are a main focus of interest in computational mathematics.

Mathematical ingredients in string compactification

- Finite group theory – classification, representation theory, group cohomology.
- Lie algebras and groups – likewise
- Theory of integral lattices.
- Algebraic geometry: construction of manifolds, especially Calabi-Yau manifolds, holomorphic vector bundles and the like.
- Toric geometry
- Singularity theory
- Quiver representation theory
- Theory of moduli spaces, arithmetic groups, automorphic functions and forms
- Theory of hypergeometric functions and PDEs.
- Diverse numerical techniques

Interdisciplinary computational mathematics

Understanding string theory and making testable predictions is a long-term project: we might find success tomorrow, or in a decade, or in the distant future.

However long this takes, we can make important contributions to human knowledge along the way. For example, this is the spirit behind the math-physics interface: string theory "knows more mathematics than we do" and its study has led to the development of a large quantity of new mathematics: mirror symmetry (enumerative formulas, homological mirror symmetry, Bridgeland stability), new invariants (Gromov-Witten, Donaldson-Thomas, Gopakumar-Vafa, *etc.*), topological field theory and topological quantum gravity, *etc.*.

The contributions of string theorists are generally not rigorous mathematics but rather conjectures, new relations, new connections, and new questions which are developed in a loose collaboration between physicists and mathematicians.

Let us make the case that the exploration of the string landscape should also be done in collaboration with computer scientists and computational mathematicians. To summarize the arguments:

- String compactification uses a wide variety of mathematical results which are of central interest to mathematicians and mathematical scientists.
- Any systematic approach to its study, especially one which involves data science, will require making databases of string vacua.
- This requires having computational implementations of many mathematical structures.
- The math itself is intricate and computational implementations will be equally so. Especially, the work to validate and maintain them as computational and other technologies develop, and as our conceptual framework improves, exceeds what most academic collaborations can sustain.

- Fortunately these mathematical concepts and their computational implementations are of interest to larger communities. Finite group theory is of interest to all mathematical scientists, as are many statistical constructs. Lattices are of central interest in coding theory. Algebraic geometry and automorphic functions are more the province of pure mathematicians, but again there is a sizable community. There are applications of algebraic geometry to statistics (see work of Sturmfels *et al*) and theoretical computer science (see Mulmuley and Sohoni, many other works). Yau and others have advocated uses of Kähler geometry and moduli spaces in applied mathematics.
- There are also major computational challenges, not just in speed of algorithms and storage of large datasets, but in organizing a very large database of formal knowledge in a distributed and collaborative way. If there were a Wikipedia of computational mathematics, we could build on it. Since there is not, there should be a project to develop one. This is beyond what anyone knows how to do, but people are starting to think along these lines.

While many of the mathematical definitions and algorithms we need to construct string vacua have been implemented, in my opinion there is a crucial missing ingredient – namely, a way to integrate them into larger projects, which is

- modular – continuous integration allowing changes in any component or even in the underlying definitions while minimizing the changes required to other components
- efficient – we have moderately large datasets with millions of items
- distributed – projects will generally be collaborative
- archival – facilitates publication and archiving of papers and datasets
- open source – not dependent on proprietary software

Many areas of scientific computation and open source software development have these requirements, so here is another goal for computational mathematics, to which formal methods can contribute.