Studies in History and Philosophy of Modern Physics 59 (2017) 6–20

Contents lists available at ScienceDirect

Studies in History and Philosophy of Modern Physics

journal homepage: www.elsevier.com/locate/shpsb

Dualities of fields and strings

Joseph Polchinski [a,b,*]

[a] Department of Physics, University of California, Santa Barbara, CA 93106, USA
[b] Kavli Institute for Theoretical Physics, University of California, Santa Barbara, CA 93106-4030, USA

ARTICLE INFO

Article history:
Received 26 December 2014
Received in revised form
24 August 2015
Accepted 29 August 2015
Available online 1 October 2015

Keywords:
Quantum field theory
String theory
Duality

ABSTRACT

Duality, the equivalence between seemingly distinct quantum systems, is a curious property that has been known for at least three quarters of a century. In the past two decades it has played a central role in mapping out the structure of theoretical physics. I discuss the unexpected connections that have been revealed among quantum field theories and string theories. Written for a special issue of Studies in History and Philosophy of Modern Physics.

When citing this paper, please use the full journal title *Studies in History and Philosophy of Modern Physics*

1. Introduction

Perturbation theory is a central part of the education of a physicist. One first learns the basic solvable systems, most notably the simple harmonic oscillator. One then learns how to approach general problems in which the Hamiltonian is of the form

$$H = H_0 + gH_1, \tag{1.1}$$

where H_0 is solvable and the parameter g is small. These approximation schemes give physical quantities as a perturbation series,

$$\mathcal{A} = \mathcal{A}_0 + g\mathcal{A}_1 + g^2\mathcal{A}_2 + \cdots . \tag{1.2}$$

Here $\mathcal{A}$ might be an energy level, a scattering amplitude, or any other quantity of interest.[1] In particular, a major focus of the standard quantum field theory (QFT) course is the development of the series (1.2) in terms of Feynman graphs.

The series (1.2) is known not to converge in most systems of interest, in particular quantum field theories (Dyson, 1952). Nevertheless, it is valuable as an asymptotic series, meaning that for g sufficiently small a few terms give accurate results. For example, in quantum electrodynamics, where the effective expansion parameter is $\alpha/2\pi \sim 10^{-3}$, this has allowed the magnetic moment of the electron to be calculated to one part in 10^{12}. However, as g increases, perturbation theory becomes increasingly inaccurate, and it can completely miss important qualitative effects. In the Standard Model, quark confinement is the most notable example of such a nonperturbative effect, but others include the spontaneous breaking of chiral symmetry by condensation of quarks, and the violation of baryon and lepton numbers by instantons and skymions in the weak interaction.

There are no general methods for studying QFT's at large g. In principle, the nonperturbative definition of QFT by means of the path integral plus the renormalization group, as given by Wilson (1983), implies that any physical quantity can be calculated on a large enough computer. In practice, the theories and observables for which this can be done are limited. Another tool in the study of QFT's is the limit of a large number of fields (Stanley, 1968; 't Hooft, 1974a; Wilson, 1973). Here the graphical expansion simplifies and in some cases can be summed, giving a description of physical phenomena that cannot be seen in the individual terms of the series. This is most successful for theories where the many fields organize into a vector ϕ_i. For matrix fields ϕ_{ij}, including the important case of Yang–Mills fields, the graphical expansion simplifies enough to allow interesting general conclusions, but usually there are still too many graphs to sum explicitly.

A new tool, which has risen to prominence in the last two decades, is weak/strong duality, also known as *S*-duality. In some cases it is possible to decompose the Hamiltonian in multiple

* Correspondence address: Department of Physics, University of California, Santa Barbara, CA 93106, USA.
E-mail address: joep@kitp.ucsb.edu

[1] To avoid confusion, it should be noted that in some cases only even powers of g appear, depending on the structure of the Hamiltonian. Also, the series for a given physical quantity may begin with a term of order g^m with nonzero m, rather than $m=0$ as is written here for simplicity.

http://dx.doi.org/10.1016/j.shpsb.2015.08.011
1355-2198/

ways,

$$H = H_0 + gH_1 = H'_0 + g'H'_1. \tag{1.3}$$

Now one has two perturbative expansions. The original H_0 will have a simple expression in terms of some fields ϕ, while H'_0 will have a simple expression in terms of some new set of fields ϕ'. The ϕ' are related to ϕ in a complicated and usually nonlocal way; we will see examples in Section 2. Typically the couplings g and g' have a relation of a form something like

$$g' = 1/g. \tag{1.4}$$

If this is so, then as g becomes very large, g' becomes very small, and the perturbation series in g' becomes an accurate description of the system just where the series in g becomes useless. Of course, if $g \approx 1 \approx g'$ then neither expansion gives a good *quantitative* description, but having an understanding of the two limits gives a powerful global picture of the physics. Phenomena that are complicated in one description are often simple in the other. In many interesting systems there are multiple coupling constants, and multiple dual representations (1.3). In Section 2.5 we will give an example where there are two coupling constants and an *infinite* number of dual descriptions.

There is another, perhaps deeper, way to think about the duality (1.3). In quantum field theories, the expansion in g is essentially the same as the expansion in $\hbar$. To see this, consider the Yang–Mills theory, whose field strength is

$$\hat{F}_{\mu\nu} = \partial_\mu \hat{A}_\nu - \partial_\nu \hat{A}_\mu + g[\hat{A}_\mu, \hat{A}_\nu]. \tag{1.5}$$

The Yang–Mills connection $\hat{A}_\mu$ is written here as a matrix. It is useful to work with a rescaled field $A_\mu = g\hat{A}_\mu$, so that

$$F_{\mu\nu} \equiv (\partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu]) = g\hat{F}_{\mu\nu}. \tag{1.6}$$

The Yang–Mills action is then

$$S_{\rm YM} = -\frac{1}{2}\int d^4x\, {\rm Tr}(\hat{F}_{\mu\nu}\hat{F}^{\mu\nu}) = -\frac{1}{2g^2}\int d^4x\, {\rm Tr}(F_{\mu\nu}F^{\mu\nu}), \tag{1.7}$$

where the matrix trace makes the action gauge-invariant. In the latter form the coupling g appears only as an overall factor in the action. Quantum amplitudes are obtained from the path integral

$$\int \mathcal{D}A\, e^{iS_{\rm YM}/\hbar} = \int \mathcal{D}A\, e^{i\int d^4x\, {\rm Tr}(F_{\mu\nu}F^{\mu\nu})/2g^2\hbar}. \tag{1.8}$$

Note that the parameters g and $\hbar$ appear only in the combination $g^2\hbar$, so that the perturbation series for typical observables is

$$\mathcal{A} = \mathcal{A}_0 + (g^2\hbar)\mathcal{A}_2 + (g^2\hbar)^2\mathcal{A}_4 + \cdots. \tag{1.9}$$

It follows that the small-g and small$-\hbar$ limits are the same: weak coupling corresponds to the classical field limit. When $g^2\hbar$ is small, the exponent in the path integral (1.8) is large and so the integral is highly peaked on configurations where $S_{\rm YM}$ is stationary; these are the solutions to the classical equations of motion. When $g^2\hbar$ is large, the path integral is not very peaked and the quantum fluctuations are large. However, when the duality (1.3) holds, we can change to the primed description, and now the expansion parameter is $g'^2\hbar = \hbar/g^2$, giving a highly peaked action. Essentially what is happening is that in the original description the fields ϕ have wild quantum fluctuations at large g, but we can find new fields ϕ' which behave classically. This is a bit like a Fourier transform, where a function that is narrow in x is wide in p, and vice versa; we will make this analogy more precise in Section 2.3. (Having made this point, we will now revert to the quantum field theorist's conventions $\hbar = c = 1$.)

The interpretation of a duality is then that we have a single quantum system that has two classical limits. Quantum mechanics is sometimes presented as a naive one-to-one correspondence between classical and quantum theories. In this view we quantize the classical theory to go in one direction, and take the classical limit to go in other. Of course there are exceptions; for example, a classical gauge theory with anomalies cannot be consistently quantized. But with dualities, a single quantum theory may have two or more classical limits. 'Quantizing' any of these produces the same quantum theory.

The original wave-particle duality already exemplified this idea: given a QFT, one can take two different classical limits depending on what one holds fixed. One limit gives classical fields, the other classical particles. It is fruitless to argue whether the fundamental entities are particles or fields. The fundamental description (at least to the extent that we now understand) is a QFT. Similarly it is fruitless to argue whether ϕ or ϕ' provide the fundamental description of the world; rather, it is the full quantum theory.

With the dualities (1.3), the functional forms of H_0 and H_1 in terms of ϕ may be the same as those of H'_0 and H'_1 in terms of ϕ'. In this case we would say that the theory is self-dual. Alternatively, the functional forms and even the nature of the fields may be quite different: in this case we have two very different ways to think about the system. The term *S*-duality is applied in either case; in some cases self-duality may be implied by the context.

It is not clear why the structure of theoretical physics is so kind to us, in providing simple description in many limits that would seem to be very complicated. It may be that there are fewer consistent quantum theories than classical ones, so that we necessarily get to the same quantum theory from multiple classical starting points.

In Section 2 we discuss dualities where both descriptions are QFT's. We begin with some classic examples, namely the Ising model, bosonization, and free electromagnetism, where the duality can be constructed rather explicitly. We then move on to richer examples, in particular supersymmetric Yang–Mills theories. For these, the duality is not proven but inferred. We discuss the evidence and the logic that supports the existence of these dualities. We also discuss the role of supersymmetry.

In Section 3 we discuss dualities between string theories. We begin with *T*-duality, which connects two weakly coupled string theories and can be demonstrated rather explicitly. It illustrates a number of remarkable features of string theory: that space is not fundamental but emergent, and that strings perceive space–time geometry in a rather different way from pointlike particles and fields. We then discuss weak/strong dualities in string theory, and the significance of branes. A notable conclusion is that there is only a single quantum theory in the end: what appear to be different string theories are different classical limits of a single quantum theory, whose full form is not yet known. The same analysis reveals the existence of new classical limits, which are not string theories at all.

In Section 4 we discuss dualities in which one description is a QFT and the other a string theory. The existence of such dualities is remarkable, because QFT's are well-understood conceptually, while string theories include quantum gravity and so present many conceptual puzzles. In fact, field-string duality currently plays a key role in providing a precise definition of the quantized theory of strings and gravity. We describe how two puzzles, the black hole entropy and the black hole information paradox, have been clarified by dualities, although important questions remain open. We also discuss the holographic principle, in which the emergent nature of space–time is even more radical. We conclude by discussing various open questions.

Apology: this is a rather sweeping subject, and I certainly have not set out to reconstruct the entire history of its development. I have tried to choose references that will be useful to the intended audience.