Mathematics and string theory

Marcos Mariño

Département de Physique Théorique et Section de Mathématiques
Université de Genève
Genève, CH-1211 Switzerland

marcos.marino@unige.ch

Contents

1 Introduction

1.1 Historical context

String theory was born in the late 1960's and the early 1970's as a conjectural description of the strong interactions which are responsible of nuclear forces. However, it was realized very soon that string theory does not provide an appropriate model for these interactions. One of the reasons for this dismissal was that string theory models predicted the existence of a massless particle of spin two which had no experimental counterpart in the the world of nuclear particles. But this problem became a virtue when John Schwarz and Joël Scherk noticed that this particle had all the properties of a graviton, the particle transmitting the gravity force. String theory became then a consistent candidate for a theory of quantum gravity, and later on, in the early 1980's, a candidate for a unified theory of all particles and their interactions– a "theory of everything.".

At the moment of this writing, it is fair to say that string theory has not found yet a clear place in our understanding of Nature. However, it has already established itself as a source of fascinating results and research directions in mathematics. String theory and some of its close cousins (like conformal field theory and topological field theory) have had an enormous impact in representation theory, differential geometry, low-dimensional topology, and algebraic geometry. Many recipients of the Fields Medal, like Shing-Tung Yau, Simon Donaldson, Vaughan Jones, Richard Borcherds, Maxim Kontsevich, and Andrei Okounkov, have been deeply influenced by ideas and results coming from string theory. Edward Witten, arguably the leading string theorist, is also the first physicist to win the Fields medal.

The relationship between string theory and mathematics has its roots in a very old tradition going back to XVIIth century. Since Newton, physics and mathematics have had many cross-fertilizing connections which are an essential part of the history of both disciplines. The two pillars of modern theoretical physics, Quantum Mechanics and General Relativity, are intimately related to functional analysis and Riemannian geometry, respectively. On the other hand, in the second third of the XXth century the link between mathematics and physics became less important, as mathematics evolved in more abstract realms and fundamental physics invested its energies in various fronts (particle physics, solid state physics) which were *a priori* far from the interests of mathematicians. There were of course exceptions to this rule. In General Relativity, geometry had always played an important role, and in the late 1960's Penrose introduced modern topological techniques in order to analyze some of the deepest aspects of the theory.

A change of direction can be appreciated in the early 1970's, in the context of particle physics, due to the advent of gauge theories of particle interactions. Geometrically, these theories are described by a fibre bundle over the spacetime manifold, and gauge fields correspond to connections on this bundle. As Witten points out in his recollections of the period, "the gauge theory revolution had created a situation in which it would be necessary and worthwhile to develop a greater mathematical sophistication than we were accustomed to" [54]. The connections established between gauge theories and modern mathematics in the 1970's display two new features as compared to prior developments. First of all, although the traditional mathematical interlocutor of quantum theory had been functional analysis, in this new period the branches of mathematics which are more deeply involved are geometry and topology. Second, these geometric and topological aspects of gauge fields turn out to be crucial for undestanding their *quantum* dynamics, in contrast to what happens for example in General Relativity, where geometry and topology had been used in a purely classical context.

An important role in these developments was played by the so-called $U(1)$ problem of Quantum Chromodynamics (QCD). QCD, a quantum gauge theory, was emerging in that period as the correct theory of strong interactions, but it seemed to have an important flaw known as the $U(1)$ problem: the theory predicted that one of the elementary particles in the family of the pions (the so-called η') should be much lighter than observed. In 1976, a solution to the $U(1)$ problem was proposed by 't Hooft, Jackiw, Rebbi, and others (with important contributions later on by Witten and Veneziano). According to this solution, the mass of this pion is of topological origin, and an appropriate treatment of the problem involves in a crucial way the second Chern class of the gauge bundle and the Atiyah–Singer index theorem.

The resolution to the $U(1)$ problem showed that, in some circumstances, quantum effects had a direct connection to rather sophisticated developments in geometry and topology. Of course, the attempts of Einstein to construct classical unified field theories relied very much on a geometrization of physics, but the new developments in quantum field theory suggested that geometry and topology could also be crucial in the quantum world of particle physics. This idea permeated many developments in theoretical particle physics in the late 1970's and early 1980's[1], and it was in this context of increasing mathematical sophistication that string theory emerged as a candidate for a unified theory of quantum fields and particles. In the late 1980's conventional particle theorists were becoming more skeptical that geometry could hold the key to the mysteries of particle physics, but by then that belief had become one of the leading ideas in string theory. This made in turn possible the surprising interactions between string theory and geometry that were to take place in the following years.

1.2 String theory and mathematics

The enormous impact of string theory on mathematics has been based on two different facts. The first one is simply that the formulation of string theory involves many structures which are mathematically very interesting, like the moduli space of Riemann surfaces, Calabi–Yau manifolds, etc. The appearance of these structures in string theory has cleary stimulated further investigation of these topics in the mathematical community. This type of relationship between string theory and mathematics is not new, and in fact it is quite similar to previous connections between physics and mathematics: when a physical theory is formulated, it is quite often the case that a precise construction of the theory requires a specific type of mathematical structure. For example, General Relativity needed tensor calculus and Riemannian geometry in order to be properly formulated, and quantum mechanics needed the theory of Hilbert spaces and of linear operators. When a mathematical structure appears naturally in a physical theory, interest in such a structure in the mathematical community typically grows. As a consequence, more mathematical results are obtained concerning this structure, and this might lead to further progress on the physical side. Unfortunately, as we pointed out before, this has not been always the case in the XXth century, where physics-inspired mathematical research increasingly had to compete with developments in mathematics arising from its own dynamics and history. As a consequence, developments in mathematics and physics have been largely independent and cross-fertilization has not been always the norm.

The second aspect of the interaction between mathematics and string theory is more

[1] In the SPIRES database, prior to 1975 there are only 50 papers with the words "topology" or "topological" in their title. This number increases by a factor of 10 in the period between 1975 and 1985.

original and unprecedented. It is true that sometimes physics suggests some interesting problems and a general pattern for their solution, but precise mathematical statements coming from physics are rare, and they tend to be more sparse as we move towards the "purest" aspects of the mathematical structure.

In contrast, string theory has provided an enormous amount of precise conjectures in disciplines that belong in principle to the most abstract parts of mathematics, like algebraic geometry and differential topology. A famous example, which we will develop in detail later in this article, is Witten's conjecture on intersection theory on the moduli space of Riemann surfaces. As we will see, string theory requires naturally in its formulation the calculation of integrals over this moduli space. According to the classical pattern of interactions between mathematics and physics, we would expect that the existence of such a connection would encourage mathematicians to develop this integration theory, and that mathematical progress in this area would make possible to obtain further results in string theory. It is harder to imagine that developments in the physics of string theory lead to a full solution to the problem of intersection theory in the moduli space of Riemann surfaces, and that the solution to this problem involves a completely unsuspected mathematical structure –the theory of classical integrable systems. However, this is precisely what happened when Witten formulated his conjecture about integration theory in the moduli space of Riemann surfaces, in 1990.

The main impact of string theory in mathematics has been made by predicting precise formulae and properties for quantities like intersection numbers in moduli spaces. These quantities are typically *topological invariants* of geometric objects, and they include Gromov–Witten invariants, knot invariants, etc. This direct impact of string theory on its mathematical counterparts has been made possible due to a combination of two ingredients. First of all, it turns out that certain quantities that appear naturally in string theory compute generating functionals of these topological invariants. The existence of this "topological sector" of string theory is surprising, and in many cases it is closely related to the existence of supersymmetry, i.e. to the existence of an additional symmetry in the theory that makes many interesting quantities sensitive only to the topological properties of the underlying spacetime.

The second ingredient is what we might call the existence of dual descriptions, or *dualities*, in string theory. This means that two very different–looking string theory models turn out to be equivalent, in the sense that a quantity computed in model A, let us say $F_A(t_A)$, and depending on a set of parameters t_A typical of model A, will be identical to a different quantity in model B, $F_B(t_B)$, and depending on a different set of parameters t_B. When this is the case, we say that the models A and B are *dual* to each other. Of course, for this statement to be useful, one needs a precise relation between the parameters t_A and t_B, which typically is also provided by a physical argument. We will see various examples of this later on.

Imagine now that $F_A(t_A)$ is a generating functional for a set of topological invariants. By duality, they can be also obtained from $F_B(t_B)$, once the appropriate dictionary between t_A and t_B is found. But since the model B is *a priori* very different, $F_B(t_B)$ might be a completely different quantity. It can be, for example, a generating functional for a *different* set of topological invariants. In particular, $F_B(t_B)$ might display properties which are not obvious when one looks at $F_A(t_A)$. We then see that the combination of these two aspects of string theory is potentially very powerful in order to produce new mathematical information. It is important to notice that the arguments establishing dualities are, very often, not proved even at the level of physical rigor, since they often use physical insights which are far from being formulated in a rigorous way. Although this is a weakness from the point of view of a logical construction, it is almost a guarantee that the mathematical translation of the duality will be a surprising conjecture in the mathematical world. Conversely, if this conjecture becomes eventually a theorem, one obtains a rigorous confirmation of the string duality one started with. In this article we will see many examples of this procedure, which we sketch in the form of a diagram in Fig. 1.

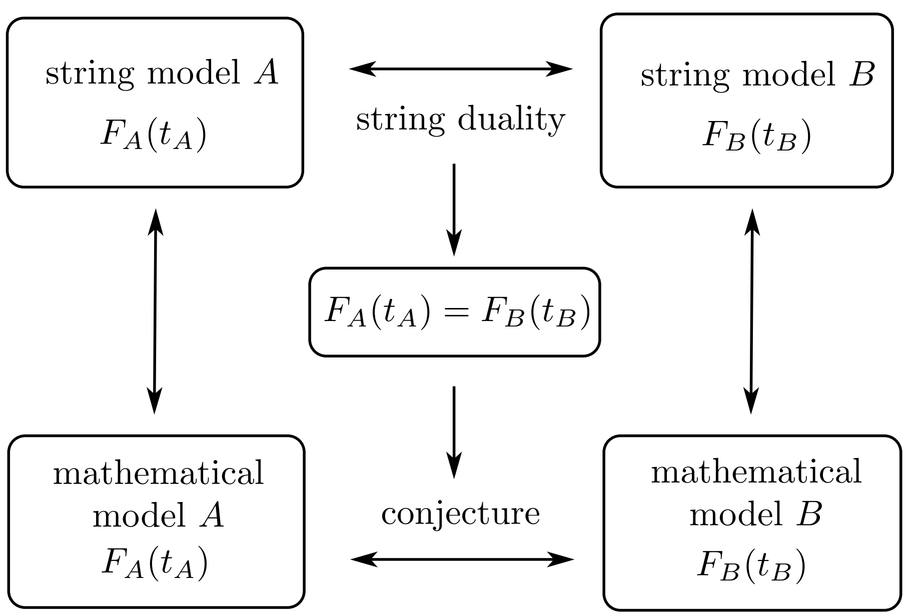

Figure 1: The unreasonable effectiveness of string theory in mathematics: two string theory models, A and B, might lead to quantities $F_A(t_A)$ and $F_B(t_B)$, respectively, with a direct mathematical interpretation (they might be generating functionals of different mathematical invariants, for example). If the models A and B are dual to each other, these two quantities will be identical, after appropriately identifying the parameters t_A and t_B. This produces a mathematical conjecture relating two *a priori* different mathematical objects.

1.3 Outline

In this article we will first give a very simple introduction to the basic features of string theory, and then we will review what are in our opinion some of the most relevant aspects of the interaction between string theory and mathematics. The examples that we will focus on are Witten's conjecture on intersection theory on Deligne–Mumford moduli space, mirror symmetry, and the Gopakumar–Vafa conjecture. This is of course not an exhaustive list. There are many other interactions between string theory and mathematics which we will not cover, like for example the remarkable connections with Monstruous Moonshine and Borcherds' work (see [14] for a recent exposition of this topic and its connection to physics). Our choice of examples is, we hope, well founded, but it is probably biased by our tastes, competences, and prejudices. Finally, in the concluding section we will make a preliminary appraisal of the subject and comment also on some of its sociological implications for the relationship between physics and mathematics.

2 What is string theory?

In order to understand in some detail the mathematical implications of string theory, it is necessary to present some of its ingredients, and in particular to explain in which sense string theory is a generalization of theories of point particles. Our presentation will be necessarily sketchy and incomplete. Good introductions to string theory for the mathematically oriented reader can be found in [10] and more recently in [14]. The standard, modern textbook on superstring theory is [45].

In a classical theory of point particles, the fundamental ingredient is the *trajectory* of the particle in a given spacetime, which is typically represented by a differentiable, Riemannian manifold X. This trajectory can be represented by an application

$$\begin{aligned} x :&\mathcal{I} \to X, \\ &\tau \mapsto x(\tau) \in X, \end{aligned} \tag{2.1}$$

where $\mathcal{I} \subset \mathbb{R}$ is an interval, and $\tau \in \mathcal{I}$ is the time parametrizing the trajectory. This specification of the trajectory provides only the kinematical data. Determining the dynamics requires as well an *action functional* $S(x(\tau))$. For example, for a free particle one has

$$S(x(\tau)) = \int \mathrm{d}\tau \, G_{\mu\nu}(x(\tau))\dot{x}^{\mu}\dot{x}^{\nu}(\tau) \tag{2.2}$$

where $G_{\mu\nu}$ is the metric of X. The classical equations of motion can be derived from the variational principle

$$\frac{\delta S}{\delta x} = 0 \tag{2.3}$$

and in the case of the free particle this means that the motion occurs along *geodesics* of the Riemannian manifold X.

Quantization of the theory is appropriately done by using Feynman's path integral formalism. In this formalism, the quantum mechanical propagator is obtained by integrating the weight $\exp(-S(x))$ over all possible trajectories $x(\tau)$ with fixed boundary conditions,

$$x(\tau_0) = x_i, \qquad x(\tau_1) = x_f, \tag{2.4}$$

where τ_0, τ_1 are the endpoints of the interval. This is written, formally, as

$$K(x_i, x_f) = \int \mathcal{D}x(\tau)\, \mathrm{e}^{-\frac{1}{\hbar}S(x(\tau))}. \tag{2.5}$$

This is the so-called *Euclidean* propagator. The weight $\exp(-S(x))$ might be regarded as a probability distribution for paths, and the propagator is closely related to the probability that a quantum particle which starts at x_i at $\tau = \tau_0$ is detected at x_f at $\tau = \tau_1$.

One can also consider *periodic* trajectories, i.e. maps of the form

$$x : S^1 \to X \tag{2.6}$$

where the circle has length β. This means that the map x is defined on a *closed* one-manifold. The path integral with these boundary conditions,

$$Z(\beta) = \int \mathcal{D}x(\tau)\, \mathrm{e}^{-\frac{1}{\hbar}S(x(\tau))} \tag{2.7}$$

gives the partition function at temperature $kT = \beta^{-1}$ (where k is the Boltzmann constant), which describes the properties of a particle in a thermal bath at this temperature.

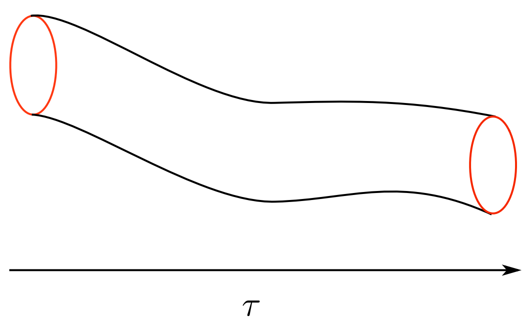

Figure 2: A closed string propagating in time.

In string theory, point-particles are replaced by one-dimensional objects. Classically, the embedding of such an object in a spacetime manifold is described by a map

$$\begin{aligned} x :&\mathcal{I} \times \mathcal{S} \to X, \\ &(\tau, \sigma) \mapsto x(\tau, \sigma) \end{aligned} \tag{2.8}$$

where $\sigma \in \mathcal{S} \subset \mathbb{R}$ parametrizes now the string. The dynamics is specified again by a classical action $S(x)$, which in the simplest case takes the form

$$S(x,h) = \frac{1}{\ell_s^2} \int \mathrm{d}\tau \mathrm{d}\sigma \, \sqrt{h} G_{\mu\nu} h^{ab} \partial_a x^\mu \partial_b x^\nu. \tag{2.9}$$

Here, h is a metric on the two-dimensional manifold $\Sigma = \mathcal{I} \times \mathbb{S}$, and it plays the role of an auxiliary field. The quantity ℓ_s, with the dimensions of a length, sets the scale of this one-dimensional extended object and for this reason it is called the *length of the string*. It is convenient to parametrize $\mathbb{S} = [0, \pi]$. A *closed string* is a loop with no free ends, and in this case the appropriate boundary condition is

$$x(\tau, 0) = x(\tau, \pi), \qquad \tau \in I. \tag{2.10}$$

A freely propagating closed string, as it propagates in time, spans the surface with the topology of a cylinder, see Fig. 2.

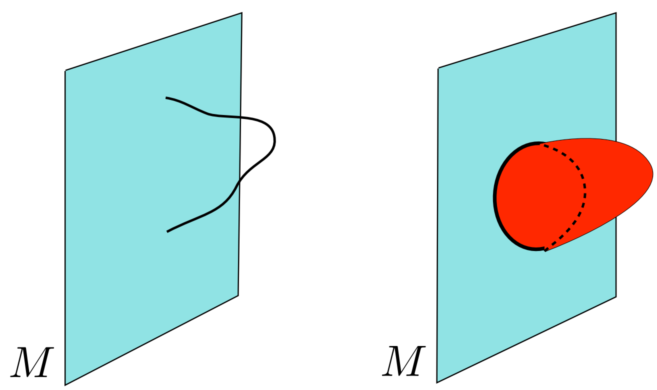

Figure 3: The left hand side of the figure shows an open string whose boundaries are attached to a submanifold M of spacetime. Such submanifolds, providing Dirichlet boundary conditions for open strings, are called D-branes. The right hand side of the figure shows an embedded disk whose boundary ends on the D-brane.

An *open* string can have many different boundary conditions. Particularly important are the so-called *Dirichlet* boundary conditions provided by a submanifold $M \subset X$. These are given by

$$x(\tau, 0), \quad x(\tau, \pi) \in M \tag{2.11}$$

and they describe a string whose ends are attached to the submanifold M of the spacetime X, which is called a *D-brane* (where D stands for Dirichlet). The left hand side of Fig. 3 shows an open string attached to a D-brane represented by a plane.

There are two types of string interaction: in a *splitting* process one single string splits into two, and in a *joining* process two stings merge into one. As in quantum physics, the strength of such an interaction is measured by a constant g_s called the *string coupling constant*. For closed strings, the basic process of joining or splitting is described by a "pair of pants" diagram as in the left hand side of Fig. 4, and it has a single power of g_s associated to it. When this process takes place n times, it has the factor g_s^n.

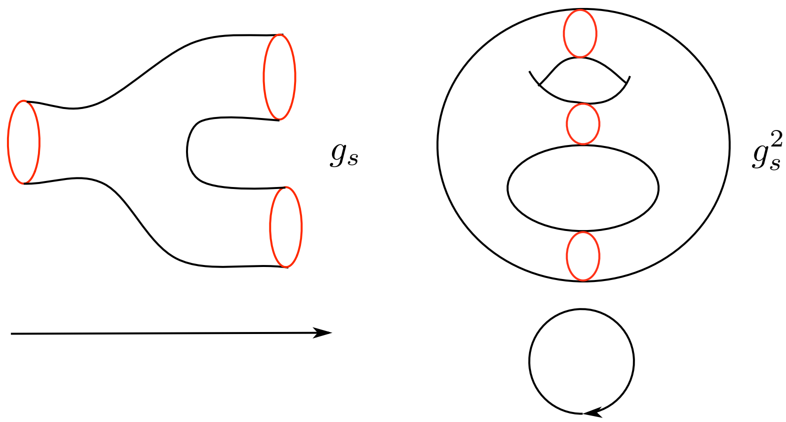

Figure 4: A closed string interaction takes place when one single string splits into two, as shown in the left hand side of the figure. Such a process is weighted by a power of g_s. In the right hand side we show a periodic configuration in which a closed string splits and then joins again before coming back to itself, spanning a Riemann surface of genus 2. Since there were two string interactions involved, this process has a weight g_s^2.

As in the theory of point particles, periodic configurations generalizing (2.6) play a very important role. For a closed string, a periodic configuration is described by a single string which, after various processes of splitting and joining, comes back to itself. This process produces a *closed, orientable Riemann surface* Σ_g, and it is easy to see that it has the weight g_s^{2g-2}, where g is the genus of the resulting Riemann surface. In the right hand side of Fig. 4 we show a periodic configuration in which a closed string evolves, splits into two strings which merge back to a single closed string, and this string goes back to the starting point. The time evolution produces a Riemann surfaces of genus 2, and since there were two string interactions (one splitting and one joining) the whole process has a factor g_s^2 associated to it.

We conclude that, in general, a periodic map is just a map from Σ_g to the spacetime X:

$$x : \Sigma_g \to X. \tag{2.12}$$

The Riemann surface Σ_g is called the *worldsheet* of the string, while the manifold X is called its *target space*. The generalization of periodic maps for open strings consists in considering maps from a Riemann surface $\Sigma_{g,s}$ with s boundaries to X,

$$x : \Sigma_{g,s} \to X. \tag{2.13}$$

The generalization of the boundary condition given by a D-brane $M \subset X$ is

$$x(\partial \Sigma_{g,s}) \subset M, \tag{2.14}$$

i.e. the boundaries of $\Sigma_{g,s}$ are mapped to M. The right hand side of Fig. 3 shows an embedded Riemann surface with genus $g = 0$ and $s = 1$ (i.e. a disk) whose boundary is fixed on the D-brane, represented by a plane.

The quantization of a theory of strings is rather delicate. Formally, one considers a path integral over all possible configurations of the fields, as in (2.5), with the appropriate boundary conditions. For simplicity we will consider periodic boundary conditions, i.e. the analogue of (2.7). Since we have to consider all possible configurations of the string, we have to take into account all possible splitting/joining processes, and this means that we should sum over all possible genera for the Riemann surfaces spanned by the string. In the computation of Z one considers *disconnected* Riemann surfaces, but in the so-called "free energy," defined by $F = \log Z$, one should sum only over connected Riemann surfaces, labelled by the genus g. For each genus we should integrate over all metrics on Σ_g and all configurations of maps x. The space of all metrics on Σ_g, after taking into account the relevant symmetries, turns out to be equivalent to the moduli space of Riemann surfaces $\overline{M}_g$ constructed in algebraic geometry, and the free energy has then the structure

$$F = \sum_{g=0}^{\infty} g_s^{2g-2} F_g, \qquad F_g = \int_{\overline{M}_g} \mathcal{D}h_{ab}\, \mathcal{D}x\, \mathrm{e}^{-S(x,h)}. \tag{2.15}$$

Note that, in the above expression, the integration over the metric has been in fact reduced to an integration over $3g-3$ complex moduli parametrizing $\overline{M}_g$.

In a quantum theory one can also calculate *correlation functions*. A correlation function is the quantum average of a product of operators, and a typical operator is a functional of the basic fields of the theory. The correlation function of the operators $\mathcal{O}_1(p_1), \cdots, \mathcal{O}_n(p_n)$ is formally defined as the path integral

$$\langle \mathcal{O}_1(p_1) \cdots \mathcal{O}_n(p_n) \rangle_g = \int_{\overline{M}_{g,n}} \mathcal{D}h_{ab}\, \mathcal{D}x\, \mathcal{O}_1(p_1) \cdots \mathcal{O}_n(p_n)\, \mathrm{e}^{-S(x,h)}. \tag{2.16}$$

Each operator is evaluated at a point p_i of Σ_g, therefore operators lead to *marked points* on the Riemann surface. The appropriate integration is now over the moduli space of Riemann surfaces with n marked points, $\overline{M}_{g,n}$.

Although the expressions (2.15) and (2.16) are purely formal, they display some properties that will be important in the following. First ot all, F depends on two parameters which are characteristic of theories based on strings: the string length ℓ_s (which appears in the string action) and the string coupling constant g_s. In particular, if we regard string theory as a two-dimensional quantum field theory, we see from (2.9) that the squared string length ℓ_s^2 plays the same role as $\hbar$. When $\ell_s \to 0$, the length of the string vanishes and we recover a theory of point particles. Of course, the coupling constant g_s can be also regarded as a quantum parameter, and in the point-particle limit of string theory it becomes the standard coupling constant governing quantum interactions of particles. The existence of these two quantum parameters makes clear that string theory is a consistent deformation of a quantum theory of point particles. The second property of the above formal expressions is that they involve an integration over the moduli space of Riemann surfaces with marked points. This opens the door to deep connections between string theory and the enumerative geometry of Riemann surfaces, as we will see.

The structure we have described is roughly speaking the so-called *bosonic string* with target space X, and it can be easily generalized after we identify its key ingredients. On one hand, the map x can be regarded as a two-dimensional quantum field described by the action (2.9). This field theory has the property of being invariant under the full conformal group in two dimensions, i.e. it is an example of a two-dimensional *conformal field theory* (CFT). On the other hand, we also introduced a metric in the two-dimensional surface where this field lives. The combination of these two ingredients is a particular example of a *two-dimensional conformal field theory coupled to two-dimensional gravity*. In more abstract terms, a string theory is just such a system, and depending on the conformal field theory involved –the so-called "matter content" of the string theory– we will have different string theories. For example, in *supersymmetric* string theories the conformal field theory is supersymmetric, i.e. there are extra fermionic fields as well as a symmetry exchanging the bosonic and the fermionic fields.

One surprising aspect of string theories is that they only make sense as quantum theories if one imposes constraints on the field content. For example, the bosonic string is only consistent if the target space has 26 dimensions, i.e. if the field x (which can be regarded as a coordinate system for X) has 26 components. Supersymmetric string theories require X to have 10 dimensions. If string theory is regarded as a model of the real world, the target space X should be identified as the physical space-time, and the operators of the CFT give rise to quantum fields propagating on X. We conclude that string theory models requires extra physical dimensions. As we will see, this also has implications for the relationship between string theory and mathematics. Another important aspect of string theory is that, for closed strings, one of the operators of the CFT describes a graviton propagating in space-time, and this makes possible to obtain a full quantum-mechanical treatment of gravity .

3 The mathematics of Riemann surfaces and string theory

3.1 The mathematics of Riemann surfaces

The theory of Riemann surfaces is a fundamental ingredient of algebraic geometry, and in its modern form it has its origins in the pioneering work of Riemann, Abel and others in the XIXth century. In the 1970's, Deligne and Mumford constructed the moduli space of Riemann surfaces $\overline{M}_{g,n}$ with genus g and n marked points (see [20] for an exposition). This is roughly the space of isomorphism classes of surfaces with marked points, and it has complex dimension $3g-3+n$ (we should notice that, strictly speaking, this space is not a smooth variety –in technical terms it is what is called an algebraic stack, see [20] for a detailed exposition). As we have seen, the computation of free energies (2.15) and correlation functions (2.16) in string theory involves in a natural way the integration over the moduli spaces of Riemann surfaces, and the work of algebraic geometers was instrumental to provide a precise formulation of these quantities. The surprising novelty was that, in turn, string theory provided a set of conjectures and concepts on the integration theory on these moduli spaces which set the agenda for its mathematical development.

In order to develop the theory of integration over moduli spaces, one has to construct first appropriate differential forms on $\overline{M}_{g,n}$. An important set of differential forms are the so-called ψ classes, which are defined as follows. We consider the line bundles $\mathcal{L}_i$ over $\overline{M}_{g,n}$, $i=1,\cdots,n$, whose fibre over a point $\Sigma \in \overline{M}_{g,n}$ is $T^*\Sigma_g l|_{p_i}$, where p_i is the i-th marked point. Then, the first Chern class

$$\psi_i = c_1(\mathcal{L}_i) \tag{3.1}$$

is a two-form on $\overline{M}_{g,n}$. The most general integral involving these classes is

$$\langle \sigma_{d_1} \cdots \sigma_{d_n} \rangle_g = \int_{\overline{M}_{g,n}} \psi_1^{d_1} \wedge \cdots \psi_n^{d_n}. \tag{3.2}$$

Such an integral is called an *intersection number* due to its geometric interpretation in terms of Poincaré duality. Conjectures arising from string theory are better formulated in terms of generating functionals, therefore we introduce an infinite vector of variables indexed by $i \geq 0$, $t=(t_0,t_1,\cdots)$, and we define

$$F(t,g_s) = \sum_{g=0}^{\infty} g_s^{2g-2} F_g(t), \tag{3.3}$$

where

$$F_g(t) = \sum_{\{n_i\}} \Big(\prod_{i=1}^{\infty} \frac{t_i^{n_i}}{n_i!} \Big) \langle \sigma_0^{n_0} \sigma_1^{n_1} \cdots \rangle_g. \tag{3.4}$$

The functional $F(g_s, t)$ encodes all possible intersection numbers defined for the ψ classes, which we can regard as topological invariants associated to the Deligne–Mumford moduli space. They encode an enormous quantity of information about these spaces, and *a priori* their calculation is not an easy task.

3.2 Witten's conjecture

However, Edward Witten, made a remarkable conjecture in 1990 which solves this problem in a closed form [48, 49]. The conjecture states that the functional $F(t, g_s)$ satisfies an infinite number of differential equations which completely capture the dependence of $F(t, g_s)$ on the variables t_i. These differential equations form a well-known integrable hierarchy called the KdV (i.e. Korteweg–de Vries) hierarchy. One can easily show that the hierarchy determines $F(t, g_s)$ uniquely, once a finite number of initial conditions are provided. These initial conditions are just the values of the intersection numbers

$$\langle \sigma_0^3 \rangle_0 = 1, \qquad \langle \sigma_1 \rangle_1 = \frac{1}{24}. \tag{3.5}$$

Equivalently, the system of differential equations provides a set of recursion relations for the intersection numbers which determines them uniquely, once the "boundary data" (3.5) are provided.

As an example of the type of results one can find with this conjecture, let us consider the particular sequence of intersection numbers

$$\langle \sigma_2^{3g-3} \rangle_g = \int_{\overline{M}_{g,3g-3}} \psi_1^2 \wedge \cdots \wedge \psi_{3g-3}^2, \qquad g \geq 2. \tag{3.6}$$

These numbers can be put together in the generating functional

$$f(z) = -\frac{4}{15} z^{5/2} - \frac{1}{48} \log z + \sum_{g \geq 2} \frac{\langle \sigma_2^{3g-3} \rangle_g}{4^g \, (3g-3)!} z^{-5g/2}. \tag{3.7}$$

This is essentially (3.3) where only $t_2 \neq 0$. Witten's conjecture implies that the formal series

$$u(z) = -f''(z) = z^{\frac{1}{2}} - \frac{1}{48} z^{-2} + \cdots \tag{3.8}$$

satisfies the differential equation

$$u^2 - \frac{1}{6} u'' = z, \tag{3.9}$$

which is nothing but the Painlevé I transcendent. This illustrates very well the type of connection found by Witten between intersection numbers in the Deligne–Mumford moduli space and differential equations with integrability properties. Notice that this particular case of Witten's conjecture reduces the calculation of the intersection numbers (3.6) to a recursion relation easily deduced from the Painlevé I equation.

What is the string theory argument behind Witten's conjecture? As we explained above, the general setting for formulating a conjecture like this is an equivalence between two different string theories. In this example, the first string theory (what we call model A) is *two-dimensional topological gravity*. Topological gravity in two dimensions can be regarded as a string theory where the target space X is just a point, and it is constructed in such a way that the total free energy (2.15) is precisely the generating functional (3.3)

The other string theory is *a priori* very different from this one. Recall that a string theory is obtained by coupling a two-dimensional conformal field theory to gravity. Among conformal field theories, the simplest ones are the so-called *minimal models* which were discussed by Belavin, Polyakov and Zamolodchikov in [5]. These minimal models are classified by two coprime integers (p, q), so for example the $(3, 4)$ model is the conformal field theory which describes the critical point of the two-dimensional Ising model. The coupling of minimal models to gravity is technically difficult and it involves in a crucial way *Liouville theory*, which describes the metric on the Riemann surface. The meaning of these theories coupled to gravity is easy to understand if the conformal field theory in two dimensions describes the critical point of a statistical system on a lattice: after coupling to gravity, they describe the critical points of statistical systems in which the underlying lattice can also fluctuate. As we approach the critical point, the fluctuating lattice is described by a fluctuating metric on a Riemann surface with the topology of the lattice, and this means that we are in the presence of two-dimensional gravity.

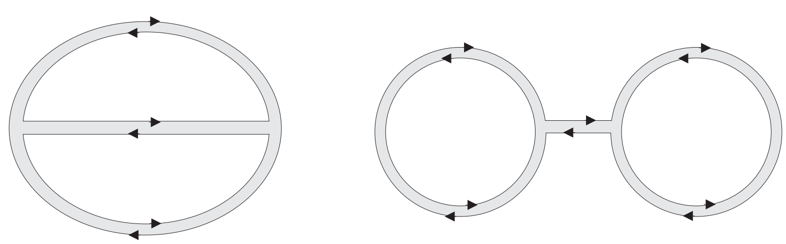

Figure 5: Two trivalent fatgraphs with $g = 0$.

The minimal models coupled to gravity are very difficult to describe in terms of standard quantum fields. Surprisingly, it turns out to be more convenient to describe first the model off-criticality, i.e. to sum over discretized fields on a discretized two-dimensional fluctuating lattice, as in lattice gauge theory. This approach to gravity, based on discretizing space-time, was pioneered by Ponzano and Regge, and it turns out to be extremely successful when applied to two-dimensional gravity. The reason is that one can easily describe fluctuations of two-dimensional Riemann surface by considering *matrix models*. A matrix model is defined by the integration over a space of matrices, which in this case are Hermitian $N \times N$ matrices. The integration is done with a standard measure and it involves, in the integrand, an exponential weight depending on the matrix. A simple matrix model which is relevant for our discussion of two-dimensional gravity is defined by the partition function

$$Z_N = \int \mathrm{d}M \, \exp\Big\{-N\mathrm{Tr}V(M)\Big\} \tag{3.10}$$

where

$$V(M) = \frac{1}{2}M^2 + \lambda M^3. \tag{3.11}$$

The partition function (3.10) depends on N, the rank of the matrix, and on the parameter λ appearing in the "potential" $V(M)$. It was long ago pointed by 't Hooft that at large N the logarithm of this integral (which is called the "free energy" of the matrix model) can be expanded as follows,

$$\log Z_N = \sum_{g=0}^{\infty} F_g(\lambda) N^{2-2g}. \tag{3.12}$$

It is not difficult to see that $F_g(\lambda)$ can be interpreted diagramatically as a sum over connected, closed, trivalent graphs with double lines (of course, these graphs are trivalent due to the cubic term in (3.11)). Graphs with double lines are called "fatgraphs" or double-line diagrams in the physics literature, and maps in the combinatorial literature. A fatgraph is characterized topologically by the number of vertices V, edges E and boundaries h, and it can be regarded as a Riemann surface of genus g with h boundaries, where g is given by Euler's formula

$$2g - 2 = E - V - h. \tag{3.13}$$

Examples of trivalent graphs are shown in Fig. 5 and Fig. 6. A trivalent fatgraph of genus g also defines a triangulation of a genus g Riemann surface by considering the "dual" graph. i.e. we associate to each vertex a triangle in such a way that the sides of the triangle are orthogonal to the edges coming out of the vertex. We then glue the faces of the triangles which are connected by edges. Therefore, the quantity $F_g(\lambda)$ can be regarded as a sum over all possible triangulations of a genus g Riemann surface, or equivalently over all possible fluctuations of a triangular lattice on a Riemann surface of genus g. More complicated matrix models can incorporate extra structures on the discrete lattice, like for example Ising spins [25].

However, we want to make contact with two-dimensional quantum gravity on a *smooth* Riemann surface, and as we discussed above this involves approaching a critical point of the theory. It turns out that $F_g(\lambda)$ has indeed such a point at a finite value of λ, λ_c. When λ is close to λ_c, the $F_g(\lambda)$ have a non-analytic behavior of the form

$$F_g(\lambda) = f_g(\lambda - \lambda_c)^{(2-\gamma)(1-g)} + \cdots \tag{3.14}$$

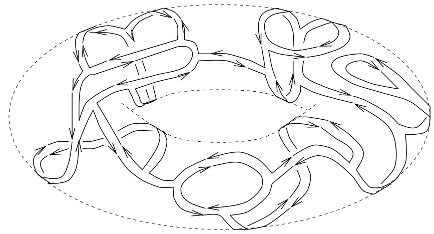

Figure 6: A trivalent fatgraph with $g = 1$.

where $\gamma = -1/2$ for the potential (3.11) (higher order potentials may have a different value of γ). One can show that, near this critical point, the area of each face in the discretization of the surface goes to zero size, but at the same time the triangulations become denser and denser so as to recover a smooth surface. We then define the free energy of two-dimensional quantum gravity by extracting the leading singularity from $F_g(\lambda)$, for all genera:

$$F_{(2,3)}(g_s) = f_0 g_s^{-2} + f_1 + \sum_{g\geq 2} f_g g_s^{2g-2} \tag{3.15}$$

$F_{(2,3)}(g_s)$ is the total free energy (2.15) of a string theory, in this case the $(2,3)$ minimal model coupled to gravity. The matter content of this string theory is trivial, but there is a nontrivial Liouville field which represents the fluctuating metric on the Riemann surface. More complicated minimal models coupled to gravity are described by more complicated matrix integrals. For example, the minimal model $(2, 2m+1)$ can be described by a Hermitian matrix model with a single matrix, like (3.10), but with a more complicated polynomial potential $V(M)$. In the case of the cubic matrix model one can show that

$$u = -f''(z), \qquad f(z) = f_0 z^{\frac{5}{2}} + f_1 \log z + \sum_{g\geq 2} f_g z^{-5g/4} \tag{3.16}$$

satisfies the differential equation (3.9) characterizing the subset of intersection numbers defined by (3.6) (see for example [11]), so in particular the free energy (3.15) is essentially the generating functional (3.7). The more general matrix models describing the full set of $(2, 2m+1)$ CFTs coupled to gravity turn out to satisfy the full system of equations of the KdV hierarchy. The time variables appearing in this hierarchy are related to the coefficients in the polynomials defining the matrix models.

Witten's conjecture is a consequence of the equivalence between two different string theories: on one hand (i.e as string model A) we have two-dimensional topological gravity, whose mathematical counterpart is intersection theory in Deligne–Mumford moduli space. On the other hand (i.e. the string model B) we have the full family of $(2, 2m+1)$ minimal models coupled to "physical" two-dimensional gravity, and whose mathematical counterparts are Hermitian matrix integrals near critical points. It is relatively easy to establish that these matrix integrals satisfy the KdV integrable hierarchy. This implies that intersection theory on the moduli space of Riemann surfaces must be governed by the same integrable structure. Witten's conjecture was first proved by Kontsevich in [26], confirming in this way the physical equivalence which originated the conjecture.

4 Mirror symmetry

Perhaps the most striking application of string theory in geometry is *mirror symmetry*, which provides a remarkable equivalence between different topological manifolds of the Calabi–Yau type, and in particular makes possible to answer difficult problems in enumerative geometry. We will start by introducing the main character of the story, Calabi–Yau manifolds. For a detailed introduction, see for example [9, 21].

4.1 Calabi–Yau manifolds

Mathematically, a Calabi–Yau (CY) manifold X is a complex manifold which admits a Kähler metric with vanishing Ricci curvature. The condition of vanishing curvature is called the *Calabi–Yau condition*. Let us spell out in detail the ingredients in this definition. Since X is complex, it has complex coordinates (in local patches) that we will denote by

$$x^I, \quad x^{\overline{I}}, \qquad I = 1, \cdots, d, \tag{4.1}$$

where d is the complex dimension of X. X is also endowed with a Riemannian metric which is *Hermitian*, i.e. it only mixes holomorphic with antiholomorphic coordinates, and in a local patch it has the index structure $G_{I\overline{J}}$. We also need X to be *Kähler*. This means that the *Kähler form*

$$\omega = \mathrm{i}G_{I\overline{J}}\mathrm{d}x^I \wedge \mathrm{d}x^{\overline{J}} \tag{4.2}$$

which is a real two-form, is *closed*:

$$\mathrm{d}\omega = 0 \tag{4.3}$$

It is easy to check that for a Kähler manifold the Christoffel symbols do not have mixed indices, i.e. their only nonvanishing components are

$$\Gamma^I_{JK}, \qquad \Gamma^{\overline{I}}_{\overline{JK}}. \tag{4.4}$$

In addition, the Calabi–Yau condition requires the metric $G_{I\overline{J}}$ to be Ricci–flat:

$$R_{IJ} = 0. \tag{4.5}$$

It was conjectured by Calabi and then proved by Yau that, for compact Kähler manifolds, Ricci flatness is equivalent to a topological condition, namely that the first Chern class of the manifold vanishes

$$c_1(X) = 0. \tag{4.6}$$

One of the most important properties of Calabi-Yau manifolds (which can actually be taken as their defining feature) is that they have a holomorphic, nonvanishing section Ω of the canonical bundle $K_X = \Omega^{d,0}(X)$. This form is unique up to multiplication by a nonzero number, and in local coordinates it can be written as

$$\Omega = \Omega_{I_1 \cdots I_d} \mathrm{d}x^{I_1} \wedge \cdots \wedge \mathrm{d}x^{I_d}. \tag{4.7}$$

Since the section is nowhere vanishing, the canonical line bundle is trivial, and this is equivalent to (4.6).

For reasons that will become clear in a moment, the most common CY manifolds appearing in string theory have complex dimension $d = 3$, and they are called CY threefolds. Using Hodge theory, Poincaré duality and the CY condition, one can see that the Hodge diamond of a CY threefold X, which encodes the Hodge numbers $h^{p,q}$, $p, q = 1, 2, 3$, has the structure

$$\begin{array}{ccccccc} & & & 1 & & & \\ & & 0 & & 0 & & \\ & 0 & & h^{1,1} & & 0 & \\ 1 & & h^{2,1} & & h^{2,1} & & 1 \\ & 0 & & h^{1,1} & & 0 & \\ & & 0 & & 0 & & \\ & & & 1 & & & \end{array} \tag{4.8}$$

and therefore it only depends on two integers $h^{1,1}(X)$, $h^{2,1}(X)$ (here we assumed that X is simply-connected). It also follows that the Euler characteristic of X is

$$\chi(X) = 2(h^{1,1}(X) - h^{2,1}(X)). \tag{4.9}$$

Calabi–Yau manifolds typically come in families, in the sense that once a Calabi–Yau manifold has been obtained, one can deform it by changing its parameters without violating the Calabi–Yau condition. There are two types of parameters in a Calabi–Yau family: the *Kähler parameters* and the *complex deformation parameters*. The first ones specify sizes (i.e. areas of embedded holomorphic curves) while the second ones specify the complex structure, i.e. the choice of holomorphic/anti-holomorphic splitting. It turns our that the number of Kähler parameters (which are real) is $h^{1,1}$, while the number of complex deformation parameters (which are complex) is $h^{2,1}$.

Calabi–Yau manifolds first entered string theory in an important paper of Candelas, Horowitz, Strominger and Witten [7], who studied *compactifications* of string theory. The idea of compactification goes back to Kaluza and Klein, who used it to unify electromagnetism and gravitation by considering Einstein's equations on a five-dimensional manifold X_5 of the form

$$X_5 = S^1 \times X_{1,3}. \tag{4.10}$$

By developing the five-dimensional metric field in harmonics with respect to the circle direction, one obtains an infinite number of fields in four dimensions. The lowest harmonic gives the standard metric field of General Relativity on $X_{1,3}$, plus a gauge connection, both with zero mass. The rest of the harmonics give an infinite tower of fields, the so-called *Kaluza–Klein modes*, with masses of the form

$$m^2 = \frac{n^2}{R^2}, \qquad n \in \mathbb{Z}, \tag{4.11}$$

where R is the radius of the S^1 along the fifth dimension. If this radius is sufficiently small, the Kaluza–Klein modes will be very massive and not easy to detect in particle accelerators. Therefore, the extra dimension will remain invisible to a large extent.

As we mentioned before, a consistent superstring theory needs a target space X with ten space-time dimensions. In order to obtain a description of physics in Minkowski spacetime $M_{1,3}$ we can mimick Kaluza–Klein's idea and assume that the ten-dimensional space-time is of the form

$$X_{10} = X_6 \times M_{1,3}, \tag{4.12}$$

where X_6 is a "small" space of six real dimensions, i.e. the appropriate generalization of a circle of small radius to string theory. A spacetime of the above form in superstring theory is called a *compactification* to four dimensions. What are the acceptable choices for X_6? In [7] an answer to this question was proposed by requiring that the resulting physics in $M_{1,3}$ preserves some of the symmetries of the original superstring theory in ten dimensions (in particular, that it preserves a certain amount of supersymmetry). Compactification of superstring theory on a generic manifold will break all supersymmetries and lead to a non-supersymmetric theory in four dimensions. It turns out that the requirement of preserving supersymmetry has a precise geometric condition: one needs X_6 to be a manifold of *special holonomy*. The holonomy of a generic, oriented $2n$-dimensional manifold is orthogonal group $\mathrm{SO}(2n)$, but a manifold of special holonomy has a smaller subgroup thereof. A manifold X_6 with $SU(3)$ holonomy or smaller (instead of the generic $\mathrm{SO}(6)$) is a Calabi–Yau threefold.

The results of [7] indicated that Calabi–Yau manifolds could lead to "phenomenological" models of superstring theory, describing the particles and interactions of the real world. According to their model, many of the features of particle physics were determined by the geometric and topological properties of the Calabi–Yau manifold used for the compactification of the superstring. For example, one of the surprising results obtained in [7] was that the number of families of elementary particles arising in four dimensions was given by half the Euler characteristic of X (in absolute value). In spite of this suggestive developments, it is still not clear if string theory is relevant to particle physics. In the meantime the interest of the string theory community on Calabi–Yau manifolds unveiled one of the most beautiful topics in modern geometry: mirror symmetry.

4.2 Examples of Calabi–Yau manifolds

The simplest examples of CY manifolds are complex tori, $X = \mathbb{T}^{2d}$. In one complex dimension, these are elliptic curves, which can be written in Weierstrass form as

$$y^2 = 4x^3 - g_2 x - g_3. \tag{4.13}$$

Notice that as we change g_2, g_3, we change the shape of the elliptic curve. These are in fact simple instances of complex deformation parameters.

A less simple example of a Calabi–Yau manifold in two complex dimensions is the *K3 surface*. The K3 surface is the unique, simply-connected complex Kähler surface which has $c_1(X) = 0$ and is therefore CY. A concrete description is in terms of a quartic hypersurface in $\mathbb{P}^3$,

$$x_0^4 + x_1^4 + x_2^4 + x_3^4 = 0. \tag{4.14}$$

A simple calculation shows that indeed, when the degree of the defining polynomial is four, one has $c_1(X) = 0$.

Of course, in string theory we need in principle Calabi–Yau threefolds. Probably the most famous Calabi–Yau threefold is the *quintic Calabi–Yau*, which can be understood as the three-dimensional analogue of the hypersurface (4.14). It is defined as a quintic hypersurface in $\mathbb{P}^4$:

$$x_0^5 + x_1^5 + x_2^5 + x_3^5 + x_4^5 = 0. \tag{4.15}$$

It is easy to see that the quintic Calabi–Yau has the Hodge diamond

$$\begin{array}{ccccccc} & & & 1 & & & \\ & & 0 & & 0 & & \\ & 0 & & 1 & & 0 & \\ 1 & & 101 & & 101 & & 1 \\ & 0 & & 1 & & 0 & \\ & & 0 & & 0 & & \\ & & & 1 & & & \end{array} \tag{4.16}$$

hence Euler characteristic $\chi = -200$.

All these examples are *compact* Calabi–Yau manifolds. In the original program of [7] these are the physically relevant ones to understand particle physics in terms of superstring theory. But one can in principle consider Calabi–Yau manifolds which are not compact as a "local limit" of a true Calabi–Yau manifold. For example, if we have a rigid two-sphere embedded in a Calabi–Yau threefold, its normal bundle will be

$$\mathcal{O}(-1) \oplus \mathcal{O}(-1) \to \mathbb{P}^1. \tag{4.17}$$

Locally, near this two-sphere, the Calabi–Yau manifold will look like the total space of the bundle (4.17). This bundle is a non-compact Calabi–Yau manifold called the *resolved conifold*, and it is perhaps the simplest Calabi–Yau manifold. A general class of non–compact Calabi–Yau manifolds including the resolved conifold are the so-called *toric* manifolds. These are manifolds which contain an algebraic torus $\mathbb{T} \subset (\mathbb{C}^*)^r \subset X$ as an open set, and they admit an action of $\mathbb{T}$ which acts on this set by multiplication. Important examples of toric Calabi–Yau threefolds, besides (4.17), are the spaces

$$K_S \to S \tag{4.18}$$

where S is an algebraic surface, and K_S is its canonical bundle. If $S = \mathbb{P}^2$ the resulting Calabi–Yau

$$\mathcal{O}(-3) \to \mathbb{P}^2, \tag{4.19}$$

is also known as *local* $\mathbb{P}^2$.

4.3 Mirror manifolds

In order to understand string theories involving Calabi–Yau manifolds, it is useful to focus on the part of the target space of the string that involves the "internal" or compactification Calabi–Yau space only. It turns out that these restricted string theories can be described by a special type of conformal field theories, called $\mathcal{N} = 2$ super-conformal field theories, since they have $\mathcal{N} = 2$ supersymmetry in two dimensions (i.e. they have two different supersymmetries). This made possible to encode the data of the Calabi–Yau manifold in a purely algebraic way, in terms of data of the underlying CFT.

In 1987-1988 it was realized that the correspondence between the abstract CFT data and the CY geometry was not unique. This type of phenomenon has been known since the early days of string theory. If we have a string whose target space X is a circle of radius R, the underlying CFT is invariant under the exchange

$$R \leftrightarrow \frac{\ell_s^2}{R}. \tag{4.20}$$

This equivalence is called T-duality. It is a simple example of a string duality in which two string theories with two different target spaces (namely circles with radii R and ℓ_s^2/R) turn out to be equivalent. The phenomenon observed for Calabi–Yau manifolds is a far-fetched generalization of T-duality, and it works as follows.

The cohomology ring of the Calabi–Yau manifold can be realized, at the level of the CFT, in terms of a ring of operators called the *chiral ring* of the $\mathcal{N} = 2$ superconformal field theory [31]. But there are in fact *two* different chiral rings in this CFT, which lead to two possible Calabi–Yau Hodge diamonds related by the exchange $h^{1,1} \leftrightarrow h^{2,1}$:

$$\begin{array}{ccccccc} & & & 1 & & & \\ & & 0 & & 0 & & \\ & 0 & & h^{1,1} & & 0 & \\ 1 & & h^{2,1} & & h^{2,1} & & 1 \\ & 0 & & h^{1,1} & & 0 & \\ & & 0 & & 0 & & \\ & & & 1 & & & \end{array} \quad \leftrightarrow \quad \begin{array}{ccccccc} & & & 1 & & & \\ & & 0 & & 0 & & \\ & 0 & & h^{2,1} & & 0 & \\ 1 & & h^{1,1} & & h^{1,1} & & 1 \\ & 0 & & h^{2,1} & & 0 & \\ & & 0 & & 0 & & \\ & & & 1 & & & \end{array} \tag{4.21}$$

This suggests that, given a single $\mathcal{N} = 2$ CFT, there exist two Calabi–Yau manifolds, X and $\widetilde{X}$, such that

$$h^{1,1}(X) = h^{2,1}(\widetilde{X}), \qquad h^{2,1}(X) = h^{1,1}(\widetilde{X}), \tag{4.22}$$

hence

$$\chi(\widetilde{X}) = -\chi(X). \tag{4.23}$$

In other words, Calabi–Yau manifolds should come in *mirror pairs* called *mirror manifolds*. This symmetry relating two topologically different Calabi–Yau manifolds is the most basic manifestation of *mirror symmetry*, but as we will see it is not the only one.

The existence of mirror symmetry was confirmed by two types of evidence, "experimental" and "theoretical." On the experimental side, Candelas, Lynker and Schimmrigk showed in [8] that, in a large class of Calabi–Yau manifolds realized as polynomials in weighted projective spaces, most of the examples could be grouped in pairs with the property (4.22). Almost simultaneously, Greene and Plesser [17] used the underlying conformal field theory to construct explicit mirror pairs from first principles. For example, they constructed the mirror of the quintic Calabi–Yau manifold (4.15). This manifold must have $h^{1,1} = 101$ and $h^{2,1} = 1$, and it can be described by a uniparametric subfamily of the quintic hypersufaces in $\mathbb{P}^4$

$$\sum_{i=0}^{4} x_i^5 - \psi \prod_{i=0}^{4} x_i = 0. \tag{4.24}$$

In addition, one has to mod out by a symmetry group $\mathbb{Z}_5^3$, which has fixed points, and resolve the resulting singularities. Since $h^{2,1} = 1$, this family has a single complex deformation parameter, namely the coefficient ψ appearing in the algebraic equation (4.24). The work of Greene and Plesser has triggered a large literature on constructions of mirror manifolds, see for example [29] for a recent review.

4.4 Mirror symmetry as a string duality

The existence of mirror manifolds is a manifestation of a new type of duality in string theory, i.e. of an equivalence between two different string theories. One useful realization of this duality involves the so-called *type II superstring theories*. These are theories of closed strings in ten space-time dimensions based on an $\mathcal{N} = 1$ supersymmetric CFT in the worldsheet of the string. There are two different theories, depending on the type of space-time fermions that appear in the theory, and they are called *type IIA* and the *type IIB* theory. But from the mathematical point of view, the most useful realization of mirror symmetry is in terms of *topological string theory*, as first pointed out in [52].

As any other string theory, topological string theory is based on the coupling of a CFT to two-dimensional gravity. In the case of topological strings, the CFT turns out to be in addition a *topological field theory* (TFT) in two dimensions. Topological field theories were introduced by Witten in [50] and underlie many of the interactions between modern physics and mathematics. TFTs are more or less conventional quantum field theories, defined on a Riemannian manifold. In general, the correlation functions of a quantum field theory (i.e. the quantum averages of products of operators) depend on the background Riemannian metric, and the study of this dependence is the subject of quantum field theory in curved space. TFTs are constructed in such a way that there is no such a dependence. In other words, the correlation functions of a TFT only depend on the diffeomorphism class of the manifold, and in this sense they are topological invariants.

Quantum field theories where the fields are maps like (2.12) are called *sigma models*. There are two types of TFTs based on sigma models and where the target space X is a CY manifold. They are called the type A topological sigma model, and the type B topological sigma model. Both contain a basic bosonic field (2.12), but they differ in the type of Grassmannian fields that they incorporate in addition to x. Both models are topological in the sense that their correlation functions do not depend on the metric on the Riemann surface Σ_g, but they have additional invariance properties: it turns out that, when formulated on a family of Calabi–Yau manifolds, the B model depends only on the complex deformation parameters, while the A model depends only on the so-called *complexified* Kähler parameters. These are the parameters needed to specify the complexified Kähler form

$$J = \omega + \mathrm{i}B, \quad B \in H^{1,1}(X), \tag{4.25}$$

where B is sometimes called the B-field. The complexification of the Kähler form occurs naturally in the context of type II superstring theory, but it is also needed to implement mirror symmetry. There are now $h^{1,1}$ *complex* Kähler parameters for a Calabi–Yau manifold, and since mirror symmetry exchanges $h^{1,1}$ with $h^{1,2}$, this makes possible to exchange as well the complex deformation parameters of X with the complexified Kähler parameters of $\widetilde{X}$. In other words, mirror symmetry relates *families* of Calabi–Yau manifolds, and the parameters determining "complexified sizes" in one family become the parameters determining "complex shapes" (the complex deformation parameters) in the mirror family.

After coupling the type A and the type B topological sigma models to two-dimensional gravity one obtains two string theories which are called the type A and the type B topological string theories, respectively. These theories inherit the invariance properties of the underlying topological sigma models. In particular, quantities computed in the type A topological string depend only on the complexified Kähler parameters of the target CY manifold, while quantities computed in the type B topological string theory depend only on the complex deformation parameters. As a consequence, the partition functions of the type A and type B topological string theories turn out to be very different. We will focus on the contribution to the free energy (2.15) of Riemann surfaces of genus zero, F_0. Let us first consider the case of the A model with target space X, and let $\{S_a\}$ be a basis of $H_2(X, \mathbb{Z})$. The "complexified" sizes of the S_a,

$$t_a = \int_{S_a} J, \quad a = 1, \cdots, h^{1,1}(X) \tag{4.26}$$

provide a useful parametrization of the complexified Kähler moduli. Positivity of the metric requires that $\mathrm{Re}\, t_a > 0$. The genus zero free energy has the form

$$F_0^{\mathrm{A}}(X; t_a) = \frac{1}{6} \sum_{a,b,c} C_{abc} t_a t_b t_c + \sum_{Q \in H_2(X,\mathbb{Z})} N_{0,Q} \mathrm{e}^{-\int_Q J/\ell_s^2}. \tag{4.27}$$

In this expression, C_{abc} is the intersection number

$$C_{abc} = S_a \cap S_b \cap S_c, \tag{4.28}$$

and the sum over Q is over two-homology classes. The numbers $N_{0,Q}$ appearing in this expression are the *Gromov–Witten invariants* of the Calabi–Yau X for genus zero and for the class Q. These invariants are symplectic invariants of X and they have an enumerative interpretation as an appropriate counting of homolorphic curves of genus zero in the homology class Q. Due to the subtleties with its definition, the Gromov–Witten invariants are not positive integer numbers (as it would happen in a true counting problem) but rather rational numbers. Gromov–Witten invariants are the cornerstone of modern enumerative algebraic geometry, and they were formulated partially as a way to make mirror symmetry mathematically precise. Note that, if we write Q as

$$Q = \sum_a n_a S_a, \tag{4.29}$$

where n_a are integer numbers, we have that

$$\mathrm{e}^{-\int_Q J/\ell_s^2} = \mathrm{e}^{-\sum_a n_a t_a/\ell_s^2}. \tag{4.30}$$

The Gromov–Witten invariants $N_{0,Q}$ vanish for negative n_a and also when all n_a are zero.

The above expression for F_0^{A} has a very interesting property: it is given by a "classical" topological invariant (namely, the intersection number (4.28)), plus an infinite series of "corrections" involving Gromov–Witten invariants. As $\ell_s \to 0$, i.e. as the length of the string goes to zero size, these corrections vanish, as it follows from the positivity of the exponent in (4.30). Since ℓ_s plays the role of $\hbar$ in the the two-dimensional quantum field theory of strings, we can think about the sum involving Gromov–Witten invariants as a "quantum" correction to the "classical" result involving the intersection number. This suggests that the genus zero free energy of type A topological string theory on a CY manifold provides a "quantum" or "stringy" deformation of classical topology, depending on a continuous parameter ℓ_s. When this parameter goes to zero and string theory becomes a theory of point particles, one recovers the classical invariant.

The type A topological string free energy $F_0^{\rm A}$ involves very rich enumerative properties of the Calabi–Yau X, and it is not easy to calculate it. What about the same quantity in the type B topological string? It turns our that $F_0^{\rm B}(X; \tilde t_a)$, where $\tilde t_a$ are now complex deformation parameters of X, is relatively easy to compute. Let A_a and B^a be a symplectic basis of $H_3(X, \mathbb{Z})$, and let us consider the *periods* of the three-form Ω in (4.7). These are defined as

$$X_a = \int_{A_a} \Omega, \quad \mathcal{F}^a = \int_{B^a} \Omega, \qquad a = 0, 1, \cdots, h^{2,1}(X). \tag{4.31}$$

The theory of variations of complex structure tells us that the X_a are projective coordinates for the space of complex structures of X, and after quotienting by a non-zero coordinate, say X_0, one finds a parametrization of the space of complex deformations by the quantities

$$\tilde t_a = \frac{X_a}{X_0}, \qquad a = 1, \cdots, h^{2,1}(X). \tag{4.32}$$

The genus zero free energy $F_0^{\rm B}(\tilde t_a)$ is then determined by the equation

$$\frac{\partial F_0^{\rm B}}{\partial \tilde t_a} = \frac{1}{X_0} \mathcal{F}^a, \qquad a = 1, \cdots, h^{2,1}(X). \tag{4.33}$$

The calculation of the periods of Ω is a well-understood problem in the theory of variations of complex structures. In particular, this calculation amounts to solving an ordinary differential equation with regular singular points, the so-called *Picard–Fuchs equation.* This leads to an explicit expression for $F_0^{\rm B}(\tilde t_a)$.

We can now give a powerful formulation of mirror symmetry as a duality between the type A topological string and the type B topological string, in the spirit of Fig. 1. Let X, $\widetilde X$ be a mirror pair of Calabi–Yau manifolds. Then, mirror symmetry leads to the conjectural equality

$$F_0^{\rm A}(X; t_a) = F_0^{\rm B}(\widetilde X; \tilde t_a) \tag{4.34}$$

where the complexified Kähler parameters of X, t_a, have to be identified with the complex deformation parameters $\tilde t_a$ of $\widetilde X$,

$$t_a = \tilde t_a, \tag{4.35}$$

and $\tilde t_a$, defined in (4.32), is an appropriate quotient of periods. Since $\tilde t_a$ is in general a complicated function of the parameters appearing in the algebraic equation defining

$\tilde{X}$, the relation (4.35) provides a complicated map between these parameters and the complexified Kähler parameters of the mirror, and it is called the *mirror map*.

The first mirror pair for which the conjectural equality (4.34) was studied in detail is the quintic CY (4.15) and its mirror (4.24). This was done by Candelas, Green, Parkes and de la Ossa in their seminal paper [6]. In the case of the quintic, there is one single complexified Kähler parameter t, since $h^{1,1} = 1$, and correspondingly its mirror has one single complex deformation parameter ψ. It is more convenient to use

$$z = \psi^{-5}. \tag{4.36}$$

The periods of the mirror CY satisfy the Picard–Fuchs equation

$$\Big[\theta^4 - 5z(5\theta+1)(5\theta+2)(5\theta+3)(5\theta+4)\Big]\Pi = 0 \tag{4.37}$$

where Π is a generic period. As we discussed above, the mirror map involves the quotient of two appropriate periods, i.e. two appropriate solutions of the differential equation (4.37). These solutions are, in this case,

$$X_0 = \sum_{n=0}^{\infty} \frac{(5n)!}{(n!)^5} z^n, \qquad X_1 = -\log z\, X_0 - 5 \sum_{n=1}^{\infty} \frac{(5n)!}{(n!)^5} \Big(\sum_{j=n+1}^{5n} \frac{1}{j} \Big) z^n, \tag{4.38}$$

and the mirror map (4.35) can be written as

$$\begin{aligned} e^{-t} &= z \exp\left[\frac{5}{X_0} \sum_{n=1}^{\infty} \frac{(5n)!}{(n!)^5} \Big(\sum_{j=n+1}^{5n} \frac{1}{j} \Big) z^n \right] \\ &= z + 770\, z^2 + 1014275\, z^3 + 1703916750\, z^4 + 3286569025625\, z^5 + \cdots , \end{aligned} \tag{4.39}$$

To calculate (4.33), we need another period, i.e. another solution of the equation (4.37). The relevant solution is of the form

$$\mathcal{F}^1 = \frac{5}{2}(\log z)^2 + \frac{5}{2} \sum_{n=1}^{\infty} \frac{(5n)!}{(n!)^5} \left[25 \Big(\sum_{j=n+1}^{5n} \frac{1}{j} \Big)^2 - 25 \sum_{j=1}^{5n} \frac{1}{j^2} + 5 \sum_{j=1}^{n} \frac{1}{j^2} \right] z^n, \tag{4.40}$$

and by using these ingredients one finally obtains the genus zero free-energy

$$F_0^B(t) = \frac{5}{6} t^3 + 2875\, e^{-t} + \frac{4876875}{8} e^{-2t} + \frac{8564575000}{27} e^{-3t} + \cdots . \tag{4.41}$$

The first term is the "classical" part, involving the intersection number. The mirror symmetry conjecture (4.34) implies that the above expression must be equal to the genus zero free energy (4.27) of the quintic, and in particular predicts the following values of the Gromov–Witten invariants $N_{0,Q}$ (in this case Q is specified by a single positive integer called the degree d)

$$N_{0,1} = 2875, \qquad N_{0,2} = \frac{4876875}{8}, \qquad \cdots \tag{4.42}$$

The value for $d = 1$ has classical enumerative content, and it counts the number of lines in the quintic. Although the higher degree Gromov–Witten invariants are not integers, but rather rational numbers, Candelas *et al.* showed that one can extract integer invariants by taking into account multicoverings of holomorphic maps [6], and these integer invariants, also labelled by the degree, turn out to be related to the classical enumerative geometry of the quintic manifold. They can be regarded as counting the number of holomorphic curves of genus zero in the quintic CY.

One remarkable aspect of the calculation of Candelas *et al.* was that it made possible to obtain these integer invariants for all degrees at once, while the traditional methods of algebraic geometry were rather painful and in practice they could only deal with the first few degrees. Therefore, the predictions of mirror symmetry for the number of holomorphic curves on the quintic created quite a sensation in the mathematical community. The predictions for the first few Gromov–Witten invariants were immediately put to test with mathematical techniques, and after some initial confusion (see [16, 13] for the full story) they were confirmed. Further work by Kontsevich [27], Givental [15] and Lian–Liu–Yau [33] finally provided a full mathematical proof of (4.34) for many CY manifolds. As in the case of Witten's conjecture, a string duality leads to a very precise mathematical conjecture which is subsequently verified with standard mathematical tools, confirming again the original physical argument.

4.5 Further developments in mirror symmetry

Mirror symmetry and topological string theory are very rich subjects, with connections to many branches of mathematics, and they have been developed vigorously in various directions. We sketch here some of them.

First of all, the original calculation of Candelas et al. used mirror symmetry for genus zero only, but in the type A topological string it is perfectly natural to consider the enumeration of holomorphic curves with genus g. This leads to the generating functionals at genus g

$$F_g^{\mathrm{A}}(X; t_a) = \frac{(-1)^g \chi(X) |B_{2g} B_{2g-2}|}{4g(2g-2)(2g-2)!} + \sum_{Q \in H_2(X,\mathbb{Z})} N_{g,Q} \mathrm{e}^{-\int_Q J/\ell_s^2}, \qquad g \geq 2. \tag{4.43}$$

In this equation, $N_{g,Q}$ are the Gromov–Witten invariants at genus g and degree Q, "counting" holomorphic curves of genus g and degree Q in X, $\chi(X)$ is the Euler characteristic of X, and B_{2g} are the Bernoulli numbers. The first term in (4.43) corresponds to curves of degree zero, i.e. maps sending the whole of Σ_g to a point in X. A similar expression holds for $g = 1$.

One could ask what is the "mirror" calculation that has to be done in the type B model in order to obtain $F_g^{\rm B}(\tilde{X}; \tilde{t}_a)$. A crucial step in that direction was made in the work of Bershadsky, Cecotti, Ooguri and Vafa, [4], who found recursive equations for the $F_g^{\rm B}$ in the type B model. These equations are based on a non-holomorphic version of the genus g free energies, depending as well on the the complex conjugate $\tilde{t}_a^*$, and they determine the anti-holomorphic dependence of these in terms of lower genus data. The original $F_g^{\rm B}$ are then recovered in the limit

$$F_g^{\rm B}(\tilde{X}; \tilde{t}_a) = \lim_{\tilde{t}_a^* \to \infty} F_g^{\rm B}(\tilde{X}; \tilde{t}_a, \tilde{t}_a^*). \tag{4.44}$$

The anti-holomorphic dependence of these quantities is due to a quantum effect in the underlying physical theory (i.e. in the topological string) and for this reason the equations of [4] are called the *holomorphic anomaly equations*. They provide a useful way to determine the free energies in the B model. By mirror symmetry one expects that

$$F_g^{\rm B}(\tilde{X}; \tilde{t}_a) = F_g^{\rm A}(X; t_a) \tag{4.45}$$

and [3, 4] were able to calculate with these techniques the genus one and two topological string free energies for the quintic CY. This led to new predictions in enumerative geometry for the counting of higher genus curves which have been only verified very recently.

Unfortunately, the holomorphic anomaly equations do not fix the answer uniquely, since they are in fact differential equations for the non-holomorphic free energies, and one has to provide appropriate boundary conditions to solve the problem. In spite of this limitation, much of what is known about the enumerative geometry of higher genus curves on compact Calabi–Yau manifolds comes from solving these equations, see for example [22] for recent results on the quintic Calabi–Yau.

For *toric* Calabi–Yau manifolds, there is an algorithm which determines the $F_g^{\rm A}(X; t)$ uniquely, for all g. This algorithm is called the *topological vertex* and it was found in [2] by using a duality between string theory and gauge theories. We will consider this new duality in more detail in the next section.

Another extension of the original framework of mirror symmetry involves considering strings with boundaries. As we explained in section 2, boundary conditions in string theory are given by submanifolds in the target manifold where the endpoints of the string can end, and they are called D-branes. Since their discovery, D-branes have played a major rôle in string theory, and they have led to very rich mathematical developments. As many other structures in string theory, D-branes have counterparts in topological string theory called *topological D-branes*. The first description of topological D-branes was made by Witten in [53]. In this work Witten studied boundary conditions in topological string theory, and he found that the natural boundary conditions in the A model are given by Lagrangian submanifolds. These are submanifolds L with the property that the Kähler form ω vanishes when restricted to L. In contrast, in the B model the natural boundary conditions are given by holomorphic submanifolds. This suggests a natural extension of mirror symmetry to the case of strings with boundaries. First, one postulates a new mirror map between boundary conditions, i.e. a map between Lagrangian submanifolds L in the Calabi–Yau manifold X, and holomorphic submanifolds S in the mirror manifold $\widetilde{X}$. Second, one postulates an equivalence between the free energies of type A topological strings with boundaries on the Lagrangian manifold $L \subset X$, and the free energies of type B topological strings with boundaries on the holomorphic submanifold $S \subset \widetilde{X}$. This type of equivalence is called *open mirror symmetry* or *homological mirror symmetry*, and it has been conjectured by various people in different forms (see for example [28] for an early and extremely influential formulation). A more detailed formulation of open mirror symmetry shows that the appropriate mathematical notion of boundary condition involves a far-reaching generalization of Lagrangian and holomorphic submanifolds in terms of *categories*: in the A model, boundary conditions form the so-called Fukaya category of X, while in the B-model they form the derived category of coherent sheaves of $\widetilde{X}$. The original formulation of homological mirror symmetry due to Kontsevich is then stated as an equivalence of these two categories.

Open mirror symmetry makes possible to obtain enumerative information about holomorphic maps with boundaries on a Lagrangian submanifold. Indeed, on the type A side, the free energies of topological strings with boundaries involve *open Gromov–Witten invariants*, which "count" holomorphic maps with Lagrangian boundary conditions. In the toric case this was first developed in [1], and in the compact case detailed results were obtained in [46]. For example, in the case of the quintic Calabi–Yau manifold, a natural Lagrangian submanifold L is the so-called real quintic. In the description of the quintic given in (4.15), the real quintic is the fixed point locus of the real involution

$$x_i \to \bar{x}_i. \tag{4.46}$$

As in the original work of Candelas et al., one can obtain the number of holomorphic disks ending on this Lagrangian submanifold by doing a relatively simple computation in the B model, see [46] for further details.

Another research direction closely related to mirror symmetry and inspired by physics involves the reformulation of Gromov–Witten invariants in terms of different invariants with enumerative content. The first such reformulation was due to Gopakumar and Vafa [18], who used a string duality to repackage the information encoded in the F_g of the type A model in terms of a new set of invariants, the so-called *Gopakumar–Vafa invariants*. These invariants, in contrast to Gromov–Witten invariants, are integer numbers, and they can be interpreted physically as counting states in a supersymmetric system. A mathematical formulation of these invariants has been proposed in [44]. Another set of invariants conjecturally equivalent to Gromov–Witten invariants are the *Donaldson–Thomas invariants*, which are based on counting sheaves [12, 40].

5 String theory and gauge theory

Tthe mathematical conjectures that we have discussed are the counterparts of equivalences, or dualities, between two string theories. However, in order to obtain Witten's conjecture for the intersection numbers of Deligne–Mumford moduli space, it was necessary to describe the minimal models coupled to gravity in terms of discretized Riemann surfaces. These were in turn described by a much simpler system, namely matrix models, in the limit in which the rank of the matrix N was very large. Since matrix models can be regarded as quantum gauge theories in zero dimensions, the description of non-critical strings in terms of double-scaled matrix models can be seen as an equivalence between a gauge theory and a string theory. The idea that gauge theories might be dual to string theories was suggested by 't Hooft in 1972 [47], but it was not implemented in more complicated examples until 1997. That year, Juan Maldacena proposed [36] that an interacting quantum gauge theory in four dimensions, Yang–Mills theory with $\mathcal{N} = 4$ supersymmetry, was equivalent to type IIB superstring theory, on the ten-dimensional space $S^5 \times \mathrm{AdS}_5$, where AdS_5 is the Anti-de Sitter space in five dimensions and S^5 is a five-dimensional sphere.

The proposal of Maldacena has become one of the richest sub-fields of string theory, and has led to an enormous amount of activity, with applications in black hole physics, nuclear physics, and condensed matter physics, among other applications. Moreover, his work inspired an equivalence between type A topological string theory on the resolved conifold and a quantum gauge theory called Chern–Simons theory. This equivalence was postulated by Gopakumar and Vafa [19] and it has rather surprising mathematical implications, since it relates two *a priori* completely unrelated mathematical objects: on one hand Gromov–Witten invariants, and on the other hand quantum invariants of knots and three-manifolds. In order to explain the conjecture of Gopakumar and Vafa, we first sketch some of the ideas underlying string/gauge theory dualities.

5.1 The string/gauge theory correspondence

In a gauge theory, the basic field is a gauge connection, i.e. a one-form $A_\mu(x)$ which takes values in the Lie algebra of a Lie group. In the case of an $SU(N)$ gauge group, the gauge fields are Hermitian matrices. The partition function of the gauge theory is formally defined as the path integral

$$Z = \int \mathcal{D}A \, \mathrm{e}^{-\frac{1}{\lambda}S(A)}, \qquad (5.1)$$

where $S(A)$ is the classical action describing the dynamics of the gauge connection. In the standard Yang–Mills theory, the action is given by

$$S(A) = \int \mathrm{d}^4x \, \mathrm{Tr}\Big(F_{\mu\nu}F^{\mu\nu}\Big) \qquad (5.2)$$

where

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu] \tag{5.3}$$

is the curvature of the gauge connection. Although (5.1) is not well-defined mathematically, it can be defined as a formal power series in λ by performing an asymptotic expansion around the trivial connection $A_\mu = 0$. In terms of the free energy we have the series

$$F = \log Z = \sum_{n \geq 0} f_n(N) \lambda^n \tag{5.4}$$

where $f_n(N)$ are functions of N (here we are simplifying the story very much, since this formal series, in order to be well-defined, requires a procedure called "renormalization"). Each of these functions can be represented as the sum of a set of diagrams, the famous Feynman diagrams of quantum field theory. 't Hooft noticed in [47] that the expansion (5.4) can be put in the form (3.12), or equivalently

$$F = \sum_{g \geq 0} F_g(t) \lambda^{2g-2}. \tag{5.5}$$

Here, t stands for the combination

$$t = \lambda N \tag{5.6}$$

which is the so-called *'t Hooft parameter*. Note that the expansion in (5.5) is in fact an expansion in inverse powers of N, since $\lambda = t/N$. The evaluation of free energies and correlation functions in a $SU(N)$ gauge theory by using an expansion in inverse powers of N is called the $1/N$ *expansion*.

As in the case of matrix models, the $F_g(t)$ can be interpreted as a sum over "fatgraphs" of genus g, and in fact it is easy to see that the structure (5.5) will be present in any theory with a gauge symmetry in which the fundamental fields are Hermitian matrices. 't Hooft noticed that the expansion in (5.5), where we sum over fatgraphs of genus g, reminds very much the structure of an amplitude in string theory, where we sum over Riemann surfaces of genus g. He then suggested that the quantities $F_g(t)$ that one obtains from a gauge theory in the $1/N$ expansion might be interpreted as the free energies at genus g of an appropriate string theory. This conjectural correspondence or duality is called the *string/gauge theory correspondence*. The original motivation of 't Hooft was to obtain some insight on the behavior of Yang–Mills theory at strong coupling. Many aspects of quantum Yang–Mills theory, like confinement, the existence of a mass gap in the spectrum, etc., are not well understood. 't Hooft hoped that when $N \to \infty$ there could be simplifications which made possible an analytic description of these phenomena, and in particular a description in terms of a string theory.

From this perspective, the fundamental question is: what is the string theory that corresponds to a given gauge theory? In most of the cases (including ordinary Yang–Mills theory), we do not have an answer, even a conjectural one, and this is why Maldacena's proposal was so important. Maldacena provided a precise, conjectural description of the string theory describing $\mathcal{N} = 4$ supersymmetric Yang–Mills theory. This gauge theory involves a gauge connection A_μ whose dynamics is described by the action (5.2), but it contains other fields in order to implement supersymmetry. The string theory is the type IIB theory that we mentioned before, with a ten-dimensional target space X given by the product of a five-sphere S^5 and Anti-de Sitter space in five dimensions AdS_5. This conjecture leads to many concrete relations between quantities calculated in the gauge theory in the $1/N$ expansion, and quantities calculated in the string theory. These relations have been tested with increasing precision in a variety of circumstances, and the conjecture is widely believed to be true.

As we mentioned before, Maldacena's conjecture has many implications for various branches of physics. Maybe its most fundamental consequence is that the quantum gravitational effects in the type IIB superstring theory side can be reliably described in terms of a gauge theory. This is a highly nontrivial statement, since quantum gravity has long been suspected to be plagued with fundamental problems, and it has been even suggested that gravity is incompatible with quantum mechanics. According to Maldacena, this is not necessarily the case, since according to his conjecture quantum gravity effects in the type IIB superstring theory on $S^5 \times \mathrm{AdS}_5$ are *a priori* described by a well-understood quantum theory, namely a quantum gauge theory.

Since Maldacena first formulated his conjecture, other string/gauge theory duals have been found by extending his ideas. One of them is of outmost relevance for mathematics: the Gopakumar–Vafa conjecture, which we now present. For a detailed exposition of this conjecture, its context and its applications, see [37].

5.2 The Gopakumar–Vafa conjecture

In order to find a mathematical application of the string/gauge theory correspondence, one should look for gauge theories which lead to interesting mathematical counterparts. One very important example is Chern–Simons gauge theory, introduced by Witten in [51]. This gauge theory can be defined on any three-dimensional manifold M, and its action is much simpler than the usual Yang–Mills action. It can be written compactly in the language of differential forms as

$$S(A) = A \wedge \mathrm{d}A + \frac{2}{3} A \wedge A \wedge A, \tag{5.7}$$

where A is a connection on M. The free energy of this theory, when expanded around the trivial connection, has the form (5.4), and as suggested by Witten in [51] it should give an invariant of the three-manifold M. In fact, it is widely believed that the resulting formal power series can be defined in a mathematically rigorous way by using the so-called LMO invariant [35].

If the gauge group defining the theory is $SU(N)$, the analysis of 't Hooft applies, and the free energy can be written in the form (5.5). In the case of $M = S^3$, the functions

$F_g(t)$ can be written down very explicitly by using the exact solution of Witten [51]. For $g \geq 2$ they have the form,

$$F_g(t) = \frac{(-1)^g |B_{2g} B_{2g-2}|}{2g(2g-2)(2g-2)!} + \frac{|B_{2g}|}{2g(2g-2)!} \mathrm{Li}_{3-2g}(\mathrm{e}^{-t}). \tag{5.8}$$

The function Li_j appearing in this equation is the polylogarithm of index j, defined by

$$\mathrm{Li}_j(x) = \sum_{n=1}^{\infty} \frac{x^n}{n^j}. \tag{5.9}$$

Is there a string theory description of the expansion (5.5) for $SU(N)$ Chern–Simons theory on S^3? Surprisingly, the $F_g(t)$ in (5.8) are precisely the genus g generating functionals (4.43) of the type A topological string when X is the resolved conifold (4.17). The 't Hooft parameter of the gauge theory has to be identified with the complexified Kähler parameter of topological string theory. This was noticed by Gopakumar and Vafa in [19], and led them to conjecture that $SU(N)$ Chern–Simons gauge theory on S^3 is equivalent to type A topological string theory on the resolved conifold, in the sense of the string/gauge theory correspondence. This equivalence implies in particular that the $F_g(t)$ which appear in the $1/N$ expansion of the $SU(N)$ Chern–Simons gauge theory are indeed generating functionals of Gromov–Witten invariants of the resolved conifold at genus g. But the equivalence has many more implications.

One of the original motivations of Witten in [51] to study Chern–Simons gauge theory was to understand the Jones polynomial of knots and links introduced in [24]. The Jones polynomial associates to any link a polynomial in one variable q which can be used to distinguish links: if two links are topologically inequivalent, their Jones polynomials are different. Witten found that there was a natural way to produce knot and link invariants in Chern–Simons theory. In any gauge theory, given a closed line $\mathcal{K}$ and a representation R of the gauge group, it is possible to construct an operator called a *Wilson loop* and given by

$$W_{\mathcal{K},R}(A) = \mathrm{Tr}_R \Big(\mathrm{Hol}_{\mathcal{K}}(A) \Big), \tag{5.10}$$

where we first calculate the holonomy of the gauge connection around $\mathcal{K}$ and then take the trace in the representation R. In Chern–Simons theory we can take $\mathcal{K}$ to be an oriented knot in S^3, and in this way we can associate a quantum operator to any knot and any representation of the group. The (normalized) quantum average of a Wilson loop operator is given by

$$\langle W_{\mathcal{K},R}(A) \rangle = \frac{1}{Z} \int \mathcal{D}A \, \mathrm{e}^{-\frac{1}{\lambda} S(A)} \, W_{\mathcal{K},R}(A). \tag{5.11}$$

Given a link $\mathcal{L}$ whose components are the knots $\mathcal{K}_1, \cdots \mathcal{K}_L$, one can form an appropriate operator by simply taking the product

$$W_{\mathcal{L},(R_1,\cdots,R_L)}(A) = \prod_{i=1}^{L} W_{R_i,\mathcal{K}_i}(A). \tag{5.12}$$

In [51], Witten showed that the quantum average

$$W_{(R_1,\cdots,R_L)}(\mathcal{L}) = \langle W_{\mathcal{L},(R_1,\cdots,R_L)}(A)\rangle \tag{5.13}$$

is a topological invariant of the link $\mathcal{L}$. Moreover, when the gauge group is $SU(2)$, and the representations R_i are the fundamental representation, this invariant is nothing but the Jones polynomial of the link. In this framework the variable q appearing in Jones' construction has to be identified with

$$q = \exp\Big(\frac{2\pi \mathrm{i}}{\lambda}\Big). \tag{5.14}$$

This surprising result gave a completely new perspective on the Jones polynomial, and suggested immediately that a whole family of invariants of links in S^3 could be found just by changing the gauge group and/or the representation. Some of these invariants had been discovered before Witten found a Chern–Simons interpretation: the HOMFLY and the Kauffman polynomial correspond to the cases where the gauge groups are $G = SU(N)$ and $SO(N)$, respectively, and the representations are the fundamental one. Other choices of representation for these groups correspond to the so-called *colored* HOMFLY and Kauffman polynomials. All these invariants, for any choice of group and representation, can also be constructed from the theory of quantum groups. Witten's point of view made clear that one should be able to define these knot invariants for any ambient three-manifold, and this led to many new insights on knot theory and three–manifold topology. The invariants related to Chern–Simons theory have evolved into a very rich sub-field of topology which goes sometimes under the name of "quantum topology" (see [42] for an exposition).

The Gopakumar–Vafa conjecture also gives a string-theory interpretation of these knot and link invariants when the gauge group is a classical gauge group [43, 30, 38]. In general, in the framewok of the string/gauge theory correspondence, including Wilson loops in the gauge theory side corresponds to introducing boundary conditions in the string theory side. Roughly speaking, the Wilson loop acts as a boundary where strings can end. Since boundary conditions in type A topological string theory are given by Lagrangian submanifolds, Ooguri and Vafa postulated in [43] that one can associate a Lagrangian submanifold $L_{\mathcal{L}}$ in the resolved conifold to any link $\mathcal{L}$ in S^3:

$$\mathcal{L} \to L_{\mathcal{L}}. \tag{5.15}$$

This correspondence is illustrated in Fig. 7 for a knot. The link invariant $W_{(R_1,\cdots,R_L)}(\mathcal{L})$ is then interpreted as a generating functional of appropriate open Gromov–Witten invariants for type A topological string theory on X with boundary conditions provided by $L_{\mathcal{L}}$.

The conjecture of Gopakumar–Vafa and its extension to link invariants produced again an unsuspected relationship between two different mathematical theories: on one hand the theory of knot, link and three-manifold invariants coming from Chern–Simons theory or quantum groups, and on the other hand the theory of open and closed Gromov–Witten invariants of non-compact Calabi–Yau threefolds. This connection is at the origin

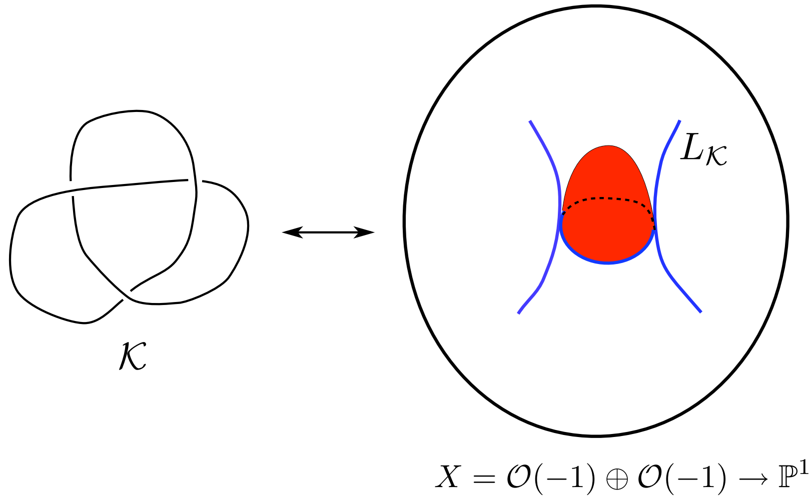

Figure 7: Given a knot $\mathcal{K}$, Ooguri and Vafa conjecture that there exists a Lagrangian submanifold $L_{\mathcal{K}}$ in the resolved conifold, providing boundary conditions for open strings.

of many remarkable conjectures, some of which have been rigorously proved. One of these conjectures is the *topological vertex* formalism, which makes possible to calculate all the generating functionals of Gromov–Witten invariants $F_g(t)$, for any toric Calabi–Yau threefold, in terms of simple Chern–Simons invariants of knots and links, and in a purely combinatorial way. This formalism, inspired by the Gopakumar–Vafa conjecture, was proposed in [2], and it has been justified rigorously in [32, 41]. A simpler version of the topological vertex formalism makes in fact possible to express all intersection numbers on the Deligne–Mumford moduli space in terms of Chern–Simons invariants of the simplest knot, namely the unknot or trivial knot [39]. Another set of conjectures [43, 30, 38] use properties of open Gromov–Witten invariants to deduce properties of the colored HOMFLY and Kauffman polynomials of links, and a proof of some of them has been recently announced [34].

6 Conclusion: a hybrid discipline

String theory is a sophisticated theoretical construction which might provide a deep understanding of many aspects of fundamental physics. It has been conjectured to enjoy some remarkable properties, like the existence of "dualities", i.e. nontrivial equivalences between different string theory models. On the other hand, the construction of string theory involves crucial ingredients of modern geometry and topology: the moduli space of Riemann surfaces, Calabi–Yau manifolds, and even knot invariants. The combination of the mathematical content of the theory with the existence of dualities has led to a continuous flow of surprising mathematical conjectures which have changed the shape of all the areas of mathematics involved in this relation. In this article, we have tried to convey some aspects of these fruitful interactions by focusing on three important examples: Witten's conjecture on intersection theory in Deligne–Mumford moduli space, mirror symmetry, and the Gopakumar–Vafa conjecture.

Although many of the conjectures formulated by string theory have been rigorously proved by mathematicians, it is still a mystery why string theory is so effective at finding them. As a matter of fact, the foundations of string theory are poorly understood, in spite of almost thirty years of intensive research. Even though we have accumulated evidence that string theory enjoys duality properties, we don't have a formulation of the theory where these dualities are manifest. Further progress in understanding these foundational issues would be immensely valuable, both for string theory itself and for its mathematical implications.

Finally, it is important to mention that the interface between mathematics and string theory has challenged the way in which physicists and mathematicians understand their own disciplines. According to Peter Galison, this interface produced a new type of theorist, part mathematician and part physicist. It also produced a new type of conceptual objects which were "not exactly physical entities and yet not quite (or not yet) fully mathematical objects," and finally it led to "a style of demonstration that did not conform either to older forms of physical argumentation familiar to particle physicists or to canonical proofs recognizable to 'pure' mathematicians" ([13], p. 24). As we have mentioned, rigorous mathematical proofs of some of the conjectures stemming from string theory have provided strong validations of string dualities. Traditionally, physics statements are validated by the confrontation with experiment, and this new relationship, in which mathematics plays the role of experimental physics, has been a source of worry for some physicists, who fear that string theory might lead to a retreat from reality. At the same time, the style of reasoning that leads string theorists to successful mathematical conjectures does not conform to the standards of the mathematical community, who relies heavily on rigor and proof. Therefore, string theorists working at the interface of string theory and mathematics have always triggered among mathematicians a "conflicted reaction," as Peter Galison puts it: "a response mixing enormous enthusiasm with grave reservations about the loss of rigor" ([13], p. 25). The anxieties of the mathematical community were made explicit in an article by Jaffe and Quinn [23] which tried to reaffirm the classical boundaries of the disciplines and suggested the name "theoretical mathematics" for the new hybrid – "a dignified name for the activity that would nonetheless isolate it from the mainline of rigorous mathematics," as Galison pointed out ([13], p. 54)[2].

As a conclusion, we can say that the interaction between string theory and mathematics has been immensely productive for both disciplines and has opened a new type of relationship between physics and mathematics. But the hybrid science which has emerged

[2]In a more joyful and less normative article [55] E. Zaslow has proposed to call it "physmatics."

from these interactions is also fragile. Its development has been fueled to a large extent by the progress of string theory itself, and it is not clear that this progress will continue in the future at the same rate that we have witnessed in the last twenty years. After all, the fate of string theory is very much linked to its ability to make contact with the real world, and if this contact is further delayed, string theory might become what it used to be in the late 1970's: a small sub-discipline with very few practitioners. The other source of fragility is rather sociological, and it has to do with the uncomfortable position of this hybrid discipline at the boundary of established traditions. In spite of all the talk about interdisciplinarity, the researcher working in the mathematics of string theory will be probably regarded as too mathematical by traditional particle physicists and as not rigorous enough by mathematicians and traditional mathematical physicists. As a result, such a researcher will have problems finding an ecological niche in the academic system. We can only hope that the sheer beauty and wonder of the mathematics of string theory will make it flourish in spite of all these difficulties.

Acknowledgments

I would like to thank all the people who have taught me about mathematics and string theory in the last fifteen years. C. Bartocci read a preliminary version of this text and made many useful suggestions, and I would like to thank him as well.

References

[1] M. Aganagic and C. Vafa, "Mirror symmetry, D-branes and counting holomorphic discs," hep-th/0012041.

[2] M. Aganagic, A. Klemm, M. Mariño and C. Vafa, "The topological vertex," Commun. Math. Phys. **254**, 425 (2005) [arXiv:hep-th/0305132].

[3] M. Bershadsky, S. Cecotti, H. Ooguri and C. Vafa, "Holomorphic anomalies in topological field theories," Nucl. Phys. B **405**, 279 (1993) [arXiv:hep-th/9302103].

[4] M. Bershadsky, S. Cecotti, H. Ooguri and C. Vafa, "Kodaira-Spencer theory of gravity and exact results for quantum string amplitudes," Commun. Math. Phys. **165**, 311 (1994) [arXiv:hep-th/9309140].

[5] A. A. Belavin, A. M. Polyakov and A. B. Zamolodchikov, "Infinite conformal symmetry in two-dimensional quantum field theory," Nucl. Phys. B **241**, 333 (1984).

[6] P. Candelas, X. C. De La Ossa, P. S. Green and L. Parkes, "A pair of Calabi-Yau manifolds as an exactly soluble superconformal theory," Nucl. Phys. B **359**, 21 (1991).

[7] P. Candelas, G. T. Horowitz, A. Strominger and E. Witten, "Vacuum Configurations For Superstrings," Nucl. Phys. B **258**, 46 (1985).

[8] P. Candelas, M. Lynker and R. Schimmrigk, "Calabi-Yau Manifolds in Weighted P(4)," Nucl. Phys. B **341**, 383 (1990).

[9] D. Cox and S. Katz, *Mirror symmetry and algebraic geometry*, American Mathematical Society, Providence, 1999.

[10] P. Deligne *et al.*, *Quantum fields and strings: A course for mathematicians*, American Mathematical Society, Providence, 1999.

[11] P. Di Francesco, P. Ginsparg and J. Zinn-Justin, "2-D Gravity and random matrices," Phys. Rept. **254**, 1 (1995) [arXiv:hep-th/9306153].

[12] S. K. Donaldson and R. P. Thomas, "Gauge theory in higher dimensions," in *The geometric universe*, p. 31, Oxford Univ. Press, Oxford, 1998.

[13] P. Galison, "Mirror symmetry: persons, values, and objects," in M. Norton Wise (ed.), *Growing explanations. Historical perspectives on recent science*, p. 23, Duke University Press, Durham, 2004.

[14] T. Gannon, *Moonshine beyond the Monster*, Cambrige University Press, Cambridge, 2006.

[15] A. Givental, "Equivariant Gromov-Witten invariants," Int. Math. Res. Notices **13** (1996), 613.

[16] B. R. Greene, *The elegant universe*, W.W. Norton, New York, 2003.

[17] B. R. Greene and M. R. Plesser, "Duality in Calabi-Yau moduli space," Nucl. Phys. B **338**, 15 (1990).

[18] R. Gopakumar and C. Vafa, "M-theory and topological strings. I, II," hep-th/9809187, hep-th/9812127.

[19] R. Gopakumar and C. Vafa, "On the gauge theory/geometry correspondence," hep-th/9811131, Adv. Theor. Math. Phys. **3** (1999) 1415.

[20] J. Harris and I. Morrison, *Moduli of curves*, Springer-Verlag, New York, 1998.

[21] K. Hori *et al.*, *Mirror symmetry*, American Mathematical Society, Providence, 2003.

[22] M. x. Huang, A. Klemm and S. Quackenbush, "Topological String Theory on Compact Calabi-Yau: Modularity and Boundary Conditions," Lect. Notes Phys. **757**, 45 (2009) [arXiv:hep-th/0612125].

[23] A. Jaffe and F. Quinn, "'Theoretical mathematics': towards a cultural synthesis of mathematics and theoretical physics," Bull. Am. Math. Soc. **29**, 3 (1993).

[24] V. F. R. Jones, "A polynomial invariant of knots and links via von Neumann algebras," Bull. Am. Math. Soc. **12**, 103 (1985).

[25] V. A. Kazakov, "Ising model on a dynamical planar random lattice: Exact solution," Phys. Lett. A **119**, 140 (1986); "The appearance of matter fields from quantum fluctuations of 2d gravity," Mod. Phys. Lett. A **4**, 2125 (1989).

[26] M. Kontsevich, "Intersection theory on the moduli space of curves and the matrix Airy function," Commun. Math. Phys. **147**, 1 (1992).

[27] M. Kontsevich, "Enumeration Of Rational Curves Via Torus Actions," in R. Dijkgraaf, C. Faber, and G. van der Geer (eds.), *The moduli space of curves*, p. 335, Birkhäuser, Boston, 1995 [arXiv:hep-th/9405035].

[28] M. Kontsevich, "Homological algebra of mirror symmetry," Proceedings of the International Congress of Mathematicians, Vol. 1, p. 120, Birkhäuser, Basel, 1995.

[29] M. Kreuzer, "Toric Geometry and Calabi-Yau Compactifications," arXiv:hep-th/0612307.

[30] J. M. F. Labastida, M. Mariño and C. Vafa, "Knots, links and branes at large N," JHEP **0011**, 007 (2000) [arXiv:hep-th/0010102].

[31] W. Lerche, C. Vafa and N. P. Warner, "Chiral Rings in N=2 Superconformal Theories," Nucl. Phys. B **324**, 427 (1989).

[32] J. Li, C.-C. M. Liu, K. Liu and J. Zhou,"A mathematical theory of the topological vertex," math.AG/0408426.

[33] B. Lian, K. Liu, and S.-T. Yau, "Mirror principle I," Asian J. Math. 4, 729 (1997).

[34] K. Liu and P. Peng, "Proof of the Labastida–Mariño–Ooguri–Vafa conjecture," math.QA/0704.1526.

[35] T.T.Q. Le, J. Murakami and T. Ohtsuki, "On a universal perturbative invariant of 3-manifolds," Topology **37**, 539 (1998).

[36] J. M. Maldacena, "The large N limit of superconformal field theories and supergravity," Adv. Theor. Math. Phys. **2**, 231 (1998) [Int. J. Theor. Phys. **38**, 1113 (1999)] [arXiv:hep-th/9711200].

[37] M. Mariño, *Chern–Simons theory, matrix models, and topological strings*, Oxford University Press, Oxford, 2005.

[38] M. Mariño, "String theory and the Kauffman polynomial," arXiv:0904.1088 [hep-th].

[39] M. Mariño and C. Vafa, "Framed knots at large N," Contemp. Math. **310**, 185 (2002) [arXiv:hep-th/0108064].

[40] D. Maulik, N. Nekrasov, A. Okounkov and R. Pandharipande, "Gromov-Witten theory and Donaldson-Thomas theory. I and II," Compos. Math. **142** (2006), no. 5, 1263 and 1286.

[41] D. Maulik, A. Oblomkov, A. Okounkov and R. Pandharipande, "Gromov-Witten/Donaldson-Thomas correspondence for toric 3-folds," math.AG/0809.3976.

[42] T. Ohtsuki, *Quantum invariants*, World Scientific, Singapore, 2002.

[43] H. Ooguri and C. Vafa, "Knot invariants and topological strings," Nucl. Phys. B **577**, 419 (2000) [arXiv:hep-th/9912123].

[44] R. Pandharipande and R.P. Thomas, "Curve counting via stable pairs in the derived category," math.AG/0707.2348.

[45] J. Polchinski, *Superstring theory*, vols. 1 and 2, Cambridge University Press, Cambridge, 1998.

[46] J. Walcher, "Opening mirror symmetry on the quintic," Commun. Math. Phys. **276**, 671 (2007) [arXiv:hep-th/0605162].

[47] G. 't Hooft, "A planar diagram theory for strong interactions," Nucl. Phys. B **72**, 461 (1974).

[48] E. Witten, "On The Structure Of The Topological Phase Of Two-Dimensional Gravity," Nucl. Phys. B **340**, 281 (1990).

[49] E. Witten, "Two-dimensional gravity and intersection theory on moduli space," Surveys Diff. Geom. **1**, 243 (1991).

[50] E. Witten, "Topological Quantum Field Theory," Commun. Math. Phys. **117**, 353 (1988).

[51] E. Witten, "Quantum field theory and the Jones polynomial," Commun. Math. Phys. **121**, 351 (1989).

[52] E. Witten, "Mirror manifolds and topological field theory," arXiv:hep-th/9112056.

[53] E. Witten, "Chern-Simons Gauge Theory As A String Theory," Prog. Math. **133**, 637 (1995) [arXiv:hep-th/9207094].

[54] E. Witten, “Michael Atiyah and the physics/geometry interface,” Asian J. Math. **3** (1999) lxi.

[55] E. Zaslow, “Physmatics,” physics/0506153.