The Interpretation of String Dualities

Dean Rickles

Unit for HPS, University of Sydney

Abstract

Many of the advances in string theory have been generated by the discovery of new duality symmetries connecting what were once thought to be distinct theories, solutions, processes, backgrounds, and more. Indeed, duality has played an enormously important role in the creation and development of numerous theories in physics and numerous fields of mathematics. Dualities often lie at those fruitful intersections at which mathematics and physics are especially strongly intertwined. In this paper I describe some of these dualities and unpack some of their philosophical consequences, focusing primarily on string-theoretic dualities. I argue that dualities fall uncomfortably between symmetries and gauge redundancies, but that they differ in that they point to genuinely new deeper structures.

Key words: Duality, String theory, Mirror symmetry, M-Theory, Underdetermination
PACS: 04.50.-h, 11.25.Mj, 11.25.-w, 01.70.+w, 02.40.-k, 11.30.Ly, 11.25.Hf, 11.15.Me

There's more to physics than nonrelativistic quantum mechanics.
Robert Weingard

1 Introduction and Motivation

Philosophers of physics are by now well-steeped in symmetries and are well aware both of their importance for the development of physical theory and of their many conceptual implications—see for example the wide-ranging essays collected in [4]. Usually, the symmetries discussed relate states, observables,

Email address: d.rickles@usyd.edu.au (Dean Rickles).

[1] I would like to thank Elena Castellani for suggesting the idea of a special issue on dualities. I would also like to thank the Australian Research Council for the award of an Australian Research Fellowship which enabled me to complete this work.

or histories (more generally, solutions) of *one and the same theory* (or some physical system described by that theory). For example, one might consider rotated and unrotated states of an individual experiment; boosted and unboosted states of a moving object; permuted and non-permuted configurations of particles; or a critical or fractal (i.e. self-similar) system viewed at a variety of scales. Such symmetries also tell us about what gets *conserved* in various processes (via Noether's theorem), and as such are crucially important in the discovery, testing, and application of laws.

More recently, philosophers have become interested in a class of transformations called "gauge symmetries," or, more accurately, gauge *redundancies.* [2] Here the orbits of the gauge symmetry group are generated by first class constraints, and the transformations thus generated are taken to be *unphysical*: unlike symmetries, gauge transformations do not map distinct physical states to one another. One might consider here transformations that amount to a reshuffling of the (indistinguishable) points of a manifold (i.e. an active diffeomorphism), or, more obviously, a multiplication of a vector potential by an arbitrary gradient in classical electromagnetism. While symmetries can result in physically distinguishable scenarios, gauge redundancies (and the preservation of invariance with respect to them) result in no physically observable differences, and so are generally (i.e. in the physics literature) removed by a quotienting procedure leaving one with a space of orbits of the gauge group whose elements are constants of the original gauge motion. Again, these gauge redundancies can tell us much about the laws (the form of interactions), and are a crucial component in the development of modern physics. They lie at the root of some philosophically interesting pieces of modern physics, including the hole argument and the problem of time in quantum gravity.

However, there is a less-well known (to philosophers of physics at least) family of symmetry, this time relating putatively distinct physical theories, rather than simply states or quantities *within* a single theory. These are more commonly referred to as 'dualities'. [3] Roughly, two theories are dual whenever they determine the same physics: same correlation functions, same physical spectra, etc. By analogy with symmetries as standardly understood, one is faced with a space whose elements are *theories*, as opposed to states or configurations,

[2] See Gordon Belot's fine survey [3] for a philosophically-oriented guide to the formal details.

[3] But note that the notion of duality is rather more general than this: dualities can also relate different 'sectors' (energies or scales, say) of a single theory. However, there is usually ambiguity here over the issue of whether we do indeed have a multiplicity (connected by a duality) or one single theory (connected by a gauge-type symmetry). This is related to the 'Landscape Problem' in string theory, according to which there is an ambiguity over whether the moduli space of string theory consists of some 10^{500} theories or a single theory with this many ground states. I return to this issue in §5.2.

so that dualities map one theory onto another in a way that preserves all 'physical' predictions. [4] One finds, for example, that certain theories at high temperatures (or energies, scales, couplings, fluctuations, ...) are physically equivalent to (other or the same: see footnote 3) theories at low temperatures (or energies, scales, couplings, fluctuations, ...). The computational value of such relations should be immediately apparent: high-energy problems are often intractable, involving extreme fluctuations and so on, but here one is at liberty to work in the dual low-energy theory and then 'translate' the results back into the high-energy lexicon. More crucially perhaps, dualities uncover new (nonperturbative) physics that is hidden from the Taylor series expansions that characterise perturbative formulations of theories.

In this paper I argue that such mappings have important consequences for scientific realists [5] , since one can find cases of dual (and therefore physically equivalent) theories that have prima facie radically (even *structurally* so, with topologically distinct spacetimes, for example) different ontologies. Duality might intuitively be viewed along the lines of gauge symmetry, with dual theories amounting to mere representational ambiguity. However, as I mentioned above, often when one has gauge symmetries, one performs an operation of 'quotienting out' to produce a 'slimmer' object with the redundancy associated with the symmetry eliminated. If we perform this operation in the case of dualities connecting theories with apparently distinct ontologies (and it is not clear that this is always a formal possibility) then it isn't clear what object we get out at the end. However, I argue that often (possibly always) something very similar to such a procedure *does* happen and introduces new, deeper physical ontologies. This both protects realist positions (from the potential underdetermination) and also answers some difficult questions over the status of dualities and those entities related by them.

Many of the dualities in appearing in string theory should have a particular res-

---

[4] There is some similarity (that might point to some profound connections) here to the way one conceives of theories in the effective field theory programme, involving renormalization group technology. A key difference between the two cases, however, is that effective field theories have their validity restricted to a particular scale, whereas (in the case of string theory at least) there is no such restriction enforced by the duality picture: the dual theories are generally applicable at all scales (since renormalizable—in fact they are conformally invariant and so are *fixed points* of the renormalization group flow), but may nevertheless become *intractable* at certain scales. Indeed, finding particular types of dual theories (those with 'strong-weak' coupling duality) is tantamount to finding theories that have good non-perturbative behaviour, and so that are demonstrably renormalizable. For philosophical examinations of some of these ideas, see: [7,19,22,23].

[5] The consequences are connected to that philosopher's 'hardy perennial': *underdetermination*. Do dualities merely point to the multiplicity of representation of one and the same thing or are we are faced with something conceptually more serious?

onance for philosophers of physics since they have the character of spacetime symmetries (there are *geometric* and *topological* dualities).[6] I shall therefore also try to draw out some potential areas of interest for philosophers of spacetime in what follows. The issues that emerge here are again connected to issues of underdetermination, but also extend into many other issues beyond this.

Finally, dualities are of wider interest in philosophy of science since they point to a mechanism for generating new theories and results. In particular, they point to the possibility of 'simulating' hard physics, in hard regimes, with simple physics. This isn't simulation in the sense of *approximation*: the dualities are **exact**. Often the 'simple physics' is classical while the 'hard' physics is quantum. This adds some confusion (or interest?) to the relationship between classical and quantum physics. One can also find that the simple physics is in lower spacetime dimensions than the hard theory. Physical equivalence in such situations puts serious pressure on our current conception of spacetime and its role in physical theories.

Though I don't aim to provide answers to these problems, or probe them in any real detail in this paper, I do wish to leave them as open problems for future work. If philosophers want to understand and interpret string theory (and many aspects of modern field theory—e.g. confinement, and so on), then they had better get to grips with the notion of duality. Besides, as I hope to show in this paper, there are many rich pickings from 'low-hanging fruit' for philosophers who are willing to investigate the subject.

2 Symmetry, Gauge, and Duality

In this section we briefly examine the differences between the prima facie rather similar concepts of symmetry, gauge redundancy, and duality. Once we have a grip on the basic notion of duality and have distinguished it from these other notions, we will aim to characterise it more precisely by looking at some specific non-string theoretic examples. With a handle on the concept we can then begin to extract some of its philosophical implications in the purely string theoretic context.

---

[6] In an early paper, Cyrus Taylor [36] suggested that such issues from string theory might be of interest to philosophers of spacetime physics. However, Taylor never went into any detail beyond this suggestion, and aside from the companion paper by Robert Weingard [40], his suggestion wasn't pursued.

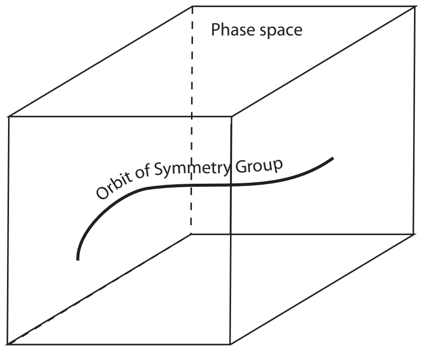


Fig. 1. The action of a symmetry group, mapping physical states onto distinct states, generating an orbit that we understand as representing a nomologically possible dynamical history of a system.

2.1 Physical Symmetries

A physical symmetry can be defined as a *structure-preserving mapping* of the space of physical states (i.e. the totality of points representing states that satisfy the relevant equations of motion) onto itself. There will always be trivial symmetries that map a state to itself, but the interesting ones (giving non-trivial dynamics) map physical states to *distinct* physical states—e.g. the action of the unitary time-evolution operator in quantum mechanics which maps states to later states. Though the states are different, they both nonetheless satisfy the laws of the theory. In other words, symmetries are transformations that keep the system within the set of physically possible states.

A symmetry will in general map a physical state to a distinct physical state. The orbit under the action of a symmetry group will, then, consist of points representing distinct physical situations. We can represent this in the phase space description as in figure 1.

In the case of gauge symmetries this is not the case: all elements within the same orbit correspond to the same physical situation. The physical state of a system is given by the orbit rather than by a point within the orbit.

2.2 Gauge Symmetries

In the case of gauge transformations, though it is true that gauge symmetries do not send physically possible states to physically impossible states, they do not map physically possible states into *distinct* physical states. Rather, the transformed and untransformed states are taken to represent one and

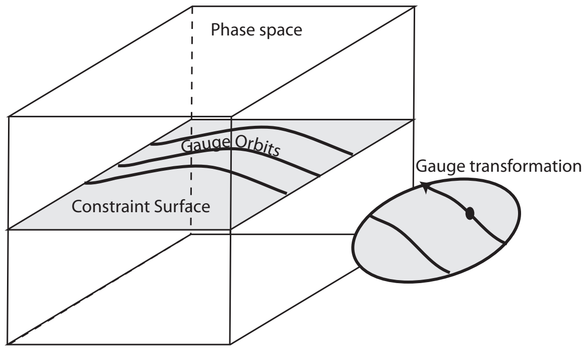


Fig. 2. In a gauge theory, physical states must satisfy constraints which pick out a submanifold in phase space, partitioned into gauge orbits representing the same physical state. Gauge transformations map between points within gauge orbits.

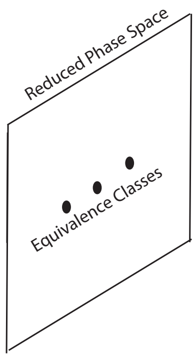


Fig. 3. The reduced phase space obtained by eliminating gauge freedom so that gauge orbits become single points.

the same physical situation. The multiplicity in representations amounts to a redundancy in the mode of representation (see fig.2). In fact, what we mean by 'a physical state' is really an equivalence class of states under the gauge symmetry, so that physical states are represented by gauge orbits rather than their elements (see fig.3).

This is easiest to see in electromagnetism in the vector potential formulation: the physical states here are not represented by single vector potentials, but by vector potentials *up to arbitrary gradients*. As Wigner so nicely expressed it:

> In order to describe the interaction of charges with the electromagnetic field, one first introduces new quantities to describe the electromagnetic field, the so-called electromagnetic potentials. From these, the components of the electromagnetic field can be easily calculated, but not conversely.

> Furthermore, the potentials are not uniquely determined by the field; several potentials (those differing by gradient) give the same field. It follows that the potentials cannot be measurable, and, in fact, only such quantities can be measurable which are invariant under the transformations which are arbitrary in the potential. This invariance is, of course, an artificial one, similar to that which we could obtain by introducing into our equations the location of a ghost. The equations the must be invariant with respect to changes of the coordinate of the ghost. One does not see, in fact, what good the introduction of the coordinate of the ghost does. ([41], p. 22)

Classically, it is true that the gauge potentials are physically redundant, but when quantum mechanical electrons interact with the electromagnetic field, as in the case of the Aharonov-Bohm effect, the potentials take on a causal role. In ([31] p. 132) Michael Redhead describes the process of 'stretching' some surplus structure by giving it a realistic interpretation. Clearly, what is deemed 'physical' and 'unphysical' is not an absolute matter, but depends on contextual factors, on what observables one has available to distinguish between elements of the theory. [7] New relational structures (in the case of the Aharonov-Bohm effect, this was the complex wave-function of a charged particle) introduced by some new system or theory can render unobservable quantities, such as the classical vector potential, physically significant and observable—see, e.g. [25] for an interpretation along these lines (the line also adopted by Aharonov and Bohm in their original presentation, though Bohm at first viewed the formal possibility of such an effect as a reductio of quantum mechanics!).

But note that in the case where some structure is stretched into reality, it is understood that just *one* potential is the *real* physical one. It does not make sense to speak of multiple simultaneous potentials; though it does make sense to speak (even in the case of the Aharonov-Bohm effect) of a gauge-invariant object that incorporates all gauge-related potentials, namely the holonomy of the potential (or, more generally, the Wilson loop)—note that this does not allow us to dispense with the vector potential as such, but the holonomies and Wilson loops are insensitive to gauge-transformations involving the potentials.

2.3 Duality Symmetries

Duality symmetries also point to a multiplicity in the descriptive machinery available for some system or phenomenon. Dualities represent freedom in the representation of some physical system, but unlike ordinary symmetries, the

[7] I would conjecture, on this basis, that the only case in which one has indistinguishability *simpliciter* is in the trivial case of self-identity.

representations connected by dualities are often surprisingly different in appearance, and they are decidedly *not* redundant. It is this marked theoretical difference (a genuine structural difference) combined with identical (observable) physics that make dualities special and distinguish them from symmetries and gauge. Dual descriptions are not in competition for 'physical reality'; rather they are considered to be complementary descriptions of one and the same physical situation.

I should note that duality is often distinguished in terms of the fact that it makes previously intractable problems tractable. It is true that it is capable of great simplifying feats. However, this is not really so different from the case in symmetry—say, where one's choice of an appropriate reference frame can make a difficult problem simpler—or in gauge theory, where choosing (or 'fixing') a particular (though strictly arbitrary) gauge can make a problem easier to work with. One might also point to the similar use of symmetry in a pure mathematics context where one can 'guess' a solution to some equation by studying its invariances—this was, essentially, Galois' great discovery (with contributions from Abel): one can distinguish between types of roots of polynomial equations by their symmetries such that functions of the roots are rational just in case they are invariant with respect to some group of permutations (see [37] for an excellent account of the development of Galois theory).

Hence, though there are clear differences in the *kinds* of problem made tractable by symmetry, gauge, and duality, [8] I don't think we can distinguish so easily in this way. The difference lies in *what* is related by symmetry, gauge, and duality transformations, as we discuss further in the next section.

3 Defining Dualities

In this section we get to grips with the notion of a duality symmetry. In simple terms, a pair of (putatively distinct) theories are said to be dual when they generate the same physics, where "same physics" is parsed in terms of having the same amplitudes, expectation values, observable spectra, and so on. The most familiar dualities will no doubt be those of the Maxwell equations and of the wave and particle pictures of quantum theory. Less familiar, but just as profound and surprising, if not more so, are those connecting the weak and

[8] For example, David Olive ([30], pp. 62–3) notes that there has been some resistance to duality amongst physicists precisely because it seems, *prima facie*, to be "so unreasonable". It has the characteristic of making very hard problems (in the non-perturbative sector of some theoretical framework) calculable in the perturbative sector.

strong coupling regimes of quantum field theories and string theories.[9]

Cumrun Vafa gives the following very useful direct characterisation of dualities, making the connection to physical observables manifest (and also introducing some of the terminology that we will use below):

> Consider a physical system $\mathcal{Q}$ ... [a]nd suppose this system depends upon a number of parameters. Collectively, we denote the space of parameters λ_i by $\mathcal{M}$, which is usually called the moduli space of the coupling constants of the theory. The parameters λ_i could for example define the geometry of the space the particles propagate in, the charges and masses of particles, etc. Among these parameters there is a parameter λ_0 which controls how close the system is to being a classical system (the analogue of what we call $\hbar$ in quantum mechanics). For λ_0 near zero, we have a classical system and for $\lambda_0 \geq 1$ quantum effects dominate the description of the physical system.
>
> Typically, physical systems have many observables which we could measure. Let us denote the observables $\mathcal{O}_\alpha$. Then we would be interested in their correlation functions which we denote by
>
> $$\langle \mathcal{O}_{\alpha 1} ... \mathcal{O}_{\alpha n} \rangle = f_{\alpha 1} ... f_{\alpha n}(\lambda_i) \tag{1}$$
>
> Note that the correlation functions will depend on the parameters defining $\mathcal{Q}$. The totality of such observables and their correlation functions determines a physical system. Two physical systems $\mathcal{Q}[\mathcal{M}, \mathcal{O}_\alpha]$, $\tilde{\mathcal{Q}}[\tilde{\mathcal{M}}, \tilde{\mathcal{O}}_\alpha]$ are dual to one another if there is an isomorphism between $\mathcal{M}$ and $\tilde{\mathcal{M}}$ [the 'moduli spaces'–DR] and $\mathcal{O} \leftrightarrow \tilde{\mathcal{O}}$ respecting all the correlation functions. Sometimes this isomorphism is trivial and in some cases it is not. [[38], p. 539–540]

We can distinguish between two general kinds of non-trivial duality: those that relate pairs of distinct theories and those that relate one and the same theory to itself. We can call these 'internal duality' (or 'self-duality') and 'external duality' respectively. Strictly speaking, since they state an equivalence between (apparently very distinct) descriptions of one and the same system, self-dualities are really just gauge symmetries in disguise, representing some interpretative ambiguity in the theory's formulation. Let us now present a selection of dualities, leading up to those appearing in string theory.

3.1 Electromagnetic Duality

Maxwell's equations describe the behaviour of a pair of (vector) fields: $\vec{B}$, the magnetic field, and $\vec{E}$, the electric field. These fields depend on the charge

[9] Of course, there is some sense in which all dualities are of the same broad type, namely a relation between large and small scales/energies.

density ρ and the current density $\vec{j}$. However, the *vacuum* Maxwell equations (with vanishing charge and current densities) exhibit a duality in a most immediate (and visual) way:

$$\nabla \cdot \vec{B} = 0 \qquad \nabla \times \vec{E} + \frac{\partial \vec{B}}{\partial t} = 0 \tag{2}$$

$$\nabla \cdot \vec{E} = 0 \qquad \nabla \times \vec{B} - \frac{\partial \vec{E}}{\partial t} = 0 \tag{3}$$

In addition to being Lorentz invariant, these equations are also conformally and gauge invariant. It is easy to see that the following pair of maps (interchanging the electric and magnetic properties, and so, one might well suppose, giving in a *different* theory) amount to a duality symmetry of the vacuum Maxwell equations:[10]

$$\vec{B} \longrightarrow \vec{E} \tag{4}$$

$$\vec{E} \longrightarrow -\vec{B} \tag{5}$$

In this case the duality points to a deeper structure into which both the electric and magnetic fields are integrated, namely the electromagnetic field. Hence, the discovery of a duality between a pair of things can be 'symptom' that the pair of things are really two aspects of one and the same underlying structure. But this duality is just as striking (though buried a little more deeply) even if we write the equations in terms of the electromagnetic field tensor $\mathcal{F}_{\mu\nu}$ (representing field strength) and the Hodge star $\star$ (taking $k-$forms to $k-1$-forms):

$$\partial_\nu \mathcal{F}^{\mu\nu} = 0 \tag{6}$$

$$\partial_\nu {}^\star\mathcal{F}^{\mu\nu} = 0 \tag{7}$$

The duality can be expressed by the mappings:

$$\partial_\nu \mathcal{F}^{\mu\nu} \longrightarrow \partial_\nu {}^\star\mathcal{F}^{\mu\nu} \tag{8}$$

$$\partial_\nu {}^\star\mathcal{F}^{\mu\nu} \longrightarrow -\partial_\nu \mathcal{F}^{\mu\nu} \tag{9}$$

Where $\star^2 = -1 =$ 'squaring the duality map (for the $3+1$ case)'. In more

[10] The duality transformation leaves physical observables, such as total energy densities $1/8\pi|\vec{E}|^2 + 1/8\pi|\vec{B}|^2$, invariant.

detail, let us define the electromagnetic field strength by $\mathcal{F}_{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$ with components:

$$\mathcal{F}^{0i} = -\mathcal{F}^{i0} = -E^i \tag{10}$$
$$\mathcal{F}^{ij} = -\epsilon_{ijk} B^k \tag{11}$$

We can write this out explicitly, highlighting the duality hidden within the field tensor, as follows

$$\mathcal{F}_{\mu\nu} = \begin{pmatrix} 0 & -\vec{E}_x & -\vec{E}_y & -\vec{E}_z \\ \vec{E}_x & 0 & \vec{B}_z & -\vec{B}_y \\ \vec{E}_y & -\vec{B}_z & 0 & \vec{B}_x \\ \vec{E}_z & \vec{B}_y & -\vec{B}_x & 0 \end{pmatrix} \tag{12}$$

Where the Hodge dual is:

$$^*\mathcal{F}_{\mu\nu} = \begin{pmatrix} 0 & \vec{B}_x & \vec{B}_y & \vec{B}_z \\ -\vec{B}_x & 0 & \vec{E}_z & -\vec{E}_y \\ -\vec{B}_y & -\vec{E}_z & 0 & \vec{E}_x \\ -\vec{B}_z & \vec{E}_y & -\vec{E}_x & 0 \end{pmatrix} \tag{13}$$

The Hodge star operation on the electromagnetic field tensor is equivalent to the application of the duality maps to the separate electric and magnetic fields.

Of course, we don't live in an electromagnetic vacuum: there are electric charges. The problem for duality is, there don't appear to be corresponding magnetic charges. This can be seen in the non-vacuum equations, which state precisely that fact: [11]

$$\nabla \cdot \vec{B} = 0 \qquad \nabla \times \vec{E} + \frac{\partial \vec{B}}{\partial t} = 0 \tag{14}$$
$$\nabla \cdot \vec{E} = \rho_e \cdot 4\pi \qquad \nabla \times \vec{B} - \frac{\partial \vec{E}}{\partial t} = \vec{j}_e \cdot \frac{4\pi}{c} \tag{15}$$

[11] In the field tensor formulation we have $\partial_\nu \mathcal{F}^{\mu\nu} = j^\nu$ and $\partial_\nu {}^\star\mathcal{F}^{\mu\nu} = k^\mu$.

Hence, electric sources spoil the duality: there are no magnetic charges since the divergence of $\vec{B}$ vanishes—by implication, there is no magnetic current (since currents amount to moving charges). Of course, the duality can be restored if we simply assume that there *are* magnetic charges (i.e. monopoles) ρ_m:

$$\nabla \cdot \vec{B} = \rho_m \cdot 4\pi \qquad \nabla \times \vec{E} + \frac{\partial \vec{B}}{\partial t} = \vec{j}_m \cdot \frac{4\pi}{c} \tag{16}$$

$$\nabla \cdot \vec{E} = \rho_e \cdot 4\pi \qquad \nabla \times \vec{B} - \frac{\partial \vec{E}}{\partial t} = \vec{j}_e \cdot \frac{4\pi}{c} \tag{17}$$

The duality is then preserved if we map all 'magnetic objects' to 'electric objects', and vice versa (with the necessary change of sign): [12]

$$(\vec{E}, \rho_e, \vec{j}_e) \longrightarrow (\vec{B}, \rho_m, \vec{j}_m) \tag{18}$$

$$(\vec{B}, \rho_m, \vec{j}_m) \longrightarrow -(\vec{E}, \rho_e, \vec{j}_e) \tag{19}$$

Of course, while this is formally a very nice situation, we don't have any direct empirical evidence for the existence of magnetic monopoles, despite quite a lot of theoretical investigation. If we want a physically realistic theory, we have to then explain why they have never been observed.

Dirac in particular worked hard to motivate the study of magnetic monopoles and discovered the quantization condition that bears his name while investigating them. His approach was to separate the north and south magnetic poles of a dipole (by an infinite distance) so that the magnetic flux between them is concentrated in the so-called 'Dirac string' connecting them. The monopole becomes unobservable (using charged particles) when the charge on the monopole obeys the 'Dirac quantization condition' (in notation where 'e' is the electric charge and 'g' is the magnetic charge): $e \cdot g = \frac{n\hbar c}{2}$ (where $n \in \mathbb{Z}$). [13] Hence the Dirac monopole is really only a 'FAPP-monopole': for all practical (i.e. experimentally observable) purposes it behaves as a true monopole.

The combination of the original electric-magnetic duality and the Dirac quantization condition gives us another aspect of electromagnetic duality. We get (dividing through by e in the first instance and g in the second) an inverse relationship between the coupling strengths of the two forces (setting $\hbar = 1 = c$):

[12] Or, in the corresponding field tensor form: $j^\mu \longrightarrow k^\mu$ and $k^\mu \longrightarrow -j^\mu$.

[13] There is much more to be said here. For example, the quantization condition is a constraint on the possible values that the charges of monopoles can possess in order to be consistent with quantum mechanics. I give a flavour of this (though just a flavour) in the next subsection.

$$e \longrightarrow g = \frac{n}{2e} \tag{20}$$
$$g \longrightarrow -e = -\frac{n}{2g} \tag{21}$$

Note that this also gives us an explanation for charge quantization since if a single magnetic monopole of charge g exists then all electrically charged particles will come with charges that are integer multiples of $1/2g$. Dirac thought this explanatory virtue was reason enough to give the existence of magnetic monopoles credence (outweighing the vice of lack of experimental evidence [14]):

> The interest of the theory of magnetic poles is that it forms a natural generalization of the usual electrodynamics and it leads to the quantization of electricity. [...] The quantization of electricity is one of the most fundamental and striking features of atomic physics, and there seems to be no explanation for it apart from the theory of poles. This provides some grounds for believing in the existence of these poles. [[11], p. 817]

The electromagnetic duality now gives us a duality between strong and weakly coupled theories. At weak coupling the electrically charged particles are well-localized quanta of the electromagnetic field and the magnetic monopoles appear as spread out, composite, bound states, or solitonic excitations. This has some clear philosophical interest since there is no principled reason why we should take one description as 'fundamental' with the other being 'supervenient.' It bears some similarity too to the notion of 'nuclear democracy' espoused by Geoff Chew in the context of his anti-field theoretic S-matrix bootstrap approach to particle physics from the 1960s. The dualities discussed here, however, are not antagonistic to field theory and the *constructive* picture of physics, but are instead 'complementary' in the sense of wave-particle duality, to be discussed below.

However, even if this duality, involving monopoles, were part of physical reality, it does not offer much by way of computational aid since electromagnetism is a weakly-coupled theory: we cannot derive much benefit from switching strongly (i.e. magnetic) and weakly (i.e. electric) coupled regimes. But if this duality could be found to hold in *strongly* coupled theories, so that the interchange rendered the theory weakly coupled, then computations would thereby be made much easier, switching a theory with very strong quantum

[14] The methodology here is certainly not what we would call 'the standard model' of how science works: there is no empirical evidence as such. Rather, what is motivating the theory of monopoles is its ability to resolve a hitherto unresolved theoretical puzzle: the quantization of charge. Then, in order to make the theory consistent with observation Dirac devises a reason for our not having observed magnetic monopoles. Namely, as a result of the strong-weak coupling duality, a quantized monopole requires an exceedingly large amount of energy to be produced.

fluctuations with one with small fluctuations. An obvious candidate is QCD, a very strongly coupled theory who's non-perturbative behaviour is still ill-understood. Just such a scheme has been suggested in QCD, and also in string theory as we will see later.[15] In QCD the difficult property to explain is confinement (the phenomenon prohibiting free quarks to be observed). 't Hooft argued that confinement can be understood as a dual, strongly coupled description of the Higgs mechanism, giving a mass-gap. However, to discuss would take us too far afield.[16]

3.2 Quantum Mechanics and Magnetic Monopoles

Let us briefly return to the reasons underlying Dirac's quantization condition. It is clear that the input of quantum mechanics seemingly spoils the nice clean duality once again, since there the electric and magnetic fields alone are not sufficient to determine the state. In the context of classical electromagnetism coupled to classical particles the basic electric and magnetic fields bear full responsibility for particle motion, via the Lorentz force law, $q(\vec{E}(x) + v \times \vec{B}(x))$. This is not the case where quantum particles are made to interact with the classical field for then the vector potential plays a crucial role in the Hamiltonian—underlying this, of course, is the fact that charged quantum particles must be described by a complex wave-function $\psi(x)$ obeying equations of motion involving gauge potentials, ϕ and $\vec{A}$ (where $\{\phi, \vec{A}\} = A$, the 4-vector potential). However, magnetic monopoles exclude the definition of the vector potential. There is a simple argument that highlights an internal inconsistency between the existence of monopoles, Maxwell's equations, and quantum mechanics—*cf.* [42], p. 28.

The equation for the magnetic monopole is $\nabla \cdot \vec{B} = \rho_m \cdot 4\pi$. However, a well-known theorem from vector calculus tells us that $\nabla \cdot (\nabla \times A) = 0$ (where A here is the vector potential, but can be any vector field for the purposes of the theorem). If we write $\vec{B}$ in terms of the vector potential, we get $\vec{B} = \nabla \times A$. Hence, $\vec{B}$ satisfies the *vacuum* Maxwell equation $\nabla \cdot \vec{B} = 0 = \nabla \cdot (\nabla \times A)$. However, if we give the weight to the vector potential *and* include monopoles we face a contradiction: with monopoles, we have $\nabla \cdot \vec{B} = \rho_m \cdot 4\pi \neq \nabla \cdot (\nabla \times A)$

[15] Indeed, though I won't discuss it here, it has been proposed that string theory (with gravity), in a special 10-dimensional (product) spacetime $AdS_5 \times S^5$, is dual to 4-dimensional conformal field theory (without gravity) defined on AdS_5's boundary. This is known as the AdS/CFT correspondence (also known as 'Maldacena duality')—see [28] for the original presentation (though there were forerunners to Maldacena's conjecture, connected to aspects of black hole physics, especially the notion of holography developed by 't Hooft, of which the Maldacena conjecture can be seen as a concrete realization).

[16] For a sampling of some relevant philosophical literature, see: [35,20,27].

(unless $\rho_m = 0$). Thus, there is a clear clash between magnetic monopoles and vector potentials.[17] However, the vector potential is seen to be demanded by quantum mechanics—as evidenced by the Aharonov-Bohm effect, for example, in which the scattering of electrons off a long, thin solenoid produces a diffraction pattern that depends on whether a magnetic field is or is not present in the solenoid. The electron feels the presence of the non-vanishing vector potential outside of the solenoid (though the measurable quantity still involves the magnetic field flux, rather than the vector potential itself). Formally, the line integral around a loop enclosing the solenoid equals the magnetic flux confined within the solenoid. Hence, charged quantum particles must be affected by the vector potential (on pain of violating local action, though one could in principle reject this).

The resolution of this apparent inconsistency is rather complex, but interesting. The trick is to first consider a magnetic monopole at the origin of a sphere the surface of which is intersected by the the magnetic field emanating from the monopole. The monopole (and so the origin of the sphere) is of course where $\nabla \cdot \vec{B} \neq 0$. The divergence of $\vec{B}$ does, however, vanish everywhere else, including on S_2, the sphere's surface. But despite the vanishing divergence on the surface, one cannot write the magnetic field as the curl of a vector potential—one cannot have an everywhere continuous vector potential for a magnetic field when one has a magnetic monopole. That is, $\vec{B} \neq \nabla \times A$, since by Stokes' Theorem we can infer:

$$(\vec{B} = \nabla \times A) \supset (\rho_m = \int_{S_2} \vec{B} \cdot dS = \int_{S_2} \nabla \times A \cdot dS = 0) \tag{22}$$

In words, the flux through the sphere would be the magnetic charge; but this would be forced to vanish if written in terms of the the curl of the vector potential, by Stokes' Theorem. Yet $\rho_m \neq 0$ by construction.[18]

The conflict can be eliminated by viewing the sphere as composed of two distinct hemispheres, N(orth) and S(outh), separated by an equator at which N and S overlap. We can consider vector potentials, A_N and A_S, restricted to N and S respectively that satisfy $\vec{B} = \nabla \times A_N$ and $\vec{B} = \nabla \times A_S$. A_N and A_S must differ on the overlap disc at the equator but possess the same curls (since

[17] Shifting to holonomies or Wilson loops here would not help, since we are still transporting particles around a curve *within a vector potential*, and so still need it—at least as a formal 'leg up'. In a little more detail, the Wilson loop is the phase that a charged particle's wavefunction is multiplied by as it traverses a loop. The phase is multiplied by an element of $U(1)$, namely: $e^{-\frac{1}{\hbar}nq \oint_\gamma A}$ (where γ is the loop and A is the vector potential).

[18] See Chapter 6 of Baez & Munian [2] for a useful discussion of this argument and its formal underpinnings.

the magnetic field must take on the same value there, and we have written this as the curl of a vector potential). The solution is that whenever one has vector potentials defined in different regions (and one has a magnetic monopole), the potentials will be related by gauge transformations ($\omega(x)$ in Dirac's notation) on their overlap (the equator in the above example), such that $e^{\frac{ie\omega}{\hbar}} = 1$. This gives us the desired result that the monopole is not observable with electrically charged particles, since its wave function *will* now be everywhere continuous.

Let us next consider a specifically quantum duality, namely wave-particle duality.

3.3 Wave-Particle Duality

The wave-particle duality is well encapsulated, in the context of quantum electrodynamics, by Dirac as follows:

> Instead of working with a picture of the photons as particles, one can use instead the components of the electromagnetic field. One thus gets a complete harmonizing of the wave and corpuscular theories of light. One can treat light as composed of electromagnetic waves, each wave to be treated like an oscillator; alternatively, one can treat light as composed of photons, the photons being bosons and each photon state corresponding to one of the oscillators of the electromagnetic field. One then has the reconciliation of the wave and corpuscular theories of light. They are just two mathematical descriptions of the same physical reality. [[10], p. 49]

This last expression is simply another way of saying that these pictures are really dual; they are connected by what Dirac called a "formal reconciliation between the wave and the light-quantum [pictures]" ([12], p. 711)—a fact that Bohr was quick to avail himself of in support of his ideas on complementarity.

Hence, we see again the ability of dualities to point to new structures beyond those things that are considered to be dual. In this case the waves and particles are but two aspects of the same underlying (and more fundamental) entity: the electromagnetic field. Mathematically, however, the duality between wave and particle descriptions is simply encapsulated in the difference between the position basis and the momentum basis of states. These are related by a Fourier transformation. Vafa in fact suggests that we can understand string dualities along the lines of a "non-linear infinite dimensional generalization of a Fourier transform" ([38], p. 537).

Many of the advances in string theory have been generated by the discovery of new duality symmetries. In fact, string theory was in effect born from a duality principle, namely the DHS [Dolan, Horn, and Schmid] duality identifying descriptions of 2-particle to 2-particle scattering in the s-channel (or 'direct' channel) with those in the t-channel (or 'transverse' channel) of the S-matrix for strongly interacting particles, or hadrons—the specific particles involved were, in this case, mesons.[19] That is, low-energy (direct) s-channel 'resonances' and high-energy (cross) $t-$channel 'Regge poles' produce equivalent (dual) physics (one representation can be analytically continued into the other).[20] Diagrams that would be added together in the Lagrangian quantum field theoretic approach to hadronic physics are considered to be representations of one and the same process in the context of DHS duality.

This duality, along with other properties of the S-matrix led Gabriele Veneziano to propose his famous (2-particles go to 2-particles, or '4-point') amplitude based on the Euler beta function (here suppressing the u-channel contribution and with linearly-rising Regge slope $\alpha(s) = \alpha_0 + \alpha' s$):[21]

$$A(s,t) = \frac{\Gamma(1-\alpha(s))\Gamma(1-\alpha(t))}{\Gamma((1-\alpha(s))-\alpha(t))} = \int_0^1 dx x^{-\alpha(s)-1}(1-x)^{-\alpha(t)-1} \tag{23}$$

Though considered empirically inadequate now, this model enjoyed considerable phenomenological success, especially with respect to low-energy proton-antiproton annihilation (cf. Rubinstein, [34], p. 55). But doing the real explanatory work is the underlying DHS duality. The connection to strings comes about since this amplitude can be derived from a theory of relativistic open strings. The poles of the amplitude (i.e. the singularities that occur when $s = m^2 = t = \frac{n-1}{\alpha'}$) correspond to the string's mass spectrum. The interpretation provides an explanation of the infinite tower $J = \alpha M^2$ of mass-energy

[19] In this scenario, $s = -(p_1 + p_2)^2$ and $t = -(p_2 + p_3)^2$. These are Lorentz invariant combinations of the momenta of the scattered particles, with subscripts labelling the particles.

[20] The earliest version of string theory was known as "dual theory" in recognition of this fact. See [24] for a collection of classic review papers on this subject. A recent festschrift [15] for Gabriele Veneziano includes some useful discussions of the earliest days of string theory, including discussions of the DHS duality principle and its role in the birth of string theory.

[21] This simultaneously satisfies what were deemed to be the crucial formal properties describing the S-matrix for strong interactions, particularly analyticity, crossing symmetry, and DHS duality. The Veneziano amplitude thus solved the bootstrap! However, there is degeneracy in the choice of α which need simply be linear.

and spin states in the Regge trajectories: they correspond to the oscillatory (and rotary) modes of the strings.

It is this notion of duality that has persisted into the development of modern QCD and string theory, and that I shall discuss in this paper. It, or something very close, lies at the heart of various pieces of conceptually important physics, such as the holographic principle, black hole information, as well as several experimental advances, including quark-gluon plasmas and high temperature superconductivity.

Indeed, duality has played an enormously important role in the creation and development of numerous theories in physics and numerous fields of mathematics. It is also a ubiquitous feature at the *interface* between physics and pure mathematics: that complex area where discoveries go both ways, from mathematics to physics and from physics to mathematics. Given the crucial role of dualities in modern physics it is incumbent on philosophers of physics to evaluate their significance, ontological, epistemological and otherwise.[22] I attempt an initial foray in this paper, with the focus squarely on string theoretic dualities. These are especially interesting from a philosophical point of view because of their close connection with geometrical considerations. However, they are also a particularly good window onto a vast array of philosophical issues emerging from the notion of duality.

[22] Richard Dawid ([9], pp. 316–7) briefly discusses string dualities. However, he focuses on T-duality and the purely perturbative theory. The conclusion he draws is that T-duality implies that "[a]n absolute limit is set on attaining new physical information below a certain scale [$2\pi\sqrt{\alpha'}$–DR]" (p. 317). Described in this way it is a fairly modest result, since if, for example, quarks were to represent the smallest possible entities, then their Compton wavelengths would set a similar limit on the physical information we could obtain, since it would be by scattering them that we get information and they themselves represent the limit of what we can scatter (in fact, the phenomenon of confinement would raise the scale somewhat). Edward Witten [43] describes T-duality in stronger terms. He argues that it points to "quantum geometry": a limit on space itself, rather than our ability to extract information. (Note that D-branes are non-perturbative objects that, in general, do not obey T-duality—though there are nonetheless suggestions that they involve quantum geometry too, since they have non-commuting position coordinates (relative to the space in which they lie) which are described by matrices. These claims are clearly in need of closer philosophical scrutiny and surely ought to be of greater interest to philosophers of spacetime physics.

4 String Dualities

There are two general kinds of duality in string theory and they map onto our earlier distinction between internal and external duality. Philosophically, the more interesting ones are the external ones. However, the internal dualities in string theory are not without interest, especially where spacetime ontology is concerned, so we shall begin with this case.

4.1 What is String Theory?

String theory, as we have seen, has its origins in strong interaction physics where it was constructed specifically to model hadrons. It wasn't able to perform this function for a variety of reasons, both formal and empirical. Empirically, it predicted the existence of particles at particular energies (namely those appropriate for hadronic interactions) that weren't seen in experiments: these were the massless spin-2 particles. Quantum chromodynamics superseded string theory putting, more or less, standard quantum field theory back in charge of elementary particle physics. In QCD the strong interaction is described by an SU(3) Yang-Mills field.

Duality has been a decisive factor throughout the entire development of string theory, from these origins in the dual resonance model for describing hadrons to the (still unknown) $\mathcal{M}-$theory. It can be found in the algebraic core of string theory too: the Leech lattice, for example, describing early string theory, is characterised by self-duality. The dualities in string theory take on a special resonance, from a philosophical point of view, because they are highly geometrical. In what I take to be the most interesting cases the dualities are mapping the physics of some string theory compactified on a manifold to another string theory on a *prima facie* very different manifold.

The easiest way to make sense of the geometrical dualities in string theory is by introducing the theory via its perturbation expansion. [23] We restrict the discussion to closed string theory, and begin with the so-called σ-model. One wants to construct an action to describe the string dynamics in spacetime. The

[23] I should perhaps point out that this perturbative 'worldsheet' formulation is somewhat outmoded. However, it is at least a well-defined theory and enables one to see in a fairly visual way how the interesting elements of mathematics (such as Riemann surfaces, modular invariance, and the like) enter into string theory. Though I don't discuss it here, the modular invariance lies beneath some of the deepest connections between physics and mathematics, and is connected also to S-duality (the strong-weak coupling duality). The technicalities would take too long to introduce here.

initial step is to consider a map Φ from a complex curve (a Riemann surface) Σ representing the 2-dimensional string worldsheet [24] into the ambient target space X (with metric G and additional background fields B^i):

$$\Phi : \Sigma \longrightarrow X \tag{24}$$

The action is then a function of this map (including the worldsheet's metric), given the background fields G and B^i:

$$S(\Phi, G, B^i) \tag{25}$$

The Φ field gives the dynamics of a 2-dimensional field theory of the worldsheet relative to the fixed background fields, one of which is the metric. The quantum theory (in 1st quantized form) is given by the path-integral (over moduli space: i.e. the space of inequivalent 2D Riemann surfaces, or Teichmüller space):

$$\mathcal{P}(X) = \sum_g \int_{moduli_g} \int \mathcal{D}\Phi e^{iS(\Phi,G,B^i)} \tag{26}$$

In terms of the interpretation of this object, there is a degree of non-separability of the kind found in loop quantum gravity, for the relevant domain is not the space of metrics on a manifold (i.e. geometries) but the loop space. However, there are consistency conditions that must be met by string models not shared by the loop models.

4.2 Compactification

Quantum superstring theory remains Lorentz invariant only if spacetime has 10 dimensions. To construct a realistic theory therefore demands that the vacuum state (i.e. the vacuum solution of the classical string equations of motion, supplying the background for the superstrings) is given by a product space of the form $\mathcal{M} \times \mathcal{K}$, where $\mathcal{M}$ is a non-compact four dimensional Minkowskian spacetime and $\mathcal{K}$ is a compact 6-real dimensional manifold. One gets the physics 'out' of this via topological invariants of $\mathcal{K}$ and gauge fields living on $\mathcal{K}$. One chooses the specific form of the compact manifold to match the observed phenomena in $\mathcal{M}$ as closely as possible. The Landscape Prob-

[24] This worldsheet has a metric $h_{\alpha\beta}$ defined on it in the so-called Polyakov version. In the original Nambu-Goto version the worldsheet was metric-free. The surface also has a genus g which plays a crucial role in the quantum theory.

lem referred to earlier is tantamount to the severe degeneracy in this space of possible classical vacua.

If one wants $\mathcal{N} = 1$ supersymmetry in the non-compact dimensions $\mathcal{M}$, then one requires a very special geometry for the compact dimensions $\mathcal{K}$, namely a Calabi-Yau manifold. This is defined to be a compact Kähler manifold with trivial first Chern class—this is just mathematical shorthand for saying that we want to get our low-energy physics (Ricci flatness [25] and the single supersymmetry) out of the compact dimensions.

There are five quantum-mechanically consistent superstring theories (in 10 dimensions): Type I, $SO(32)$-Heterotic, $E_8 \times E_8$-Heterotic, Type IIA and Type IIB. The Type I theory and the heterotic theories differ from the Type II theories in the number of supersymmetries, and therefore in the number of conserved charges.

One is able to compute physical quantities from the these theories using perturbation expansions in the string coupling constant. Given the extended nature of the strings, there is just a single Riemann surface for each order of the expansion (that is, the initially distinct diagrams can be topologically deformed into one another since there are no singularities representing interaction points: interactions are determined by global topological considerations of the world sheet, rather than local singularities). [26]

4.2.1 T-Duality

T-duality results from the combination of compactified dimensions and strings. [27] T-duality is a kind of scale-invariance: it says that a theory at one size is equivalent to a theory at another size. It is essentially a duality that arises in conformal field theory. For superstring theories (i.e. with fermions and supersymmetry relating bosons and fermions) we find that the Type IIA and

[25] The first Chern class $c_1(\mathcal{X})$ of a metric-manifold is represented by the 2-form $1/2\pi\rho$ (with ρ the Ricci tensor $R_{i\bar{j}}dz^i \wedge d\overline{z^{j}}$). Calabi and Yau determined the various interrelations between Chern classes, Kählericity, and Ricci forms. If one has a Ricci flat metric then one also gets the desired single supersymmetry since Ricci flatness is a sufficient condition for an $SU(3)$ holonomy group. Any textbook on complex algebraic geometry will explain these matters in detail—[1] and [29] are good sources of information.

[26] Note, this is true for all but the Type I theory since its strings can be opened up. However, this does not need to concern us in what follows.

[27] There are various options for the referent of the 'T'. Some take it to refer to (T)arget space, some to the fact that it is similar to the Kramers-Wannier (T)emperature duality of the Ising model, others to the fact that the letter 'T' was used to refer to a low-energy field in early string theory.

Type IIB theories are dual, as are the two heterotic theories. In the context of bosonic string theory it is a self-duality and can therefore be viewed as a gauge symmetry.

T-duality is very simply expressed: given two manifolds, with different compact geometries, a circle of R and of radius $\tilde{R}$, and string length scale α', we have (schematically):

$$\text{String Theory on } R \quad \xleftrightarrow{\textit{isomorphic}} \quad \text{String Theory on } \tilde{R} = \frac{\alpha'^2}{R}$$

This isomorphism can be seen by considering the case where we have compactified one of the dimensions onto a circle. When this is done, the momentum is quantized around the circle according to the relation $p = n/R$ (where $n \in \mathbb{N}$). If we then consider the mass-energy of a system in such a compactified configuration then we must add a term corresponding to these so-called Kaluza-Klein modes:

$$E^2 = M^2 + \frac{n}{R}^2 \tag{27}$$

So far everything we have said applies just as well to particles. Strings have the additional property that they can wind around the compact dimension. This brings with it another term (the winding modes, where m counts the number of such windings) that must be added to the total energy-mass:

$$\frac{1}{2\pi}\alpha' \times 2\pi R \cdot m = \left(\frac{mR}{\alpha'}\right)^2 \tag{28}$$

This gives us the following equation for computing the mass-energy:

$$E^2 = M^2 + \frac{n}{R}^2 + \left(\frac{mR}{\alpha'}\right)^2 \tag{29}$$

If we then make the following (duality) transformations we leave the energy invariant:

$$R \longrightarrow \frac{\alpha'}{R} \tag{30}$$

$$m \longleftrightarrow n \tag{31}$$

since we then have:

$$E^2 = M^2 + \frac{m^2}{\frac{\alpha'}{R}} + \left(\frac{n\frac{\alpha'}{R}}{\alpha'} \right)^2 \tag{32}$$

This can be converted back into the original by simply multiplying the numerator and denominator of the 2nd and 3rd terms by R and cancelling the α's in the 3rd term.

Though this is a very elementary account, it serves to highlight the curious nature of strings and compact dimensions: *from the stringy perspective there is no difference between a space with a large radius and one with a small radius!* If we consider a theory to be an equivalence class of structures (with the equivalence given by the determination of identical observables) then what we took to be four distinct theories—type IIA and IIB on the one hand, and $SO(32)$ and $E_8 \times E_8$ on the other—are really just two.

Physical sense can be made of this by viewing T-duality through the lens of the uncertainty principle: the attempt to localize a closed string at very small scales increases its energy-momentum. This increase in energy as one localizes to smaller and smaller length scales increases the size of the string.

In a nutshell, T-duality tells us that it is only some deeper intrinsic properties of the backgrounds for string propagation that matter in terms of 'the physics'. Different background spaces are identical from the point of view of the strings. Since, in a string theory, everything is assumed to be made of strings, then in a purely string theoretic world, these backgrounds are indiscernible. This is very similar to the implications of diffeomorphism invariance in general relativity. There the *localization* of the fundamental objects relative to the manifold is a gauge freedom in the theory: the physics is therefore insensitive to matters of absolute localization. Quantities that are defined at points of the manifold are clearly not diffeomorphism-invariant, and therefore not gauge-invariant. The physics should not depend on such gauge-variant local properties. In the case of string theory, the physics should not depend on the size of the compact dimensions.

4.2.2 S-Duality

S-duality is not a purely string-theoretic symmetry. It maps a strongly-coupled theory to a weakly-coupled theory. It is, therefore, a *nonperturbative* duality. Combined with T-duality, it shows that the five apparently distinct string theories (and, in fact, an additional 11 dimensional theory) are dual. Invoking our principle from earlier, that dualities point to some underlying structure, we can assume that there is a deeper theory of which these various 'theories' are aspects. This is indeed what has been conjectured, with the hidden theory

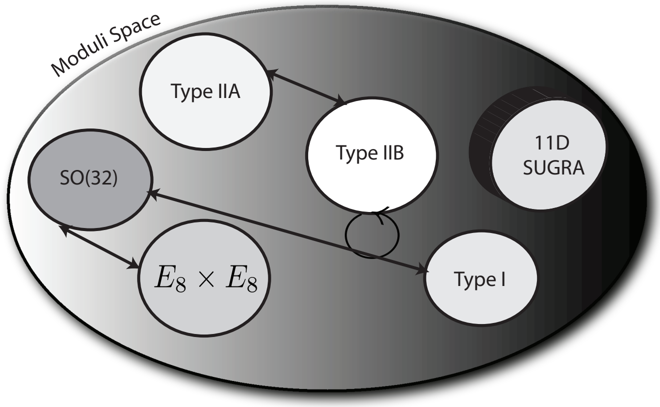


Fig. 4. Web of dualities between the supposedly distinct superstring theories.

labeled M-theory—see [42] for an elementary account.

S-duality identifies Type I and $SO(32)$ and different sectors (in the moduli space) of type IIB in much the same way that T-duality related the other string theories. Schematically, we have the following duality relations (where g is the string coupling constant):

$$\text{Type I with } g \gg 1 \quad \xleftrightarrow{\textit{isomorphic}} \quad \text{SO(32) with } g \ll 1$$

$$\text{Type IIB with } g \gg 1 \quad \xleftrightarrow{\textit{isomorphic}} \quad \text{Type IIB with } g \ll 1$$

Since Type IIB theory is 'internally' S-dual, or self-S-dual, we might more properly refer to it as a gauge symmetry in the standard sense.

Though I won't discuss it here, an interesting phenomenon occurs when we consider the Type IIA and $E_8 \times E_8$ theories at strong coupling. The theories 'grow' an additional dimension, with the size of the dimensions given by $g\sqrt{\alpha'}$. These theories are dual to some other quite different 11 dimensional theory that is not apparent in the weak-coupling regime studied in perturbation theory. This is a clear case where new physics lies hidden from view when one restricts the analysis to the purely perturbative regime.

4.3 Mirror Symmetry

Mirror symmetry is possibly the most conceptually curious aspect of string theory. It is essentially a generalization of T-duality (which holds only for homeomorphic manifolds) to topologically inequivalent manifolds.

Recall that a phenomenologically respectable string theory requires that six of the 10 dimensions be hidden from view somehow. Compactification is the process that achieves this (at least formally). As we saw earlier, this involves writing the 10 dimensional spacetime $\mathcal{M}_{10}$ (required by quantum consistency) as a product space of the form $M^4 \times K^6$, where M^4 is flat Minkowski spacetime and K^6 is some compact 6 real-dimensional space. $M^4 \times K^6$ then forms the background space (the ground state in fact) for the classical string equations of motion. One chooses K^6 in such a way so as to use its geometrical and topological structure to determine the physics in the four non-compact spacetime dimensions (i.e. the low-energy physics). By choosing in the right way one can get explanations for a host of previously inexplicable features of low-energy physics, such as the numbers of generations of particles in the standard model, the various symmetry groups of the strong, electroweak, and gravitational forces, and the masses and lifetimes of various particles.

Calabi-Yau manifolds were found to be of importance in string theories since they allow for $\mathcal{N} = 1$ supersymmetries in four spacetime dimensions and other nice properties. Calabi-Yau manifolds are compact spaces satisfying the conditions of Ricci-flatness (to accommodate general relativity at the phenomenological 4D level) and Kählericity (generating the $\mathcal{N} = 1$ supersymmetry in the non-compact dimensions). The problem is, there is a huge number of Calabi-Yau spaces (in D=6) meeting the required conditions, so the selection of one is a difficult task. However, what I want to discuss here is the identification of various of these, seemingly very different, manifolds via mirror symmetry.

To characterize manifolds one needs to know about their topological structure. To pick out this structure one looks for the invariants, of which there are various kinds. For example, a real 2-dimensional manifold is specified by its genus. In string theory, the topological and complex structure of the compact manifold determines the low energy physics in the real, four non-compact dimensions. What was required by the string theorists, in order to consistent the observed particle physics, was a Calabi-Yau space with an Euler characteristic χ of ± 6. These can be found (and were found by Yau himself). However, there is an entire family of 'mirror' Calabi-Yau spaces with opposite Euler number. These look distinct from a topological and complex structure perspective, but from the point of view of the string theory (or, more precisely, the 2D conformal field theory) living on these spaces, the difference is merely apparent: the field theory is insensitive to the mirror mapping and is, in this sense, background independent.

4.3.1 The Hodge Diamond

The concept of the Hodge diamond makes the phenomenon of mirror symmetry easy to see in a visual way, and was in fact discovered and named as

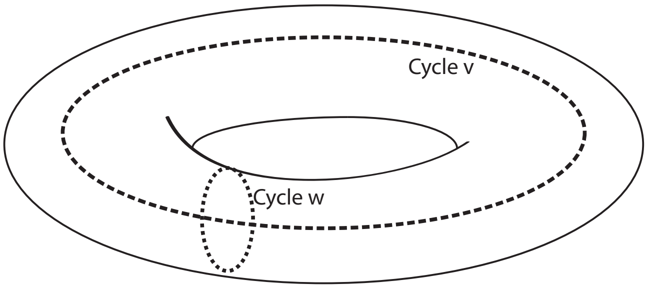


Fig. 5. The two independent cycles of a torus.

a result of this visual appearance. Hodge numbers are to (complex) Kähler manifolds what Betti numbers are to real manifolds: they specify topological invariants of the manifold and correspond to the dimension of the relevant cohomology group. The Betti numbers count the number of irreducible n-cycles of some manifold—see fig.5.

The n-cycles themselves are defined as 'chains' without boundary, where chains are sums of (oriented) submanifolds of the manifold. So, for example, $H_{n=0}$, a 0-cycle is a 0-chain and and is simply a point—note, cycles are considered equivalent if they differ by a boundary; so, for example, for a *connected* manifold, all points are deemed equivalent. The Hodge numbers do the same, but for complex cycles p and their complex conjugates $\overline{p} = q$. Schematically:

DeRham Cohomology Group $H^n_D \Rightarrow$ Betti number $b_n = \dim(H^n_D)$
Dolbeault Cohomology Group $H^{p,q} \Rightarrow$ Hodge number $h^{p,q} = \dim(H^{p,q})$

The Betti number and the Hodge number are related (by the Hodge decomposition) as:

$$b_n = \sum_{p+q=n} h^{p,q} \tag{33}$$

The Hodge diamond encodes these various Hodge numbers as follows:

$$
\begin{array}{ccccccc}
 & & & h^{0,0} & & & \\
 & & h^{1,0} & & h^{0,1} & & \\
 & h^{2,0} & & h^{1,1} & & h^{0,2} & \\
h^{3,0} & & h^{1,2} & & h^{2,1} & & h^{0,3} \\
 & h^{1,3} & & h^{2,2} & & h^{3,1} & \\
 & & h^{2,3} & & h^{3,2} & & \\
 & & & h^{3,3} & & &
\end{array}
\tag{34}
$$

For a complex 3-dimensional manifold, we can compute the Hodge numbers via the Hodge decomposition, giving:

$$
\begin{aligned}
b_0 &= 1 \\
b_1 &= 0 \\
b_2 &= h^{1,1} \\
b_3 &= 2(1 + h^{2,1}) \\
b_4 &= h^{2,2} = h^{1,1} \\
b_5 &= 0 \\
b_6 &= 1
\end{aligned}
\tag{35}
$$

The only independent Hodge numbers of the 3-manifold (with non-vanishing Euler characteristic—see below) are $h^{1,1}$ (roughly describing, via a number of real parameters, the size, or radius, and shape of the manifold) and $h^{2,1}$ (roughly the number of complex parameters to describe the complex structures that can be defined on the manifold). The other numbers are set by various mathematical identities and properties: $h^{p,q} = h^{q,p}$ by complex conjugation; $h^{p,q} = h^{3-p,3-q}$ by Poincaré duality (giving us the identity $h^{1,1} = h^{2,2}$ above); and the condition of vanishing first Chern class sets up an isomorphism between $h^{0,p}$ and $h^{0,3-p}$. Hence, we have:

$$\begin{array}{ccccccc} & & & 1 & & & \\ & & 0 & & 0 & & \\ & 0 & & h^{1,1} & & 0 & \\ 1 & & h^{2,1} & & h^{2,1} & & 1 \\ & 0 & & h^{1,1} & & 0 & \\ & & 0 & & 0 & & \\ & & & 1 & & & \end{array} \tag{36}$$

Since the Euler number χ for a real manifold is computed via the Betti numbers as:

$$\chi = \sum_n (-1)^n b_n \tag{37}$$

The Euler characteristic for a complex Kähler manifold can be computed, again invoking Hodge decomposition, as:

$$\chi_k = \sum_{p,q} (-1)^{p+q} h^{p,q}. \tag{38}$$

This number is, as mentioned above, crucial in the mapping to real-world, low-energy physics.

4.3.2 The Mirror Principle

It is a claim of algebraic geometry, having its origin in string theory, that every space described by such a Hodge diamond has a mirror (with the axis of reflection lying along the diagonal). The phenomenon of mirror symmetry then refers to an isomorphism between pairs of conformal field theories (worldsheet string theories) defined on prima facie very distinct Calabi-Yau manifolds, differing even with respect to their topology. In this case the manifolds have their Hodge numbers switched as:

$$H^{p,q}(M) \xleftrightarrow{isomorphic} H^{n-p,q}(\tilde{M}) \tag{39}$$

Where n is the (complex) dimension of the manifold. In the case where this is 3, we find that the remaining Hodge numbers $h^{1,1}$ and $h^{2,1}$ are isomorphic. These numbers parametrize the size and shape of the compact space, along with its

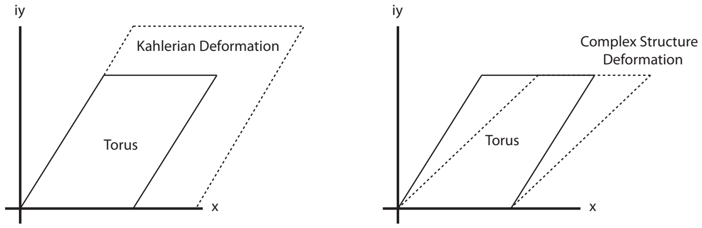


Fig. 6. The torus (with top and bottom and left and right identified) is an example of a 1-dimensional Calabi-Yau manifold. Deformations of the Kähler form of the torus change the volume while leaving the shape invariant (that is, the angles between the independent cycles are constant). A complex structure deformation does the opposite: it changes the shape (the angles) while leaving the volume invariant. (Adapted from Greene [16], p. 25.)

complex structural properties—see fig.6. [28] Mirror symmetry tells us that the physics (of relativistic *quantum* strings) is invariant when these, apparently very different (with different corresponding *classical* theories), features are exchanged. That is, there is quantum equivalence despite a marked difference at the classical level.

For example, the Euler character is equal to twice the number of particle generations. It can be connected to these shape and size parameters as follows:

$$\frac{|\chi|}{2} = |(h^{1,1} + h^{2,1})| = |(h^{1,1} - h^{2,1})| = \frac{|-\chi|}{2} = \text{No. Gen.} \tag{40}$$

To achieve a realistic string theory, then, one needs to find a Calabi-Yau manifold with $h^{1,1}$ and $h^{2,1}$ satisfying:

$$|(h^{1,1} + h^{2,1})| = 3 \tag{41}$$

Gang Tian and Shing-Tung Yau discovered such a manifold [44]. Though there is degeneracy here too, with multiple candidates available.

4.3.3 Using Mirrors to Count Curves

This setup was used to great (and surprising) effect to resolve a problem in pure mathematics, in the field of enumerative geometry. Briefly, Gromov-Witten invariants were used to calculate the number of curves of a given degree

[28] They correspond to topologically nontrivial 2-cycles and 3-cycles respectively.

of a particular surface.[29] Candelas *et al.* [6] developed a generating function to find the number of curves for all degrees n through a surface (a well-known Calabi-Yau manifold) known as a quintic, defined by the equation:

$$x_o^5 + x_1^5 + x_2^5 + x_3^5 + x_4^5 = 0 \supset \mathbb{P}^4 \tag{42}$$

The function they came up with was based on string perturbation theory (that is, a sum-over-Riemann-surfaces approach):

$$K(q) = 5 + \sum_{d=1}^{\infty} n_d d^3 \frac{q^d}{1-q^d} \tag{43}$$

Mathematically, n_d is the number of rational curves of degree d, and $q = e^{2\pi it}$. In terms of the physics, n_d is the 'instanton number', pertaining to the quantum corrections.[30] Each curve of degree d adds $d^3 \frac{q^d}{1-q^d}$ to the Yukawa coupling. This gives the various intersection numbers as coefficients in the series:

$$K(q) = 5 + 2875 \frac{q}{1-q} + 609250 \cdot 2^3 \frac{q^2}{1-q^2} + 317206375 \cdot 3^3 \frac{q^3}{1-q^3} + \cdots \tag{44}$$

That is

$$\begin{aligned} n_1 &= 2875 \\ n_2 &= 609250 \\ n_3 &= 317206375 \\ &\vdots \end{aligned}$$

The $d = 1$ and $d = 2$ cases were already well-known. But $d = 3$ was under investigation. The string theoretic calculation turned out to be correct, giving strong evidence that the formula was giving the correct values.[31]

[29] Full and very readable accounts of mirror symmetry, including the application discussed in the subsection, can be found in: [8] and [21].

[30] In more rigorous accounts, n_d is taken to represent the Gromov-Witten invariants of the space. These, roughly, correspond to the structure that is left invariant under deformations of the complex structure (i.e. those infinitesimal deformations parametrized by the cohomology group $H^{2,1}$).

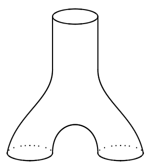
[31] I discuss the methodological ramifications of this scenario (*vis-à-vis* the concept of *evidence* for string theory) in [33]. Peter Galison has a related, though more

The application of the duality (mirror symmetry) here amounts to the simulation of the difficult *quantum* corrections (with yield the desired intersection numbers as instanton corrections) using aspects of the *classical* geometry in the dual theory. As Vafa explains, using the concepts introduced in §3:

> [W]hat happens is that a parameter which controls quantum corrections λ_0 on one side gets transformed to a parameter $\tilde{\lambda}_k$ with $k \neq 0$ describing some classical aspect of the dual side. This in particular implies that quantum corrections on one side have the interpretation on the dual side as to how correlations vary with some classical concept such as geometry. [[38], p. 540]

In the case of the string theoretic enumerative geometry, what is going on is that the Yukawa coupling (here, the 3-point vertex function or correlation function) is giving the count of the curves. This function contains both a classical (easy) piece and a quantum corrected (hard) piece. Following the prescription sketched by Vafa above, one can compute the quantum part using elements of the classical geometry and then convert back.

The mirror theories are equivalent for $n-$point functions, not just 3-point functions, since all string diagrams can be constructed from the basic 'pair of pants' vertex:



Given the remarkable nature of this application of the duality conjecture, one might not unreasonably view the positive results as offering evidence for the correctness of the duality. [32]

I think it is fair to say that mirror symmetry is the most philosophically interesting part of string theory *vis-à-vis* spacetime ontology. It is tantamount to the claim that very different manifolds are physically equivalent. Not only that, it involves an isomorphism between classical geometry and quantum phenomena. We unpack the meaning of such isomophisms in the final sections.

historical article, covering similar themes: [14].

[32] The formula of Candelas *et al.* was in fact made more rigorous using a variety of techniques external to string theory, see [26] for example. The various *proofs* of mirror symmetry can be found in [21].

5 Duality and Physical Content

Theories related by dualities can appear very different while making exactly the same predictions about *all* observable phenomena. Indeed, the theories can look sufficiently different that would-be interpreters of the theories would surely consider them to be representations of very different possible worlds indeed. Different elements and relations: one quantum, one not; one gravitational, one not; one 5-dimensional, one 10-dimensional; one large, one small, and so on. Yet duality symmetries may hold the key to extracting physical content from string theories. But what could this physical content possibly be, given these dualities?

Strongly coupled theories can have as 'duals' weakly coupled theories. This physical equivalence can then be used as an exploratory resource to enable the probing of the, practically unsolvable, strongly coupled theory. Thus, much as the method of simulation opened up new ways of tackling otherwise intractable physical system, so duality symmetries offer a similar possibility. Indeed, they appear to be very similar in terms of 'representational style', since a simpler system enables one to describe a more complicated system. The difference is, however, that the duality-based 'simulation' is 'non-lossy'. It is not that one description approximates another (say by abstracting some details away). Rather, there is a curious exact congruence in their physical predictions despite an often extreme superficial incongruity. However, though this aspect of duality is important and rather remarkable, it is the philosophical consequences of this situation that concern us more here. Exactly what do the existence of these dualities between apparently distinct theories tell us about these theories, about representation, and about interpretation?

5.1 Plain Vanilla Underdetermination?

Dual theories provide distinct but ultimately physically equivalent representations. Do they thereby amount to underdetermination? I would argue that there are crucial and subtle differences. The dual theories are not in competition: they are complementary. They are both true in a sense, and the practice of physics suggests, in many cases at least, a pluralistic stance with respect to the dual theories. However, in many cases (perhaps, ultimately, all) the cases point to a deeper underlying structure, and in that sense *neither* dual theory is true in anything but an approximate sense.

Duality symmetries appear to land us in familiar philosophical waters: we have, it seems, a multiplicity of representational schemes that, simply interpreted, support widely differing ontologies. A positivist might well be no more fazed by dual descriptions than standard underdetermined cases (such as Poincaré's example of the curved space versus distorting forces that have equivalent empirical geometrical consequences). But are dualities really just different ways of saying the same thing? In a sense, of course, they are, for duality is defined in terms of an isomorphism of physically meaningful consequences. In another sense they clearly aren't equivalent, for we often cannot use the dual theories in the same way to do the same things.

The syntactic view of scientific theories will clearly view the dual 'theories' as distinct *simpliciter*, with different basic axioms. However, on the semantic view of theories, matters are not so simple. As van Fraassen explains:

> The essential job of a scientific theory is to provide us with a family of models, to be used for the representation of physical phenomena. On the one hand, the theory defines its own subject matter—the kinds of system that realize the theory; on the other hand, empirical assertions have a single form: the phenomena can be represented by the models provided. [[39], p. 310]

In the case of string dualities, we have genuinely distinct topological manifolds that serve as background manifolds for the compact dimensions. These are structurally different. They would seem to be describing very different systems. In terms of models, they amount to different relational structures. Hence, even on van Fraassen's own 'state space' approach (with symmetries factored out), these will be distinct theories since the states will involve the backgrounds as part of their definition, not to mention a variety of different parameters (or moduli). The dualities point to a 'double (or n-fold) counting' in moduli space where we would not expect to see it. That is to say, dual theories occupy what appear to be very different regions of parameter space (or theory space). The duality identifies them on the grounds that they have a perfect matching with respect to their physical predictions.

Constructive empiricism would distinguish dual theories too, since it adopts a direct, literal interpretation of the formalism representing 'unobservable' things, properties and structures. Of course, this stance includes a freedom of choice component when it comes to empirically equivalent, empirically adequate alternatives. But recall that dual descriptions are not alternatives in the usual sense: they are indispensable modes of representing particular physical systems in different situations.

I don't wish to attempt to resolve the relationship between dualities and underdetermination here, but I do wish to emphasise that they *is* a very clear connection to be probed. What we have with dualities is something that does not quite fit the usual case studies presented as examples of underdetermination, and for that reason they should be of vital interest to those investigating underdetermination and issues of scientific realism more generally.

5.2 Theory, Duality, and Moduli Space

A central question of string theory, and one that has led to a great deal of controversy over string theory's scientific prowess and its ability to make sensible predictions, concerns *how many* string theories there in fact are. In a recent retrospective of string theory, David Gross, answering a question he had posed nearly two decades earlier, writes:

> *Do there exist more consistent string theories than the known five—the two forms of the closed superstring, the* $SO(32)$ *open superstring and the two forms of the heterotic string? Do there exist fewer in the sense that some of the above might be different manifestations (different vacua?) of the same theory?*
>
> This question has been definitively answered. All the five 'string theories' referred to above are different manifestations of one and the same theory. [...]
>
> The various forms of string theory are related by an intricate and beautiful web of dualities that relate one form of the theory, most appropriate for its description in a given region of parameter space, to another form, more appropriate for its description in a different region of parameter space. [[17], p. 102]

In other words, the various consistent superstring theories are distinguished points in the moduli space of vacua of some underlying theory. They represent *solutions* of the underlying theory—not forgetting the 11-dimensional theory, which amounts to a sixth solution. The web of dualities is taken to restore the uniqueness that was thought to characterise the earliest incarnation of string theory.

The problem with this 'solutions rather than theories' interpretation is that it is purely verbal as it stands, until the 'underlying theory' of which these 'theories' are solutions is found. If the underlying theory *is* found, with these limits, then the dualities will be converted into gauge symmetries of this new theory:[33]

> In practice, there is frequently a natural notion of what one should consider as a theory, so that any further analysis of the difference between these types of symmetries [gauge versus duality—DR] would be a rather academic

[33] However, in the cases discussed in this paper, relating solutions with very different topological spaces (as in mirror symmetry) so that they are taken to represent 'one and the same physical situation' is highly non-trivial in interpretive terms.

exercise. On the other hand, it can be a major breakthrough in science to discover a symmetry between a priori different theories and then promote it to a symmetry of the first type [i.e. relating states *within* a theory—DR]. Such a paradigm change is currently advocated in supersymmetric gauge theories and superstring theory with extended supersymmetry. ([13], p. 14)

As mentioned earlier, there is a relationship here with the Wilsonian conception of a renormalization group flow on 'theory space' (in the context of quantum field theory this is really just space of Lagrangians). Here the idea is that the energy scale at which some physical theory is studied and applied matters. In some theories one finds that the physics at one energy (or, inversely, distance) scale is dynamically decoupled (save for a few global parameters) from the scales below. [34] The nature of this dynamical insensitivity between various energy scales is precisely described by Wilson's theory. What it describes is the way in which the coupling constants and masses of some theory (i.e. in the Lagrangian) must be varied as the energy scale is varied so as to keep the values of known observables fixed. As David Gross puts it, "If there was no decoupling, it would be necessary for Newton to know string theory [valid at 10^{-33}—DR] to describe the motion of a viscous fluid [valid at 1cm—DR]" ([18], p. 553).

However, there is clearly a crucial difference with dualities. The renormalization group is really only a semi-group: it is irreversible. So one cannot work out the details of lower scales from the physics at higher scales: micro-information is lost at higher scales. In the case of dualities one is seemingly able to evade this informational embargo, studying strongly coupled, high energy, small scale situations using weakly coupled theories (and vice versa).

However, without some clear definition of 'theory' on the table, the claim that $\mathcal{M}$-theory is a unique theory with various limits (satisfying the various dualities) looks like a case of enforcing uniqueness by fiat. We also need an argument telling us that when a pair of theories are related by a duality transformation, they can be viewed as merely describing different regions of parameter space of an underlying theory (rather than those different regions of parameter space themselves amounting to distinct theories). Though this seems intuitive and appealing, I have seen no definite argument showing why it should be so—one suspects that the lure of 'unification' is playing some role in the claim that the various string theories really amount to some freedom in the way one describes a deeper (unique) theory. There is, in any case, clearly

[34] This is basically what renormalizability amounts to in the modern theory. It is simply the fact that the details of the physics at scales lower than some scale of interest can be bundled up into a finite number of parameters. If we can then measure and understand these parameters we can understand the theory at all scales. Non-renormalizable theories require infinitely many such parameters who's measurement would clearly be impossible.

good work for philosophers of science in clarifying what notion of theory is in operation here, and how it stands up to the various debates in philosophy of science over how theories ought to be viewed.

5.3 Duality Contra *Fundamentalism*

There are many surprising ontological consequences that flow from duality. For example, as we saw earlier, the notion of 'fundamentality' is impacted in a serious way by S-dualities exchanging strong and weakly coupled theories: there is no distinction between elementary (or what would standardly be labelled 'fundamental') and composite descriptions. That is, one can have a pair of representations, one in which some phenomenon is described by composite particles and another with 'elementary' entities. This is easiest (and most surprising, given quarks' supposedly elementary nature) to see in the case of quarks—though it holds for other dualities, including the electromagnetic duality. Quarks have a generalised notion of charge, known as colour, and the motion of such charges generates 'colour magnetic fields'. Quarks are able to combine to form a composite monopole with its own charge (i.e. colour). S-duality exchanges strong and weak coupling, however, which enables us to view the monopole as a non-composite object that itself forms quarks (which are then viewed as composite objects themselves).

If we apply our 'duality rule' to this case, then it seems even more puzzling than the geometrical cases. What structure could possibly underlie this duality, expressing as it does an equivalence between so seemingly different a pair of descriptions as 'fundamental' and 'composite'? Very similar aspects of duality involving quantum black hole physics (connected to the so-called holographic principle) have led 't Hooft to conjecture that this particular duality points to some deeper structure beneath quantum mechanics so that the quantum fluctuations are really statistical fluctuations in an underlying deterministic system who's degrees of freedom are neither field nor particles and who's picture of space and time is revised.

Ultimately, of course, duality isn't such a good case study for those who wish to deny fundamentalism, since it points to *deeper* structures that 'might' have fundamental status. Certainly, dualities are not inconsistent with the fundamentalist world picture.

Though I will not defend the claim here, it appears that the consequences of duality may be the structural realists best 'physics-motivated case' for their position. For the underdetermination is a part of our best physics and it is of a form where the dual descriptions are complementary, rather than in competition (that is, they are both thought to be true in some sense). [35] The differences in the representations (involving 'composite/elementary'-duality, dualities between distinct topologies, and so on) are significant enough to cause genuine trouble for non-structural realists. It seems that both descriptions in cases of duality are 'successful,' and yet they describe very different possible worlds at the level of an ontology of individual objects and properties (such as fields and field values, spacetime points and regions, and so). The fact that the dualities have been used to discover genuinely new and unexpected physics are enough to pose a problem for anti-realists who will need to provide an explanation for how this is possible.

Hence, the typical structural realist route to evading underdetermination involves the commitment to those structural aspects that are common to both of the pair of underdetermined descriptions. Here, we don't have underdetermination as such, but the same strategy is applicable: the dual pictures are dual in precisely the sense that they share their observable structure (they match up with respect to all physically measurable properties). As such , they seem like ready-made exemplars for the structural realist position.

A potential problem that would need to be considered, however, is that the underdetermination is itself structural (with different topologies and such like). This would seem to put the structural realist in as much trouble as standard realists. However, the resulting 'underlying' structure that the dual pictures point to are of the sort that provide more grist to the structural realists mill: they are of a distinctly non-local, relational nature since precisely what is at stake in the dual pictures are distinct individualistic or local elements. Hence, the idea of duality is one that the structural realist can accommodate quite naturally. Dual pictures amount to a (highly constrained, non-trivial) multiplicity in the ways that one can realise some system of physically observable (in the physicist's sense) relations.

6 Conclusion

Dualities are at the root of many difficult and profound debates in contemporary theoretical physics. They are essential to a proper understanding of nonperturbative physics and by implication to a proper understanding of our best physical theories, including the quantum gauge field theories comprising the standard model and string theory. They are connected to a plethora of problems that philosophers can profitably and constructively engage with and, in addition to posing *new* questions, have the potential to reinvigorate many old issues in the philosophy of physics and science. This paper merely

[35] I discuss the possible support that dualities offer to structural realism in [32].

skimmed the surface of a handful of these[36] in the hope that they will spark some interesting new lines of research.

References

[1] Ballman, W. (2006), *Lectures on Kähler Manifolds*. ESI Lectures in Mathematics and Physics. American Mathematical Society.

[2] Baez, J. C. and J. P. Munian (1994), *Gauge Fields, Knots and Gravity*. World Scientific.

[3] Belot, G. (2002). Symmetry and Gauge Freedom. *Studies In History and Philosophy of Modern Physics* **34**(2): 189—225.

[4] Brading, K. and E. Castellani (*eds.*), *Symmetries in Physics: Philosophical Reflections*. Cambridge University Press, 2003.

[5] Candelas, P., G. Horowitz, A. Strominger, and E. Witten. (1985), Vacuum Configurations for Superstrings. *Nuclear Physics B* **258** 46–74.

[6] ——————-, X C. de la Ossa, P. S. Green, and L. Parkes. (1991). A Pair of Calabi-Yau Manifolds as an Exactly Soluble Superconformal Theory. *Physics Letters B* **258**:118–126.

[7] Castellani, E. (2002). Reductionism, Emergence, and Effective Field Theories. *Studies In History and Philosophy of Modern Physics* **33**(2):251–267.

[8] Cox, D. A. and S. Katz (1999). *Mirror Symmetry and Algebraic Geometry*. American Mathematical Society.

[9] Dawid, R. (2006), Underdetermination and Theory Succession from the Perspective of String Theory. *Philosophy of Science* **73**: 298—322.

[10] Dirac, P. A. M. (1983), The Origin of Quantum Field Theory. In L. Brown and L. Hoddeson (eds.), *The Birth of Particle Physics* (pp. 39–55). Cambridge University Press, 1983.

[11] ——————————- (1948). The Theory of Magnetic Poles. *Phys Rev* **74** 817–830.

[36] One serious omission, briefly mentioned in §4.3.2, concerns the *testing* of duality conjectures. This is a highly nontrivial task (on account of the strong coupling aspect) involving supersymmetry and BPS states (these states are known to be stable under the renormalization group flow). Roughly, one utilises certain special situations that enable one to determine the whole series from the leading term in the Taylor series expansion of some physical quantity. One can then perform the computation in a pair of theories that are conjectured to be dual to see if they do indeed match up. The tests depend on whether these special cases can indeed be extrapolated from the special BPS states to *all* physical quantities of interest.

[12] ———————- (1927). The Quantum Theory of Dispersion. *Proceedings of the Royal Society of London. Series A* **114**(769): 710-728

[13] Fuchs, J. and C. Schweigert (1997), *Symmetries, Lie Algebras and Representations: A Graduate Course for Physicists.* Cambridge University Press.

[14] Galison, P. (1999) Mirror Symmetry: Persons, Values, and Objects. In M. Norton Wise *et al.* (ed.), *Growing explanations: Historical Perspectives on Recent Science* (pp. 23–61). Duke University Press.

[15] Gasperini, M. and J. Maharana (eds.). (2008) *String Theory and Fundamental Interactions. Gabriele Veneziano and Theoretical Physics: Historical and Contemporary Perspectives.* Springer.

[16] Greene, B. (1999) Aspects of Quantum Geometry. In D. H. Phong, L. Vinet, and S-T Yau (eds.), *Mirror Symmetry III* (pp. 1–67). American Mathematical Society.

[17] Gross, D. (2005) Where do we Stand in Fundamental (String) Theory? *Physica Scripta* **T117** 102–105.

[18] Gross, D. (1999) Renormalization Groups. In P. Deligne *et al.* (eds.), *Quantum Fields and Strings: Course for Mathematicians. Volume 1* (pp. 551–593). American Mathematical Society.

[19] Hartmann, S. (2001). Effective Field Theories, Reductionism and Scientific Explanation. *Studies In History and Philosophy of Modern Physics* **32**(2): 267–304.

[20] Healey, R. (2007). *Gauging What's Real: The Conceptual Foundations of Contemporary Gauge Theories.* Oxford University Press.

[21] Hori, K., S. Katz, A. Klemm, C. Vafa, R. Vakil, and E. Zaslow (2003). *Mirror Symmetry.* American Mathematical Society.

[22] Huggett, N. (2002). Renormalization and the Disunity of Science. In M. Kuhlmann, H. Lyre, and A. Wayne (eds.), *Ontological Aspects of Quantum Field Theory* (pp. 255277). New Jersey: World Scientific.

[23] Huggett, N. and R. Weingard (1995) The Renormalization Group and Effective Field Theories. *Synthese* **102**: 171–194.

[24] Jacob, M. (*ed.*). (1974), *Dual Theory.* Physics Reports Reprints Book Series, Volume 1. North Holland: Elsevier.

[25] Leeds, S. (1999). Gauges: Aharonov, Bohm, Yang, Healey. *Philosophy of Science* **66**(4): 606-627.

[26] Lian, B. H. and K. Liu, and S-T Yau. (2000). The Candelas-de la Ossa-Green-Parkes Formula. *Nuclear Physics B. Proceedings Supplements* **67**(1-3): 106–114.

[27] Lyre, H. (2008). Does the Higgs Mechanism Exist? *International Studies in the Philosophy of Science* **22**(2): 119–133.

[28] Maldacena, J. (1998). The Large N Limit of Superconformal Field Theories and Supergravity. *Advances in Theoretical and Mathematical Physics* **2**: 231–252.

[29] Morrow, J. and K. Kodaira. (1971), *Complex Manifolds.* American Mathematical Society.

[30] Olive, D. I. (1999), Introduction to Duality. In D. I. Olive & P. C. West (*eds.*), *Duality and Supersymmetric Theories.* Cambridge University Press, 1999: 62–94.

[31] Redhead, M. (2003), The Interpretation of Gauge Symmetry. In K. Brading and E. Castellani (*eds.*), *Symmetries in Physics: Philosophical Reflections* (pp. 124–139). Cambridge University Press, 2003.

[32] Rickles, D. (forthcoming). Duality, Underdetermination, and Structuralism. In E. Landry and D. Rickles (eds.), *Structure, Object, and Causality.*

[33] Rickles, D. (to appear). Leviathan and the Air Pump *Redux*: Miracles, Mirror Symmetry, and String Theory.

[34] Rubinstein, H. (2008). The Birth of the Veneziano Model and String Theory. In M. Gasperini and J. Maharana (eds.), *String Theory and Fundamental Interactions. Gabriele Veneziano and Theoretical Physics: Historical and Contemporary Perspectives* (pp. 47–58). Springer, 2008.

[35] Smeenk, C. (2006). The Elusive Higgs Mechanism. *Philosophy of Science* **73**: 487–499.

[36] Taylor, C. C. (1988). String Theory, Quantum Gravity, and Locality. In Fine & Lepin (*eds.*), *Proceedings of the Philosophy of Science Association 1988: Volume 2*, 1989: 107–111. Philosophy of Science Association.

[37] Tignol, J-P. (2001), *Galois' Theory of Algebraic Equations.* World Scientific.

[38] Vafa, C. (1998). Geometric Physics. In G. Fischer and U. Rehmann (*eds.*), *Proceedings of the International Congress of Mathematics. Volume 1* (pp. 537–556). Documenta Mathematica J. DMV.

[39] van Fraassen, B. (1972). A Formal Approach to the Philosophy of Science. In R. Colodny (*ed.*), *Paradigms and Paradoxes: The Philosophical Challenge of the Quantum Domain.* Pittsburgh University Press, 1972: 303–366.

[40] Weingard, R. (1988). A Philosopher Looks at String Theory. In Fine & Lepin (*eds.*), *Proceedings of the Philosophy of Science Association 1988: Volume 2*, 1989: 95-106. Philosophy of Science Association. Reprinted in *Physics Meets Philosophy at the Planck Scale*, C. Callender & N. Huggett (*eds.*): 138-151.

[41] Wigner, E. (1964). Symmetry and Conservation Laws. Reprinted in E. Wigner (*ed.*), *Symmetries and Reflections: Scientific Essays.* University of Indiana Press, 1967: 14–27.

[42] Witten, E. (1997). Duality, Spacetime and Quantum Mechanics. *Physics Today* May 1997: 28–33.

[43] Witten, E. (1996). Reflections on the Fate of Spacetime. *Physics Today* April 1996: 24–30.

[44] Yau, S-T. (1985). Compact Three Dimensional Kähler Manifolds with Zero Ricci Curvature. In W. A. Bardeen and A. White (eds.), *Proceedings of the Symposium on Anomalies, Geometry and Topology: Argonne* (pp. 395–406). World Scientific.